Genomic and phenotypic divergence unveil microgeographic adaptation in the Amazonian hyperdominant tree *Eperua falcata* Aubl. (Fabaceae)

Louise Brousseau[1,2,6], Paul V. A. Fine[3], Erwin Dreyer[4], Giovanni G. Vendramin[5], Ivan Scotti[6]

[1.] INRA, UMR EcoFoG Ecologie des Forêts de Guyane, Campus Agronomique BP709, 97387 Kourou cedex, French Guiana

[2.] UMR AMAP, IRD, CIRAD, CNRS, INRA, University of Montpellier, Bd de la Lironde, 34398 Montpellier, France.

[3.] Department of Integrative Biology, University of California, Berkeley, 1005 Valley Life Science Building, Berkeley, California 94720-3140, USA.

[4.] INRA, Université de Lorraine, UMR EEF Ecologie et Ecophysiologie Forestière, 54280 Champenoux, France.

[5.] Institute of Biosciences and BioResources, National Research Council (IBBR-CNR), Division of Florence, Via Madonna del Piano 10, 50019 Sesto Fiorentino (FI), Italy.

[6.] INRA, URFM Ecologie des Forêts Méditerranéennes, Domaine Saint-Paul, Site Agroparc CS 40509, 84914 Avignon cedex 9, France.

Corresponding Author:

Louise Brousseau. IRD, UMR DIADE, 911 Avenue Agropolis, BP 64501, 34394 Montpellier, France. louise.brousseau@ird.fr

Abstract

Plant populations can undergo very localized adaptation, allowing widely distributed populations to adapt to divergent habitats in spite of recurrent gene flow. Neotropical trees - whose large and undisturbed populations often span a variety of environmental conditions and local habitats - are particularly good models to study this process. Here, we carried out a genome scan for selection through whole-genome sequencing of pools of populations, sampled according to a replicated sampling design, to evaluate microgeographic adaptation in the hyperdominant Amazonian tree *Eperua falcata* Aubl. (Fabaceae). A high-coverage genomic resource of ~250 Mb was assembled *de novo* and annotated, leading to 32,789 predicted genes. 97,062 bi-allelic SNPs were detected over 25,803 contigs, and a custom Bayesian model was implemented to uncover candidate genomic targets of divergent selection. A set of 290 divergence outlier SNPs was detected at the regional scale (between study sites), while 185 SNPs located in the vicinity of 106 protein-coding genes were detected as replicated outliers between microhabitats within regions. Outlier genomic regions are involved in a variety of physiological processes, including plant responses to stress (e.g., oxidative stress, hypoxia and metal toxicity) and biotic interactions. Together with evidence suggesting microgeographic divergence in functional traits, the discovery of genomic targets of microgeographic adaptation in the Neotropics is consistent with the hypothesis that local adaptation is a key driver of ecological diversification, operating across multiple spatial scales, from large- (i.e. regional) to microgeographic- (i.e. landscape) scales.

Key words: Whole-genome shotgun sequencing, Neotropics, Pool-Seq, genome scan, Bayesian modelling, local adaptation.

Introduction

Plant populations can undergo extremely localized adaptation, allowing populations of widely distributed species to adapt to divergent habitats in spite of recurrent gene flow. This phenomenon has been repeatedly observed in annual plants growing in strongly contrasting yet geographically adjacent habitats, particularly in cases with large differences in soil composition (Antonovics, 2006; Antonovics & Bradshaw, 1970; Bradshaw, 1960; Gould, McCouch & Geber, 2014; Jain & Bradshaw, 1966); in some cases, molecular evidence also supports adaptation to habitat mosaics (Fustier et al., 2017; Turner, Bourne, Von Wettberg, Hu, & Nuzhdin, 2010). In a study of *Howea* palm trees, Savolainen et al. (2006) provided strong evidence of a link between adaptation to habitat mosaics and sympatric speciation. The process underlying these instances of adaptation with gene flow has been dubbed 'microgeographic adaptation', whereby adaptive divergence occurs at geographical scales of the same order of magnitude as gene dispersal distance (Richardson, Urban, Bolnick, & Skelly, 2014). Adaptation to multiple optima within a continuous population has been identified as a major mechanism for the maintenance of genetic diversity (Delph & Kelly, 2014). Moreover, conditions under which adaptive divergence can occur with gene flow have been explored theoretically with models including two environmental patches connected by varying levels of gene flow (Bulmer, 1972; Hendry, Day, & Taylor, 2001; Schmid & Guillaume, 2017; Yeaman & Guillaume, 2009; Yeaman & Otto, 2011). These simulations show that a wide range of migration-selection equilibrium states can lead to adaptation to each patch with maintenance of polymorphism (i.e. maintenance of patch specialists, as opposed to the emergence of a generalist genotype).

Data from perennial plants, and from trees in particular, suggest that this may be a general phenomenon, that can be readily observed at the phenotypic level (Brousseau, Bonal, Cigna, & Scotti, 2013; Lind et al., 2017; Vizcaíno-Palomar, Revuelta-Eugercios, Zavala, Alía, & González-Martínez, 2014; Yeaman & Jarvis, 2006) (but see Latreille & Pichot (2017) for a case of a complete lack of microgeographical adaptation). At the molecular level, association between alleles (or genotypes) and habitats has been detected in a variety of studies (see review by Savolainen et al., 2007), suggesting an important role for intra-specific and intra-populational variation in the evolution of habitat specialization (Scotti, González-Martínez, Budde, & Lalagüe, 2016). The advent of genome-wide approaches has increased detection power, allowing to more precisely evaluate the genetic bases of microgeographic adaptation (Eckert et al., 2015; Fustier et al., 2017; Izuno et al., 2017; Lind et al., 2017; Turner et al., 2010). In general, genome scan studies have found only a minority of loci (on the order of a few percent

of the total) that exhibit microgeographic disruptive selection. Yet genome-wide, sequencing-based approaches can go beyond a finer estimation of the number of loci involved in the process, because they permit formulating functional interpretations of the divergence process as sequence annotation can suggest putative function of the target loci (Eckert et al., 2015, 2010; Fustier et al., 2017; Turner et al., 2010).

Microgeographic adaptation involves the filtering of phenotypes and genotypes that favor adaptation to habitat patches, and therefore can be viewed as an initial step towards the diversification of populations, ecotypes and eventually species divergence. In particular, environmental filtering has frequently been invoked as a driver of niche partitioning and species diversification in the Neotropics (Baraloto, Morneau, Bonal, Blanc, & Ferry, 2007; Fine, 2015; Fine, Daly, & Cameron, 2005; Fine, Mesones, & Coley, 2004; Fine et al., 2013; Kraft, Valencia, & Ackerly, 2008); intra-populational diversification processes across environmental gradients and/or habitat patches thus may play a key role in the generation of functional diversity. Moreover, species' short-term adaptive potential (De Kort & Honnay, 2017; Harrisson, Pavlova, Telonis-Scott & Sunnucks, 2014) largely depends on standing genetic variation, and microgeographic adaptation can therefore play a role in the maintenance of adaptive genetic diversity. Indeed, highly heterogeneous landscapes create the conditions for microgeographic adaptations to various environmental conditions and would increase inter-habitat diversity and global diversity in meta-populations considered at higher spatial scales, analogous to beta- and gamma-diversity across species pools in ecology. Measuring the extent of microgeographic adaptation in the wild is thus of major importance, as it could provide insights into the ability of populations to evolve towards new phenotypic and genetic optima under global climate change.

Trees are particularly good models to study microgeographic adaptation: many wild tree species have large populations with widespread distribution spanning a variety of microhabitats, harbor large amounts of genetic diversity, and can disperse over long distances (Fetter, Gugger, & Keller, 2017; Hamrick & Godt, 1990; Neale & Kremer, 2011; Petit & Hampe, 2006; Sork et al., 2013). Trees should therefore not display major departures from the conditions approximating selection-migration equilibrium; nevertheless, their long generation span, combined with continuously changing environmental conditions, may prevent them from actually reaching equilibrium. Moreover, methods to detect locus-specific divergence (and infer selection) should work quite well with microgeographic processes in populations of the most common forest trees, because their relatively shallow (spatial) population structures and large

effective population sizes should minimize the introduction of biases (e.g., because of drift) and the generation of false positives.

Lowland Amazonian rainforests, whose landscapes show fine patchiness of contrasting topographic and edaphic conditions varying over microgeographic scales (**Fig. 1a,c,d**), represent an ideal study system to investigate microgeographic adaptation processes. In Amazonian rainforests, the recent study by ter Steege et al. (2013) reported 227 hyperdominant tree species accounting for half of all stems. Some of these widespread species are also ecological generalists, whose preferences could be explained by individual phenotypic plasticity or by microgeographic adaptation to habitat patchiness (Fortunel et al., 2016). We focus here on dense stands of the hyperdominant tree *Eperua falcata* Aubl. (Fabaceae) in French Guiana that occur across environmentally heterogeneous areas, with the goal of identifying genomic loci undergoing divergence at the microgeographic scale between habitats. To do this, a genome scan was carried out through a replicated sampling design combining regional and microgeographic spatial scales (i.e. replicates of microhabitats, ~300 m apart, in two study sites, ~300 km apart). We analyzed patterns of genomic and phenotypic divergence from large (i.e. regional) to microgeographic spatial scales. A custom Bayesian framework (Brousseau et al., 2016) was designed to identify locus-specific departures from hierarchical genome-wide divergence in replicated sampling designs. This method was preferred over other outlier detection methods because it is suitable for analyzing Pool-Seq data and it is specifically designed to identify adaptive divergence in nested sampling designs under an explicit hierarchical model of neutral divergence, taking advantage of the cumulative divergence effects occurring in replicated population pairs. Results obtained through the custom Bayesian framework were compared to standard methods for verification. The results of tests for genomic divergence at nested geographical scales were supplemented with trait divergence data, obtained from a large-scale reciprocal transplant experiment involving the very same populations and sites studied at the genomic level. The combination of these two approaches allowed us to formulate detailed interpretations of the observed diversity patterns.

Materials and Methods

Species model

Eperua falcata Aubl. (Fabaceae, Caesalpinioideae, **Fig. S1**) is a reportedly non-nodulating (Villadas, Fernández-López, Ramírez-Saad, & Toro, 2007) tree species. It is widespread in Northeast Amazonia, from North Brazil to Venezuela through the Guiana shield (https://demo.gbif.org/species/2943522), and it is a 'hyperdominant' species, the 13th most abundant tree species of the Amazonian tree flora (ter Steege et al., 2013). It is an evergreen, late-successional canopy tree species and it is often emergent. It is shade hemi-tolerant (Bonal, Barigah, Granier, & Guel, 2000) and can tolerate both drought (Bonal, Atger, et al., 2000; Bonal, Guehl, & Champenoux, 2001) and water-logging (Baraloto et al., 2007). Its distribution is very clumped and populations are commonly found at high densities. It is a generalist with respect to microgeographic edaphic conditions, although it is most abundant in water-logged forests of bottomlands (Baraloto et al., 2007). *E. falcata* flowers and fruits during the rainy season, and seeds reach maturity in April-May, synchronously among trees in all microhabitats. Its large pollen grains (>100µm) are dispersed by animals such as bats (Cowan, 1975), while seed dispersal is autochorous: heavy seeds are dispersed around crowns of mother trees through explosive pod dehiscence.

For this species, previous studies have already reported high heritability in quantitative growth and leaf traits in a common garden experiment (Brousseau et al., 2013), suggesting high levels of standing genetic variation for functional phenotypic traits. In addition, seedlings native to different microhabitats displayed genetically-driven differences in phenotypic traits probably due to a combination of neutral processes (i.e. short-distance seed flow and local inbreeding) and local adaptation (Audigeos, Brousseau, Traissac, Scotti-Saintagne, & Scotti, 2013; Brousseau, Foll, Scotti-Saintagne, & Scotti, 2015).

Sampling design and study sites

We sampled populations that replicated the microgeographic environmental contrast 'bottomland *versus* terra-firme' (~300 m apart) at two regional locations (i.e. study sites) in French Guiana (~300 km apart), **Fig. 2a**. Laussat (western French Guiana) and Regina (eastern French Guiana) study sites are located on the coastal Guiana Shield and experience different rainfall regimes (~2,500 and 3,500 mm/year, respectively) (Brousseau et al., 2015). Each regional location harbors different microhabitats (hygromorphic sandy bottomlands and ferralitic *terra-firme* plateaus or hills). In Laussat, a plateau of low elevation is adjacent to a permanently water-logged bottomland, while Regina is composed of higher elevation hilltops bordered by abrupt slopes surrounding a seasonally flooded bottomland forest, **Fig. S2**. *E.*

179 *falcata* trees are found through the study area, ensuring gene flow between tree populations
180 occupying different microhabitats (Brousseau et al., 2015), **Fig. 2b**. *E. falcata* stem (> 20 cm
181 diameter at breast height) density varies between 30 and 50 trees / ha at these locations
182 (Brousseau et al., 2015). Twenty adult trees for each of the four *regional location* ×
183 *microhabitat* combination were sampled, totaling 80 individuals.
184 Details on soil characteristics of the two sites and on growth and functional characters
185 of the sampled trees are reported in **File S1**. Soil properties (including soil type composition,
186 clay/silt/sand composition, fertility, and water table depth) vary between sites and
187 microhabitats, providing a basis for population divergence. The sampled trees have similar
188 diameter at breast height (d.b.h.) and have similar canopy status (almost all are dominant trees)
189 across sites and microhabitats, indicating that sampled populations are demographically
190 homogeneous, although individuals exhibited phenotypic divergence in some traits.
191
192 *Molecular methods*
193
194 Genomic DNA was extracted from fresh leaves according to a modified CTAB protocol
195 (Doyle & Doyle, 1987): CTAB buffer was supplemented in proteinase K in concentrations of
196 80 µg/ml, and the deproteinization step was realized twice (once with
197 Dichloromethane:Chloroform:Isoamylalcool 24:24:1 and once with Chloroform:Isoamylalcool
198 24:1). Genomic DNA was eluted in Tris-EDTA 0.5X with RNAse A in concentration 10 µg/ml,
199 and quantified twice using a Nanodrop 2000. Individuals belonging to the same regional
200 location and microhabitat were pooled together in equimolarity, resulting in four libraries of
201 pooled individuals. Sequencing libraries were prepared with a TruSeq Kit (Illumina, San Diego,
202 CA) following manufacturer's instructions. The four libraries were sequenced (paired-end) on
203 three lanes of an Illumina HiSeq2500. Sequencing library preparation and sequencing were
204 carried out by IGA (Udine, Italy).
205
206 *Bioinformatics pipeline and computation of basic diversity statistics*
207
208 The bioinformatics pipeline is described in **File S2**. Nei's molecular diversity π (Nei,
209 1987) was computed for each of the four sub-populations at the contig level with unphased SNP
210 data, under the hypothesis of random mating, as the sum of unbiased estimates of heterozygosity
211 at the nucleotide (SNP) level (Nei, 1987):

$$\pi = \sum_{i}^{S} H_{e,i}$$

with S = number of polymorphic loci in a contig, and

$$H_{e,i} = H'_{e,i} \frac{n}{n-1}$$

where n = haploid sample size and $H'_{e,i}$ = expected heterozygosity at the ith polymorphism in the contig. Because different contigs have different lengths, to compare across contigs we present $\pi_L = \pi/L$, where L is contig length, that is, molecular diversity per base.

Bayesian modelling of genomic divergence and outlier detection

A Bayesian modelling approach was used to identify outlier SNPs under selection (Brousseau et al., 2016) has described this Bayesian framework in detail and its power has been assessed through simulations, revealing low false discovery and false non-discovery rates. Briefly, allele frequencies were inferred from read counts and pairwise locus-specific G_{ST} was estimated in the model, which integrates this way coverage variability across the genome and libraries.

$$n_{i,m} Bin(p_{i,m}, N_{i,m})$$

$$q_{i,m} = 1 - p_{i,m}$$

where $n_{i,m}$ is the read count of the reference allele and $N_{i,m}$ the total number of read counts at marker m in population i. $p_{i,m}$ and $q_{i,m}$ are allele frequencies of the reference and the alternative alleles respectively.

$$GST_{(i,j),m} = \left(Ht|(i,j),m - Hs_{(i,j),m}\right)/Ht_{(i,j),m}$$

$G_{ST(i,j),m}$ is the Nei's fixation index (i.e. genetic differentiation) between populations i and j at marker m. $H_{T(i,j),m}$ and $H_{S(i,j),m}$ are the expected heterozygosities in the total metapopulation and within subpopulations respectively:

$$Hs_{(i,j),m} = (p_{i,m} q_{i,m}) + (p_{j,m} q_{j,m})$$

$$Ht_{(i,j),m} = 2\acute{p}_m \acute{q}_m = 2\left(\frac{p_{i,m} + p_{j,m}}{2}\right)\left(\frac{q_{i,m} + q_{j,m}}{2}\right)$$

Pairwise G_{ST} was further partitioned into genome-wide and locus-specific components according to a hierarchical framework: neutral (genome-wide) divergence both between study sites and between microhabitats within each study sites (**Fig. 2a**), and adaptive (locus-specific) divergence both between sites and between microhabitats common to both sites.

$$GST_{(i,j),m} N\left(\mu_{(i,j),m}, \sigma_R^2\right)$$

$$\mu_{(i,j),m} = kr_{(i,j)}(\mu r + \theta r_m) + kl_{laussat_{(i,j)}}(\mu l_{laussat} + \theta l_m) + kl_{regina_{(i,j)}}(\mu l_{regina} + \theta l_m)$$

$$kr_{(i,j)} = \begin{cases} 0 \\ 1 \end{cases}; kl_{laussat_{(i,j)}} = \begin{cases} 0 \\ 1 \end{cases}; kl_{regina_{(i,j)}} = \begin{cases} 0 \\ 1 \end{cases};$$

μr and μl are the global mean of differentiation (over all markers) at regional and local scales respectively: μr is the global mean of differentiation between sites, $\mu l_laussat$ is the global mean of differentiation between microhabitats within the study site of Laussat, and μl_regina is the global mean of differentiation between microhabitats within the study site of Régina.

$kr_{i,j}$ and $kl_{i,j}$ are binary matrices describing whether the populations *i* and *j* belong to the same study site (regional scale) and to the same microhabitat within each study site (local scale): $kr_{i,j} = 1$ if the two populations *i* and *j* belong to the different study sites, $kr_{i,j} = 0$ otherwise; $kl_laussat_{i,j} = 1$ if the two populations *i* and *j* belong to different microhabitats within the study site of Laussat, $kl_laussat_{i,j} = 0$ otherwise; $kl_regina_{i,j} = 1$ if the two populations *i* and *j* belong to different microhabitats within the study site of Régina, $kl_regina_{i,j} = 0$ otherwise.

θr_m and θl_m are locus-specific parameters describing the departure from the global means of differentiation at regional (i.e. between sites) and local (i.e. between microhabitats) scales. Posteriors of θr_m and θl_m are normally distributed allowing thus to attribute an empirical Bayesian p-value (eBP-value) to each SNP. SNPs with an eBP < 1% were considered as outliers under divergent (upper tail) or homogenizing (lower tail) selection between sites (θr_m) or microhabitat within sites (θl_m).

Bayesian Model specifications are described in **File S3**.

For comparison, two additional methods for divergence outlier detection were also applied to our data: BayPass (Gautier, 2015) and G2D (Nielsen et al., 2009). The detailed methods associated to each analysis, as well as the results, are reported in **Files S4 and S5**, respectively. The values of the *XtX* and of the G2D statistics (respectively from the BayPass and the G2D method) from different sets of SNPs (neutral *versus* outlier, at each geographic scale) detected through our custom Bayesian framework were compared by a non-parametric Wilcoxon rank sum test with R (function 'wilcox.text'). It has to be noted that the G2D method also required the identification of the most-likely historical demography model for the sampled populations, thus providing additional characterization of the sample.

Phenotypic divergence in reciprocal transplants

Reciprocal transplants are ideal to test whether genetic divergence is associated with phenotypic divergence *in situ*, i.e. in real although uncontrolled conditions. A large reciprocal

transplant experiment across sites and microhabitats was set-up in April-May 2011 to partition genetic and environmental sources of phenotypic differentiation in the two study sites of Laussat and Régina. The methodology is described in **File S6**.

Results

Genomic divergence

Bioinformatics results of *de novo* assembly and annotation are presented in **File S7**. Almost one hundred thousand (97,062) bi-allelic SNPs were identified in 25,803 contigs and 23,276 scaffolds, according to a stringent procedure, for a SNP density of 0.003/base (*sd*=0.004) with a Ti/Tv=1.80, **Table S2** and **Fig. S4**. Normalized molecular diversity π_L varied between 9×10^{-6} and 8×10^{-3} (**Table S3**), with little variation among populations. Among all SNPs detected, 41,050 (42.3%) were located in scaffolds with at least one predicted gene, and 56,012 SNPs (57.7%) in scaffolds without any predicted gene (total sequence length of the two scaffold types: 97828840 and 167106160 bp respectively. The average scaffold length is 3292 bp for gene-containing scaffolds and 687 bp for scaffolds without genes, and the difference between means is significant: Wilcox rank sum test, W = 6.586e+9, p-value < 2.2e-16). At last, 26,435 SNPs (27.2%) lie within a predicted gene, of which 10,863 (11.2%) within exons, while 14,615 SNPs were close to a predicted gene with a mean distance to a predicted gene of 737 bp (with min = 1bp and max = 27.2 kb).

The Bayesian model revealed low genome-wide differentiation between populations and significant departures from neutral expectations for a small subset of SNPs. Genome-wide genetic differentiation (average G_{ST} across loci) ranged between 0.026 and 0.032 at both geographical scales (credible intervals are shown in **Fig. S5**). Accordingly, genome-wide parameters of divergence estimated through G_{ST} partitioning were not higher between regions than between microhabitats within regions, as confirmed by overlapping credible intervals, **Table S4**. Coalescent modelling (**File S5**) suggested that the divergence between sites is about twice as old as divergence between sub-populations within sites. Additionally, it showed that populations at both sites globally expanded, although bottomland populations underwent expansion while hilltop populations underwent contraction. It finally provided fine estimates of migration rates between the studied populations that were in the order of 1% between sites and 20-30% between microhabitats within site.

Locus-specific departures from genome-wide divergence (i.e. outlier SNPs) were identified from the distribution of locus-specific parameters (θr_m and θl_m). Between 375 and

2,045 outlier SNPs were detected to be under selection with an eBP threshold of 1%: 375 SNPs under homogenizing selection at regional scale, 2,045 SNPs under divergent selection at regional scale, 510 SNPs under homogenizing selection at local scale and 1,770 SNPs under divergent selection at local scale. After multiple-testing correction, the number of significant SNPs fell to 475 outlier SNPs (0.49% of all SNPs screened): 290 SNPs (in 268 contigs) under divergent selection at the regional scale and 185 SNPs (in 168 contigs) under divergent selection at the local scale, with no outlier shared between the regional and the local scales. Outlier SNPs under divergent selection were characterized by clines in allele frequencies between regional locations and/or local microhabitats. Clines in allele frequency across sites and microhabitats are illustrated in **Fig. 4** for different kinds of neutral and outlier SNPs.

The results of the BayPass and G2D analyses are reported in **Files S4** and **S5**, respectively, providing support to the findings of our custom hierarchical model. While only six contigs showed up as outliers at both Laussat and Régina sites by the G2D, the value of the G2D statistic was significantly higher for the contigs containing outliers than for the remaining loci (**File S5**). Similarly, comparing the distribution of locus-specific *XtX* between the different sets of SNPs revealed that SNPs identified as outliers by our custom Bayesian model neutral were significantly more divergent than neutral SNPs at both regional and microgeographic scales (W statistics = 27866000 and 14542000 respectively, with p-values < 2.2e-16), **Fig. 5**. *BayPass* detected 283 outlier SNPs at the regional scale under a significance threshold of 99%, 131 of them corresponding to SNPs detected as outliers through our custom Bayesian approach (**File S4**). No microgeographic scale outliers were found by BayPass.

A total of 140 outlier SNPs (29.5% out of 475 outlier SNPs) detected through our custom Bayesian framework were located within predicted genes, among which 39 (8.2%) were within exons, and 86 are in the neighborhood (< 5kb) of one or several predicted genes. However, outlier SNPs were not more likely to be located within exons (X^2 = 3.9368, df = 1, p-value = 0.05), within predicted genes (X^2 = 1.0846, df = 1, p-value = 0.30), or in the neighborhood of predicted genes (X^2 = 3.1964, df = 1, p-value = 0.07), compared to any SNP in the entire dataset. Nevertheless, 87 GO-terms were significantly enriched in the subset of genes surrounding outlier SNPs (84 genes) compared to all predicted genes of the entire reference at the regional scale, and 100 GO-terms were significantly enriched in the subset of genes surrounding outlier SNPs (106 genes) at the local scale, **Table S5**, **Files S8 and S9**.

Phenotypic divergence in reciprocal transplants

Patterns of genomic divergence between sites and microhabitats co-occurred with phenotypic divergence between populations occupying different regions and microhabitats. A significant provenance effect was detected for fourteen traits at the regional scale, and for ten traits at microgeographic scale, **Table 1**, **File S10** and **Fig. S7**.

Seedlings originating from the study site of Laussat grew slowly in height but developed wider stems in diameter at early stages (in 2011), but these differences became non-significant after 2012. They however developed significantly more growth units, a trait which became significant in 2014. Leaf properties also varied between regional provenances: seedlings from Laussat established significantly more leaves (significant in 2014) and leaflets (significant in 2012, 2013 and 2014); these traits being associated with a significantly higher leaf area in September 2012. They also displayed lower chlorophyll content and more predation in 2013 and 2014 although they were less attacked by predators in 2012.

The microhabitat of provenance was also significant for several traits: seedlings from hilltops survived better in 2011 but this result became non-significant at later stages. They also grew slowly in both height (September 2012 and 2013) and diameter (September 2012) but were significantly less attacked than seedlings originating from bottomlands (2011, 2012 and 2014). Although significant in 2014, the microhabitat of provenance effect was not very pronounced regarding the number of growth units, leaves and leaflets.

Significant environmental effects of the transplant site and microhabitat were also detected. Fewer traits (four) were significant at the regional scale than at the micro-geographic scale (twelve), thus indicating a stronger influence of microenvironmental variations on seedling growth. Seedling transplanted in Laussat survived better and established more growth units (significant in 2014) and more leaflets that were also more attacked (significant in 2011). At the microgeographic scale, seedlings transplanted in bottomlands less survived (September 2011) but grew faster in both height (September 2011 and April 2012) and diameter (September 2011, April 2012 and September 2013) at early stages. They also established less leaves and leaflets (significant in 2014) that were less attacked (2011, 2012 and 2014) and had a lower chlorophyll content.

Significant Provenance-by-Environment interactions were detected providing additional evidence of local adaptation at the regional scale. A significant 'local advantage' in Laussat was detected on seedlings' survival at early stages (September 2011 and September 2012), on the development of growth units (2013 and 2014), on the number of leaflets (September 2014) and on chlorophyll content. These interactions were accompanied by a 'home advantage' for seedlings originating from Laussat regarding the number of leaflets, the

chlorophyll content, the number of growth units in 2014, and for both provenances regarding the number of growth units in 2013 and survival. Predation measurements, for their part, revealed a local disadvantage in Laussat associated with a home disadvantage for both provenances in 2011 and for seedlings originating from Laussat only in 2013 and 2014.

Discussion

This study analyzed patterns of genomic and phenotypic divergence in wild and undisturbed forest stands of the Amazonian hyperdominant tree *Eperua falcata* in French Guiana. Previous studies (Audigeos et al., 2013; Brousseau et al., 2013, 2015) had already reported signatures of microgeographic divergence both in traits and in molecular data; nevertheless, they were not able to exhaustively investigate its evolutionary drivers, because they were based on datasets restricted to few (e.g., candidate gene approach, Audigeos et al., 2013) or anonymous markers (AFLPs, Brousseau et al., 2015), and the study of microgeographic adaptation lacked a comprehensive survey including both genomic and phenotypic levels. Here, we provide new evidence of microgeographic adaptation associated with (topographic and edaphic) habitat patchiness in this species, by identifying hundreds of SNP loci that display much higher levels of divergence at the microgeographic and regional scale than expected by pure chance, and by identifying sets of phenotypic traits that significantly diverge between populations for each geographic scale.

Patterns of genome-wide divergence.

Genome-wide (supposedly neutral) differentiation was low at both the local and regional geographic scales, ranging from 2 to 3% (**Fig. S5**), similar to what was previously found by Brousseau et al. (2015), who reported F_{ST} values ranging from 2 and 3% between microhabitats within sites, and about 3% between these study sites (based on a screening of about 1,100 AFLPs). In two additional study sites in French Guiana (Paracou and Nouragues), another study (Audigeos et al., 2013) provided comparable estimates: $F_{ST} \cong 1\%$ between study sites and F_{ST} from 1 to 2% between microhabitats within sites. The SNP density that we estimated here (a mean of 3 SNP every 1,000 bp, **Table S2**) was quite low: Neale and Kremer (2011) for example reported that SNP frequency in trees is generally in the range 1 SNP per 100 bp. However, in

our study the stringent SNP filtering probably filtered out many actual polymorphisms, leading to an under-estimation of genome-wide SNP density.

Low genome-wide differentiation suggests that high levels of gene flow in *E. falcata* regularly homogenize (neutral) genome-wide genetic diversity both between regions and between microhabitats within regions. Even though in absolute terms gene flow across sites may be relatively small, as suggested by coalescent modeling, rare long-distance seed dispersal is a common property of tree populations, and appears to have a disproportionate effect on genetic connectivity, potentially even facilitating local adaptation (Kremer et al., 2012). In general, the amount of divergence observed here is comparable to what has been found in more intensively studied temperate or tropical tree species (e.g., Alberto et al., 2010; Dick & Heuertz, 2008; Scalfi et al., 2014). The consequences of recurring inter-habitat gene flow on population adaptability could be paradoxical. On one hand, extensive gene flow can allow the spread of advantageous alleles between populations, thus granting high adaptive potential for populations established in heterogeneous environments. On the other hand, gene flow may also counteract the effect of natural selection, with migration load halting adaptive divergence of populations into locally adapted subpopulations (De Kort, Vandepitte, Mergeay, Mijnsbrugge, & Honnay, 2015; Le Corre & Kremer, 2012; Lenormand, 2002; Pujol et al., 2018). The deleterious effect of gene flow on adaptive divergence could explain the low proportion of outliers detected (see below); another explanation may be the highly polygenic nature of adaptive traits, which would prevent individual underlying genes to diverge in frequency (Le Corre & Kremer, 2012). Nevertheless, high gene flow (and resulting low genome-wide differentiation) did not completely swamp the signal of adaptive divergence between populations.

Local adaptation across geographic scales.

Locus-specific (putatively adaptive) departures from genome-wide divergence were detected for 0.49% of the 97,062 SNPs analyzed. Results at the single SNP level were confirmed by trends at the contig level, showing similar trends of microgeographic differentiation. Moreover, a comprehensive reciprocal transplant experiment also corroborated evidence of local adaptation by revealing phenotypic differences at both regional and microgeographic scales in seedlings at early developmental stages, when natural selection is expected to be stronger (Petit & Hampe, 2006).

Significantly more outliers were detected at the regional than at the microgeographic scale: 290 *versus* 185 outlier SNPs ($X^2 = 22.826$, df = 1, $p = 1.77 \times 10^{-6}$). This result can be

445 explained as the logical consequence of an order of magnitude difference in the geographical
446 distances between populations: 300 meters between microhabitats within sites versus 300 km
447 between study sites. Indeed, ecological conditions are probably more variable, and divergent
448 selection stronger, at the regional scale (mainly driven by climatic differences associated with
449 different rainfall regimes) than divergent selective pressures between microhabitats separated
450 by hundreds of meters (driven by differences in soil properties), and the effect of divergence
451 would be reinforced by more limited gene flow. The proportion of outliers detected over
452 microgeographic scales, although smaller, is nevertheless relevant. Because the Bayesian model
453 focuses on replicated footprints of adaptation only, it is extremely unlikely that the detected
454 outliers could have been structured between microhabitats in two different regions
455 independently purely by chance (Savolainen, Lascoux, & Merila, 2013). Moreover, no outlier
456 was detected jointly at both regional and microgeographic spatial scales, as shown by the heart-
457 shaped joint distribution of regional- and local-scale locus-specific parameters (**Fig. 3**). Instead,
458 outliers were always specific to only one spatial scale, suggesting that different portions of the
459 genome respond to different types of environmental constraints. Reciprocal transplants also
460 indicated that different sets of traits diverge between populations at regional and
461 microgeographic scales. Although it is impossible to link the analyzed phenotypic traits to one
462 or several SNPs detected without a comprehensive association study, these results suggest that
463 different environmental drivers would be at stake and that, in future studies, local adaptation in
464 wild populations should be investigated by integrating different spatial scales, including
465 microgeographic habitat patchiness, into the sampling design.
466
467 *Methodological considerations*
468
469 Whole-genome sequencing of population pools has opened a new window of
470 opportunity, especially for non-model species. Their analysis however remains complicated as
471 it raises many methodological and statistical challenges, even more in the case of complex –
472 presently *replicated* – sampling designs.
473 The Bayesian modelling framework implemented here was specifically designed to
474 detect replicated footprints of divergent selection between microhabitats under an explicit
475 hierarchical model of neutral divergence both between and within study sites, thus taking into
476 account variable levels of gene flow across spatial scales. This custom framework however
477 suffers from several limitations, including computation time, estimation uncertainty and
478 detection power.

479 Computation time limitation was overcame by splitting the SNP dataset into ten
480 independent subsets. The fact that neutral parameters were similar across subsets suggests that
481 locus-specific parameters were estimated under similar conditions of hierarchical neutral
482 divergence between populations across subsets (**Fig. S6**). This is a natural consequence of the
483 largely fragmented reference assembled *de novo*: SNPs were randomly scattered across the
484 genome and they can therefore be considered as independent. Although SNPs located on the
485 same scaffolds are physically linked, they did not cause major discrepancies in the estimation
486 of neutral parameters across subsets.

487 Assessing the chances of F_{ST}-based selection tests to detect false-positives and to miss
488 true-positive outliers is another challenge, and this issue strongly depends on the size of the
489 dataset. While the power of our custom Bayesian model was not estimated for very large
490 numbers of loci by Brousseau et al. (2016), we included in the preset study an additional
491 correction for multiple-testing (with a FDR threshold of 10%) to deal with the large number of
492 SNPs analyzed. Although linking directly FDR with FNDR, and subsequently with the power
493 of our model is difficult, there exist theoretical and numerical expectations of the False Non-
494 Discovery Rates (FNDR) of multiple tests based on the chosen FDR threshold. In particular,
495 Genovese & Wasserman (2002) show that, for an FDR threshold of 5%, FNDR is always below
496 20% for cases where most tests are expected to be true negatives. We can expect that only a
497 small minority of loci is under divergent selection at our small phylogenetic and spatial scales.
498 Finally, Sarkar (2006) defines the power $\pi_0 = 1 - FDR - FNDR$; therefore, in our case, power
499 should be in the order of 1- 0.1 - O(0.2), that is, no larger than 70%, and probably somewhat
500 smaller.

501 Another source of bias lies in the inference of population-specific allele frequencies and
502 genetic differentiation from read counts in population pools because of the double sampling
503 uncertainties, (1) when sampling individuals from a population, and (2) when sampling reads
504 from the pool of haploid sequences during sequencing (Ferretti, Ramos-Onsins, & Pérez-
505 Enciso, 2013; Hivert, Leblois, Petit, Gautier, & Vitalis, 2018). This limit was tackled by
506 applying a library-level coverage threshold to detect SNPs: a threshold of 20X was used, that
507 is equal the number of diploid individuals (or half the haploid sample size) in each population
508 pool. Applying library-level coverage threshold to infer library-specific allele frequency is
509 mandatory to reduce the probability of missing alleles at low frequency, although fluctuations
510 in the sampling process are unavoidable and sequencing all distinct lineages in a pool of 20
511 diploid individuals remains unlikely even at 40X coverage (Ferretti et al., 2013). Aware of these

512 limitations, uncertainties in allele frequencies were however further integrated in estimates of
513 pairwise genetic differentiation (G_{ST}) between libraries under the Bayesian framework.
514 It is also important to stress that our Bayesian approach may miss footprints of polygenic
515 adaptation, if adaptation acts in different ways in the two study sites (through different selection
516 strengths and/or different loci affected) which is known to be common in wild trees (Berg &
517 Coop, 2014; Harrisson et al., 2017; Lind et al., 2017). Even if the Bayesian model employed
518 here proved its robustness to detect footprints of polygenic adaptation caused by weak selection
519 strengths (s=0.05) from genotype data in simulations (Brousseau et al., 2016), it may still miss
520 footprints of polygenic adaptation if the strength of selection is variable among loci (Sork et al.,
521 2013). Indeed, F_{ST}-based methods are more prone to detect adaptation caused by loci of large
522 effects than by loci with small effects, especially since adaptation occurs with ongoing gene
523 flow (Hedrick, 2006; Le Corre & Kremer, 2012; Savolainen et al., 2013; Yeaman & Whitlock,
524 2011). In other words, the set of outliers detected is probably enriched with loci of large effects.
525 Moreover, it is possible that the genetic architecture of traits selected at different geographical
526 scales differs, thus explaining both the differences in numbers of outliers detected and lack of
527 overlap. Indeed, the reciprocal transplant experiment showed that different sets of traits
528 inherently diverged between sites and between microhabitats within sites. It is also likely that
529 traits closer to metabolic pathways (such as chlorophyll content, possibly chemical defense
530 against grazers) are less polygenic than complex growth traits, and that therefore the chances
531 of identifying single-locus outliers vary between these different sets of traits.
532 Beyond theoretical considerations, the results obtained were further compared with two
533 other widespread selection tests suitable for analyzing Pool-Seq data, G2D and *BayPass*,
534 respectively. Applying different methods often produce partially different results, because of
535 inherent conceptual and mathematical differences in methods formulation and implementation.
536 For example, G2D can detect replicated footprints of selection taking into account the
537 demographic history of the studied populations at the contig-level, not at the SNP level.
538 *BayPass v2.1*, for its part, takes into account complex neutral structuring through the inference
539 of a genome-wide variance-covariance matrix and allows detecting SNPs that are over-
540 differentiated between standing populations and the ancestral one, without however
541 distinguishing outliers caused by crossed-factors nor repeated binary contrasts. The identity and
542 quantity of outlier contigs or SNPs detected in the present study thus vary between these
543 different selection tests. Comparing the distribution of locus-specific differentiation (G-values
544 and *XtX*) between the different sets of SNPs detected by our custom Bayesian framework
545 however revealed a shared trend: the extent of differentiation was always higher for outlier

546 SNPs than for neutral SNPs, indicating that the sets of SNPs detected as outliers at both regional
547 and microgeographic scales were enriched in SNPs influenced by divergent selection although
548 they possibly hide false-positive outliers.
549
550 *Microgeographic adaptation*
551
552 This study was primarily designed to test the hypothesis of replicated microgeographic
553 adaptation within two independent study sites, and the regional scale was thus taken into
554 account to avoid confounding footprints of adaptation between large-scale and small-scale
555 factors, and to compare the strength of adaptive divergence between spatial scales (in terms of
556 number of SNPs affected). For these reasons, our interpretation thus focuses on footprints of
557 microgeographic adaptation only thereafter.
558 Although detecting adaptive divergence between regions at geographic scales on the
559 order of hundreds kilometers is not surprising, detecting patterns of microgeographic adaptive
560 divergence on the order of hundreds meters is less common in tree populations. By providing
561 evidence of adaptation to microgeographic habitat patchiness at both genomic and phenotypic
562 levels, our results agree with previous studies showing soil variation as a driver of adaptation
563 in trees and, more specifically, in the Neotropics (Audigeos et al., 2013; Brousseau et al., 2015).
564 Topography-linked microgeographic patchiness is likely driven by a complex
565 interaction of abiotic and biotic factors (including soil types (Misiewicz & Fine, 2014), water-
566 logging and flooding events, water availability (Clark, Palmer, & Clark, 1999), seasonal soil
567 drought (Markewitz, Devine, Davidson, Brando, & Nepstad, 2010), soil fertility (John et al.,
568 2007), light availability (Ferry, Morneau, Bontemps, Blanc, & Freycon, 2010), competition and
569 predation (Fine et al., 2004)). At the community level, the study by Fortunel et al. (2016)
570 already suggested a mechanism through which microgeographic adaptation may take place:
571 seedling mortality of *terra firme* species was high when planted in seasonally flooded habitats.
572 At the intra-specific level, previous studies have reported examples of adaptation to soil nutrient
573 composition such as potassium deficiency in *Cornus florida* (Cornaceae) (Pais, Whetten, &
574 Xiang, 2017), to aluminum and phosphate ion concentrations in *Pinus contorta* (Pinaceae)
575 (Eckert, Shahi, Datwyler, & Neale, 2012), to substrate age in *Metrosideros polymorpha*
576 (Myrtaceae) (Izuno et al., 2017), to soil types in *Abies religiosa* (Pinaceae) (Méndez-González,
577 Jardón-Barbolla, & Jaramillo-Correa, 2017), or to soil water availability in *Pinus albicaulis*
578 (Pinaceae) (Lind et al., 2017). Preliminary evidence of microgeographic adaptation has been
579 recently reported in Neotropical tree species, including *E. falcata* (Audigeos et al., 2013;

580 Brousseau et al., 2013, 2015; Torroba-Balmori et al., 2017). We provide here additional
581 evidence of microgeographic adaptation in *E. falcata* at the genomic scale, also corroborated
582 by significant effects of the microhabitat of provenance at the phenotypic level.

583 It is, however, difficult to identify the exact genomic targets of natural selection,
584 particularly in non-model species. First, because linking complex phenotypes (growth and leaf
585 traits) with the SNPs analyzed is impossible without an accurate association genomics
586 approach. Second, because of potential hitchhiking between causal mutations and the analyzed
587 SNPs, and because of the lack of chromosome-level genome assembly as well as no exhaustive
588 genome-wide estimates of linkage disequilibrium in this species (but linkage disequilibrium is
589 nearly zero beyond few hundred base pairs in coding sequences, Audigeos et al., 2013). Despite
590 these limitations, the identity of the genes neighboring outlier SNPs may provide useful
591 information, as many 'divergence' outlier SNPs detected here were located within or close to
592 (< 5 kb) protein-coding genes involved in a variety of biological functions (**Files S8 and S9**),
593 with 106 genes neighboring a genomic footprint of microgeographic divergence (i.e. one or
594 several outlier SNPs) and 100 Gene Ontology terms (corresponding to 49 genes) significantly
595 enriched in the subset of genes neighboring outlier SNPs, **Table S5**. Some of them are
596 particularly meaningful regarding their biological significance in terms of survival, growth,
597 defense and reproduction (**Files S8 and S9**). Functional annotation and subsequent enrichment
598 tests are however to be taken with caution. Although they can provide useful clues about the
599 biological functions of the genes identified, they do not necessarily provide a proof of the results
600 functional meaning (Pavlidis, Jensen, Stephan, & Stamatakis, 2012).

601 We detected several genes involved in plant response to stress close to SNPs detected
602 as outliers at the microgeographic scale. For example, a glutaredoxin-C11 (predgene_010110)
603 involved in managing the level of reactive oxygen species (ROS) in plant cells (Mittler,
604 Vanderauwera, Gollery, & Van Breusegem, 2004), a H_2O_2-responsive gene encoding a
605 metallopeptidase M24 family protein which is up-regulated under stress in *Arabidopsis thaliana*
606 (Ordoñez et al., 2014), and a gene encoding a plant cysteine oxidase 4-like (predgene 003921)
607 involved in response to flooding and soil hypoxia (Weits et al., 2014; White et al., 2017) were
608 detected. Oxidative stress is very common in bottomlands, where water-logging and seasonal
609 flooding events cause soil hypoxia and oxygen deprivation in root cells, leading to the
610 production of ROS (Blokhina, Virolainen, & Fagerstedt, 2003). This observation is concordant
611 with a previous candidate-gene approach in which a catalase (i.e. a gene involved in ROS
612 detoxification) has been reported under selection in this species (Audigeos et al., 2013). We
613 also detected five genes with possible implications in lipid biosynthesis (predgenes 006075,

614 019500, 002577, 001066, 000689), as for example a palmitoyl-acyl carrier protein thioesterase
615 which is known to play an essential role for plant growth and seed development (Bonaventure,
616 Salas, Pollard, & Ohlrogge, 2003; Dörmann, Voelker, & Ohlrogge, 2000). Lipid metabolism is
617 of high importance for plant development, from early germination to reproduction. Indeed,
618 lipids are the primary material for cutin, and they are largely involved in stress signaling and
619 resistance (Okazaki & Saito, 2014). We also found two genes involved in proline amino acid
620 biosynthesis (predgenes 007023 and 007024), which is accumulated in response to drought
621 stress in plants (Fu, Ma, Chen, Gu, & Gong, 2018). A calmodulin-like protein (probable
622 CML48, predgene 022002) was also detected in the neighborhood of a microgeographic outlier.
623 Calmodulin and calmodulin-like plant calcium sensors have evolved unique functions of high
624 importance in the regulation of plant development and plant response to stress (Perochon,
625 Aldon, Galaud, & Ranty, 2011). We also found a probable cellulose synthase A (predgene
626 000745) whose mutations can alter morphology or induce resistance to cellulose biosynthesis
627 inhibitors as reported in *A. thaliana* (Richmond, 2000), a gene for plant metal tolerance
628 (predgene 014347), and a mRNA-decapping enzyme-like (predgene 011170) which is of
629 importance for early seedling development (Goeres et al., 2007).

630 Two genes mediating biotic interactions - through plant defense against pathogens and
631 recognition of symbiotic bacteria – were also detected as potential targets of microgeographic
632 adaptation. In particular, plant chemical defense against herbivores and pathogens is likely to
633 play a key role in microgeographic adaptation; previous studies have experimentally
634 demonstrated that a trade-off between growth and antiherbivore defenses strengthens habitat
635 specialization between congeneric species within large Neotropical plant families (e.g.,
636 Annonaceae, Burseraceae, Euphorbiaceae, Fabaceae, Malvaceae) (Fine et al., 2004, 2013,
637 2006). Here, we detected a chitinase (predgene 000254), a family of enzymes known for their
638 role in resistance to pathogens (Grover, 2012; Punja & Zhang, 1993). This is fully consistent
639 with the finding of significant provenance effects of seedling predation (estimated by the
640 number of leaflets attacked) in the present study, which suggests biotic selective pressures
641 exerted by herbivore communities should contribute to microgeographic adaptation. We also
642 discovered a gene involved in recognition of symbiotic bacteria (LysM receptor-like kinase,
643 predgene 002974) which enables the model legume *Lotus japonicus* to recognize its bacterial
644 symbiont (Madsen et al., 2003; Radutoiu et al., 2003; Ryals et al., 2008). Symbiotic nitrogen
645 fixation is a specific feature of legumes, and nutrient availability largely varies across
646 microhabitats (Ferry et al., 2010). The presence of genes involved in the perception of symbiotic
647 bacteria close to genomic footprints of microgeographic adaptation in this supposedly non-

648 nodulating legume (Villadas et al., 2007) is particularly interesting. However, accurate
649 functional characterizations of the proteins surrounding outlier SNPs are required to draw
650 robust conclusions about the physiological processes involved in microgeographic adaptation.
651
652 *Microevolution in the (Neo)tropics.*
653
654 Although largely exploratory, our results open a new window into the evolutionary
655 processes underlying the maintenance of genetic and functional diversity, and support the
656 hypothesis of niche evolution as a driver of habitat specialization and diversification (Leibold,
657 2008). Microgeographic adaptation driven by a complex interaction of abiotic and biotic factors
658 may be an initial step towards habitat specialization and ecological (sympatric) speciation by
659 trees (Feder, Egan, & Nosil, 2012; Savolainen et al., 2006). While it is true that
660 microgeographic adaptation is not at all restricted to (Neo)tropical trees (e.g., Lind et al., 2017;
661 Scotti et al., 2016), the extreme patchiness of micro-environmental conditions and the overall
662 levels of competition for resources and space observed in tropical ecosystems may amplify the
663 specialization of trees to specific microhabitats. In this study we provide support for this process
664 at the intra-populational level in an Amazonian tree species. However, the efficacy of
665 microgeographic adaptive divergence at a small but non-negligible fraction of loci in driving
666 ecological speciation (as opposed to the effect of genome-wide neutral divergence, caused by
667 strong restrictions in gene flow, in assisting allopatric speciation) requires further investigation.
668 In particular, a comprehensive genome-wide association study in our reciprocal transplant
669 would provide detailed information on the functional role of the SNPs and genes potentially
670 under selection.
671 It is also possible that microgeographic adaptation contributes to the generalist and
672 hyperdominant distribution of *E. falcata* (ter Steege et al., 2013). In this view, local adaptation
673 to habitat patches would maintain multiple adaptation optima (i.e. multiple genetic
674 combinations of alleles at different loci), thus pre-adapting the populations to a multitude of
675 environmental situations. Indeed, two evolutionary processes can theoretically allow
676 populations to cope with environmental changes, such as global climate change, and avoid
677 extinction (Aitken, Yeaman, Holliday, Wang, & Curtis-McLane, 2008; Harrisson et al., 2014):
678 populations can migrate to new areas to track ecological niches, causing shifts in the species'
679 distribution, and/or adapt locally to the new conditions through the action of natural selection
680 (Hoffmann & Sgro, 2011). These processes are likely to occur simultaneously and populations'
681 response to global climate change can be viewed as a 'race where populations are tracking the

682 moving optima both in space and time' (Kremer et al., 2012). From the standpoint of the entire
683 metapopulation of *E. falcata* in the region, local adaptation to multiple optima would be
684 tantamount to population-level balancing selection, which is considered to be a major
685 component of adaptive potential (Barrett & Schluter, 2008; Delph & Kelly, 2014). Such a
686 reservoir of adaptively useful genetic diversity may help the species deal with environmental
687 challenges.

688 To conclude, here we provide new, genomic-scale, evidence of divergent selection at
689 both regional and microgeographic scales in the Amazonian hyperdominant *Eperua falcata*,
690 suggesting that microgeographic environmental heterogeneity caused by variations in
691 topography and edaphic factors creates the conditions for local adaptation in lowland
692 Neotropical rainforests. We notably detected signals of adaptive divergence at many SNPs,
693 whose surrounding genes may ultimately constitute the genomic basis for local adaptation.
694 Accurate functional annotation further provided some interesting clues as to the biological
695 processes targeted by divergent selection, emphasizing the potential roles of abiotic stresses
696 related to water-logging and drought, and of biotic interactions, in driving local adaptation. We
697 firmly believe that new advances in next-generation sequencing technologies and in Bayesian
698 methods will contribute to fill the gaps in our understanding of trees evolution in the Neotropics,
699 and we strongly recommend the integration of different spatial scales (including
700 microgeographic habitat patchiness) in future studies addressing the processes of adaptation
701 and diversification in lowland rainforests of Amazonia.

702

703 # Acknowledgements

704 This study has been funded by the ANR '*BIOADAPT FLAG'* (ref. ANR-12-ADAPT-0007-01)
705 and benefited from an '*Investissement d'Avenir*' grant (CEBA, ref. ANR-10-LABX-25-01),
706 both managed by the French *Agence Nationale de la Recherche*. Louise Brousseau has been
707 funded by a young investigator grant '*Contrat Jeune Scientifique*' (CJS) of the *French National*
708 *Institute for Agronomical Research* (INRA). We are grateful to the *Genotoul bioinformatics*
709 *platform Toulouse Midi-Pyrenees* (Bioinfo Genotoul) for providing computing and storage
710 resources. We thank Saint-Omer Cazal (INRA, UMR EcoFoG) and Julien Engel (IRD, UMR
711 AMAP) for technical assistance and botanical identifications. We thank Bruno Ferry
712 (AgroParisTech, UFR FAM), Hendik Davi (INRA, URFM) for sharing soil and functional data
713 of adult tree populations. We thank Dr Anna Stavrinides for helpful discussions about the

functional interpretation of the results. We are also grateful to Victoria Sork, Andrew Eckert and two anonymous reviewers for their very useful comments on this study.

Author contribution

LB, ED and IS designed the experiment. LB set up the reciprocal transplant experiment and analyzed related phenotypic data under the supervision of IS and ED. LB sampled leaf material, extracted and conditioned DNA, implemented the bioinformatics pipeline and analyzed SNPs data with the custom Bayesian framework and *BayPass* under the supervision of IS and GVV. IS computed basic diversity statistics and implemented G2D analyses. All authors wrote the article.

References

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, *1*(1), 95–111. doi: 10.1111/j.1752-4571.2007.00013.x

Alberto, F., Niort, J., Derory, J., Lepais, O., Vitalis, R., Galop, D., & Kremer, A. (2010). Population differentiation of sessile oak at the altitudinal front of migration in the French Pyrenees. *Molecular Ecology*, *19*(13), 2626–2639. doi: 10.1111/j.1365-294X.2010.04631.x

Antonovics, J. (2006). Evolution in closely adjacent plant populations X: long-term persistence of prereproductive isolation at a mine boundary. *Heredity*, *97*(1), 33–37. doi: 10.1038/sj.hdy.6800835

Antonovics, Janis, & Bradshaw, A. D. (1970). EVOLUTION IN CLOSELY ADJACENT PLANT POPULATIONS. VIII. Clinal patterns at a mine boundary. *Heredity*, *25*, 349–362.

Audigeos, D., Brousseau, L., Traissac, S., Scotti-Saintagne, C., & Scotti, I. (2013). Molecular divergence in tropical tree populations occupying environmental mosaics. *Journal of Evolutionary Biology*, *26*(3), 529–544. doi: 10.1111/jeb.12069

Baraloto, C., Morneau, F., Bonal, D., Blanc, L., & Ferry, B. (2007). Seasonal water stress tolerance and habitat associations within four neotropical tree genera. *Ecology*, *88*(2), 478–489. doi: 10.1890/0012-9658(2007)88[478:SWSTAH]2.0.CO;2

Barrett, R. D. H., & Schluter, D. (2008). Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, *23*(1), 38–44. doi: 10.1016/j.tree.2007.09.008

Berg, J. J., & Coop, G. (2014). A Population Genetic Signal of Polygenic Adaptation. *PLoS Genetics*, *10*(8), e1004412. doi: 10.1371/journal.pgen.1004412

Blokhina, O., Virolainen, E., & Fagerstedt, K. V. (2003). Antioxidants, oxidative damage and oxygen deprivation stress: A review. *Annals of Botany*, *91*(2), 179–194. doi: 10.1093/aob/mcf118

Bonal, D., Atger, C., Barigah, T. S., Ferhi, A., Guehl, J. M., & Ferry, B. (2000). Water acquisition patterns of two wet tropical canopy tree species of French Guiana as inferred from H218O extraction profiles. *Ann For Sci*, *57*, 717–724.

Bonal, D., Barigah, T. S., Granier, A., & Guel, J. M. (2000). Late-stage canopy tree species with extremely low d13C and high stomatal sensitivity to seasonal soil drought in the tropical rainforest of French Guiana. *Plant, Cell and Environment*, *23*, 445–449.

Bonal, D., Guehl, J. M., & Champenoux, F. (2001). Contrasting patterns of leaf water potential and gas exchange responses to drought in seedlings of tropical rainforest species. *Functional Ecology*, *15*, 490–496.

Bonaventure, G., Salas, J. J., Pollard, M. R., & Ohlrogge, J. B. (2003). Disruption of the FATB gene in Arabidopsis demonstrates an essential role of saturated fatty acids in plant growth. *The Plant Cell*, *15*(4), 1020–1033. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12671095

Bradshaw, A. (1960). Population differentiation in *Agrostis tenuis* Sibth. III. Populations in varied environments. *New Phytol*, *59*, 92–103.

Brousseau, L., Bonal, D., Cigna, J., & Scotti, I. (2013). Highly local environmental variability promotes intrapopulation divergence of quantitative traits: An example from tropical rain forest trees. *Annals of Botany*, *112*(6), 1169–1179. doi: 10.1093/aob/mct176

Brousseau, L., Foll, M., Scotti-Saintagne, C., & Scotti, I. (2015). Neutral and adaptive drivers of microgeographic genetic divergence within continuous populations: The case of the neotropical tree *Eperua falcata* (aubl.). *PLoS ONE*, *10*(3), e0121394. doi: 10.1371/journal.pone.0121394

Brousseau, L., Postolache, D., Lascoux, M., Drouzas, A. D., Källman, T., Leonarduzzi, C., … Vendramin, G. G. (2016). Local Adaptation in European Firs Assessed through Extensive Sampling across Altitudinal Gradients in Southern Europe. *PLOS ONE*, *11*(7), e0158216. doi: 10.1371/journal.pone.0158216

Bulmer, M. G. (1972). Multiple Niche Polymorphism. *American Naturalist*, *106*, 254–257.

Clark, D. B., Palmer, M. W., & Clark, D. A. (1999). Edaphic fators and the landscape-scale distribution of tropical rain forest trees. *Ecology*, *80*, 2662–2675.

Cowan, R. S. (1975). *A monograph of the genus Eperua (Leguminosae-Caesalpinioideae).* Smithsonian Institution Press.

De Kort, H., Vandepitte, K., Mergeay, J., Mijnsbrugge, K. V., & Honnay, O. (2015). The population genomic signature of environmental selection in the widespread insect-pollinated tree species *Frangula alnus* at different geographical scales. *Heredity*, *115*(5), 415–425. doi: 10.1038/hdy.2015.41

De Kort, Hanne, & Honnay, O. (2017). Assessing Evolutionary Potential in Tree Species Through Ecology-Informed Genome Screening. In P. Pontarotti (Ed.), *Evolutionary Biology: Self/Nonself Evolution, Species and Complex Traits Evolution, Methods and Concepts* (pp. 313–327). doi: 10.1007/978-3-319-61569-1_17

Delph, L. F., & Kelly, J. K. (2014). On the importance of balancing selection in plants. *New Phytologist*, *201*(1), 45–56. doi: 10.1111/nph.12441

Dick, C. W., & Heuertz, M. (2008). The complex biogeographic history of a widespread tropical tree species. *Evolution*, *62*(11), 2760–2774. doi: 10.1111/j.1558-5646.2008.00506.x

Dörmann, P., Voelker, T. A., & Ohlrogge, J. B. (2000). Accumulation of palmitate in Arabidopsis mediated by the acyl-acyl carrier protein thioesterase FATB1. *Plant Physiology*, *123*(2), 637–644. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10859193

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem Bull*, *19*, 11–15.

Eckert, A. J., Maloney, P. E., Vogler, D. R., Jensen, C. E., Mix, A. D., & Neale, D. B. (2015). Local adaptation at fine spatial scales: an example from sugar pine (*Pinus lambertiana*, Pinaceae). *Tree Genetics & Genomes*, *11*(3), 42. doi: 10.1007/s11295-015-0863-0

Eckert, A. J., Shahi, H., Datwyler, S. L., & Neale, D. B. (2012). Spatially variable natural selection and the divergence between parapatric subspecies of lodgepole pine (*Pinus contorta*, Pinaceae). *American Journal of Botany*, *99*(8), 1323–1334. doi: 10.3732/ajb.1200055

Eckert, A. J., van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S. C., … Neale, D. B. (2010). Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics*, *185*(3), 969–982. doi: 10.1534/genetics.110.115543

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342–350.

Ferretti, L., Ramos-Onsins, S. E., & Pérez-Enciso, M. (2013). Population genomics from pool sequencing. *Molecular Ecology*, *22*(22), 5561–5576. doi: 10.1111/mec.12522

Ferry, B., Morneau, F., Bontemps, J.-D. D., Blanc, L., & Freycon, V. (2010). Higher treefall rates on slopes and waterlogged soils result in lower stand biomass and productivity in a tropical rain forest. *Journal of Ecology*, *98*(1), 106–116. doi: 10.1111/j.1365-2745.2009.01604.x

Fetter, K. C., Gugger, P. F., & Keller, S. R. (2017). *Landscape Genomics of Angiosperm Trees: From Historic Roots to Discovering New Branches of Adaptive Evolution*. doi: 10.1007/7397_2016_19

Fine, P. V. A. (2015). Ecological and Evolutionary Drivers of Geographic Variation in Species Diversity. *Annual Review of Ecology, Evolution, and Systematics*, *46*(1), 369–392. doi: doi:10.1146/annurev-ecolsys-112414-054102

Fine, P. V. A., Daly, D. C., & Cameron, K. M. (2005). The contribution of edaphic heterogeneity to the evolution and diversity of Burceraceae trees in the Western Amazon. *Evolution*, *59*(7), 1464-1478. doi: 10.1111/j.0014-3820.2005.tb01796.x

Fine, P. V. A., Mesones, I., & Coley, P. D. (2004). Herbivores promote habitat specialization by trees in Amazonian forests. *Science*, *305*(5684), 663–665. doi: 10.1126/science.1098982

Fine, P. V. A., Metz, M. R., Lokvam, J., Mesones, I., Zuniga, J. M. A., Lamarre, G. P. A., … Baraloto, C. (2013). Insect herbivores, chemical innovation, and the evolution of habitat specialization in Amazonian trees. *Ecology*, *94*(8), 1764–1775. doi: 10.1890/12-1920.1

Fine, P. V. A., Miller, Z. J., Mesones, I., Irazuzta, S., Appel, H. M., Stevens, M. H. H., … Coley, P. D. (2006). The growth-defense trade-off and habitat specialization by plants in Amazonian forests. *Ecology*, *87*(sp7), S150–S162. doi: 10.1890/0012-9658(2006)87[150:TGTAHS]2.0.CO;2

Fortunel, C., Paine, C. E. T. T., Fine, P. V. A. A., Mesones, I., Goret, J.-Y. Y., Burban, B., … Baraloto, C. (2016). There's no place like home: seedling mortality contributes to the habitat specialisation of tree species across Amazonia. *Ecology Letters*, *19*(10), 1256–1266. doi: 10.1111/ele.12661

Fu, Y., Ma, H., Chen, S., Gu, T., & Gong, J. (2018). Control of proline accumulation under drought via a novel pathway comprising the histone methylase CAU1 and the transcription factor ANAC055. *Journal of Experimental Botany*, *69*(3), 579–588. doi: 10.1093/jxb/erx419

Fustier, M.-A., Brandenburg, J.-T., Boitard, S., Lapeyronnie, J., Eguiarte, L. E., Vigouroux,

Y., … Tenaillon, M. I. (2017). Signatures of local adaptation in lowland and highland teosintes from whole-genome sequencing of pooled samples. *Molecular Ecology*, *26*(10), 2738–2756. doi: 10.1111/mec.14082

Gautier, M. (2015). Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics*, *201*(4), 1555–1579. doi: 10.1534/genetics.115.181453

Genovese, C., & Wasserman, L. (2002). Operating characteristics and extensions of the false discovery rate procedure. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(3), 499–517. doi: 10.1111/1467-9868.00347

Goeres, D. C., Van Norman, J. M., Zhang, W., Fauver, N. A., Spencer, M. Lou, & Sieburth, L. E. (2007). Components of the Arabidopsis mRNA decapping complex are required for early seedling development. *The Plant Cell*, *19*(5), 1549–1564. doi: 10.1105/tpc.106.047621

Gould, B., McCouch, S., & Geber, M. (2014). Variation in soil aluminium tolerance genes is associated with local adaptation to soils at the Park Grass Experiment. *Molecular Ecology*, *23*(24), 6058–6072. doi: 10.1111/mec.12893

Grover, A. (2012). Plant Chitinases: Genetic Diversity and Physiological Roles. *Critical Reviews in Plant Sciences*, *31*(1), 57–73. doi: 10.1080/07352689.2011.616043

Hamrick, J. L., & Godt, M. J. W. (1990). Allozyme diversity in plant species. *Plant Population Genetics, Breeding, and Genetic Resources.*, 43–63. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19901612624

Harrisson, K. A., Amish, S. J., Pavlova, A., Narum, S. R., Telonis-Scott, M., Rourke, M. L., … Sunnucks, P. (2017). Signatures of polygenic adaptation associated with climate across the range of a threatened fish species with high genetic connectivity. *Molecular Ecology*, *26*(22), 6253–6269. doi: 10.1111/mec.14368

Harrisson, K. A., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, *7*(9), 1008–1025. doi: 10.1111/eva.12149

Hedrick, P. W. (2006). Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics. *Annual Review of Ecology, Evolution, and Systematics*, *37*(1), 67–93.

Hendry, A. P., Day, T., & Taylor, E. B. (2001). Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. *Evolution; International Journal of Organic Evolution*, *55*(3), 459–466. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11327154

Hivert, V., Leblois, R., Petit, E. J., Gautier, M., & Vitalis, R. (2018). Measuring Genetic Differentiation from Pool-seq Data. *Genetics*, *210*(1), 315–330. doi: 10.1534/genetics.118.300900

Hoffmann, A. A., & Sgro, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479–485.

Izuno, A., Kitayama, K., Onoda, Y., Tsujii, Y., Hatakeyama, M., Nagano, A. J., … Isagi, Y. (2017). The population genomic signature of environmental association and gene flow in an ecologically divergent tree species *Metrosideros polymorpha* (Myrtaceae). *Molecular Ecology*, *26*(6), 1515–1532. doi: 10.1111/mec.14016

Jain, S. K., & Bradshaw, A. D. (1966). Evolutionary divergence among adjacent plant populations. I. The evidence and its theoretical analysis. *Heredity*, *21*(3), 407–441. Retrieved from http://dx.doi.org/10.1038/hdy.1966.42

John, R., Dalling, J. W., Harms, K. E., Yavitt, J. B., Stallard, R. F., Mirabello, M., … Foster, R. B. (2007). Soil nutrients influence spatial distributions of tropical tree species. *Proceedings of the National Academy of Sciences of USA*, *104*(3), 864–869.

Kraft, N. J., Valencia, R., & Ackerly, D. D. (2008). Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, *322*, 580–582.

Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., … Schueler, S. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, *15*(4), 378–392. doi: 10.1111/j.1461-0248.2012.01746.x

Latreille, A. C., & Pichot, C. (2017). Local-scale diversity and adaptation along elevational gradients assessed by reciprocal transplant experiments: lack of local adaptation in silver fir populations. *Annals of Forest Science*, *74*(4), 77. doi: 10.1007/s13595-017-0673-7

Le Corre, V., & Kremer, A. (2012). The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, *21*(7), 1548–1566. doi: 10.1111/j.1365-294X.2012.05479.x

Leibold, M. A. (2008). Ecology: Return of the niche. *Nature*, *454*(7200), 39–41. doi: 10.1038/454039a

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, *17*(4), 183–189. doi: 10.1016/S0169-5347(02)02497-7

Lind, B. M., Friedline, C. J., Wegrzyn, J. L., Maloney, P. E., Vogler, D. R., Neale, D. B., & Eckert, A. J. (2017). Water availability drives signatures of local adaptation in whitebark pine (Pinus albicaulis Engelm.) across fine spatial scales of the Lake Tahoe Basin, USA.

Molecular Ecology, *26*(12), 3168–3185. doi: 10.1111/mec.14106

Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., … Stougaard, J. (2003). A receptor kinase gene of the LysM type is involved in legumeperception of rhizobial signals. *Nature*, *425*(6958), 637–640. doi: 10.1038/nature02045

Markewitz, D., Devine, S., Davidson, E. A., Brando, P., & Nepstad, D. C. (2010). Soil moisture depletion under simulated drought in the Amazon: Impacts on deep root uptake. *New Phytologist*, *187*(3), 592–607. doi: 10.1111/j.1469-8137.2010.03391.x

Méndez-González, I. D., Jardón-Barbolla, L., & Jaramillo-Correa, J. P. (2017). Differential landscape effects on the fine-scale genetic structure of populations of a montane conifer from central Mexico. *Tree Genetics & Genomes*, *13*(1), 30. doi: 10.1007/s11295-017-1112-5

Misiewicz, T. M., & Fine, P. V. A. (2014). Evidence for ecological divergence across a mosaic of soil types in an Amazonian tropical tree: *Protium subserratum* (Burseraceae). *Molecular Ecology*, *23*(10), 2543–2558. doi: 10.1111/mec.12746

Mittler, R., Vanderauwera, S., Gollery, M., & Van Breusegem, F. (2004). Reactive oxygen gene network of plants. *Trends in Plant Science*, *9*(10), 490–498. doi: 10.1016/J.TPLANTS.2004.08.009

Neale, D. B., & Kremer, A. (2011). Forest tree genomics: growing resources and applications. *Nature Review Genetics*, *12*(2), 111–122.

Nei, M. (1987). *Molecular Evolutionary Genetic*. Columbia University Pres, New York.

Nielsen, R., Hubisz, M. J., Hellmann, I., Torgerson, D., Andrés, A. M., Albrechtsen, A., … Clark, A. G. (2009). Darwinian and demographic forces affecting human protein coding genes. *Genome Research*, *19*(5), 838–849. doi: 10.1101/gr.088336.108

Okazaki, Y., & Saito, K. (2014). Roles of lipids as signaling molecules and mitigators during stress response in plants. *The Plant Journal*, *79*(4), 584–596. doi: 10.1111/tpj.12556

Ordoñez, N. M., Marondedze, C., Thomas, L., Pasqualini, S., Shabala, L., Shabala, S., & Gehring, C. (2014). Cyclic mononucleotides modulate potassium and calcium flux responses to H_2O_2 in Arabidopsis roots. *FEBS Letters*, *588*(6), 1008–1015. doi: 10.1016/j.febslet.2014.01.062

Pais, A. L., Whetten, R. W., & Xiang, Q. (2017). Ecological genomics of local adaptation in *Cornus florida* L. by genotyping by sequencing. *Ecology and Evolution*, *7*(1), 441–465. doi: 10.1002/ece3.2623

Pavlidis, P., Jensen, J. D., Stephan, W., & Stamatakis, A. (2012). A Critical Assessment of

Storytelling: Gene Ontology Categories and the Importance of Validating Genomic Scans. *Molecular Biology and Evolution*, *29*(10), 3237–3248. doi: 10.1093/molbev/mss136

Perochon, A., Aldon, D., Galaud, J.-P., & Ranty, B. (2011). Calmodulin and calmodulin-like proteins in plant calcium signaling. *Biochimie*, *93*(12), 2048–2053. doi: 10.1016/J.BIOCHI.2011.07.012

Petit, R. J., & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annu Rev Ecol Evol Syst*, *37*(1), 187–214.

Pujol, B., Blanchet, S., Charmantier, A., Danchin, E., Facon, B., Marrot, P., … Winney, I. (2018). The Missing Response to Selection in the Wild. *Trends in Ecology & Evolution*. doi: 10.1016/J.TREE.2018.02.007

Punja, Z. K., & Zhang, Y. Y. (1993). Plant chitinases and their roles in resistance to fungal diseases. *Journal of Nematology*, *25*(4), 526–540. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19279806

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., … Stougaard, J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature*, *425*(6958), 585–592. doi: 10.1038/nature02039

Richardson, J. L., Urban, M. C., Bolnick, D. I., & Skelly, D. K. (2014). Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution*, *29*(3), 165–176. doi: 10.1016/j.tree.2014.01.002

Richmond, T. (2000). Higher plant cellulose synthases. *Genome Biology*, *1*(4), reviews3001.1. doi: 10.1186/gb-2000-1-4-reviews3001

Ryals, J. A., Neuenschwander, U. H., Willits, M. G., Molina, A., Steiner, H. Y., Hunt, M. D., … Stacey, G. (2008). A LysM Receptor-Like Kinase Plays a Critical Role in Chitin Signaling and Fungal Resistance in Arabidopsis. *The Plant Cell Online*, *20*(2), 471–481. doi: 10.1105/tpc.8.10.1809

Sarkar, S. K. (2006). False discovery and false nondiscovery rates in single-step multiple testing procedures. *Annals of Statistics*, *34*(1), 394–415. doi: 10.1214/009053605000000778

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Review Genetics*, *14*(11), 807–820. Retrieved from http://dx.doi.org/10.1038/nrg3522

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Annu. Rev. Ecol. Evol. Syst.*, *38*, 595–619.

Savolainen, V., Anstett, M.-C. C., Lexer, C., Hutton, I., Clarkson, J. J., Norup, M. V, … Baker, W. J. (2006). Sympatric speciation in palms on an oceanic island. *Nature*, *441*(7090), 210–213.

Scalfi, M., Mosca, E., Di Pierro, E. A., Troggio, M., Vendramin, G. G., Sperisen, C., … Neale, D. B. (2014). Micro- and Macro-Geographic Scale Effect on the Molecular Imprint of Selection and Adaptation in Norway Spruce. *PLoS ONE*, *9*(12), e115499. doi: 10.1371/journal.pone.0115499

Schmid, M., & Guillaume, F. (2017). The role of phenotypic plasticity on population differentiation. *Heredity*, *119*(4), 214–225. doi: 10.1038/hdy.2017.36

Scotti, I., González-Martínez, S. C., Budde, K. B., & Lalagüe, H. (2016). Fifty years of genetic studies: what to make of the large amounts of variation found within populations? *Annals of Forest Science*, *73*(1), 69–75. doi: 10.1007/s13595-015-0471-z

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., & Neale, D. B. (2013). Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, *9*(4), 901–911. doi: 10.1007/s11295-013-0596-x

ter Steege, H., Pitman, N. C. A., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., … Silman, M. R. (2013). Hyperdominance in the Amazonian tree flora. *Science*, *342*(6156). Retrieved from http://www.sciencemag.org/content/342/6156/1243092.abstract

Torroba-Balmori, P., Budde, K. B., Heer, K., González-Martínez, S. C., Olsson, S., Scotti-Saintagne, C., … Heuertz, M. (2017). Altitudinal gradients, biogeographic history and microhabitat adaptation affect fine-scale spatial genetic structure in African and Neotropical populations of an ancient tropical tree species. *PLoS ONE*, *12*(8), e0182515. doi: 10.1371/journal.pone.0182515

Turner, T. L., Bourne, E. C., Von Wettberg, E. J., Hu, T. T., & Nuzhdin, S. V. (2010). Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. *Nature Genetics*, *42*(3), 260–263.

Villadas, P. J., Fernández-López, M., Ramírez-Saad, H., & Toro, N. (2007). Rhizosphere-bacterial community in *Eperua falcata* (Caesalpiniaceae) a putative nitrogen-fixing tree from French Guiana rainforest. *Microbial Ecology*, *53*(2), 317–327. doi: 10.1007/s00248-006-9158-1

Vizcaíno-Palomar, N., Revuelta-Eugercios, B., Zavala, M. A., Alía, R., & González-Martínez, S. C. (2014). The Role of Population Origin and Microenvironment in Seedling Emergence and Early Survival in Mediterranean Maritime Pine (*Pinus pinaster* Aiton).

1019 *PLoS ONE*, *9*(10), e109132. doi: 10.1371/journal.pone.0109132
1020 Weits, D. A., Giuntoli, B., Kosmacz, M., Parlanti, S., Hubberten, H.-M., Riegler, H., …
1021 Licausi, F. (2014). Plant cysteine oxidases control the oxygen-dependent branch of the
1022 N-end-rule pathway. *Nature Communications*, *5*, 3425. doi: 10.1038/ncomms4425
1023 White, M. D., Klecker, M., Hopkinson, R. J., Weits, D. A., Mueller, C., Naumann, C., …
1024 Flashman, E. (2017). Plant cysteine oxidases are dioxygenases that directly enable
1025 arginyl transferase-catalysed arginylation of N-end rule targets. *Nature Communications*,
1026 *8*, 14690. doi: 10.1038/ncomms14690
1027 Yeaman, S., & Guillaume, F. (2009). Predicting adaptation under migration load: the role of
1028 genetic skew. *Evolution*, *63*(11), 2926–2938. doi: 10.1111/j.1558-5646.2009.00773.x
1029 Yeaman, S., & Jarvis, A. (2006). Regional heterogeneity and gene flow maintain variance in a
1030 quantitative trait within populations of lodgepole pine. *Proceedings Biological Sciences*,
1031 *273*(1594), 1587–1593. doi: 10.1098/rspb.2006.3498
1032 Yeaman, S., & Otto, S. P. (2011). Establishment and maintenance of adaptive genetic
1033 divergence under migration, selection, and drift. *Evolution*, *65*(7), 2123–2129. doi:
1034 10.1111/j.1558-5646.2011.01277.x
1035 Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under
1036 migration-selection balance. *Evolution*, *65*(7), 1897–1911. doi: 10.1111/j.1558-
1037 5646.2011.01269.x
1038
1039

Figures

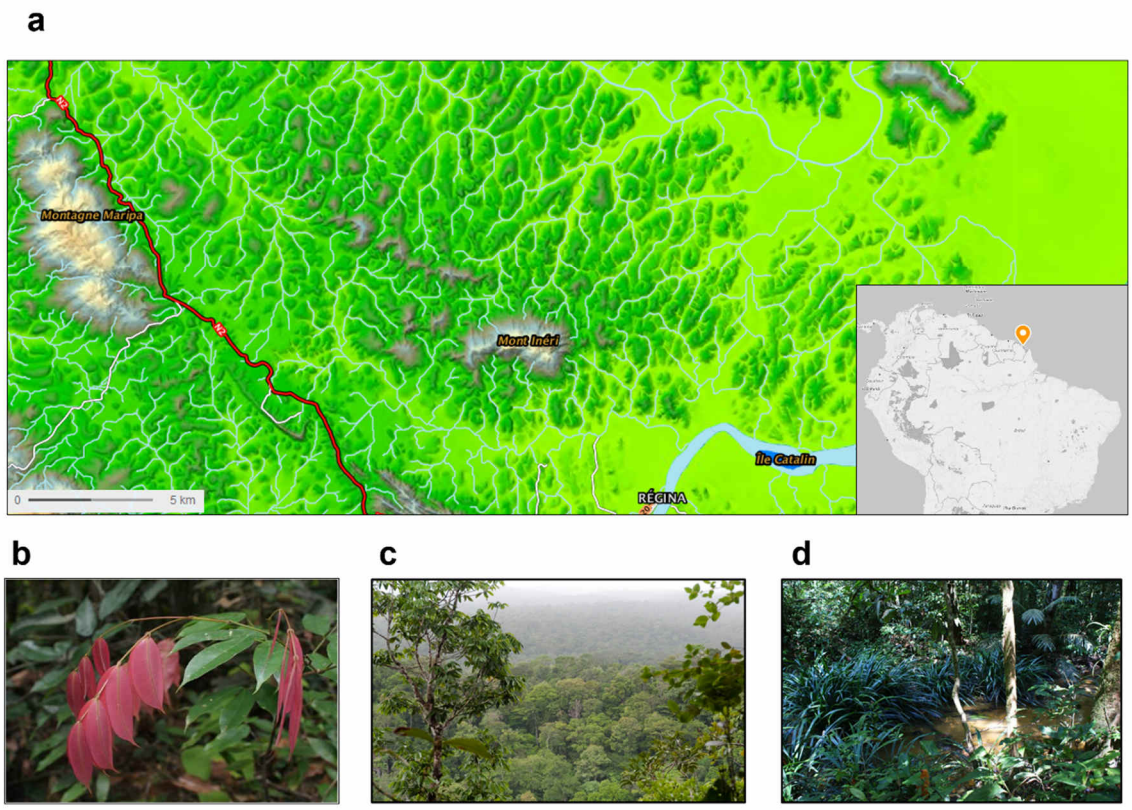


Fig. 1. (a) Relief map of forest landscapes in the region of Régina, French Guiana. The map was created through the French 'Géoportail' (https://www.geoportail.gouv.fr/): official geospatial data were provided by the French Institute of Geographic and Forest Information ©IGN 2017. (b) *Eperua falcata* sapling. (c) and (d) Forest communities associated with topography and related soil types: *terra-firme* forests on ferralitic soils at the top of hills (c), water-logged forests on hydromorphic soils in bottomlands (d).

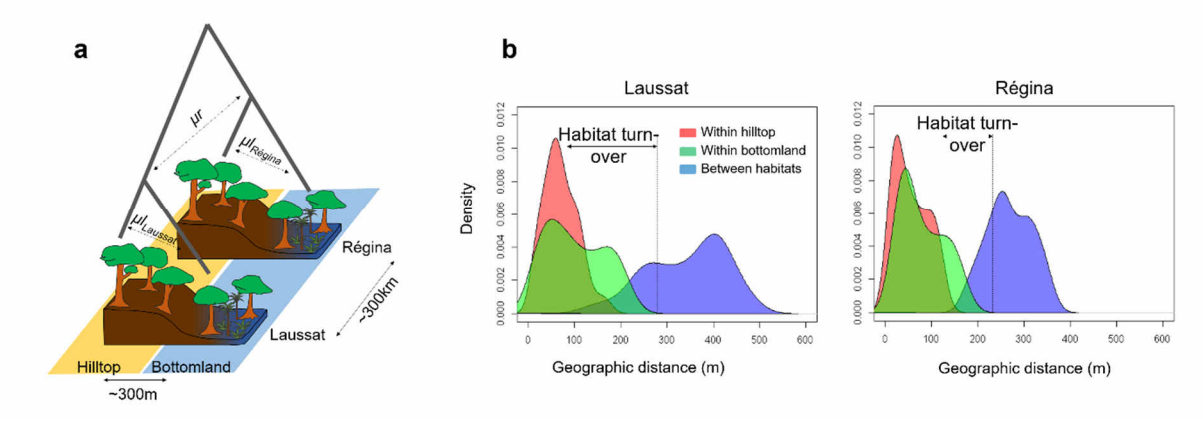


Fig. 2. (a) Sampling design scheme and correspondence with genome-wide parameters of divergence estimated through the hierarchical Bayesian model at regional (μr) and local scales ($\mu l_{Laussat}$ and $\mu l_{Régina}$ respectively). (b) Distribution of geographic distances between trees in the study sites of Laussat (Western French Guiana) and Régina (Eastern French Guiana). Vertical dotted bars indicate the range of habitat turnover in each study site.

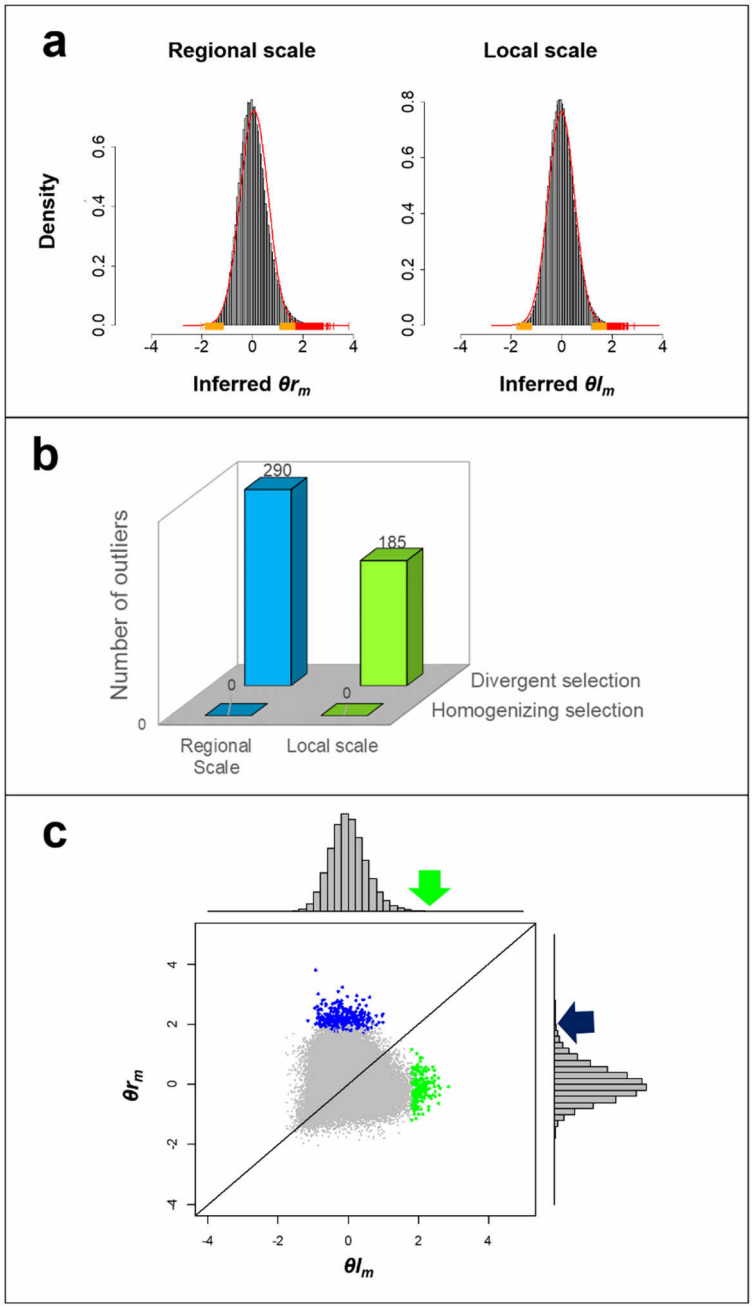


Fig. 3. (a) Locus-specific parameters (θr_m and θl_m) and normal distribution inferred by the Bayesian model. Vertical bars indicate the location of outliers under homogenizing (lower tail) and divergent (upper tail) selection. Left: regional scale (θr_m), right: local scale (θl_m). (b) Proportion of outlier SNPs under selection at regional (i.e. between study sites) and at local (i.e. between microhabitats within sites) scales. (c) Joint distribution of regional and local-scale locus-specific parameters inferred through the Bayesian model. Blue and green dots indicate outliers under divergent selection at the regional and local scale, respectively.

Regional scale	Neutral	Divergent selection	Neutral
Local scale	Neutral	Neutral	Divergent selection

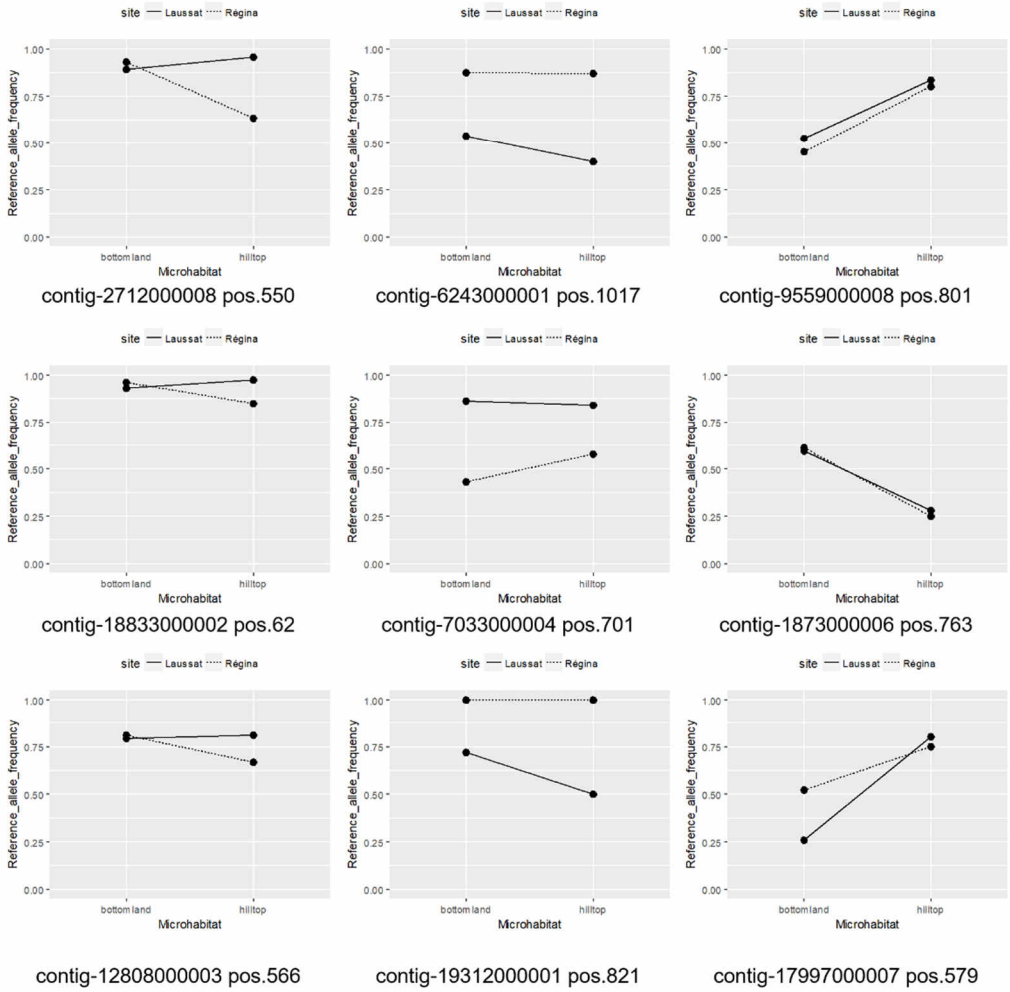
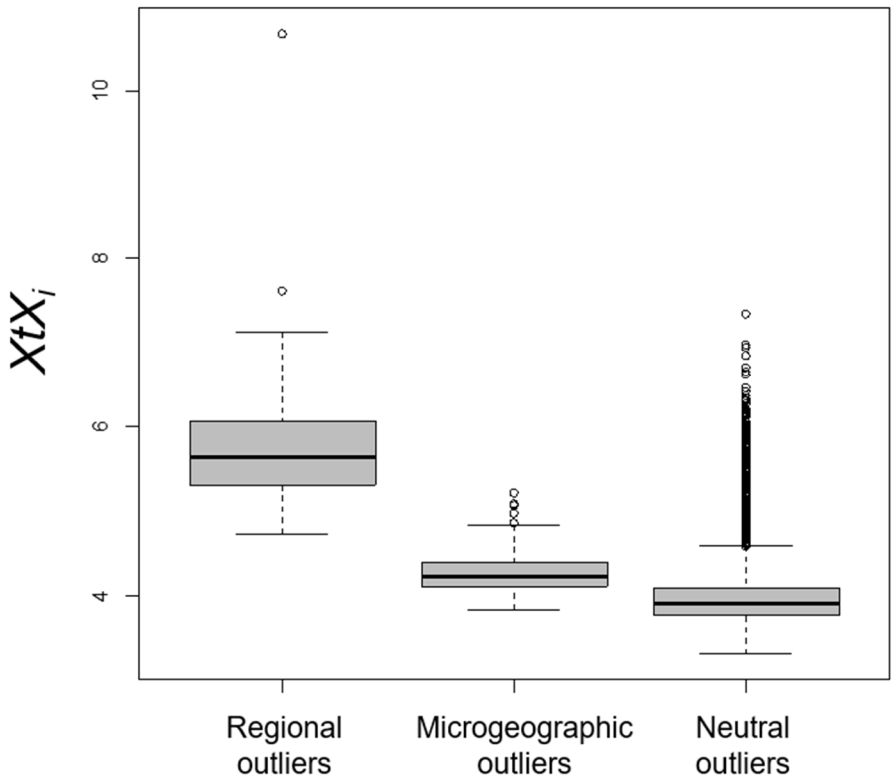


Fig. 4. Clines in allele frequency between regional locations (i.e. study sites) and microhabitats for the different kinds of (randomly sampled) neutral and outlier SNPs at different geographical scales.



1072

1073 **Fig. 5.** Comparison of the *XtXi* statistics estimated using BayPass across sets of SNPs: those
1074 detected as outliers at regional and microgeographic scales, respectively, and those detected as
1075 neutral by our custom Bayesian modelling framework.

1076

1077

Tables

Table 1. Results of reciprocal transplants - linear models of phenotypic value decomposition.

		Effects (p-values)								
Trait	**Date**	**Seed_mass**	**Site prov**	**Site trans**	**Topo prov**	**Topo trans**	**Site prov x trans**	**Topo prov x trans**	**AIC**	**Retained model**
Survival	june 2011	0.227	0.969	0.288	0.044	0.255	NA	NA	389.38	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Survival	september 2011	0.111	0.406	0.21	0.156	0.025	0.027	0.508	574.31	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Survival	april 2012	0.022	0.456	0.145	0.101	0.104	NA	NA	673.43	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Survival	september 2012	0.004	0.99	0.088	0.34	0.061	0.034	0.693	725.71	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Survival	september 2013	0.241	0.425	0.059	0.419	0.09	0.1	0.293	804.61	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Survival	september 2014	0.027	0.267	0.046	0.953	0.236	NA	NA	872.21	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
N Growth Units	september 2011	0	0.9744	0.7382	0.9596	0.5188	NA	NA	1744.67	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
N leaves	september 2011	1.00E-04	0.6557	0.4786	0.5822	0.4829	NA	NA	1963.09	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
N leaflets	september 2011	0	0.0575	0.047	0.5591	0.3378	NA	NA	3105.56	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
N leaflets attacked	september 2011	0	0.509	0.041	0.008	0	0	0.8	2521.05	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N Growth Units	april 2012	5.00E-04	0.3917	0.1237	0.4413	0.597	NA	NA	2374.30	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
N leaves	april 2012	0	0.6879	0.6725	0.6272	0.2792	NA	NA	2081.12	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
N leaflets	april 2012	0	0.001	0.65	0.502	0.479	0.425	0.841	3286.33	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N leaflets attacked	april 2012	0	0.008	0.921	0.051	0.011	NA	NA	2479.99	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)

N Growth Units	september 2012	0	0.37	0.403	0.558	0.789	NA	NA	2728.67	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
N leaves	september 2012	4.00E-04	0.6198	0.4812	0.6739	0.1758	NA	NA	2196.86	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
N leaflets	september 2012	0	0.002	0.674	0.606	0.359	0.317	0.978	3536.14	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N Growth Units	september 2013	0	0.535	0.402	0.243	0.48	0.04	0.863	2637.24	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N leaves	september 2013	0	0.061	0.905	0.467	0.366	NA	NA	2366.96	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
N leaflets	september 2013	0	0	0.582	0.029	0.565	NA	NA	4072.96	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
N leaflets attacked	september 2013	0	0	0.564	0.486	0.153	0	0.199	3183.24	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N Growth Units	september 2014	0	0.011	0.042	0.011	0.141	0.014	0.567	5893.73	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N leaves	september 2014	0	0	0.168	0.019	0	NA	NA	4682.05	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
N leaflets	september 2014	0	0	0.358	0	0	0.001	0.962	10890.30	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
N leaflets attacked	september 2014	0	0	0.338	0.017	0.005	0.009	0.977	7905.87	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Height	september 2011	0	0.004	0.901	0.322	0	0.098	0.136	3614.54	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo
Height	april 2012	0	0.412	0.709	0.475	0.085	0.053	0.907	5272.12	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Height	september 2012	0	0.112	0.494	0.022	0.317	NA	NA	3974.47	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Height	september 2013	0	0.067	0.669	0.04	0.257	NA	NA	4105.51	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Height	september 2014	0	0.892	0.098	0.14	0.32	0.1	0.886	6579.71	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Diameter	september 2011	0	0.0653	0.4053	0.1628	3.00E-04	NA	NA	1401.39	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo
Diameter	april 2012	0	0.722	0.532	0.064	0.039	NA	NA	1566.20	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Diameter	september 2012	0	0.956	0.548	0.09	0.12	NA	NA	1336.35	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)

Diameter	september 2013	0	0.709	0.477	0.309	0.004	NA	NA	1445.02	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Diameter	september 2014	0	0.216	0.181	0.602	0.459	NA	NA	3503.87	Trait~ Prov_site + Transplant_site + Prov_topo + Transplant_topo + (1 | Garden)
Leaf area	september 2011	0	0.133	0.431	0.092	0.595	0.314	0.618	6904.20	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Leaf area	april 2012	0	0.047	0.318	0.159	0.894	0.312	0.114	10518.37	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
Leaf area	september 2012	0	0.014	0.773	0.586	0.325	0.234	0.726	10539.89	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)
SPAD	April 2013	0.001	0.004	0.054	0.254	0	0	0.111	4293.16	Trait~ Prov_site * Transplant_site + Prov_topo * Transplant_topo + (1 | Garden)

Supplementary figures and tables

Supplementary figures S1 to S7.

Supplementary tables S1 to S5.

1090

Online supporting information

Supplementary file S1. Soil characteristics of the two sites and on growth and functional characters of the sampled trees

Supplementary file S2. Bioinformatics pipeline

1095 **Supplementary file S3.** Bayesian model specifications

Supplementary file S4. BayPass – methodology & results

Supplementary file S5. G2D – methodology & results

Supplementary file S6. Reciprocal transplants – methodology

Supplementary file S7. Bioinformatics results

1100 **Supplementary file S8.** BLASTp and functional annotation of genes neighboring (< 5 kb) 'divergence' outlier SNPs at the microgeographic scale.

Supplementary file S9. Interactive gene network to easily navigate through the functional annotation of genes neighboring 'divergence' outlier SNPs (< 5 kb) at the microgeographic scale. Green bubbles indicate predicted genes ; blue and pink bubbles indicate Gene ontology

1105 terms, with size depending on the number of predicted genes neighboring an outlier SNP flagged with each GO term ; pink bubbles indicate GO terms significantly enriched between the subset of predicted genes neighboring outlier SNPs compared to the all predicted genes.

Supplementary file S10. Reciprocal transplants – linear models raw outputs

1110