Whole genome sequencing reveals Fagus sylvatica tree genetic heterogeneity between mono-and multi-specific plots in mixed forests

Stéphane Lobréaux, Marion Jourdan, Christian Miquel, Eric Bazin, Xavier X. Morin

Whole genome sequencing reveals *Fagus sylvatica* tree genetic heterogeneity between mono- and multi-specific plots in mixed forests

Stéphane Lobréaux[a*], Marion Jourdan[b,c], Christian Miquel[a], Eric Bazin[b] & Xavier Morin[2]

[a]Laboratoire d'Ecologie Alpine, UMR 5553 CNRS, Université Grenoble-Alpes, Grenoble, France. [b]Centre d'Ecologie Fonctionnelle et Evolutive, Université de Montpellier, UMR 5175 CNRS, EPHE, IRD, Université Paul Valéry Montpellier 3, Montpellier, France. [c]Present address: ESE, CNRS, AgroParisTech, Université Paris-Saclay, Orsay, France.

*Corresponding author : Stéphane Lobréaux

Email: stephane.lobreaux@univ-grenoble-alpes.fr

Abstract

Forests are ecosystems where naturally multiple tree species may coexist. Assessing possible tree behavior to ongoing climate change is key to forecast forest's future, and better understanding the links between tree community assembly and functioning is essential to this aim. Here we compared genomes of beech trees growing either in monospecific or multispecific plots, in different sites in the French Alps. We found a high genetic association at specific loci between SNP profiles and the monospecific or multispecific tree community composition. The studied mixed forests appear as genetic patchworks structured at the sub-population level. The most significant SNPs identified point out a key metabolic pathway, pathogen defense related genes and stress response. As mixed forests are likely to better promote biodiversity conservation and forest ecosystem functions provision and to show a stronger resilience to new climate conditions, our findings strengthens the need of understanding multispecific tree community assembly at the genetic level to forecast and improving forest adaptation to climate change.

1. Introduction

Forests are essential ecosystems harboring an important amount of species [1][2]. A positive correlation between tree species composition and biodiversity richness has been shown for several taxonomic groups: flora, fauna, microorganisms [3][4]. If biodiversity thus differs between multispecific (or mixed, ie, with at least two dominating tree species [5] at the plot level - between 10^3 and 10^4 m²) and monospecific forests, this is also the case for forest functioning - notably ecosystem productivity [1][6] [7]. In fact an increasing number of studies – using various approaches (observations, experiments, process-based simulations) - pointed out that multispecific forests or plantations showed a higher productivity [7][8][9][10] but also a stronger resilience and resistance to biotic and abiotic stress [6][11][12], although these patterns are not universal as several counter-examples have been highlighted [13] [14]. Species interactions that mostly drive these biodiversity effects, actually depend on many factors, such as species identity [15], environmental conditions and resource availability [16]. However, ongoing climate change strongly impacts tree physiology and therefore the outcome of intra- and interspecific interactions [17]. Therefore, although promoting mixed forests is more and more considered as a promising approach to sustain both biodiversity and ecosystem functioning under harsher

environmental conditions [18][19][20], knowledge remains scarce about how mixed forests may respond to new climatic conditions relatively to monospecific stands.

To better understand differences between mono- and multi-specific tree communities, most studies have focused on phenotypic aspects by relying on functional trait variations that are supposed to reflect plant response to local conditions [21][22]. These variations may be related to phenotypic plasticity [23][24] but also to genetic adaptation [25]. In this latter context, beneficial alleles are favored by local environmental (ie. abiotic and biotic) conditions leading to an enhanced fitness of the species [26]. Such a process has been mainly investigated in relation to abiotic environmental conditions so far [27]. Yet, it has been shown that genetic variations within a species may influence plant community composition, but also the composition of other trophic levels [28]. This is the case for example through chemical compounds for Scots pine [29] whose needle monoterpene composition affects the understorey vegetation composition. *Baccharis pilularis*, a coastal dune shrub, has a genetically-dependent architecture dimorphism (prostrate or erect morph) that has been shown to affect the understory environment leading to distinct plant community assembly [30]. Yet, beyond these few examples, the interplay between community composition, species coexistence and genetic structure of species remains mostly unknown. For instance, the effect of the biotic context on the genetic structure of a species remains poorly investigated, especially for interspecific competition, and so are the consequences on the local abiotic conditions. Advances in DNA sequencing techniques enable now to investigate such processes at the whole genome level.

In this work, we investigated the genetic association of European beech (*Fagus sylvatica*) trees with their local environmental conditions. This species is a major deciduous tree in Europe, one of the most dominant late-successional tree species, and is of key interest at the economic level. We aimed at testing whether the structuration of genetic diversity of European beech trees varied with the species composition of their local environment. To do so, we assessed the structuration of genetic diversity of beech trees in both monospecific beech and multispecific beech-oak (*Quercus pubescens*) stands *in situ* across several sites, relying on whole genome analysis.

2. Materials and Methods

2.1 Sampling

Sampling sites were located in three natural French forests: Luberon-Lagarde (Lagarde d'Apt), Grand Luberon (St Martin de Castillon), Ste Baume (Plan-d'Aups-Ste-Baume) [15] (SI Appendix, Fig. S1).). The sampled forest are natural regenerations with no plantations or tree selections from what we know about their history. The sites have been choosen with forest managers of the French National Forest Office for this purpose. The Lagarde and Grand Luberon forests have not been managed for more than 50 years, and as long as we know from the history of these forests. The Sainte Baume forest (protected with a national protection status from the French Forest National Office) has not been managed for more than 150 years (except cutting trees that were possibly dangerous for visitors). All sites have a limestone bedrock and a northwest exposition. In each site, two plots have been selected: a monospecific plot dominated by beech (*Fagus sylvatica* and a downy oak (*Quercus pubescens*) -beech multispecific plot with ca. 50% of each species in terms of total basal area (ie. cross-sectional area of

trees at breast height). The two kinds of plots have been sampled as close as possible from each other to limit topographic and climatic variations. Each plot corresponds to a 17.5m radius circle (ca 1000 m²). Four plots were sampled in the Ste Baume and Grand Luberon forests, and two in the Luberon-Lagarde forest, for a total of ten plots. For genomic analysis, leaves were harvested from three to five mature trees per plot and dehydrated in silica gel, in which were stored until DNA extraction.

2.2 DNA extraction and sequencing

25 milligrams of dried leaf samples were grounded and total DNA was extracted using DNeasy Plant Kit (Qiagen) according to the manufacturer instructions. DNA quality was checked using agarose gel and DNA quantification was performed on a Qubit fluorometer (Invitrogen, USA) with the Qubit dsDNA BR assay kit.

Genomic DNA libraries were prepared and sequenced by Fasteris (Geneva, Switzerland). After DNA breaking, bead size selection enabled recovering inserts of 350bp and libraries were prepared using Illumina Genomic Nano kit. Sequencing was performed for the 45 samples on a HiSeq4000 sequencer as paired-end reads of 150 bp to get an average 10x coverage for each sample.

2.3 SNP mapping

Raw reads were filtered using trimmomatic [31] to remove low quality sequences and length below 50 bases. Read mapping on *Fagus sylvatica* scaffolds [32] was performed using BWA 0.7.17 [33]. Samtools 1.9 and bcftools 1.9 enabled SNP calling and filtering [34]. Indels were removed together with SNPs located within a distance of 10 bases

around an indel. Low quality variants (qual<20) and low quality mapping (MQ<20) were discarded. Only biallelic sites were retained for which data were available for all populations, so the dataset do not contain missing data. A minor allele frequency filter was applied (5%). Allele counts from individuals were finally grouped for a plot to get plot population data.

2.4 SNP detection

A GLMM model with logit link and a binomial error distribution was estimated between each SNP of the complete dataset and the plot type [35]. The R package lme4 was used to calculate the models [36]. Using the allelic proportions of the alleles at a defined genomic position, the number of occurrences of each allele was calculated for a total count of 30, and these occurrence data were used as input for model determination. This strategy enabled to perform such calculations based on an equal total count of alleles at each genomic position. The plot composition was coded as 0 or 1 respectively for monospecific and multispecific plots. To evaluate model performances, the likelihood ratio (LR) was calculated using the ANOVA function from the R car package with the model and null model as input. A Bonferroni correction was applied to the LR p-values as previously described [35]. Wald1 and Wald2 coefficients were also used to evaluate model performances and were recovered from the GLMM model data produced [37]. As a control, we ran similar analysis on the complete dataset with three random variables in which the plot composition status was randomly assigned for each site..

The LFMM 1.5 program was used to confirm highly significant SNPs detected with the GLMM approach. It enables testing for the association of genomic loci with

environmental variables while taking into account the population structure through unobserved latent factors([38] http://membres-timc.imag.fr/Olivier.Francois/lfmm/index.htm). The number of latent factors was fixed to three with the K parameter. Genotyping data were provided as allelic proportions per population. LFMM implement an MCMC algorithm, the number of iterations in the Gibbs sampling algorithm was set to 30,000, and the number of burning iterations to 15,000. We performed 6 in- dependent runs, and the Z-scores were combined using the Fisher-Stouffer method. The p-values were adjusted as described by LFMM manual (http://membres-timc.imag.fr/Olivier.Francois/lfmm/index. htm).

2.5 Principal Component Analysis

Population analysis was performed using Principal Component Analysis (PCA) in R with the FactoMineR and factoextra packages [39]. A covariance matrix was first calculated based on data from a randomly selected set of 10,000 SNPs across the beech genome and PCA was then performed. Different random datasets of different sizes where produced and teste. We also tried to filtered these datasets from potentially linked SNPs according to the linkage disequilibrium and get the same result.

2.6 Tree growth analysis

To determine *Fagus sylvatica* growth between mono- and multispecific plots, we calculated the plot productivity for each plot. For this purpose, productivity was investigated through basal area increments (BAI), based on the sampling of all beech trees in a plot and growth dynamic was evaluated using tree rings over 18 years during

the period from 1995 to 2013 (for detailed explanation, see supplementary data) [15][40][41].

3. Results

3.1. Genomic sequencing

Studied plots were located in three forests in the south of France (Fig. S1). In each forest, monospecific and multispecific plots were located as close as possible to ensure similar environmental conditions [40], while climate conditions between the three forests differed (Fig. S2).

Genomic DNA from 50 beech trees collected in the three different forests have been sequenced and mapped onto *Fagus sylvatica* draft genome [32]. After filtering, 1,163,710 Single Nucleotide Polymorhisms (SNPs) were retained, corresponding to a density of 1 SNP/437 bp. These SNPs were located on 5217 scaffolds of the genome assembly, which contains a total of 6491 scaffolds (N50=145kb), and they cover 94% of the genome estimated size (542 Mb). In order to get population data from the samples, a PCA was performed on different random sets of SNPs leading to identical groups. The results (Fig. 1) indicated a structuration in three distinct groups, strongly linked to the three sampled forests. All three groups contain trees from both multispecific and monospecific stands, and no clustering was observed depending on the tree community composition from which the samples originated (Fig. 1).

3.2. Detection of SNP allele association

To investigate a potential genomic association with tree community composition, a search for the association of SNP alleles with the community type was performed. We took into account for these analyses the population structure that was identified by PCA in order to reduce false positive detection. Such population effects were also controlled by the presence of both multispecific and monospecific stands at each sampling site.

General Linear Mixed Models (GLMM) detection enabled to identify the most significant SNPs for their association with multispecific or monospecific beech communities. At a p-value threshold of 0.05 for the LR and the Wald2 test, we get 24,619 positive hits, the lower corrected p-value for the LR being $4.4x10^{-22}$. Among the dataset, 43 SNPs allelic profile led to p-values lower than 10^{-16}. Since a large number of mapped SNPs were available and therefore a high SNP density, we searched for the most significant regions associated to the stand status. For this purpose, a sliding 1.5kb window approach was performed and identified regions with at least two highly significant SNPs with p-values below 10^{-16} [35]. The window size was chosen according to the estimated LD in *Fagus sylvatica* [42], which was up to 1kb. We detected six regions (Table 1) located on different scaffolds of the genome assembly.

For the SNPs detected with the plot status (ie. monospecific or multispecific stand) variable into the region 119,197-120,737 on scaffold 635, a consistent result was obtained among the different sites with a similar behavior (e.g. Fig. 2 shows two of these results for SNPs 119,197 and 119,226). In monospecific beech plots, two alleles were detected at an average ratio of 0.5 (Fig. 2 and Fig. 3A). In contrast, in multispecific beech-oak plots, only one allele was present. Such data thus indicate a clear shift in allele proportion for these SNPs in the plots according to their composition, which was detected at an equal level in all studied sites. A similar behavior was observed for SNP

128,466 on scaffold 655, with one major allele in multispecific plots while in monospecific plots two alleles were present at an average ration of 0.5, but with a variability higher than for the region 119,197-120,737 of scaffold 635 (Fig. 3B). In contrast, among the SNPs significantly associated to the community type, a different situation was observed on scaffold 2 and 2088. For this second class of SNPs, only one allele was detected in trees in monospecific beech plots, while two alleles were present in multispecific plots (Fig. 3C and 3D). A clear and consistent community effect was therefore detected in the different forests, showing that allele proportions at specific SNPs strongly vary according to community composition at the plot level. As a control, we performed a similar association detection with three random variables where the plot status (ie. monospecific or multispecific stand), was assigned randomly in each site. We obtained between 17,895 and 18,200 positive hits at a threshold of 0.05 for Wald2 and LR p-values. The profiles of the SNPs with the lowest pvalues didn't reveal reproducible results among the different sites with these random variables (Fig. S3). If some low pvalues were in the range of those obtained with the true variables (e.g. Fig. S3 shows the SNP profile of one of them corresponding to a p-value of 9.83×10^{-22}), this was linked to the weight attributed by the model to a high difference in two sites only. Low differences are observed in the 3 other sites (Fig. S3). We obtained similar results with the different random variables.

3.3. Genomic annotations into the highly significant genomic regions

Using the beech genome draft annotations and re-annotations using *Arabidopsis thaliana* or oak data as references, we investigated the occurrence of genes in the

detected regions for their association to the tree community composition. In trees the linkage disequilibrium has been shown to be low and reaching up to 1kb in *Fagus sylvatica*. Seven of the nine significant SNPs mentioned in Table 1 were located into genes, the three SNPs on scaffold 635 being in the same KHG-KDPG aldolase gene. On scaffold 2 the detected region matches with genes encoding a wall associated receptor kinase and a sterol 3-beta-glucosyltransferase. Genes related to ion transport and plant development or defense against pathogens were pointed out by significant SNPs on scaffold 655 and 2068. Different class of genes are therefore located into beech genomic regions associated with the community status.

3.4. Fagus sylvatica growth analysis

Since our results highlight genetic differences between beech trees in monospecific *vs.* multispecific plots, we then questioned whether such differences affect the growth of beech trees in between monospecific and multispecific plots. For this purpose, we compared the growth patterns of beech trees in the two kinds of plots across 18 years, using basal area increment (ie. increment of cross-sectional area of trees at breast height) as a proxy for growth. In all sites, the measured growth was close between trees in multispecific *vs.* monospecific plots, as only small differences were observed (Fig. 4). Therefore, genotypes that are filtered by mono- or multispecific biotic environments showed similar growth patterns in their respective community composition conditions.

4. Discussion

In this work, using a genetic association approach at the whole genome level in beech trees, we observed a genetic differentiation at some specific SNPs between the monospecific and multispecific plots. This association was found reproducibly in all the studied sites, which shows the consistence of this association and strongly suggests that local environmental conditions into these plots act to filter the present genotypes of beech trees. A consistent result of genetic profiles for associated SNPs was observed in different forests with different beech populations and distinct climatic and soil environments. It is known that environmental parameters are involved in plant local adaptation, and studies of this process have actually a particular interest to understand the impact of climate parameters variations due to the ongoing climate change [27]. To explain the distinct genetic profiles observed according to the plot community composition, a hypothesis could be that micro-environmental variations act to favor some genotypes and lead to an allelic enrichment at some SNP positions. This phenomenon would be at a sub-population level since these observations were made in specific plots in the population and would be linked to micro-variations that we have been unable to characterize so far.

Indeed, the trees of the two species composing the study forests are not uniformly distributed in these forests. They occur in patches of various sizes and with various levels of species mixing. In this study, the plots have been chosen in the field to compare two extreme situations: monospecific plots dominated by a single species or

multispecific plots with ca. 50 % of each species [41]. These forests have not been managed for at least 50 years, the observed patterns are thus naturally occurring, which means that the species composition is the result of long-term filtering from local environmental conditions [43]. It is noteworthy that it is very unlikely that the observed heterogeneity between plots is the result of differential seed dispersion, as the paired monospecific and multispecific plots were very close to each other (50 m maximum). To explore putative causes of these short-distance variation, the climatic parameters of the different plots have been investigated using both climatic databases and local measurements. Soil data have been also collected and analyzed for different parameters. Using these environmental data, we did not find any significant environmental factors that would explain the forest heterogeneity in terms of community composition. We cannot exclude, however, that micro-environmental variation occur among the plots that were impossible to properly measure. In this context, some local conditions may either favor one of the two species, leading to a monospecific plot, or allowing the two species to coexist, resulting in a multispecific plot. Such micro-spatial environmental variations have been reported in specific conditions such as between normal and contaminated soils in mine boundaries [44]. Another example was reported between hilltops and bottomland plots in tropical forest [45] leading to divergence in quantitative traits of forest trees, but with topographical variations and in a range of distance larger than what we observed in our sampling sites (300m).

The genetic structuration that we have characterized is based on the comparison of mono- and multispecific plots. Therefore, apart from some potential micro environmental variations between the plots as stated above, another possible explanatory factor is the difference in biotic environments between plots. In

multispecific plots, the presence of oak trees as neighbors for beech trees may influence genetic filtering of alleles of beech trees. Beech and oak have different physiological characteristics, and thus tend to respond differently to competition and climate conditions, as already shown in these same forest plots [40]. When present side by side in a plot, tree-tree interaction or indirect effects on the environment of the neighboring tree may influence what tree genotypes are filtered. Tree-tree interactions have already been shown to affect growth patterns of individual trees [40] or their architecture [46]. Regarding indirect effects, beech trees are more sensitive to dry conditions than downy oak [41]. However, such changes have mostly been discussed in the light of plasticity [46][47] and long-term legacy of the biotic environment (plurispecific *vs.* monospecific interactions in this case) on the genetic structure of a tree species has never been shown, to our knowledge.

A third hypothesis could be also that the observed association of SNPs with the plot composition is linked to a combination of the two first hypotheses with both micro-environmental variations and biotic factors linked to the species of the tree neighbors.

The highly significant SNPs detected were mostly located in genes, and since LD is low in trees, we checked the function of these genes. Highly positive signals were obtained for the region 119,197-120,737 on scaffold 635. The three positive SNPs are located within a KHG KDGP aldolase gene and a clear shift in allelic proportion was measured when beech grow with downy oak as neighbor at the community level. This aldolase is the key enzyme of the Entner-Doudoroff pathway (ED) [48]. Common pathways to produce ATP and reducing power from glucose are the glycolysis (Embden-Meyerfod-Parnas pathway) and the oxidative pentose phosphate pathway.

The ED pathway is an alternate way to perform this process that uses the KHG KDGP aldolase to produce Glyceraldehyde-3-Phosphate from 2-keto3-deoxy6-phosphogluconate. This metabolic pathway has been described in prokaryotes and more recently in cyanobacteria and plants [48]. In plants, the ED pathway has been shown to be active in tobacco or *Hordeum vulgare* for example, but not all species tested. Analysis of the drought stress response in tobacco at the metabolite level showed a 70 fold increase in 4-hydroxy-2-oxoglutaric acid (KHG) in such condition and KHG aldolase was suggested to play a role in stress recovery as in *E. coli* [49].

A second set of candidate genes is known to be developmentally regulated and involved in pathogen response in other species. A first gene is a member of the cell wall-associated kinase family encoding a receptor-like protein kinase which is part of the pathogen response in *Arabidopsis thaliana* [50][51]. We also identified a Major Latex Protein (MLP) encoding gene [52]. In *Morus multicaulis,* a gene of this family protect against pathogens, but a role in drought resistance has also been reported in *Arabidopsis* [53]. Mixing tree communities have been shown as a way to improve tree protection against pest and pathogens [11][12][54]. The last two candidate genes are a cation chloride transporter and a sterol-3-beta-glucosyl-transferase. Both of them have been shown to take part in plant salt tolerance in other species [55][56]. The detected genes are therefore involved in stress responses. So far, genetic association studies for the trees of the *Fagaceae* family have been mainly focused on climate parameters or functional traits [27], and are supported by increasing genomic data becoming available for these species with the recent progresses in sequencing [57]. Some studies have highlighted the role of plant genotypes into the composition of ecological communities, affecting different trophic levels like for example insects, soil microbial communities, arthropods [58], but intraspecific diversity can also have some influence at the same

trophic level [28]. In this study focusing on the genomic level, we highlighted that tree community composition may influence species local genetic composition. These data provide candidate genes for such a process that may help improving our knowledge about tree interactions and may be valuable in the understanding of mixed forest assembly and functioning.

5. Conclusion

With ongoing climate change, warmer and drier conditions are expected to occur in Southern Europe. Therefore a better understanding local adaptation to environmental conditions (eg. drought stress) appears as a key stake for long-lived organisms such as trees [17]. If our findings should be confirmed in other conditions and for other species, they nevertheless strongly suggest that genetic issues should also be considered to better understand diversity effects on community ecology and ecosystem functioning, since the genetic pool may differ according to the tree community composition in natural forests. As mixed tree communities are viewed as a way to sustain biodiversity [59], and maintain forest ecosystem functioning, notably by promoting resistance and resilience against climate change [60], our results highlight how genetic effects may also contribute to creating differences in resistance and resilience to climate change as well as in ecosystem functioning between mixed and monospecific forests.

They may also question the use of tree plantations as key nature-based solutions against climate change. In fact, most of these plantations are monospecific, and our findings may thus call for a better consideration of genetic issues in the massive tree planting programs currently under development.

Competing Interest Statement

The authors declare no competing interest.

Data availability

DNA sequences are being deposited in SRA database.

Author Contribution :

Conceptualization: Eric Bazin, Xavier Morin, Stéphane Lobréaux

Technical assistance: Christian Miquel

Investigation: Marion Jourdan, Xavier Morin, Stéphane Lobréaux

Writing: Xavier Morin, Stéphane Lobréaux

Funding : This work was supported by a research grant from the Agence Nationale de la Recherche (APPATS_ANR_15_CE2_0004). The plot design used in this study is part of the GMAP plot network (https://oreme.org/observation/foret/gmap/), partly funded by the OSU OREME in Montpellier.

Acknowledgments

We dedicate this paper to Eric Bazin. We are grateful to Jonas Baudry, Lana Coste, Jean-Marc Ourcival, Soline Martin-Blangy and Sylvain Gérard for their help during field sampling.

Appendix A. Supplementary data

Supplementary data to this article can be found online at .

References

[1] Gamfeldt L, Snäll T, Bagchi R, Jonsson M, Gustafsson L, Kjellander P, et al, Higher levels of multiple ecosystem services are found in forests with more tree species, Nat. Commun. 4 (2013) 1340, doi:10.1038/ncomms2328

[2] The State of the World's Forests 2020, FAO and UNEP (2020) doi:10.4060/ca8642en

[3] Ampoorter E, Barbaro L, Jactel H, Baeten L, Boberg J, Carnol M, et al, Tree diversity is key for promoting the diversity and abundance of forest-associated taxa in Europe, Oikos 129 (2020) 133–146, doi:10.1111/oik.06290

[4] Fornoff F, Klein A-M, Blüthgen N, Staab M, Tree diversity increases robustness of multi-trophic interactions, Proc. R. Soc. B. 286 (2019) 20182399. doi:10.1098/rspb.2018.2399

[5] Bravo-Oviedo A, Pretzsch H, Ammer C, Andenmatten E, Barbati A, Barreiro S, et al, European Mixed Forests: definition and research perspectives, Forest Syst. 23 (2014) 518-533, doi:10.5424/fs/2014233-06256

[6] Pretzsch H, Schütze G, Uhl E, Resistance of European tree species to drought stress in mixed *versus* pure forests: evidence of stress release by inter-specific facilitation: Drought stress release by inter-specific facilitation, Plant Biology 15 (2013) 483–495, doi:10.1111/j.1438-8677.2012.00670.x

[7] Liang J, Crowther TW, Picard N, Wiser S, Zhou M, Alberti G, et al, Positive biodiversity-productivity relationship predominant in global forests, Science 354 (2016) aaf8957–aaf8957. doi:10.1126/science.aaf8957

[8] Paquette A, Messier C, The effect of biodiversity on tree productivity: from temperate to boreal forests: The effect of biodiversity on the productivity, Glob. Ecol. Biogeogr. 20 (2011) 170–180, doi:10.1111/j.1466-8238.2010.00592.x

[9] Zhang Y, Chen HYH, Reich PB, Forest productivity increases with evenness, species richness and trait variation: a global meta-analysis: *Diversity and productivity relationships*, J. Ecol. 100 (2012) 742–749, doi:10.1111/j.1365-2745.2011.01944.x

[10] Morin X, Fahse L, Scherer-Lorenzen M, Bugmann H, Tree species richness promotes productivity in temperate forests through strong complementarity between species: Species richness promotes forest productivity, Ecol. Lett. 14 (2011) 1211–1219, doi:10.1111/j.1461-0248.2011.01691.x

[11] Jactel H, Brockerhoff EG, Tree diversity reduces herbivory by forest insects, Ecol. Lett. 10 (2007) 835–848. doi:10.1111/j.1461-0248.2007.01073.x

[12] Guyot V, Castagneyrol B, Vialatte A, Deconchat M, Jactel H, Tree diversity reduces pest damage in mature forests across Europe, Biol. Lett. 12 (2016) 20151037. doi:10.1098/rsbl.2015.1037

[13] Vilà M, Vayreda J, Gracia C, Ibáñez JJ, Does tree diversity increase wood production in pine forests?, Oecologia 135 (2003) 299–303, doi:10.1007/s00442-003-1182-y

[14] Grossiord C, Granier A, Ratcliffe S, Bouriaud O, Bruelheide H, Checko E, et al, Tree diversity does not always improve resistance of forest ecosystems to drought, Proc. Natl.

Acad. Sci. USA 111 (2014) 14812–14815, doi:10.1073/pnas.1411970111

[15] Jourdan M, Lebourgeois F, Morin X, The effect of tree diversity on the resistance and recovery of forest stands in the French Alps may depend on species differences in hydraulic features. For. Ecol. Manage. 450 (2019) 117486. doi:10-1016/j.foreco-2019-117486.

[16] Jactel H, Gritti ES, Drössler L, Forrester DI, Mason WL, Morin X, et al, Positive biodiversity-productivity relationships in forests: climate matters, Biol. Lett. 14 (2018) doi:10.1098/rsbl.2017.0747

[17] Senf C, Buras A, Zang CS, Rammig A, Seidl R, Excess forest mortality is consistently linked to drought across Europe, Nat. Commun. 11 (2020) 6200, doi:10.1038/s41467-020-19924-1

[18] Forrester DI, The spatial and temporal dynamics of species interactions in mixed-species forests: From pattern to process, For. Ecol. Manage. 312 (2014) 282–292, doi:10.1016/j.foreco.2013.10.003

[19] Knoke T, Ammer C, Stimm B, Mosandl R, Admixing broadleaved to coniferous tree species: a review on yield, ecological stability and economics, Eur. J. Forest Res. 127 (2008) 89–101, doi:10.1007/s10342-007-0186-2

[20] Morin X, Fahse L, Jactel H, Scherer-Lorenzen M, Garcia-Valdés R, Bugmann H, Long-term response of forest productivity to climate change is mostly driven by change in tree species composition, Sci. Rep. 12 (2018) doi:10.1038/s41598-018-23763-y

[21] Pretzsch H, Bielak K, Block J, Bruchwald A, Dieler J, Ehrhart H-P, et al, Productivity of mixed versus pure stands of oak (Quercus petraea (Matt.) Liebl. and Quercus robur L.) and European beech (Fagus sylvatica L.) along an ecological gradient, Eur. J. Forest Res. 132 (2013) 263–280, doi:10.1007/s10342-012-0673-y

[22] Martin-Blangy S, Charru M, Gérard S, Jactel H, Jourdan M, Morin X, et al, Mixing beech with fir or pubescent oak does not help mitigate drought exposure at the limit of its climatic range, For. Ecol. Manage. 482 (2021) 118840. doi:10.1016/j.foreco.2020.118840

[23] Pérez-Ramos IM, Matías L, Gómez-Aparicio L, Godoy Ó, Functional traits and phenotypic plasticity modulate species coexistence across contrasting climatic conditions, Nat. Commun. 10 (2019) 2555. doi:10.1038/s41467-019-10453-0

[24] Pritzkow C, Williamson V, Szota C, Trouvé R, Arndt SK, Phenotypic plasticity and genetic adaptation of functional traits influences intra-specific variation in hydraulic efficiency and safety, Tree Physiol. 40 (2020) 215–229, doi:10.1093/treephys/tpz121

[25] Sork VL, Genomic Studies of Local Adaptation in Natural Plant Populations, J. Hered. 109 (2017) 3–15, doi:10.1093/jhered/esx091

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

[27] Müller M, Gailing O, Abiotic genetic adaptation in the Fagaceae, Plant Biol. J. 21 (2019) 783–795, doi:10.1111/plb.13008

[28] Ehlers BK, Damgaard CF, Laroche F, Intraspecific genetic variation and species coexistence in plant communities, Biol. Lett. 12 (2016) 20150853.

doi:10.1098/rsbl.2015.0853

[29] J. Pakeman R, K. Beaton J, Thoss V, J. Lennon J, D. Campbell C, White D, et al, The extended phenotype of Scots pine *Pinus sylvestris* structures the understorey assemblage, Ecography. 29 (2006) 451–457, doi:10.1111/j.2006.0906-7590.04617.x

[30] Crutsinger GM, Strauss SY, Rudgers JA, Genetic variation within a dominant shrub species determines plant species colonization in a coastal dune ecosystem, Ecology 91 (2010) 1237–1243, doi:10.1890/09-0613.1

[31] Bolger AM, Lohse M, Usadel B, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics 30 (2014) 2114-2120, doi:10.1093/bioinformatics/btu170

[32] Mishra B, Gupta DK, Pfenninger M, Hickler T, Langer E, Nam B, et al, A reference genome of the European beech (Fagus sylvatica L.), GigaScience. (2018) 7, doi:10.1093/gigascience/giy063

[33] Li H, Durbin R, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics. 25 (2009) 1754–1760, doi:10.1093/bioinformatics/btp324

[34] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al, The Sequence Alignment/Map format and SAMtools, Bioinformatics. 25 (2009) 2078–2079, doi:10.1093/bioinformatics/btp352

[35] Lobréaux S, Melodelima C, Detection of genomic loci associated with environmental variables using generalized linear mixed models, Genomics. 105 (2015) 69–75, doi:10.1016/j.ygeno.2014.12.001

[36] Bates D, Mächler M, Bolker B, Walker S, Fitting Linear Mixed-Effects Models Using lme4, J. Stat. Soft. (2015) 67, doi:10.18637/jss.v067.i01

[37] Joost S, Bonin A, Bruford MW, Després L, Conord C, Erhardt G, et al, A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation, Mol. Ecol. 18 (2007) 3955–3969. doi:MEC3442 [pii] 10.1111/j.1365-294X.2007.03442.x

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

[39] Lê S, Josse J, Husson F, FactoMineR : An R Package for Multivariate Analysis, J. Stat. Soft. (2008) 25. doi:10.18637/jss.v025.i01

[40] Jourdan M, Kunstler G, Morin X, How neighbourhood interactions control the temporal stability and resilience to drought of trees in mountain forests, J. Ecol. 108 (2020) 666–677, doi:10.1111/1365-2745.13294

[41] Jourdan M, Piedallu C, Baudry J, Defossez E, Morin X, Tree diversity and the temporal stability of mountain forest productivity: testing the effect of species composition, through asynchrony and overyielding, Eur. J. Forest Res. 140 (2021) 273–286, doi:10.1007/s10342-020-01329-w

[42] Lalagüe H, Csilléry K, Oddou-Muratorio S, Safrana J, de Quattro C, Fady B, et al, 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 (2014) 15–26, doi:10.1007/s11295-013-0658-0

[43] Bernard-Verdier M, Navas M-L, Vellend M, Violle C, Fayolle A, Garnier E, Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland, J. Ecol. 100 (2012) 1422–1433, doi:10.1111/1365-2745.12003

[44] Antonovics J, Bradshaw AD, Evolution in closely adjacent populations, Heredity 25 (1970) 349-362, doi:10.1038/hdy.1970.36

[45] Brousseau L, Bonal D, Cigna J, Scotti I, Highly local environmental variability promotes intrapopulation divergence of quantitative traits: an example from tropical rain forest trees, Ann. Bot. 112 (2013 1169–1179, doi:10.1093/aob/mct176

[46] Pretzsch H, Canopy space filling and tree crown morphology in mixed-species stands compared with monocultures, For. Ecol. Manage. 327 (2014) 251–264, doi:10.1016/j.foreco.2014.04.027

[47] Jucker T, Bouriaud O, Coomes DA, Crown plasticity enables trees to optimize canopy packing in mixed-species forests, Funct. Ecol. 29 (2015) 1078–1086, doi:10.1111/1365-2435.12428

[48] Chen X, Schreiber K, Appel J, Makowka A, Fähnrich B, Roettger M, et al, The Entner–Doudoroff pathway is an overlooked glycolytic route in cyanobacteria and plants, Proc. Natl. Acad. Sci. USA. 113 (2016) 5441–5446, doi:10.1073/pnas.1521916113

[49] Rabara RC, Tripathi P, Reese RN, Rushton DL, Alexander D, Timko MP, et al, Tobacco drought stress responses reveal new targets for Solanaceae crop improvement, BMC Genomics. 16 (2015) 484. doi:10.1186/s12864-015-1575-4

[50] Verica JA, He Z-H, The Cell Wall-Associated Kinase (*WAK*) and *WAK* -Like Kinase Gene Family, Plant Physiol. 129 (2002) 455–459, doi:10.1104/pp.011028

[51] Meier S, Ruzvidzo O, Morse M, Donaldson L, Kwezi L, Gehring C, The Arabidopsis Wall Associated Kinase-Like 10 Gene Encodes a Functional Guanylyl Cyclase and Is Co-Expressed with Pathogen Defense Related Genes, PLoS ONE. 5 (2010) e8904, doi:10.1371/journal.pone.0008904

[52] Gai Y-P, Zhao Y-N, Zhao H-N, Yuan C-Z, Yuan S-S, Li S, et al, The Latex Protein MLX56 from Mulberry (Morus multicaulis) Protects Plants against Insect Pests and Pathogens, Front Plant. Sci. 8 (2017) 1475. doi:10.3389/fpls.2017.01475

[53] Wang Y, Yang L, Chen X, Ye T, Zhong B, Liu R, et al, *Major latex protein-like protein 43* (*MLP43*) functions as a positive regulator during abscisic acid responses and confers drought tolerance in *Arabidopsis thaliana*, J. Exp. Bot. 67 (2016) 421–434, doi:10.1093/jxb/erv477

[54] Hantsch L, Bien S, Radatz S, Braun U, Auge H, Bruelheide H, Tree diversity and the role of non-host neighbour tree species in reducing fungal pathogen infestation, J Ecol. 102 (2014) 1673–1687, doi:10.1111/1365-2745.12317

[55] Chen ZC, Yamaji N, Fujii-Kashino M, Ma JF, A Cation-Chloride Cotransporter Gene Is Required for Cell Elongation and Osmoregulation in Rice, Plant Physiol. 171 (2016) 494–507, doi:10.1104/pp.16.00017

[56] Mishra MK, Chaturvedi P, Singh R, Singh G, Sharma LK, Pandey V, et al, Overexpression of WsSGTL1 Gene of Withania somnifera Enhances Salt Tolerance, Heat Tolerance and Cold Acclimation Ability in Transgenic Arabidopsis Plants, PLoS ONE. 8 (2013) e63064. doi:10.1371/journal.pone.0063064

[57] Kremer A, Abbott AG, Carlson JE, Manos PS, Plomion C, Sisco P, et al, Genomics of Fagaceae, Tree Genet. Genomes. 8 (2012) 583–610, doi:10.1007/s11295-012-0498-3

[58] Lamit LJ, Busby PE, Lau MK, Compson ZG, Wojtowicz T, Keith AR, et al, Tree genotype mediates covariance among communities from microbes to lichens and arthropods, J Ecol. 103 (2015) 840–850, doi:10.1111/1365-2745.12416

[59] Sabatini FM, de Andrade RB, Paillet Y, Ódor P, Bouget C, Campagnaro T, et al, Trade-offs between carbon stocks and biodiversity in European temperate forests, Glob. Change Biol. 25 (2019) 536–548, doi:10.1111/gcb.14503

[60] Morin X, Fahse L, de Mazancourt C, Scherer-Lorenzen M, Bugmann H, Temporal stability in forest productivity increases with tree diversity due to asynchrony in species dynamics, Ecol. Lett. 17 (2014) 1526–1535, doi:10.1111/ele.12357

Figures and Tables

Fig. 1. Analysis of beech population structure into the sampling area.

A random set of 10,000 SNPs was used to define the population structure through Principal Component Analysis (PCA). Plot names are mentioned indicating the massif where the plot is located : SB (Sainte Baume), GL (Grand Luberon), LA (Luberon lagarde). Mono and multispecific community plot names contain respectively _m and _ph. The first two dimensions of the PCA were plotted.

Fig. 2. Allelic profiles of two SNPs highly associated to the community composition status.

Barplots of allelic proportion as a percentage of total allele count at the position were drawn for SNPs 119,197 and 119,226 located on scaffold 635 of the draft *Fagus sylvatica* genome assembly. Sampling site names are indicated. Black bars correspond to multispecific beech-oak plots and grey bars to beech monospecific community plots.

Fig. 3. Average allelic proportion of four positive SNPs.

Barplot of average allelic proportion were drawn from the calculated average proportion at a position over the 5 mono or multispecific community studied plots. Standard deviations are indicated. Data are provided for 4 SNPs significantly associated to the community status : A) scaffold635 position 119,197 B) scaffold655 position 128,466 C) scaffold2 position 24219 D) scaffold2 position 68595.

Fig. 4. Beech productivity into the studied sites.

Basal area increment was calculated as indicated for multispecific plots and the associated monospeficic plot of the same sampling site as a productivity estimate. Site names mentioned the corresponding sampling massif SB (Sainte Baume), GL (Grand Luberon), LA (Luberon lagarde). Black bars correspond to multispecific beech-oak plots and grey bars to beech monospecific community plots.

Table 1 *Fagus sylvatica* SNPs highly associated with the tree community composition. Candidate genes are indicated when available.

Scaffold	SNP position	corrected_pvalue	gene
Scaffold_635	119197	4.14e-22	KHG-KDPG aldolase
Scaffold_635	119226	5.84e-19	KHG-KDPG aldolase
Scaffold_635	120737	6.05e-18	KHG-KDPG aldolase
Scaffold_2	24219	3.16e-20	Wall associated receptor kinase
Scaffold_2	71526	8.16e-18	Sterol 3-beta-glucosyltranferase
Scaffold_2	90214	2.38e-19	-
Scaffold_2	90865	4.12e-19	-
Scaffold_655	128466	1.84e-18	Cation chloride cotransporter
Scaffold_2068	39162	4.19e-18	MLP 31 defense response

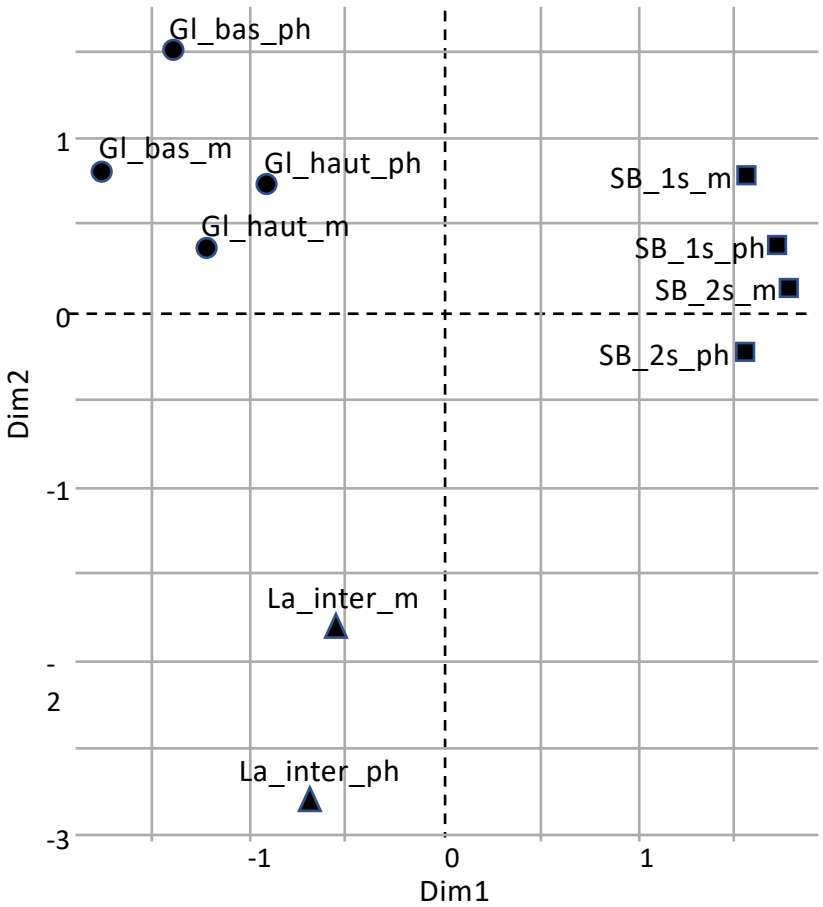
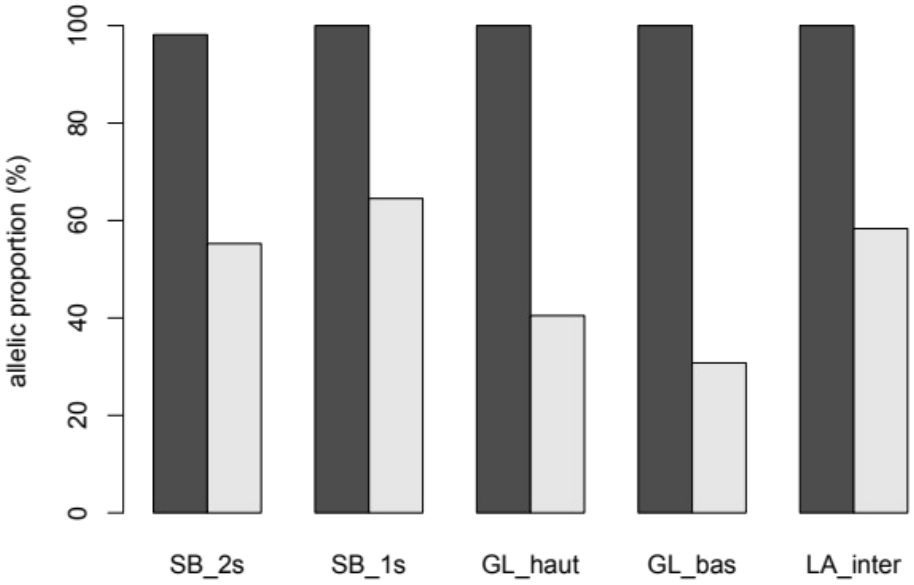

Gl_bas_ph
Gl_bas_m
Gl_haut_ph
Gl_haut_m
SB_1s_m
SB_1s_ph
SB_2s_m
SB_2s_ph
La_inter_m
La_inter_ph
1
0
-1
-2
-3
-1
0
1
Dim1
Dim2

Fig. 1.

SNP_119197



SNP_119226

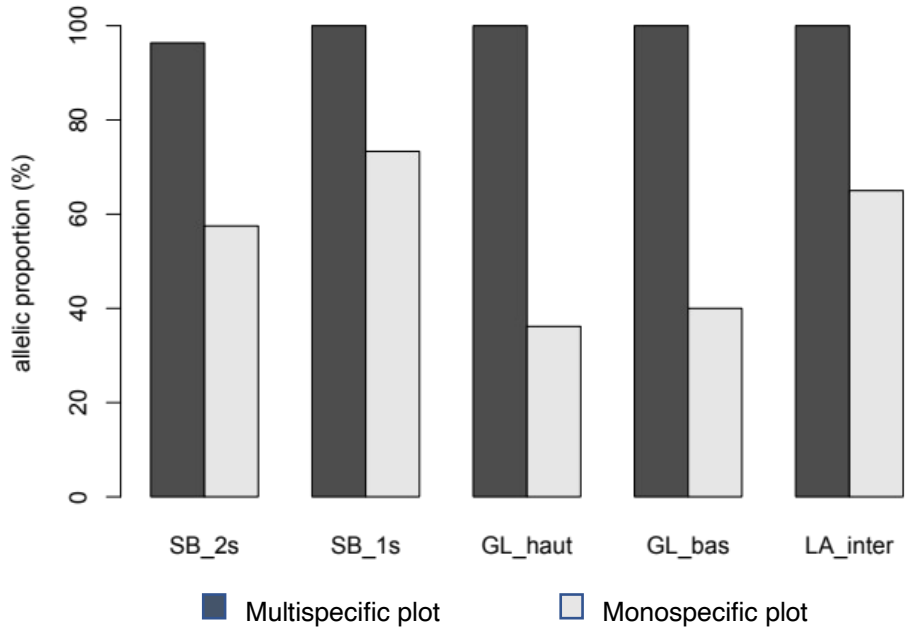


Fig. 2.

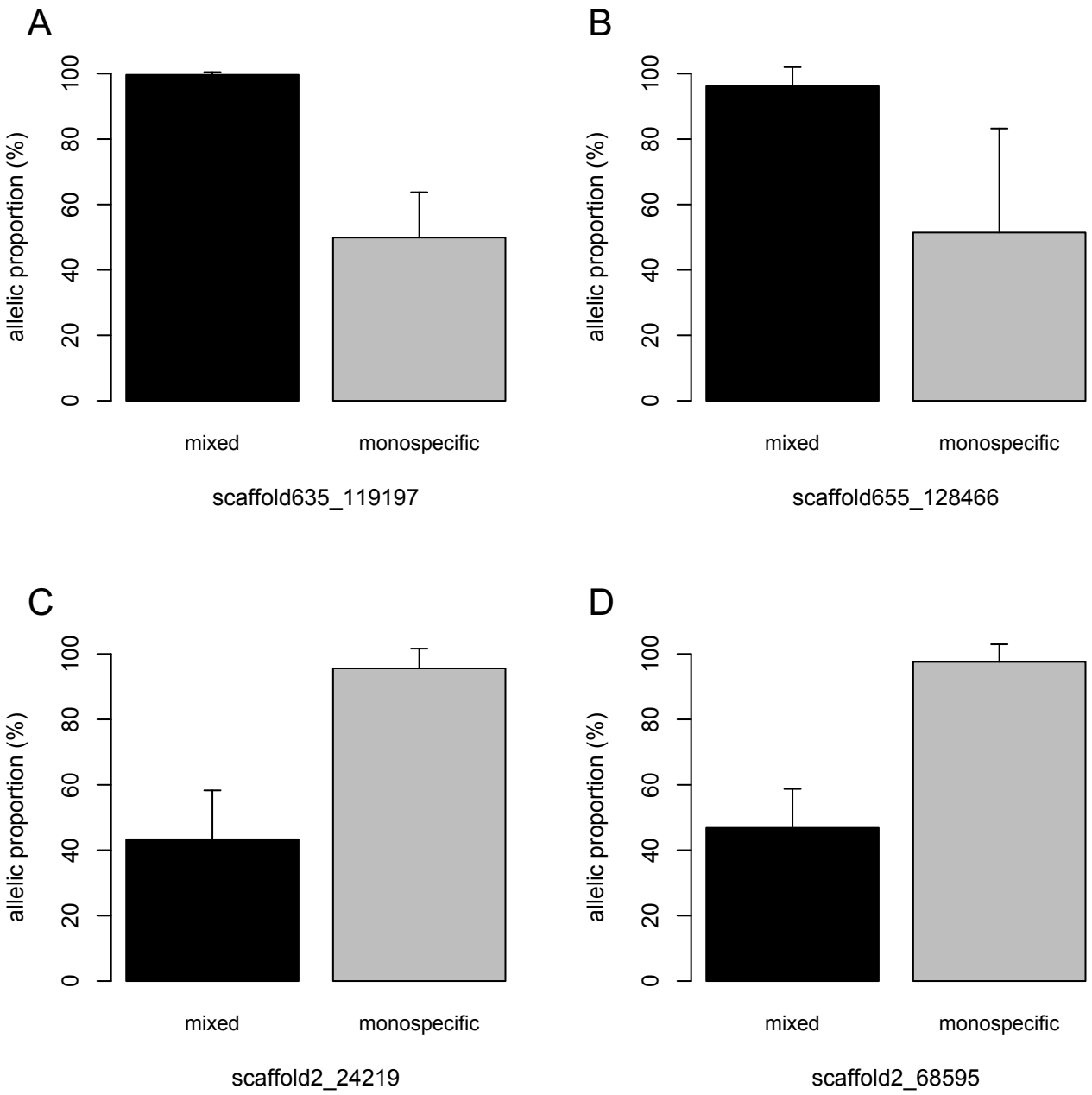

A
allelic proportion (%)
0
20
40
60
80
100
mixed
monospecific
scaffold635_119197
B
allelic proportion (%)
0
20
40
60
80
100
mixed
monospecific
scaffold655_128466
C
allelic proportion (%)
0
20
40
60
80
100
mixed
monospecific
scaffold2_24219
D
allelic proportion (%)
0
20
40
60
80
100
mixed
monospecific
scaffold2_68595

Fig. 3.

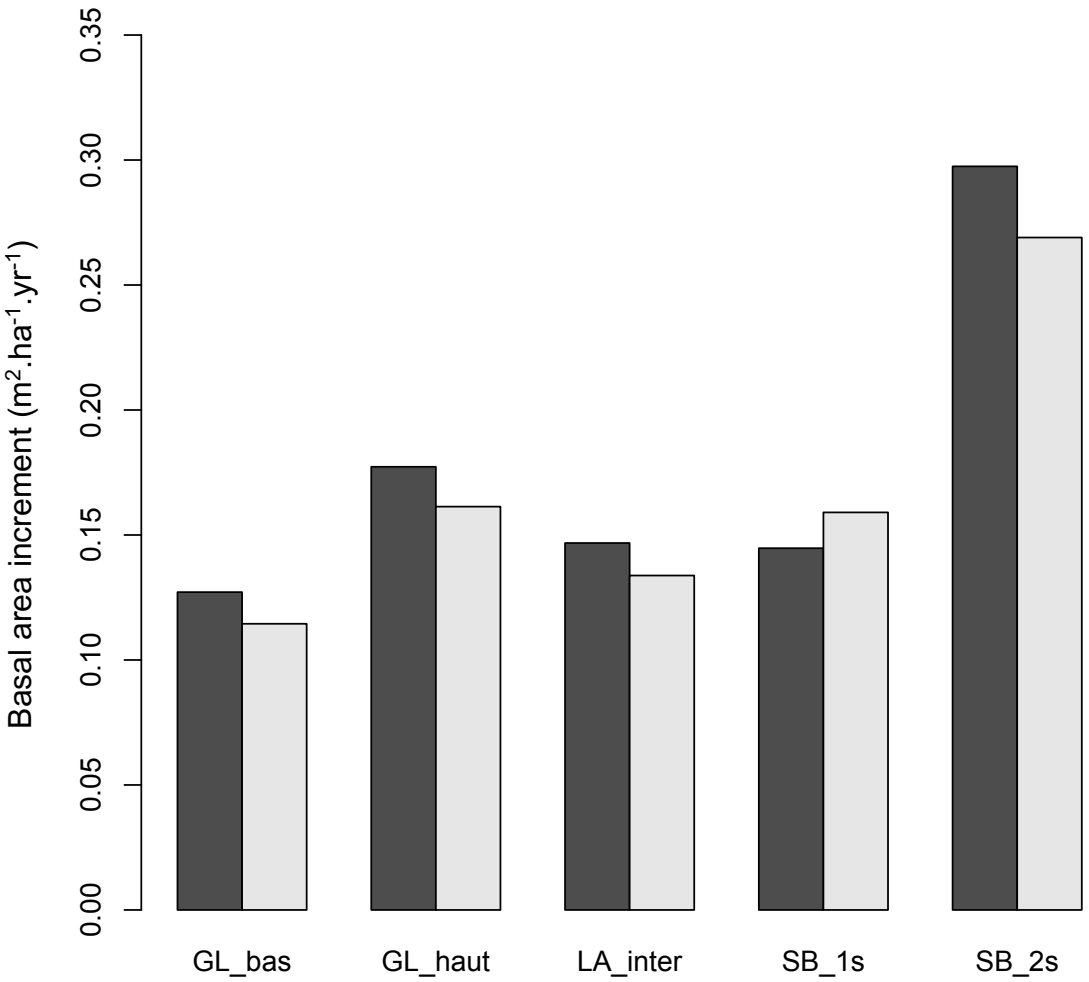

Basal area increment (m².ha⁻¹.yr⁻¹)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
GL_bas
GL_haut
LA_inter
SB_1s
SB_2s

Fig. 4.

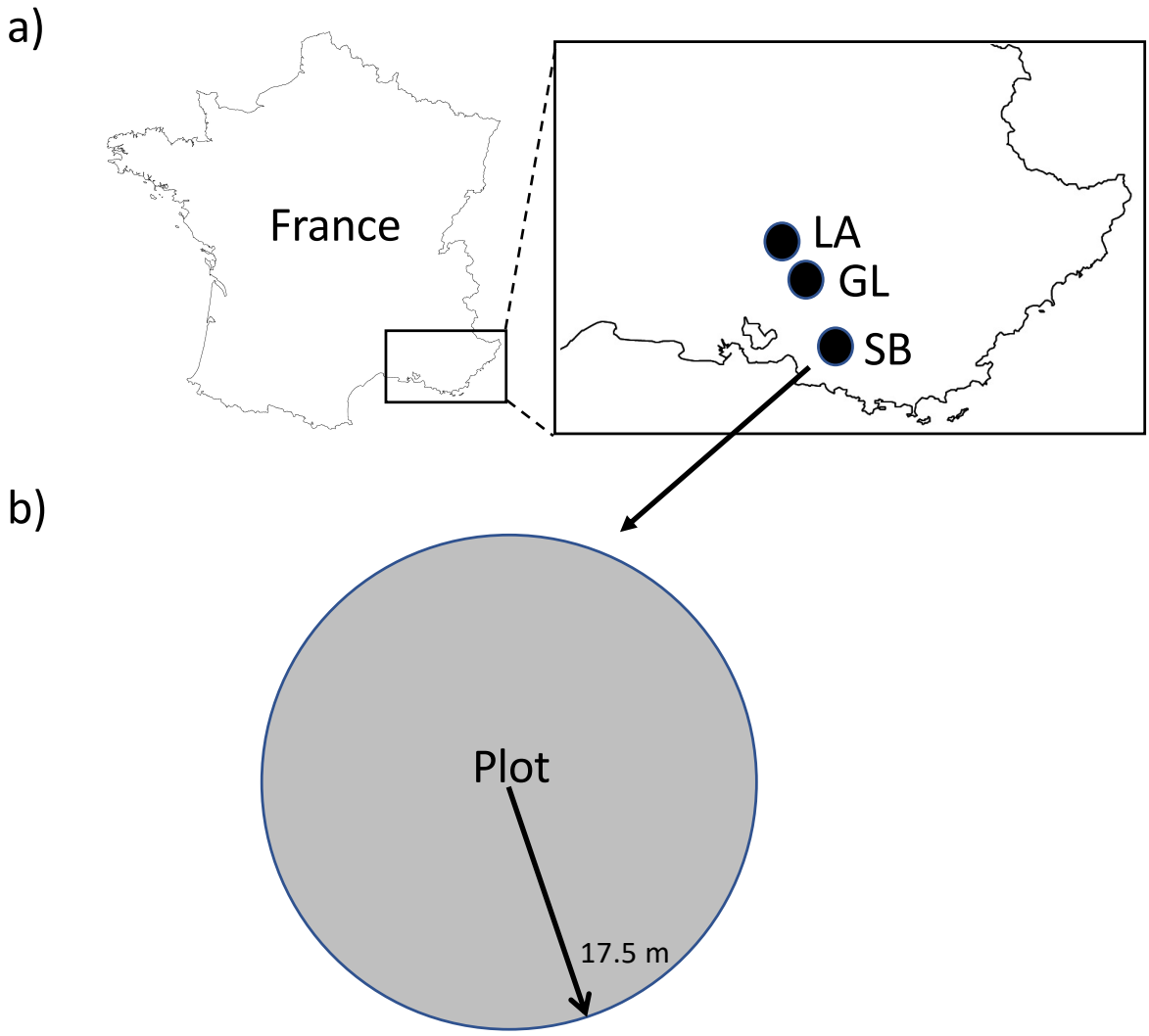


Fig S1

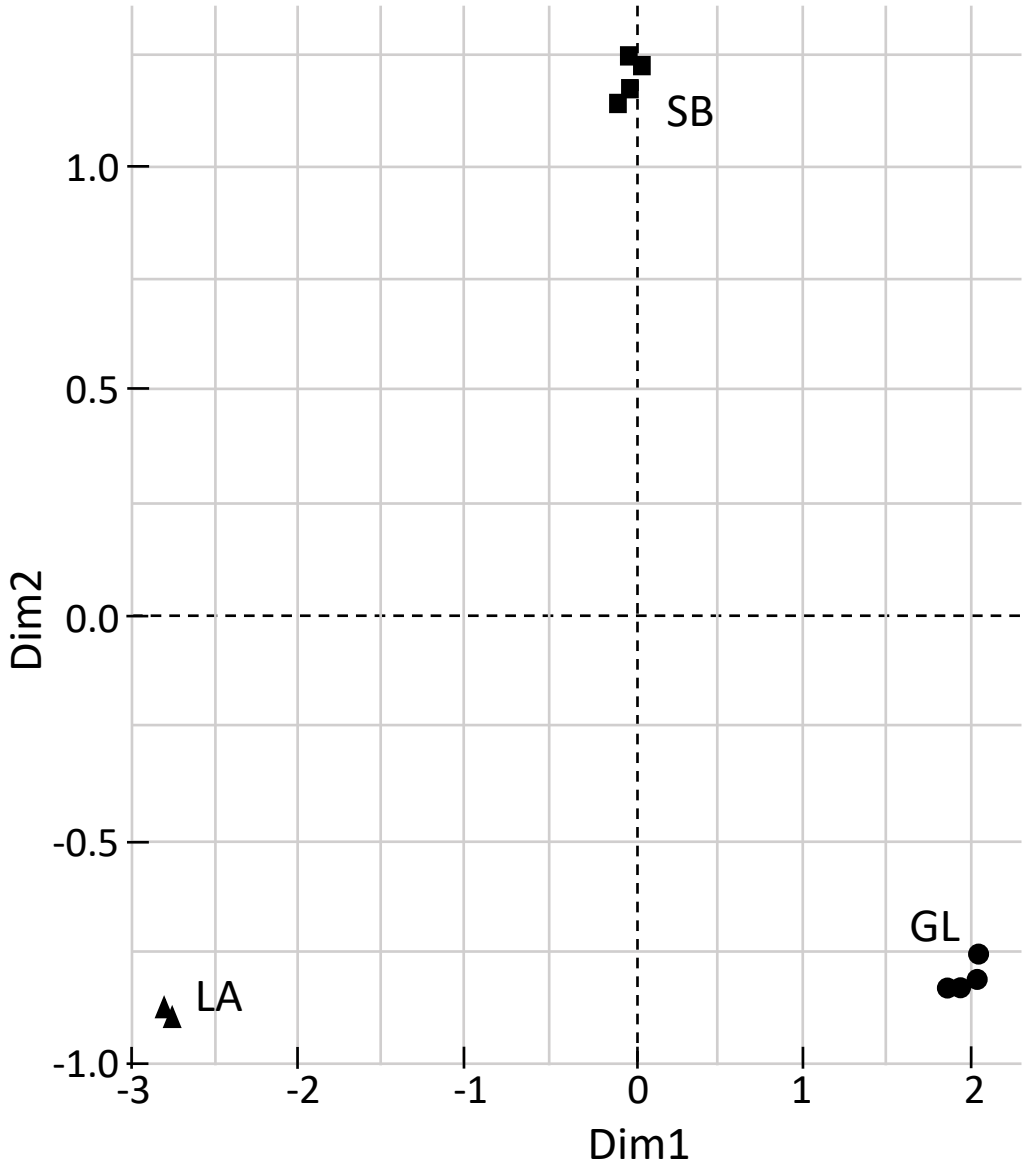


Fig S2

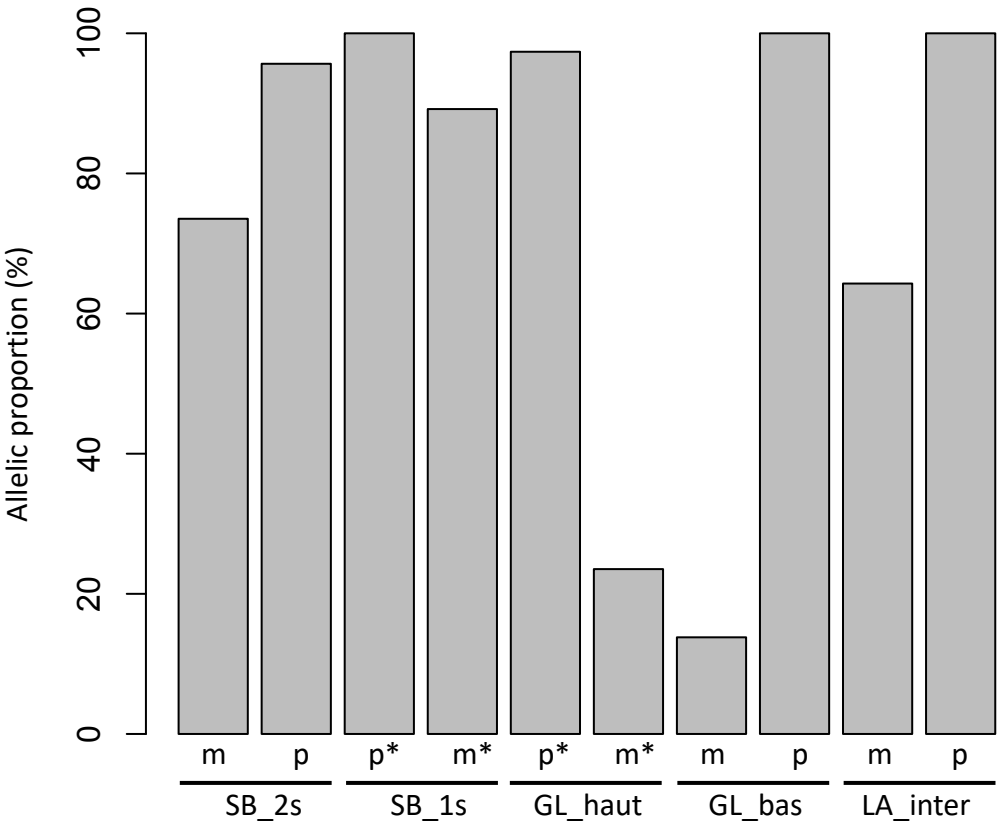

Allelic proportion (%)
0
20
40
60
80
100
m
p
p*
m*
p*
m*
m
p
m
p
SB_2s
SB_1s
GL_haut
GL_bas
LA_inter

Fig S3

Material and Method SI

Tree growth analysis

We studied growth dynamics using tree rings for the last 18 years before sampling, i.e. from 1995 to 2013 (Jourdan et al 2019). All cores were cross-dated using specific species pointer years, as described in Lebourgeois and Merian (2012), but without series standardization. This analysis allowed ensuring that chronologies were synchronized, without correcting for interannual variations amplitude. Here, productivity was represented by basal area increment *BAI* instead of diameter increment in Lloret et al. (2011). Diameter increments were transformed into *BAI* using measured DBH (Biondi and Qeadan, 2008). Some cores were too difficult to read accurately (blurred distinction between rings) and were thus not reliable and have been removed for the analyses.

We used annual *BAI* at the stand level. The annual productivity at year *y* (BAI_y) was calculated by summing the *BAI* of the trees for each plot:

$$BAI_y = \sum_{i}^{n} BAI_{iy} \qquad (1)$$

with *n* being the number of trees in the plot and BAI_{iy} being the basal area increment of tree *i* at year *y*.

References

M. Jourdan, F. Lebourgeois, X. Morin, The effect of tree diversity on the resistance and

recovery of forest stands in the French Alps may depend on species differences in hydraulic features. *For. Ecol. Manage.* **450**, 117486 (2019).

F. Lebourgeois, P. Merian, Principes et méthodes de la dendrochronologie. LERFOB AgroPariTech Cent Nancy, 85 (2012).

F. Lloret, E. G. Keeling, A. Sala, Components of tree resilience: effects of successive low-growth episodes in old ponderosa pine forests. *Oikos* **120**, 1909-1920 (2011).

F. Biondi, F. Qeadan, A Theory-Driven Approach to Tree-Ring Standardization: Defining the Biological Trend from Expected Basal Area Increment. *Tree-Ring Res.* **64**, 81-96 (2008).