This is the accepted version of the article:
Veresoglou, Stavros; Peñuelas, Josep. Variance in biomass-allocation fractions is explained by distribution in European trees. DOI 10.1111/nph.15686

This version is available at https://ddd.uab.cat/record/218158
under the terms of the © COPYRIGHT license
Variance in biomass-allocation fractions is explained by distribution in European trees

Stavros D Veresoglou\textsuperscript{1,2*}, Josep Peñuelas\textsuperscript{3,4}

\textsuperscript{1} Freie Universität Berlin, Institut für Biologie, Plant Ecology, D-14195 Berlin, Germany
\textsuperscript{2} Faculty of Agriculture, Laboratory of Ecology and Environmental Protection, Aristotle University of Thessaloniki, 541 24 Thessaloniki, Greece.
\textsuperscript{3} CSIC, Global Ecology Unit CREAT-CSCIC-UAB, Bellaterra, 08193, Catalonia, Spain
\textsuperscript{4} CREAT, Cerdanyola del Vallès, 08193, Catalonia, Spain
Intraspecific variability in ecological traits confers the ability of a species to adapt to an ever-changing environment. Fractions of biomass allocation in plants (BAFs) represent both ecological traits and direct expressions of investment strategies and so have important implications on plant fitness particularly under current global change.

We combined data on BAFs of trees in more than 10000 forest plots with their distributions in Europe. We aimed to test whether plant species with wider distributions have more or less variable intraspecific variance of BAFs foliage-woody biomass and shoot-root ratios than species with limited distribution.

Irrespective of corrections for tree age and phylogenetic relatedness, the standard deviation in BAFs was up to three times higher in species with most extensive distributions than in those with least extensive distribution due to a higher genetic diversity. Variance in BAFs also increased with latitude.

We show that a combination of 36% tree genetic diversity and 64% environmental variability explains variance in BAFs and imply that changes in genetic diversity occur quickly. Genetic diversity should thus play a key role in regulating species responses to future climate change. Loss of habitat, even if transient, could induce a loss of genetic diversity and hinder species survival.

Key-words: allometry; effective population size; extinction; genetic diversity; plant allometry; plant morphology; woody plants
Introduction

The strategies of fitness and growth of sessile organisms are largely determined by biomass allocation (Hodge, 2004; Poorter et al., 2012; Veresoglou et al., 2017), which in turn determines the long-term morphology of the individual. The morphology of sessile organisms has been studied the most in plants (Kokko, 2007). Plant morphology can be summarised in several ways, each with distinct strengths and limitations. Poorter and Sack (2012) reviewed these methods and concluded that fractions of biomass allocation, especially after correcting for confounding allocation parameters such as size, represent a particularly effective measure. Plant morphology has high functional importance, so the influence of environment on morphology has been well studied. Some representative parameters that have been studied extensively for their influence on plant allometry are the size of individual plants (Reich et al., 2005; Poorter et al., 2012), historical environmental abiotic parameters such as stressors (e.g. uniform stress hypothesis) (Mogan & Gannell, 1994; Dean et al., 2002), shading (Lusk et al., 2008; Duursma et al., 2010; Forster et al., 2011), temperature (Reich et al., 2014), precipitation (McCarthy & Enquist, 2007), biotic interactions such as competition and diversity (Forrester et al., 2017a); the growth form of plants (Wyka et al., 2013) and species-specific ontogeny (Forrester et al., 2017b). Plant age, however, is an influential factor that is often not available for integration into such analyses (Nikklas, 1997; Nikklas et al., 2003; Barthélémy & Caraglio, 2007; Duursma et al., 2010). Analysts often use plant size as a proxy of age (Nikklas, 2004; Bowman et al., 2013).

Biomass-allocation fractions (BAFs) describe biomass ratios, usually at a logarithmic scale, of plant organs. BAFs are measurable to an individual level and thus constitute plant traits (Müller et al., 2000; Reich et al., 2003). Trait variance is often comparable to or even more important than trait means (Messier et al., 2010; Bolnick et al., 2011; Violle et al., 2012). High variance in trait values could facilitate, for example, the evolution and adaptation of a species to new environmental settings (Bolnick et al., 2011). Quantifying variation in BAFs could improve modelling uncertainty in the standing biomass of woody habitats because most of our estimates of standing biomass are projections of allometric equations (Chave et al., 2005; Muukkonen, 2007).

Plant morphology has been described as an equilibrium between constraints to plant growth and exogenous environmental stressors (Barthélémy & Caraglio, 2007). Constraints to
plant growth and other intrinsic factors can negate the influence of the environment to varying
degrees, i.e. environmental conditions tend to shift BAFs against the stabilising influence of
intrinsic factors. We would then expect that BAFs would be more variable in plant species that
experience extreme environmental conditions more often. Reyer et al. (2013) argued that
extreme conditions occur mainly at the edges of the distribution of a species. Source-sink
population dynamics describe instances where species only occur in an area because of a
constant influx of propagules from surrounding areas where the species grows better. At the
edges of the distribution of a species it is more likely to observe source-sink population
dynamics (Remeš, 2000) than in the kernel of the distribution which could induce unique
phenotypes. Plant species that have a small distribution should more frequently experience such
“extreme edge” conditions so plants with smaller distributions may have the most variable BAFs.
The reason is that environmental heterogeneity mainly increases the variance in BAFs in these
species (Fig. 1a). We thus hypothesize that plant species with smaller distributions have more
variable BAFs (hypothesis one). Alternatively, the higher effective population sizes and genetic
variability could allow plant species that have extensive distributions to be those showing the
highest intraspecific trait variance, including BAFs. The reason could be that environmental
variability acts independently of distribution but more populous species show more diverse
BAFs due to a higher genetic diversity (Fig. 1b). As a result, plant species with smaller
distributions could vary less in their BAFs (hypothesis two). A final expectation is that variance
in BAFs should indicate the ability of species to adapt to local environmental conditions, so a
high variance suggests rapid evolution. Stapley et al. (2017) reported considerably lower
recombination rates in gymnosperms than angiosperms. The variance in BAFs should thus be
lower in conifers, compared to angiosperms (hypothesis three).

Most of our existing understanding on the way traits of woody species vary with genetic
diversity and the environment originates from provenance tests, where plants differing in their
origin are grown under common environmental conditions (Thompson et al. 2008; Wang et al.
2010). Some of the limitations of provenance test relate to the choice of the common
environment (Leites et al. 2012), the time they require for long-lived species to grow and thus the
logistics of destructively harvesting them (e.g. to assess BAFs). Here we worked on an
observational approach which synthesizes across existing BAF measurements from records of
destructive tree harvests in Europe (Fig. 1). Schepaschenko et al. (2017) have recently released
two large data sets detailing the biomass fractions of many trees that had been destructively harvested between 1930 and 2014 in Eurasia. We combined this data set with information on the distributions of many of these plant species in Europe from Mauri et al. (Mauri et al., 2017) to address the above mentioned three hypotheses.

Materials and Methods

Sources of data

We used the two data sets published by Schepaschenko et al. (2017) for our main analysis. The Biomass-tree data set provided information on biomass fractions following destructive harvesting for 9613 trees, mainly in Europe. The Biomass-plot data set provided information on biomass fractions for 10351 plots distributed throughout Eurasia, each of which provided information for cumulative biomass of two or more trees. The two data sets synthesized across approximately 1200 experiments over the period 1930-2014 and contained information on location and age of the trees as well as biomass information on several different fractions. We used biomass information on foliage vs. woody aboveground biomass for our main analyses. We also used root:shoot ratios as part of our sensitivity analyses. We extracted distributions of tree species in Europe from Mauri et al. (Mauri et al., 2017). Mauri et al. (2017) only described tree distributions in Europe, so we limited our analysis to trees that occur mainly in Europe. We used QGis v 2.12.3 to estimate the size of the polygonal envelopes provided by Mauri et al. (2017). We used Phylocom v 4.2 to reconstruct the phylogenetic relationships of the tree species in our analysis.

We worked with an aggregate of 80 species in our analysis of European trees. A list of the species and the phylogenetic reconstruction used to correct for phylogenetic relatedness can be found in Table S1. The three data sets varied considerably in terms of their resolution and their suitability for the different analyses. The cumulative analysis of the Biomass-tree data set used biomass and age information for 4719 trees from 42 species (median observations per species, 10.5; 1st quartile, 7; 3rd quartile, 78.5). The analysis of the Biomass-plot data set used information for 3898 plot entries describing 63 species (median observations of plot per species, 15; 1st quartile, 8.5; 3rd quartile, 46) making it particularly suitable for assessing how BAFs are influenced by environmental heterogeneity. The analysis of individual stands in the Biomass-tree
data set used an aggregate of 1854 tree observations from 40 species (median observations per species, 8.5; 1st quartile, 7; 3rd quartile, 33.25). Because for the analysis on individual stands (i.e. plots) in the Biomass-tree data set we only used per species information of trees in a single stand (i.e. plot with most tree observations in the Biomass-tree data set), this made it particularly suitable for quantifying the influence upon BAFs of genetic diversity. The Biomass-tree data set contained proportionally more observations located in Europe than Biomass-plot data set did (Fig. 2 – inserts).

**Rationale for the analyses**

We modelled three sources of variance in BAFs, age, environment and genetic diversity. We inferred genetic diversity from genetic variability (i.e. the degree to which the genetic characteristics of populations vary) using a phenomenological approach that determined the variance of morphological characteristics in populations after controlling for sources of variance that were not of genetic origin such as age, environment and latitude. To control for age-related differences in BAFs we used the slopes of linear models with BAFs as response variables and age as predictor. To assess the fraction of BAFs that is explained by genetic diversity we assessed how variance in BAFs scaled with tree species distribution in tree individuals which belonged to the same species and stand after correcting them for tree age (individual stands in the Biomass-tree data set; Fig. 1b). To assess the fraction of BAFs that is explained by environmental variability we assumed that environmental variability is mainly due to latitudinal range (de Frenne et al. 2013; we provide more details at the section poling effect sizes) and quantified how much the variance in BAFs changed with each degree of latitudinal range (i.e. latitudinal breadth – to avoid an overrepresentation of stands we used the Biomass-plot dataset).

**Statistical analyses**

Our analyses only considered tree species (which were the unit of our analysis) for which a minimum of five observations of age, foliage biomass and aboveground woody biomass were available, because we needed a minimum of three observations to fit a linear model and then two additional residuals to yield meaningful estimates of variance for the fit of the regression line (results from a sensitivity test on the inclusion threshold are presented in Appendix 1 and Fig.
To extract the fraction of variance in BAFs that was not due to age differences we fitted linear models with BAFs as a response variable and age as predictor and then quantified the standard deviation of the residuals. We explain the procedure in higher detail below (also Appendix 2; Table S2):

1. We calculated the natural log response ratio (logRR) of foliage over aboveground woody biomass \( \log RR = \ln \left( \frac{m_{\text{foliage}}}{m_{\text{woody}}} \right) \), where \( m \) stands for biomass) for each individual tree in the case of the Biomass-tree data base or plot in the case of the Biomass-plot data base (Appendix 2). In the form of a sensitivity analysis we also analysed logRR of root over shoot BAFs. LogRR represents a widely used nonstandard effect size in synthesis studies. A large logRR indicates a higher investment in foliage than woody biomass.

2. We fitted a linear model with logRR as a response variable and age as the sole predictor. The single most important idiosyncratic cause of variability in BAFs is age (Nikklas et al., 2003; Nikklas, 2004; Barthélémy & Caraglio, 2007; Duursma et al., 2010). To correct our data for this source of variance, we extracted the residuals of the linear model and assessed their standard deviations. We assumed a first-order linear correlation between logRR and age using the standard deviation of the residuals as a measure of the variance, in agreement with preliminary analyses (Appendix 2).

3. We correlated our metric of variance of the relationship between BAFs and age with the distribution of the plants.

In the first of our three analyses (Table S2) we used all observations of individual trees in the Biomass-tree data set, which consisted of multiple trees per stand and multiple stands per species. We then used information for all plots in the Biomass-plot data set, consisting of a single plot per stand, which allowed us to correct for spatial autocorrelation. Finally, we reanalysed the Biomass-tree data set but only using the information for each species in a single stand, with the sole criterion that the stand had provided most observations for that species. This analysis addressed the concern that the cause of higher variance in BAFs was due to a larger distance across observations and thus a higher variability in environmental conditions. We conducted this analysis both with and without phylogenetic correction. We used an analysis of phylogenetically independent contrasts to correct for phylogeny.
We further compared our estimates of variance between angiosperms and gymnosperms. We thus used $t$-tests, assuming unequal variance.

Sensitivity analyses

We tested the specificity of our observations to logRR between foliage and aboveground woody biomass by replicating the analysis for logRR between total root biomass vs. total aboveground biomass. The two data sets contained unequal numbers of observations of total root biomass, and the number of species which we could analyse was low. We directly assessed how the variance in biomass fractions scaled with distribution (i.e. independent of age) by directly assessing the standard deviation of biomass-allocation ranges and correlating it with distribution (Appendix 2).

We further replicated the analysis with the Legacy Tree Data data set (Radtke et al., 2016). To estimate the distribution of North American trees we used information from the United States Department of Agriculture (USDA 2017). This data set did not contain any information on tree age. We used data from the Legacy_Tree table and used the variability of logRR foliage over aboveground woody biomass for single locations. Sufficient information was available for only three species, which were assessed for four locations. All three species belonged to the Pinaceae family, had nested distributions and occurred in southeastern North America. We assessed distribution by extracting information from the USDA for the number of states in which the species occurred and regressed this metric against the variance of logRR. The amount of information we extracted from the Legacy Tree Data base and the resolution, given that the size of US states varies considerably, were much lower than in our main analysis so we only used as a means of supporting the main results with a different dataset. Because we did not have accurate distribution range data, we used here non-parametric statistics.

To assess the degree to which our observations were sensitive to the inclusion of invasive tree species we classified tree species in native vs invasive as in Veresoglou et al. (2018; Table S3). We repeated the analysis separately for the subsets of native and invasive species.

The influence of latitude

Latitude is an influential predictor of variance in plant traits (Heibo et al., 2005; Aerts et al., 2012), and as such it was important for us to show that it was not driving our results. We
assessed the degree to which our patterns were explained by latitude using two analyses. (a) We first extracted the average latitudes of the distributions of the woody species in Europe reported by Mauri et al. (2017). We compared these values with the variances in BAFs corrected for plant age (i.e. the standard deviation of the residuals of the linear model described above). (b) We then divided Europe into northern and southern Europe. We thus used the average of the two latitudinal extremes of Europe: 82°N for the northern region and 35°N for the southern region, averaging 58.5°N. We used the subset of points west of 69°W. We separately calculated the variance of BAFs corrected for age for the woody species for which the Biomass-tree data set contained a minimum of five observations both north and south of the latitudinal average threshold, which we then combined in a new logRR (i.e. northern variance over southern variance). We calculated the variance for these two sites and used them to recalculate logRR.

Pooling effect sizes to quantify genetic and environmental variance in BAFs

We partitioned here variance in BAFs into a fraction due to genetic variability and a fraction due to environmental variability. We used a phenomenological approach for this (Fig. 1; Appendix 4). We made the following three assumptions:

1. The key factor contributing to environmental variability was latitudinal range of the distribution of an organism. In support to this simplistic statement, see de Frenne et al. (2013) demonstrating that some key environmental parameters such as temperature, precipitation, soil pH and human influence covary with latitude. We assumed that there is a first order linear relationship between latitudinal range and environmental variability in BAFs.

2. Genetic variability in BAFs increased with distribution at a slope equivalent to that in Fig. 2b.

3. The relative importance of the two fractions was assessed for an imaginary species with an “average” latitudinal range which in this case was 19.4° latitude and a mean distribution range of 10,000 km².

We use the additive principle of variances in statistics to add the two fractions on the assumption that the observed variance was independent of our sample sizes. Our variance partitioning was carried out for an idealized species having an average latitudinal range of 58.9° – 39.5° = 19.4° latitudinal range and a distribution of 10,000 km² which represented averages in our dataset and were also consistent with (Tkach et al., 2008). We estimated the overall form of the model and
we calculated relative effect sizes for mean parameters of distribution and latitude in Europe. We provide more information on the analysis in Appendix 4.

Results

The main results are on records throughout Eurasia but these were supported by a reanalysis exclusively to the records from Europe. Our main analysis gave consistent results for all three data sets (Fig. 2), namely that variance in BAFs increases with the distribution size of a species. LogRR of foliage vs aboveground woody biomass was significantly positively correlated with distribution in all three tests (Fig. 2). The correlation was strongest (i.e. highest Kendall’s tau) for individual trees in the same stand (tau=0.51, Fig. 2c). We found yet stronger correlations when we narrowed down the analysis on tree records occurring in Europe which we here define as those with a longitude smaller than 69°W (Fig. S2). Correlations would have been considerably weaker if we had not corrected for age of the trees (Fig. S3). No phylogenetic conservatism was found in either the variance of biomass fractions (Blomberg’s K in the cumulative analysis of the Biomass-tree data set was 0.37 with $P = 0.95$) or the distribution of the species (Blomberg’s $K$ in the same data set was 0.53 with $P = 0.12$; Appendix 3). The correlations were weaker after correction for phylogeny with phylogenetically independent contrasts, except for the Biomass-tree data set (Fig. 3). The variances of the BAFs did not differ significantly between angiosperms and gymnosperms in any of the three data sets (Fig. 4 – evidence against hypothesis three).

The analysis using root:shoot ratios produced comparable results but with lower statistical power (Fig. 5a,b). The results of our analysis of the biomass data from America agreed with those for European trees (Fig. 5c; tau=0.91, $P=0.07$). Most of the trees in the datasets were classified as native and there were minimal differences in the additional analyses targeting natives (Fig. S4). There was insufficient statistical power to reach robust conclusions for invaders (Fig. S4).
The influence of latitude

There was no relationship between mean latitude and the variance of BAFs between foliage over woody aboveground biomass in the Biomass-tree data set (Fig. 6a). The five tree species that were observed both North and South of the latitudinal threshold of 58.5°N did not display any patterns with regards to their BAFs (Fig. 6a – insert). *Picea obovata* showed higher BAFs in the southern extent of its distribution, *Pinus sylvestris* and *Betula alba* in the northern extent of their distribution whereas *Larix sukaczewii* and *Pinus sibirica* showed relatively balanced BAFs in both extents. The lack of a relationship was even more apparent when we narrowed down our analysis to individual stands in the Biomass-tree data set (data not shown). By contrast we could observe a strong relationship between mean latitude and the variance of BAFs in the Biomass-plot data set (Fig. 6b).

Relative effect sizes of genetic and environmental variability

We had 40 observations and a mean slope of $3.324 \times 10^{-5}$ per thousand km$^2$ of distribution in the case of the individual stands in the Biomass-tree data set assessing genetic variability and 18 observations and a mean slope of $3.05 \times 10^{-2}$ per degree latitudinal range in the case of the Biomass-plot data set measuring environmental variability. The resulting expression of variance (Appendix 4) was $s_p = 10^{-4} \sqrt{(0.11D^2 + 93.025\phi^2)}$ with $D$ representing distribution range and $\phi$ degrees latitudinal range. For average European settings of $D \approx 10,000$ thousands km$^2$ (also consistent with (Tkach et al., 2008)) and $\phi$ differences of $58.9^\circ – 39.5^\circ$ latitudinal range the two factors inside the parenthesis take the values $11.10^6$ and $35.10^6$ suggesting that latitudinal range (i.e. here used as a proxy of environmental variability) exerts on average a 48% stronger (i.e. the resulting variances for genetic and environmental variability, if that the other factor was zero, would be 0.33 and 0.59, respectively, giving relative proportions of 36% and 64%) influence on variance in logRR of BAFs than distribution (i.e. here genetic variability) does. Genetic diversity and latitudinal range, both induce strong changes on the variance in BAFs and the respective standard deviations were multifold higher for species with extensive distributions (Fig. 2c) and high latitudes (Fig. 6b).

Discussion
Living organisms are constantly challenged to optimally allocate their finite resources to maximise fitness. This challenge leads to multiple investment trade-offs, many of which have been extensively studied (Poorter et al., 2006; Cadotte, 2007; Huot et al., 2014). BAFs represent direct expressions of some of these trade-offs and the phenotypes of the organisms. We experimented with two such BAFs, foliage over woody aboveground biomass and the root:shoot ratio, to show that both these fractions were more variable in tree species with extensive distributions, even after limiting our analysis to neighbouring trees (Fig. 2c), in agreement with hypothesis two. Our results were also valid after correcting for phylogenetic relatedness (Appendix 3). A higher ability to adapt to a changing environment is one of the implications of higher trait variance (Bolnick et al., 2011). The causality of this relationship is unclear and could imply either that tree species that experience a higher phenotypic variability tend to have larger distributions (e.g. van Vallen, 1973) or that some species are phenotypically more diverse in response to a larger distribution and most likely a larger effective population size. Environmental heterogeneity has been identified as a mechanism that facilitates genetic variation in plants (Delph & Kelly, 2013), and a larger distribution implies a higher environmental heterogeneity. Many tree species have extensive distributions, but individual trees can also disperse over very large distances (Bacles et al., 2006; Petit & Hampe, 2006; Kremer et al., 2012) and cross-fertilise with individuals experiencing differing environmental conditions. We thus favour the interpretation that a larger distribution most likely induces a higher variance in BAFs. BAFs in plants represent expressions of an equilibrium between stabilising intrinsic/genetic factors such as ontogeny and the destabilising influence of the environment (Barthélémy & Caraglio, 2007), i.e. any biomass-allocation fraction is an aggregate of these two mechanisms. We thus present an argument detailing why our results were likely due to a higher genetic variability of trees with more extensive distributions and not because of the environmental conditions that the trees experienced (in support of hypothesis two and against hypothesis one). The influence of the environment is expected to be only moderately important at sites close to the centres of species distributions (compared to the edges). Most of the trees in our data sets had been harvested near the centres of their distributions, because moderate environmental conditions facilitate silvicultural practices. Most importantly, our results were valid when using individual sites per species in which case the differences were most likely due
to rapid evolution (i.e. the absence of systematic differences between angiosperms and
gymnosperms provided evidence against hypothesis three). Co-occurring tree species experience
comparable environmental conditions, so the resulting phenotypic variance should best represent
the genetic diversity (Fig. 2c).

Does latitude influence BAFs? We found no relationship in the Biomass-tree data set but
there was a strong positive relationship in the Biomass-plot data set (Fig. 6). A reason why the
results across the two data sets were incongruent probably relates to the way these were
standardized. The Biomass-tree data set contained multiple tree records per sampling site and
there were few replicate sites per tree species. Co-occurring woody species are likely to
experience comparable environmental conditions and the resulting BAFs should mainly reflect
genetic diversity. Because of the few replicate sites per tree species any influences of latitude
were masked in the data set by those of genetic diversity. This was even more pronounced when
we narrowed the analysis to a single site per tree species. By contrast in the Biomass-plot data set
via being limited to a single record per sampling site, we could effectively investigate differences
arising from environmental variability. Based on this analysis, trees that typically occur closer to
the poles exhibit higher variances of BAFs. Even though it has been disputed in the past
(Vázquez & Stevens, 2004), it is widely appreciated that environmental variability increases with
latitudinal range (MacArthur, 1972). This result is in accordance with our first hypothesis that
environmental variability should increase BAFs variance (Fig. 1a). The effect size we estimated
for environmental variability exceeded that for genetic diversity.

The survival challenges of tree species to the accelerating pace of global change is a key
topic in the biology of global change (Lenoir et al., 2008; Bertrand et al., 2016; Veresoglou &
Halley, 2018). Identifying tree species at a high risk of extinction is important. Several traits such
as longevity (Morin & Thuiller, 2009) and seed size (Veresoglou & Halley, 2018) might be
informative in terms of tree susceptibility to extinction. The loss of habitat is a key factor that
drives the eventual risk of extinction, but the relationship between habitat size and genetic
variability is poorly understood. The loss of genetic diversity following habitat loss can further
limit the ability of a species to cope with environmental conditions (Sexton et al., 2009) and
eventually accelerate extinction. The implications of extensive distributions have been debated.
A review of 31 studies by Lowe et al. (Lowe et al., 2005) found that habitat loss usually did not
significantly affect the genetic variability of tree species. Another meta-analysis by Vranckx et
al. (2012), however, found that habitat loss induced losses in the genetic diversity of species. An extensive synthesis by Morueta-Holme et al. (2013) reported that plant species with extensive geographic ranges were more genetically variable, and Kremer et al. (Kremer et al., 2012) argued that long-distance gene flow amongst trees likely conferred an evolutionary advantage. In contrast, the implications of habitat loss in trees may be fully reversible if the former habitat of these species can be restored before extinction (Newmark et al., 2017).

Our results suggest that it is tree genetic diversity that induces variance in BAFs (Fig. 1b; we found higher variance in BAFs in species with large distributions). Our analysis makes the assumption that genetic variability is a good proxy of genetic diversity, which despite being a common and well supported assumption in the literature (e.g. Avolio et al., 2012; Jöqvist and Kremp, 2016), remains less robust than direct estimates of genetic diversity. Most importantly, our general models did not discriminate between native and non-native plant species and used real distributions to assess effect sizes. We often observe that the distributions of non-native (i.e. invasive) species are not at equilibrium with their environment and that they possess a lower than expected genetic diversity (Beaumont et al., 2009; Bradley et al., 2010). We would have expected, as a result, relatively weak relationships between genetic diversity and realized distributions for the subset of non-native trees which was not the case (Fig. S4). This finding implies that changes in genetic diversity can occur quickly and develop after a few generations of growth in isolation. Genetic diversity should thus play a key role in regulating the response of species to future climate change, because of the extensive fitness implications of BAFs on the ability of a species to adapt. The loss of habitat, even if transient, could lead to a loss of genetic diversity, which would hinder species survival.

Human activity has sped up gene flow in almost all types of ecosystems and this should continue in the near future (Wilson et al. 2009). On the short term, assisted gene flow homogenizes populations (i.e. and their genetic diversity), allowing species to more effectively cope with unfavourable environmental conditions (e.g. through acquiring more favourable BAFs). Assisted gene flow should thus steepen the positive relationship between variance in BAFs and distribution (Fig. 1b), benefiting disproportionally species with a large distribution at the risk of species that maintain a limited genetic diversity. It has been documented, for example, that invasions, which represent an alternative form of introductions of species with a high competitive ability, induce extinctions of native species (Colautti et al. 2017; Catford et al.
Even though we did not directly model assisted gene flow (as for example in Adams et al. 1998), the strong relationship between distribution and genetic diversity should be instructive for forest management. Silvicultural practices such as the artificial regeneration of stands from commercial genetic material (e.g. Bradshaw 2004; Finkeldey Ziehe 2004) might, therefore, be precarious for the native species diversity.

In summary, BAFs were more variable in trees with extensive distributions, i.e. our results supported only the second of our hypotheses. Most notably, we made the point that it was a higher genetic variability that resulted in more variable BAFs for tree species with extensive distributions. We thus present evidence that the loss of habitat for tree species through a quick loss of genetic diversity could lower the ability of a species to modify its architecture (BAFs) in response to changes in the environment (Fig. 1). It is thus likely that any loss of habitat may not be as reversible as many believe.

Acknowledgements
We want to thank two anonymous reviewers and the editor for constructive criticism on our paper. JP was funded by the European Research Council Synergy grant ERC-SyG-2013-610028 IMBALANCE-P.

Contributions
Conceived the study and carried out the analysis: SDV. The two authors wrote together the article and approved the final version.

Data availability
All the data we analyse here are already in the public domain. We present intermediate data in our analysis in the form of Appendix 8.

Competing interests
The authors declare no conflicts of interest.
Supplementary Information

Fig. S1: Sensitivity analysis on the inclusion criterion of five full records

Fig. S2: Relationships between the variance in BAFs in records from Europe

Fig. S3: Weighted regressions between the variance in BAFs and the distribution of the species without correcting for age

Fig. S4: Sensitivity analysis on how relationships differed between native and invasive tree species

Table S1: List of the plant species that we used in our analyses.

Table S2: Overview of all analyses.

Table S3: Classification of tree species into native vs invasive

Appendix 1: Inclusion Criteria

Appendix 2: Effect sizes and model specifications

Appendix 3: Strength of BAF relationships

Appendix 4: Partitioning of BAFs into environmental and genetic variability

Appendix 5: Biases due to differences in coverage of the databases

Appendix 6: Phylogenetic tree

Appendix 7: Supplementary references

Appendix 8: Raw data
References


Aerts R, van Bodegom PM, Cornelissen JHC. 2012. Litter stoichiometric traits of plant species of high-latitude ecosystems show high responsiveness to global change without causing strong variation in litter decomposition. *New Phytologist* 196: 181-188.


Fig. 1 Conceptual diagram illustrating the three hypotheses we address in the manuscript. Four hypothetical species, the broadleaves blue triangle and yellow circle (represented with an elliptical leaf) and the conifers red rhomb and green square (represented by an acicular leaf) are each sampled at four locations (map on top; note that distribution envelopes differ). The ranking of their distributions is as follows: red rhomb, blue triangle, yellow circle and green square. In Hypothesis One (a) we expect that the fraction of variance in biomass allocation fractions that is explained by genetic variability (purple discontinuous lines) is independent of distribution and that environmental factors increase the variance (green arrow) more in those species that have a small distribution (larger arrow) resulting in a negative relationship between observed variance and distribution (green discontinuous lines). In Hypothesis Two (b) we expect that a larger distribution results in a higher genetic diversity which is depicted with a purple line. The environment increases variance irrespective of distribution (green arrows) resulting in a positive relationship between observed variance and distribution (green discontinuous lines). In Hypothesis Three (c) the biomass allocation fractions vary independently of distribution and can be explained by the evolutionary history of the plant (here angiosperms vs gymnosperms).

We could assess genetic variance in BAFs by comparing conspecific trees in the same stand after correcting for age-differences and the sum of genetic and environmental variability by comparing across stands, also after correcting for age.

Fig. 2 Relationships between the variance in biomass-allocation fractions (foliage over woody aboveground biomass) and distribution of the tree species. (a) Biomass-tree data set, all possible observations (multiple trees per location) of the specific data set (lower than in Biomass-plot); (b) Biomass-plot data set, one observation per location (at a plot level); (c) Biomass-tree data set, observations per species only describe trees in the stand (plot) that contained the most tree observations. The dashed lines represent the best fits. Numbers next to the data points indicate the number of observations per species used to calculate the variance. Overlaid map shows in red the location of the sites where the data
originated from – we only analysed woody species with an extensive distribution (over 2
million km$^2$) in Europe. Phylogenetic correction was not applied. Relationships were
stronger when we repeated the analyses with the subset of sites located in Europe (Fig.
S2).

Fig. 3 Relationships between variance in biomass-allocation fractions (foliage over woody
aboveground biomass) and distribution of the tree species for phylogenetically corrected
data using phylogenetically independent contrasts (PIC). (a) Biomass-tree data set, all
possible observations; (b) Biomass-plot data set, one observation per location; (c)
Biomass-tree data set, observations per species are from the stand that contained most
observations. The dashed red lines represent the best fits. The overlaid phylogenetic tree
depicts in the form of squares information (blue: first quartile; white middle two
quartiles; red 4th quartile) on variance in biomass allocation fractions (left) and
distribution (right) of the woody species included in the analysis and their phylogenetic
relationships (tree). Note the absence of phylogenetic signal which was confirmed with
Blomberg K tests.

Fig. 4 Differences in variance in biomass-allocation fractions (foliage over woody aboveground
biomass) between angiosperms (elliptical leaf in yellow) and gymnosperms (acicular leaf
in green). (a) Biomass-tree data set, all possible observations; (b) Biomass-plot data set,
one observation per location; (c) Biomass-tree data set, observations per species are from
the stand that contained most observations. None of the t-tests were significant.

Fig. 5: (a; b) Relationships between the variance in biomass-allocation fractions (root over shoot
fractions) and distribution of the tree species. (a) Biomass-tree data set, all possible
observations; (b) Biomass-tree data set, observations per species are from the stand that
contained most observations; (c) variance in biomass-allocation ratios for Pinus echinata
Pinus taeda (observed in 20 states, from two studies) and Pinus palustris (observed in 10 states) across four studies in the Legacy Tree Data database (USA). The three species have nested distributions in America. The dashed lines represent the best fits. Numbers next to the data points indicate the number of observations per species used to calculate the variance. Overlaid map shows in red the location of the sites where the data originated from.

Fig. 6 Influence of latitude on variance in biomass-allocation fractions (foliage over woody biomass). The x-axis describes the average latitude of the distribution range of the woody species whereas the y-axis the variance in biomass-allocation fractions. Panel (a) describes the Biomass-tree data set whereas panel (b) the Biomass-plot data set. The insert in panel (a) shows how variance differs in the north and south distribution of five tree species in the Biomass-tree data set (opaque triangles). North here describes latitudes in Europe higher than 58.5°N (mean of the latitudinal extremes of European territory) and South lower than this value. These five species were the only ones that met the inclusion criterion, namely being represented with a minimum of five entries both in North and South Europe. The continuous opaque line is an isocline where variance in the North equals that in the South.
Variance in biomass allocation fractions

(a) Environment
(b) Environment + Genetic
(c) Distribution
Variance in biomass-allocation fractions

(a) \( \tau = 0.48, \ P = 0.03 \)

(b) \( \tau = 0.34, \ P = 0.028 \)

(c) \( \tau = 0.51, \ P = 0.047 \)
PIC Variance in biomass allocation fractions

- (a) $r=0.75, P=0.008$
- (b) $r=0.28, P=0.24$
- (c) $r=0.61, P=0.08$
Variance in biomass-allocation fractions

(a) 

(b) 

(c) 

Variance in biomass-allocation fractions
Variance in biomass-allocation fractions

(a) $r=0.8$, $P=0.11$

(b) $r=0.87$, $P=0.05$

(c) $\tau=0.91$, $P=0.07$
Variance in biomass-allocation fractions

Mean latitude

(a) $r=0.48, \quad P=0.19$

(b) $r=0.72, \quad P<0.001$