ARTICLE

Quercus species divergence is driven by natural selection on evolutionarily less integrated traits

Jaroslav Klápště^{1,8} · Antoine Kremer^{2,3} · Kornel Burg⁴ · Pauline Garnier-Géré^{2,3} · Omnia Gamal El-Dien⁵ · Blaise Ratcliffe⁶ · Yousry A. El-Kassaby⁶ · Ilga Porth⁷

Received: 16 February 2020 / Revised: 7 October 2020 / Accepted: 7 October 2020 / Published online: 27 October 2020
© The Author(s), under exclusive licence to The Genetics Society 2020

Abstract

Functional traits are organismal attributes that can respond to environmental cues, thereby providing important ecological functions. In addition, an organism's potential for adaptation is defined by the patterns of covariation among groups of functionally related traits. Whether an organism is evolutionarily constrained or has the potential for adaptation is based on the phenotypic integration or modularity of these traits. Here, we revisited leaf morphology in two European sympatric white oaks (*Quercus petraea* (Matt.) Liebl. and *Quercus robur* L.), sampling 2098 individuals, across much of their geographical distribution ranges. At the phenotypic level, leaf morphology traditionally encompasses discriminant attributes among different oak species. Here, we estimated in situ heritability, genetic correlation, and integration across such attributes. Also, we performed Selection Response Decomposition to test these traits for potential differences in oak species' evolutionary responses. Based on the uncovered functional units of traits (modules) in our study, the morphological module "leaf size gradient" was highlighted among functionally integrated traits. Equally, this module was defined in both oaks as being under "global regulation" in vegetative bud establishment and development. Lamina basal shape and intercalary veins' number were not, or, less integrated within the initially defined leaf functional unit, suggesting more than one module within the leaf traits' ensemble. Since these traits generally show the greatest species discriminatory power, they potentially underwent effective differential response to selection among oaks. Indeed, the selection of these traits could have driven the ecological preferences between the two sympatric oaks growing under different microclimates.

Introduction

The biological complexity of organisms can be characterized by examining the strength, developmental patterns, and evolutionary processes of covariation among phenotypic traits, which are captured in the concepts of phenotypic integration and its counterpart, modularity (Olson and

Associate editor: Olivier Hardy

Supplementary information The online version of this article (<https://doi.org/10.1038/s41437-020-00378-6>) contains supplementary material, which is available to authorized users.

✉ Jaroslav Klápště
Jaroslav.Klapste@scionresearch.com

¹ Department of Genetics and Physiology of Forest Trees, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences in Prague, Kamýcká 129, 165 21 Prague 6, Czechia

² INRA, UMR Biodiversité Gènes et Communautés, 69 route d'Arcachon, 33612 Cestas Cedex, France

³ University of Bordeaux, UMR 1202, Biodiversité Gènes et Communautés, F-33400 Talence, France

⁴ Department of Health and Environment (Bioresources), AIT Austrian Institute of Technology, Konrad-Lorenz-Straße 24, 3430 Tulln, Austria

⁵ Pharmacognosy Department, Faculty of Pharmacy, Alexandria University, Alexandria, Egypt

⁶ Department of Forest and Conservation Sciences, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada

⁷ Département des sciences du bois et de la forêt, Université Laval, 1030, Avenue de la Médecine, Québec, QC G1V 0A6, Canada

⁸ Present address: Scion (New Zealand Forest Research Institute Ltd.), 49 Sala Street, Whakarewarewa, Rotorua 3010, New Zealand

Miller 1958; Berg 1960; Cheverud 1982; Murren 2002; Pigliucci 2003; Murren 2012; Wagner et al. 2007; Klingenberg 2008; Armbruster et al. 2014). Phenotypic integration generally refers to some degree of cohesion and coordination between phenotypic traits. Modularity, by contrast, focuses on the relative independence among sets of variables (defined as functional units or modules if they show strong integration), and differences in the degree of integration within and between groups of traits (Wagner et al. 2007; Klingenberg 2008). Depending on the biological context, the underlying research hypotheses, and approaches, these concepts have been extended and partitioned, according to their nature and the underlying processes connecting them (Klingenberg 2008; Armbruster et al. 2014). Among these, the developmental integration includes a large range of physiological, environmental, and molecular genetic processes affecting the developmental pathways causing covariation, for example, in morphological traits (Klingenberg 2008; Murren 2012); while the functional integration refers to the underlying processes allowing a functional unit to perform particular functions but the links with developmental integration are not straightforward (Young and Badyaev 2006; Armbruster et al. 2014).

The concept of phenotypic integration has historically been of great interest to evolutionary biologists, because, on one hand, it has served to explain the evolutionary changes that can occur under genetic constraints. For example, between-trait covariances are thought to prevent the independent evolution of traits when the selection direction is opposite to the correlation among these traits. This has been shown to be the case in *Lycaeides* butterflies, where the biological diversity of wing patterns is constrained and evolutionary changes can only occur in the directions of the genetic covariances of the wing pattern elements (Lucas et al. 2018). Another example is the evolution of the human body shape, where genetic covariation of limb lengths provides a compelling case for how genetically complex traits co-evolved (Savell et al. 2016). On the other hand, phenotypic integration can serve as an adaptive feature that can further promote the acceleration of evolutionary responses. This will be the case, when traits in one module can evolve independently of traits in other modules, and thereby, a more effective response to changes in selection direction can be provided by phenotypic integration (e.g., Lewontin 1978; Pigliucci 2003; Armbruster et al. 2014; Healy et al. 2018).

Strong integration of floral traits in several plant species has been argued to provide an effective functional solution for maintaining their reproductive function (Berg 1960; Conner and Via 1993; Brock and Weinig 2007). Subsequent work often found tightly integrated modules within floral traits, suggesting that such integration had been driven by the efficient natural selection in pollinator specialization

(see Ordano et al. 2008 in *Rosaceae*, and Bissell and Diggle 2010 across different *Nicotiana* species and their hybrids). A recent comparison of floral/reproductive traits and vegetative traits (leaf size) in two sympatric *Primulina* species growing in different microhabitats also showed that these characteristics evolved independently due to different selection pressure (Feng et al. 2019).

Leaf morphological traits were previously identified as significant determinants of species differentiation within *Quercus* spp. (Kremer et al. 2002; Viscosi et al. 2009) and are considered to be genetically highly complex based on results from quantitative trait locus (QTLs) analysis (Gailing 2008; Saintagne et al. 2004). Most of the previously conducted analyses have focused on the construction of synthetic variables through principal component analysis (PCA) or related multivariate approaches to provide evidence for the use of certain leaf morphology characteristics to distinguish among oaks. However, there is a current lack of deeper knowledge about the integration and evolvability of these traits.

Phenotypic/genetic correlations are measures of the (preexisting) standing genetic variation and provide valuable inferences about the statistical integration of traits. Here, we adopted a multivariate linear model (de los Campos and Pérez-Rodríguez 2013) to estimate genetic correlations between the studied oak characters. Nevertheless, such correlations should not be used to make inferences about developmental or evolutionary integration (Armbruster et al. 2014). Therefore, we also employed evolvability-based estimates of phenotypic/genetic integration of all oak phenotypic traits under investigation to interpret their concerted evolution (Hansen and Houle 2008).

Addressing the role of phenotypic integration on species' adaptive potential to novel environments, Hermant et al. (2013) found endemic species occupying narrower ecological niches generally show higher levels of phenotypic integration and stronger trait-environment correlations along the different environmental gradients, suggesting that their adaptability to new environments is lower. Leaf morphological characteristics have long been recognized for their physiological significance in plant-environment interactions (Lewis 1972), as exemplified by leaf pubescence (Press 1999) or sclerophylly (Lo Gullo and Salleo 1988) for better drought adaptation in species (De Micco and Aronne 2012). Moreover, the sizes and shapes of fossil leaves have been serving as proxies for paleoclimatic and paleoecological variables (Royer et al. 2005), and in fact, whole ecosystem functioning has been related to variations in leaf structures and function (Press 1999).

Changes in leaf structure and their functional significance are best studied using reverse genetics (mutagenesis) by comparing the generated transgenic lines

Fig. 1 *Q. petraea* (gray stripes area) and *Q. robur* (black stripes area) natural distributions across Europe.

The seven here studied naturally regenerating mixed *Q. petraea*/*Q. robur* forests are indicated as follows: 1—Petite Charnie, 2—Escherode, 3—Dalkeith, 4—De Meinweg, 5—Sigmundsherberg, 6—Roudsea Wood, 7—Arlaban. The source files for species distribution presentation originated from www.euforgen.org.

against wild-type landraces. Plant ecology integrated with taxonomy can indicate general adaptive leaf structures, and therefore, the conserved molecular pathways underlying the same structures in phylogenetically distant groups.

If tight or long-term co-evolution represents a set of stabilizing selection pressures, an analogy can be made with the environmental pressures to which long-lived organisms, such as forest tree species are submitted, such as different developmental stages and unpredictable heterogeneous environmental conditions across their lifetime and space. These could promote higher levels of modularity than in short-lived organisms passing through simpler development without the need for complex modularity (Kashtan and Alon 2005; Lipson et al. 2002). Forest tree species are therefore an excellent model for testing predictions of higher levels of modularity across functionally related traits and comparing patterns observed across different species.

The examples cited above (and many others, see Esteve-Altares 2017 for a review on morphological modularity research since 1958 to present) illustrate the diversity of research questions related to traits' phenotypic integration studies, consistent with the possible ecological or environmental causes which can act as selection pressures for the

efficient functioning of an organism's developmental system. However, alternative, nonselective processes have also been suggested that could explain the great flexibility in morphological structures for performing similar functions (Young and Badyaev 2006). Whatever the processes and their interaction, this is encompassed in the concept of evolutionary integration which was understood as the disposition/ability of two or more traits to evolve jointly during the divergence of populations or species (Armbruster et al. 2014).

Trait integration and modularity can be statistically assessed directly from the preexisting standing genetic variance of the investigated samples through genetic correlations among different traits. These genetic correlations can represent the evolutionary constraints which can limit traits' ability to change in the direction of selection (i.e., evolvability, defined as the capacity of an organism for adaptive evolution; Wang et al. 2010; Wagner and Zhang 2011; Hill and Zhang 2012), especially in cases where there is stabilizing selection acting in different directions on other traits (conditional evolvability). Therefore, the level to which evolvability of the investigated traits would be conditionally reduced by stabilizing selection can be estimated

by integrating indices of traits relative to other traits studied, as proposed in recent methods (Hansen and Houle 2008).

Several methods and conceptual frameworks have been developed over the years to describe a whole organism's multidimensional genotype-phenotype maps, and which integrate traits' genetic architecture and quantitative genetics concepts such as pleiotropy and epistasis (genetic architecture's evolution: Hansen 2006; the integrated phenotype: Murren 2002, 2012; adaptation and constraints: Armbruster et al. 2014; phenotype to fitness: Laughlin and Messier 2015; modular genomic organization and evolvability: Pepper 2003). Concerning the physical (nonrandom) ordering of genes on chromosomes for improved evolvability and increased rate of adaptation, J.W. Pepper (2003) concluded that "such epistatic clustering may reduce the interference present due to the selection acting on different traits, and it may allow for simultaneous optimization of different recombination rates for gene pairs with additive and epistatic fitness effects." With the advent of high-throughput technologies, using genome-wide scans, we can now pinpoint the actual quantitative trait nucleotide functional variants that underly highly integrated (pleiotropic) traits (e.g., for *Populus*: Porth et al. 2014). In general, the current development of vast genomic resources along with intensive phenotyping (phenomics) allows for a better understanding regarding the evolution of functional relationships among morphological traits (Murren 2012). However, for plants, such resources are mostly available for model species such as the annual *Arabidopsis thaliana*, which also benefits from a large collection of genetic lines. Still, there is no genome-wide coverage of phenotypes, even for *A. thaliana* (Murren 2012), that could elucidate entire phenotypic integration in a plant organism.

The plant species of the present study, *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L., are closely related and belong to the section *Quercus*, white oaks (Hubert et al. 2014). Both species can be found across Europe largely in sympatry as illustrated in Fig. 1. However, these oaks occupy slightly different micro-environmental niches, with *Q. robur* being more tolerant of flooding and *Q. petraea* more drought resistant (Lévy et al. 1992; Porth et al. 2005a, b).

Traditionally, foresters use leaf characteristics to discriminate oak species when they occur in sympatry; an operational method built on multi-trait discriminant analyses and based on a large-scale sample across Europe was proposed by Kremer et al. (2002) to help avoid bias and to standardize species identification across distant populations. Here, we examined phenotypic integration for these leaf morphology characteristics in two different European oak species, utilizing the same metrics (Kremer et al. 2002, 2008) collected across a Europe-wide sample of natural populations. The commonly accepted metrics in

multi-factorial analysis yielding the bimodal distribution to assign *Q. petraea* and *Q. robur* individuals include petiole length (PL), presence of intercalary veins, and occurrence of pubescence. Moreover, three leaf dimensional characteristics provide a synthetic variable, the "leaf size gradient," with a continuous distribution, to help distinguish between species.

The objective of our study on wild stands of co-occurring *Q. petraea* and *Q. robur* was to (a) characterize, at different levels, the modular architecture among leaf morphology traits, (b) estimate in situ genetic parameters such as heritability and genetic correlations for these traits, (c) examine how these parameters can further highlight differences in trait integration patterns across species, and (d) discuss species integrity concerning known reproductive barriers between the studied species.

Material and methods

Plant material/study site

Natural oak forests with co-occurring and naturally regenerating *Q. petraea* and *Q. robur* were investigated among a total of nine intensively studied plots (ISPs) (Kremer et al. 2008) within the OAKFLOW project. In our study, we analyzed a total of 2098 individuals (896 *Q. petraea*, 1138 *Q. robur*, and 64 morphologically unclassified *Q. petraea/robur* individuals) from seven *Q. petraea/robur* mixed forests representing ISPs from northern Spain (Basque country) to the UK (Scotland), Fig. 1. Among all ISPs available within the OAKFLOW project, these seven stands were chosen for the availability of common marker genotypic data (Mariette et al. 2002; Scotti-Saintagne et al. 2004). The oak species identity referred to a previous study by Kremer et al. (2002), who used the first canonical component from analyses performed on leaf morphology measurements (<http://www.pierroton.inra.fr/Fairoak/Protocoles/OAKMORPH.html>). For the present study, all phenotypic data on up to ten leaves per tree (5.54 leaves on average) were already collected in this previous work (Kremer et al. 2002).

Genotyping

We used microsatellites, genetic markers, which we found adequate (see Hodel et al. 2016; Ramírez-Valiente et al. 2009) to test our hypotheses and investigate the large population sample (c. 2k individual oak trees) in our survey. Genotypic data were used to estimate pairwise relationship coefficients among individuals (see below) from six simple sequence repeats (SSRs) originating from two different sources: (a) MSQ4 and MSQ13 originally developed by

Dow et al. (1995) for *Quercus macrocarpa* and subsequently described in Streiff et al. (1998) for *Q. petraea* and *Q. robur*, as well as (b) *ssrQpZAG9*, *ssrQpZAG36*, *ssrQpZAG104*, and *ssrQpZAG1/5* developed by Steinkellner et al. (1997). Genotypic data originated mostly from Mariette et al. (2002), who assessed genetic diversity within and among these oak populations, while data from the Arlaban, Spain, population were taken from Scotti-Saintagne et al. (2004). Importantly, none of these SSRs differentiated the two species (Mariette et al. 2002), and moreover, they were selectively neutral (Scotti-Saintagne et al. 2004; Porth et al. 2016).

Brief description of leaf morphology phenotyping data

Variation in eight leaf morphological primary descriptors was investigated: lamina length (LL), PL, lobe width (LW), sinus width (SW), length of the lamina from the lamina base to the widest part (WP), number of lobes (NL), number of intercalary veins (NV), and basal shape of the lamina (BS); among these, PL, BS, SW, NV, and NL were the traits with the highest species discriminatory power (Porth et al. 2016). Since BS was originally scored on a scale basis, we transformed this trait to a normal score (Gianola and Norton 1981) for all quantitative analysis. All phenotypic data were collected in a previous study (Kremer et al. 2002). Figure 2 visualizes each measured phenotypic trait in detail.

Genetic clustering analyses and treatment of putatively introgressed individuals

Using the molecular markers, genetic structure was assessed through discriminant analysis of principal components (DAPC) to identify genetically differentiated clusters (Jombart et al. 2010) across all stands including both species and interspecific hybrids. The principle of DAPC is based on the search for synthetic variables (discriminant functions) to maximize differences between genetic clusters and minimize variances within clusters. Genotypes of samples were treated altogether since a priori defined morphological clusters are not relevant for population structure inferred from selectively neutral markers as utilized here (Scotti-Saintagne et al. 2004; Porth et al. 2016). Raw data transformation through principal component analysis was performed a priori of any cluster algorithm application as follows:

$$X^T DXU = UA, \quad (1)$$

where X is the $n \times p$ matrix of p relative allelic frequencies for n individuals (missing data were replaced by mean allelic frequencies), D is the diagonal matrix where each

Fig. 2 Variation of eight oak leaf morphological primary descriptors. To the left: pedunculate oak (*Q. robur*); to the right: sessile oak (*Q. petraea*). Display samples were taken from Czech Republic natural oak populations (courtesy: Department of Genetics and Physiology of Forest Trees, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences in Prague). Lamina length (LL), petiole length (PL), lobe width (LW), sinus width (SW), length of lamina from the lamina base to the widest part (WP), number of lobes (NL), number of intercalary veins (NV), and basal shape of the lamina (BS; scoring system 1–9) were used; among these, PL, BS, SW, NV, and NL were the traits with the highest species discriminatory power (Porth et al. 2016).

diagonal element represents the weight of $1/n$, U is the $p \times r$ matrix of eigenvectors, and A is the diagonal matrix of non-zero eigenvalues.

The prior groups required for the DAPC routine were identified by using the function “find.clusters” based on the k-mean clustering algorithm implemented in the R package “adeget” (Jombart 2008) as follows:

$$\text{Var}(XU) = B(XU) + W(XU), \quad (2)$$

where XU is the matrix of principal components obtained in the previous step, $\text{Var}(X) = \text{tr}(A)$, $B(XU)$ is between-group variance defined as $B(XU) = \text{tr}(U^T X^T P^T D P X U)$, where P is defined as $P = H(H^T D H)^{-1} H^T D$, where the H matrix assigns genetic groups to each individual, and $W(XU)$ is

the within-group variance defined as $W(XU) = \text{tr}(U^T W U)$, where W is the within-group covariance matrix. The optimal number of clusters was defined by the lowest Bayesian information criterion (BIC) value as follows:

$$\text{BIC} = n \log(W(X)) + g \log(n), \quad (3)$$

where n is the number of individuals and g is the number of genetic groups. The results from DAPC are obtained by eigen-analysis of matrix $PXU(U^T W U)^{-1}U^T X^T P^T D$. The function “optim.a.score” implemented in the same R package was used to find the optimal number of principal components further used in the discriminant analysis. The analysis was performed in the R package “adeigenet” (Jombart 2008).

In situ quantitative genetics analyses

The univariate mixed linear model implemented in “ASReml-R” R package (Butler et al. 2009) was used on standardized phenotypes, and within each stand separately, to estimate genetic parameters such as variances components and narrow-sense heritability as follows:

$$y = X\beta + Zu + e, \quad (4)$$

where y is the vector of measurements across all individuals, β the vector of fixed effects (intercept and genetic clusters defined by DAPC analysis to capture any “local” specific demographic processes which could confound additive genetic effects), u the vector of genotypic effects following $\text{var}(u) \sim N(0, G\sigma_u^2)$, where σ_u^2 is the genotypic variance and G the marker-based relationship matrix.

First, as the Lynch and Ritland (1999) estimator was found to be superior in populations with low levels of relatedness (Van de Castele et al. 2001), the composite multi-locus estimate was estimated as follows:

$$r_{xy} = \frac{1}{\sum_{l=1}^L w_{r,x}(l)} \sum_{l=1}^L w_{r,x}(l) \hat{r}_{xy}(l), \quad (5)$$

where $\hat{r}_{xy}(l)$ is the single locus relationship estimator following:

$$\hat{r}_{xy} = \frac{p_a(S_{bc} + S_{bd}) + p_b(S_{ac} + S_{ad}) - 4p_a p_b}{(1 + S_{ab})(p_a + p_b) - 4p_a p_b}, \quad (6)$$

and $w_{r,x}(l)$ is the single locus weight defined as follows:

$$w_{r,x}(l) = \frac{(1 + S_{ab})(p_a + p_b) - 4p_a p_b}{2p_a p_b}, \quad (7)$$

where S represents the similarity index coded 1, when the pair of considered alleles is identical, and 0 otherwise, and p is the allelic frequency.

Additionally, we also used the Queller and Goodnight (1989) estimator since it was identified as informative for populations under more complex relatedness (Van de Castele et al. 2001):

$$\hat{r}_{xy} = \frac{0.5(S_{ac} + S_{ad} + S_{bc} + S_{bd}) - p_a - p_b}{1 + S_{ab} - p_a - p_b}. \quad (8)$$

This estimator was obtained using the R package “related” (Pew et al. 2015). Since the Queller and Goodnight estimator does not account for inbreeding, the diagonals were set to 1 in all marker-based relationship matrices following Yu et al. (2006). Often, molecular marker-based pairwise relationship matrices are not positive definite due to internal inconsistency resulting from a lack of markers, genotyping errors, or missing values. This might prevent the marker-based relationship matrix to be inverted and cause failed model convergence. Thus, we used the “nearPD” function in the R package “Matrix” which computes the nearest positive definite matrix to an approximate one (Klápště et al. 2014). The vector e represents residual effects following $\text{var}(e) \sim N(0, I\sigma_e^2)$, where σ_e^2 is the residual variance, X and Z are incidence matrices assigning fixed (cluster) and random (genotype) effects in the vectors β and u to the measurements in vector y . Since the marker-based relationship matrices were estimated based on six SSR markers, a simulation was performed to compare the ability of LR and QG estimators to estimate different classes of relatedness using the implemented SSR markers. The simulation was performed using the marker allele frequencies observed in the studied populations, and 200 pairwise cases were simulated for each relatedness class (parent–offspring, full sibs, half sibs, and unrelated pairs). The simulation was performed using the R package “related” (Pew et al. 2015).

All quantitative genetic analyses were performed on a given species within each stand and by masking phenotypes for the alternative species and leaf morphological hybrids as missing data. Narrow-sense heritability for each trait was estimated as follows:

$$h^2 = \frac{\sigma_w^2}{\sigma_w^2 + \sigma_e^2}. \quad (9)$$

Multivariate models

Since the Multiple-trait Model (MTM) software package does not accommodate repeated measurements, genotypic values were estimated for each individual and leaf trait. The genotypic values were estimated using the above described mixed linear model, except for using the identity matrix I instead of the marker-based relationship matrix G . The genetic correlation estimates were obtained through

Multivariate mixed linear models implemented in the MTM R package (de los Campos and Grüneberg 2016) using Gibbs sampling implemented in the R package “BGLR” (de los Campos and Pérez-Rodríguez 2013) as follows:

$$Y = X\beta + Za + e, \quad (10)$$

where Y is the matrix of genotypic values for each combination of trait (columns) and genotype (rows), β is the vector of fixed effects (overall means), a is the matrix of random additive genetic effects following $\text{var}(a) \sim N(0, G1)$, where $G1$ is the variance–covariance structure for additive genetic effects following:

$$G1 = \begin{bmatrix} \sigma_{a_1}^2 & \cdots & \sigma_{a_1 a_n} \\ \vdots & \ddots & \vdots \\ \sigma_{a_n a_1} & \cdots & \sigma_{a_n}^2 \end{bmatrix} \otimes G,$$

where $\sigma_{a_1}^2$ and $\sigma_{a_n}^2$ are the additive genetic covariances between them, $\otimes$ is the Kronecker product and G is the marker-based relationship matrix, e is the matrix of random residual effects following $\text{var}(e) \sim N(0, R)$, where R is the variance–covariance structure for the residual effects following:

$$R = \begin{bmatrix} \sigma_{e_1}^2 & \cdots & \sigma_{e_1 e_n} \\ \vdots & \ddots & \vdots \\ \sigma_{e_n e_1} & \cdots & \sigma_{e_n}^2 \end{bmatrix} \otimes I,$$

where $\sigma_{e_1}^2$ and $\sigma_{e_n}^2$ are the residual variances in the 1st and n th trait, $\sigma_{e_1 e_n}$ and $\sigma_{e_n e_1}$ are the residual covariances between them. The model used uninformative priors, and the number of iterations was set to 500,000, with a burnIn of 200,000, and the thin interval was set to 10. The analysis was performed within each species by masking phenotypes of the other species and interspecific hybrids. Genetic correlations were estimated as follows:

$$r_G = \frac{\sigma_{a_x a_y}}{\sqrt{\sigma_{a_x}^2 \sigma_{a_y}^2}}, \quad (11)$$

where $\sigma_{a_x a_y}$ is the posterior mode for the additive genetic covariance between the x th and the y th traits, $\sigma_{a_x}^2$ and $\sigma_{a_y}^2$ are the posterior modes for the additive genetic variance of the x th and the y th traits, respectively. Additionally, single-site analyses were conducted for each species and Mantel tests were performed between each pair of site-specific

correlation matrices to investigate the impact of genotype-by-environment interactions ($G \times E$) on the general pattern of the correlation structure. Similarly, the Mantel test was performed between species' site-specific correlation matrices to investigate species integrity across the sampled species' natural distribution.

Furthermore, following Hansen and Houle (2008), we estimated each trait's integration, the relative reduction in evolvability of a given trait due to its correlations with other traits based on nonindependent additive genetic variation is estimated as follows:

$$i(x) = 1 - a(x), \quad (12)$$

where $a(x)$ represents the autonomy (independent proportion of additive genetic variation) estimated by:

$$a(x) = \frac{(x^T G^{-1} x)^{-1}}{x^T G x}, \quad (13)$$

where G is the additive variance–covariance matrix and x is the vector of 1 for the trait under study and 0 for the other traits. The analysis was performed using the R package “evolvability” integrating Hansen and Houle's theory.

Next, to obtain a more detailed insight into variance/covariance matrices differences to identify traits that cause matrices dissimilarity (ultimately, identify traits' differential response to selection), we compared genetic correlation matrices between species through the Selection Response Decomposition (SRD; Marroig et al. 2011) approach using the modified random skewers procedure. The method is built on the decomposition of evolutionary responses inferred from selection gradient vectors simulated through random skewers (Cheverud 1996), which are then compared regarding direct and indirect response to selection for each considered trait. Briefly, the correspondence in response to selection between two considered matrices is estimated as the average correlation between vectors of evolutionary responses derived from the same simulated vectors of selection gradients and is plotted as the dotted line in SRD-related figures (see “Results”). The multivariate response to selection is estimated as follows:

$$\begin{bmatrix} \Delta_{z_1} \\ \Delta_{z_2} \\ \vdots \\ \Delta_{z_n} \end{bmatrix} = \begin{bmatrix} G_{11}\beta_1 + G_{12}\beta_2 + \cdots + G_{1n}\beta_n \\ G_{21}\beta_1 + G_{22}\beta_2 + \cdots + G_{2n}\beta_n \\ \vdots \\ G_{n1}\beta_1 + G_{n2}\beta_2 + \cdots + G_{nn}\beta_n \end{bmatrix}, \quad (14)$$

where G is the additive genetic variance/covariance matrix, β is the simulated vector of directional selection gradients, and Δ_z is the inferred vector of response to selection. The

Fig. 3 Distribution of relatedness and self-relatedness coefficients. Histograms of relatedness coefficients estimated in *Q. petraea* (left plots (a, c)) and *Q. robur* (right plots (b, d)) using LR (upper plots (a, b)) and QG (bottom plots (c, d)) estimators as well as self-relatedness inferred through search for the positive definite relationship matrix closest to the estimated relationship matrices using the “nearPD” function.

above-mentioned product of the variance/covariance matrix $\mathbf{G}$ and the vector of directional selection gradients $\boldsymbol{\beta}$ is then decomposed to trait-specific vectors of response as follows:

$$\begin{bmatrix} G_{11}\beta_1 + G_{12}\beta_2 + \dots + G_{1n}\beta_n \\ G_{21}\beta_1 + G_{22}\beta_2 + \dots + G_{2n}\beta_n \\ \vdots \\ G_{n1}\beta_1 + G_{n2}\beta_2 + \dots + G_{nn}\beta_n \end{bmatrix} \Rightarrow \begin{bmatrix} (G_{11}\beta_1 + G_{12}\beta_2 + \dots + G_{1n}\beta_n) \\ (G_{21}\beta_1 + G_{22}\beta_2 + \dots + G_{2n}\beta_n) \\ \vdots \\ (G_{n1}\beta_1 + G_{n2}\beta_2 + \dots + G_{nn}\beta_n) \end{bmatrix}, \quad (15)$$

where the trait-specific vectors are correlated across the tested matrices, and the average correlation across all simulated scenarios is called the SRD score. In case the $\mathbf{G}$

matrices are different, the traits will respond differently to direct or indirect selection, and the SRD score will be low, and the variance of correlations will be large. In the opposite case, the SRD score will be high and the variance in correlations will be small. The analysis was performed using the R package “EvolQG” (Melo et al. 2015).

Results

The genetic population structure analysis performed by DAPC on *Q. petraea*, *Q. robur*, and few hypothetical hybrid individuals identified 24 genetic clusters (Fig. S1).

However, due to extensive historical interspecific hybridization (Leroy et al. 2020), the observed pattern follows a continuous transition between species with few “species-specific clusters” rather than strong species differentiation (Fig. S1). Additionally, the level of introgression was more obvious in *Q. petraea* compared to *Q. robur* (often showing one dominant cluster), but these differences disappeared toward the North (Fig. S2).

The pairwise relatedness estimated by the LR method ranged from -0.196 to 1.035 in *Q. petraea* and from -0.240 to 1.205 in *Q. robur* and the estimates for the QG method ranged from -0.276 to 1.007 in *Q. petraea* and from -0.288 to 1.262 in *Q. robur*. The self-relatedness derived through the process of searching for the positive definite matrix closest to the marker-based relationship matrix reached values for the LR estimator ranging from 1.009 to 2.597 (mean 1.095) in *Q. petraea* and from 1.007 to 2.211 (mean 1.069) in *Q. robur*, and for the QG estimator from 1.001 to 1.988 (mean 1.041) in *Q. petraea* and from 1.001 to 2.174 (mean 1.026) in *Q. robur* (Fig. 3). The relatedness classes simulations using the implemented markers showed that the six SSR markers from our study can recover relatedness but produced high dispersion around the expected values, at the same time, the QG estimator was identified as superior due to lower dispersion of the estimated coefficients of relatedness around their expected values for each class (Fig. S3). The two different marker-based relationship matrices (LR and QG) were also explored to obtain site-specific trait heritability estimates with mixed results regarding model fit (see Akaike’s information criterion provided in Tables S1a and S2a for LR, and Tables S1b and S2b for QG marker-based relationships). Generally, the LR marker-based relationship estimator provided lower and less dispersed heritability estimates across all investigated sites, and for both species (Fig. 4). The lowest heritability estimates were found for SW and WP traits reaching values only between ~ 0.2 and 0.3 , while the highest heritability was achieved for the PL trait ranging between ~ 0.7 and 0.8 in both species (Fig. 4 and Tables S1a, S2a, S1b, and S2b).

Genetic correlations were highest between groups of traits including LL, LW, and WP reaching from 0.48 to 0.60 within *Q. petraea*, and slightly higher from 0.47 to 0.66 within *Q. robur* using the LR relationship matrix (Table S3a). Yet, higher genetic correlations from 0.46 to 0.58 were found in *Q. petraea* compared to their estimates in *Q. robur* reaching values from 0.21 to 0.46 when the QG relationship matrix was implemented (Table S3b). The remaining genetic correlations were either negative or moderately positive, reaching from -0.38 to 0.24 within *Q. petraea*, and from -0.17 to 0.39 within *Q. robur* when the LR relationship matrix was used (Table S3a and Fig. 5), and from -0.30 to 0.29 in *Q. petraea* and from -0.19 to 0.18 in

Fig. 4 Boxplot representation of site-specific heritability estimates for *Q. petraea* (upper plot) and *Q. robur* (bottom), respectively, using the LR marker-based matrix (white boxes) and the QG marker-based matrix (gray boxes), respectively. The box in each boxplot shows the lower quartile, and the median and the upper quartile values, and whiskers indicate the ranges of the variation in each trait’s heritability estimates across sites within a given species. See Table 1 for trait codes, and Fig. 2 for trait specifics.

Q. robur, respectively, when the QG relationship matrix was implemented (Table S3b and Fig. 5). The deviance information criterion (DIC)-based model fit values in the multivariate analysis revealed consistent results when comparing the LR marker-based relationship estimator with the QG marker-based relationship estimator across both species (Table S3a, b). We note here that DAPC-based clustering removed species-specific population structure effects prior to variance decomposition (*Q. petraea* populations usually exhibit higher spatial genetic structure which is attributed to differences in seed dispersal characteristics between the two species, see Streiff et al. 1998). The investigation of $G \times E$ found strong differences in the correlation matrices between sampled forest stands reaching correlations from 0.2 to 0.8 . However, the comparison of $G \times E$ (Mantel correlation between sites) and species’ integrity (Mantel correlation between species) found that $G \times E$ decreases with increasing species integrity. Therefore, the correlation matrices differ mainly due to the level of species integrity (Fig. S4).

The genetic integration estimates across leaf morphology traits suggested different levels of integration (Table 1). Using the LR method-based relationship matrix, values ranged from 0.140 (NL) to 0.537 (LL) in *Q. petraea*, and from 0.072 (PL) to 0.620 (LL) in *Q. robur*; using the QG method-based relationship matrix, values ranged from 0.109 (SW) to 0.521 (LL) in *Q. petraea* and from 0.042 (SW) to 0.605 (LL) in *Q. robur*. Overall, trait integration was consistent for highly integrated traits across oak species and whether the LR or the QG method was used in obtaining genetic relatedness, while results were slightly inconsistent

Fig. 5 Network of genetic correlations for *Q. petraea* (left) and *Q. robur* (right), respectively, using the LR (upper row) and the QG (bottom row) marker-based relationship matrix, respectively. See Table 1 for trait codes, and Fig. 2 for trait specifics. The thickness of connecting lines corresponds to the strengths of correlations; the dashed lines indicate the negative correlations.

Table 1 Coefficients of genetic integration (estimated using the LR and QG marker-based relationship matrix) for each trait within *Q. petraea* and *Q. robur*, respectively.

Trait	LR estimator		QG estimator	
	<i>Q. petraea</i>	<i>Q. robur</i>	<i>Q. petraea</i>	<i>Q. robur</i>
LL	0.537	0.620	0.521	0.605
PL	0.187	0.072	0.225	0.083
LW	0.409	0.477	0.375	0.481
SW	0.176	0.089	0.109	0.042
WP	0.430	0.467	0.399	0.441
NL	0.140	0.119	0.129	0.139
NV	0.245	0.346	0.239	0.373
BS	0.204	0.242	0.247	0.250

Leaf morphological primary descriptors: lamina length (LL), petiole length (PL), lobe width (LW), sinus width (SW), length of lamina from the lamina base to the widest part (WP), number of lobes (NL), number of intercalary veins (NV), and basal shape of the lamina (BS), see also Fig. 2 for specifics.

for all other traits (Table 1). The SRD analysis of genetic correlation matrices further revealed for highly integrated traits (LL, LW, and WP) also high agreement in their

response to selection. At the same time, SRD uncovered substantial differences for the less integrated traits, NV, and BS (Fig. 6).

Discussion

Challenges in detecting phenotypic integration

Inferences about phenotypic integration of quantitative characters essentially depend on patterns we obtain from constructing networks of their genetic correlations. Therefore, the accuracy of quantitative genetic parameter estimates such as additive genetic variances and covariances is crucial. Since genetic correlations represent the relationships between any two traits, both attributes should exhibit decent genetic variance, and hence, correlation analysis requires samples with a high genetic divergence between these traits. The heritabilities estimated for the traits included in our study showed moderate to high values which is in agreement with studies in other species such as quaking aspen with a heritability of 0.47 for the leaf length/width

Fig. 6 Selection Response Decomposition (SRD) performed on the genetic correlation matrices within *Q. petraea* and *Q. robur* using the LR (upper plot), the QG (middle), and the best DIC (bottom) marker-based relationship matrices, respectively. Left panels show, for each trait, the average SRD scores (points) and their two-times standard deviations (error bars) as well as the overall average SRD score (dashed line). Right panels show the trait-specific SRD score deviation differences from the overall average, given on the Y-axis, for each marker-based relationship matrix. Empty circles represent the traits found significantly different in SRD analysis. See Table 1 for trait codes, and Fig. 2 for trait specifics.

ratio (Kanage et al. 2008) or oriental beech with heritabilities from 0.22 (PL) to 0.99 (WP) (Bijarpasi et al. 2019).

Here, we tested several strategies to estimate correlation/variance–covariance matrices (results not shown). We experienced challenges in estimating evolvability-based integration from the variance–covariance matrices that were constructed based on the bivariate mixed linear model using the REML algorithm with packages such as ASReml-R (Butler et al. 2009) or breedR (Muñoz and Sanchez 2017). This was due to the internal inconsistencies of the variance–covariance matrix which produced singularity and, therefore, unstable estimates of integration. There are statistical procedures such as the “nearPD” function in the “Matrix” R package to make the G matrix positive definite, and thus, its inversion numerically stable. However, we did not pursue the strategy based on REML algorithm in our

study. Moreover, including all traits in the model to estimate all genetic covariances at once, resulted in the inability of tools including the REML algorithm to reach convergence. Additionally, Thomas (2005) found that the MCMC-based algorithms are superior to the regression-based algorithms, which produced biased and unrealistic estimates of genetic parameters. Hence, considering the quantity/quality of the implemented genomic resources and the population structure (wild populations), we recommend using Bayesian approaches (Hadfield 2010; de los Campos and Pérez-Rodríguez 2013) to avoid such constraints in the construction of variance–covariance matrices. Additionally, the accuracy of genetic covariances is affected by the level of stratification of the investigated sample and the level of precision in tracking genealogy. The classical genetic analyses are based on using expected relatedness inferred from

documented pedigrees; Bijma and Bastiaansen (2014) found that around 200 half-sib families are needed to achieve statistically significant genetic covariances/correlations.

For studies such as ours, employing a relatively sparse set of genetic markers (six SSRs), it is recommended to compare several relatedness estimators to test their performances with regards to the inherent genetic structure of the studied population for which pedigree information is not available (Van de Castele et al. 2001). Here, the comparison of the two estimators found more reliable estimates of relatedness for the QG compared to the LR estimator (Fig. 3), which was further corroborated in our simulation study through less dispersion of the relatedness coefficients for QG within each relatedness class (Fig. S3). Recent advancements in DNA sequencing have allowed for the generation of genome-wide dense marker arrays (Elshire et al. 2011), which, in turn, permit the estimate of realized relatedness (VanRaden 2008). On the one hand, the most precise estimates of genetic parameters are achieved when only causal variants are used to construct the marker-based relationship matrix (Lippert et al. 2013). On the other hand, some part of the genetic variance/covariance estimates can be lost due to imperfect linkage to QTLs when only sparse marker arrays are implemented (de los Campos et al. 2015). Through further advancements in analytical tools (such as Bayesian multivariate analysis), however, a marker's contribution to either a single trait or to multiple traits (via estimated genetic pleiotropy) can now be uncovered (Karaman et al. 2018).

From an evolutionary perspective, the variance–covariance genetic matrix (G matrix) changes across generations through the interplay of mutation, recombination, and selection (Lande 1980). Theoretical simulation studies that investigated the stability of the G matrix found that a large population size is needed to achieve matrix' stability in size (i.e., their eigenvalues), and strong correlational selection is required to achieve matrix' stability in shape (i.e., their eigenvectors) (Jones et al. 2003, 2004; Arnold 2005). Therefore, weakly integrated attributes will ultimately lead to the G matrix' instability and genetic covariance fluctuations between such traits over generations. Moreover, the G matrix can change due to evolutionary responses and follow the adaptive landscape to reach an adaptive optimum peak (Arnold et al. 2001). The speed of such movement is given by the orientation of the G matrix in the adaptive landscape and by the topography of the landscape defined in the γ matrix (containing information about curvature and orientation of the adaptive optimum peak) (Phillips and Arnold 1989). However, as in our study, the least phenotypically integrated traits are contributing to species diagnostics, we do not expect large changes in genetic covariances between the investigated leaf attributes over generations.

Phenotypic integration of morphological attributes as a response to spatial optimization

The present study identified the main leaf size attributes in oak (LL, LW, and WP, Fig. 2) to be genetically highly integrated characters in both studied species (Table 1). This is consistent with results from multivariate analyses (PCA or multiple correspondence analysis) that demonstrated high phenotypic correlations among the same traits across several *Quercus* species (*Q. petraea*, *Q. robur*, *Quercus pubescens*, and *Quercus pyrenaica* studied in Viscosi et al. 2009). These traits also correspond to the so-called “leaf size gradient” (corresponding to the second synthetic component in the Kremer et al. 2002 study). Indeed, our results are also biologically meaningful as the statistical integration reflects the developmental integration of characters that are mechanistically correlated, and therefore, demonstrates a clear allometric relationship. Such strong phenotypic integration could directly be linked to the leaves' development within the bud which is known to be driven by global regulation of leaf growth and is subject to steric constraints (Couturier et al. 2009). Indeed, the final shape of the developing leaves depends on how these were initially folded within their buds. For *Quercus* spp., each bud contains three types of leaves, the laminate (photosynthetic), the aborted lamina, and the scale leaves, with, on average, eight laminate leaves developing within each bud (Barnola et al. 1990). Therefore, the selection pressure to optimize the space allocated to each leaf within the bud, and the spatial limitations observed during the developmental process, could result in developmental constraints (Lickliter 2014) and the observed allometric relationships. In turn, these constraints would be consistent with the phenotypes' stability across generations by reducing possible leaf shape variation under differential natural selection for a given oak species within this species complex (Alberch 1982).

Moreover, Hammond (1941) found that during leaf enlargement, the ratio between length and width in relative leaf growth extensions remains constant. Therefore, these allometric relationships for mature leaves are already pre-defined during bud formation. Indeed, the above-mentioned global regulation hypothesis (Couturier et al. 2009) can explain the substantial leaf traits phenotypic integration found for both oak species. In our study, this was corroborated by the estimated genetic correlations (Table S3a, b) and genetic integration (Table 1), presumably driven by pleiotropy. Pleiotropy is considered one of the underlying principles of phenotypic integration (Klingenberg 2008) allowing for organ (e.g., floral) evolution in cases where covariation between traits is aligned with the direction of natural selection (Schluter 1996). Conversely, pleiotropy can be considered an evolutionary constraint when directional selection would cause adverse effects on the affected

traits and force the system to converge toward an optimum trade-off (Wagner and Zhang 2011; Smith 2016); such can be expected for oaks with highly integrated leaf morphology traits. Since high integration of these traits was already observed across multiple oak species (Viscosi et al. 2009); it can be assumed to be the product of a long-term stabilizing selection pattern. Such a pattern usually causes the alignment of the $\mathbf{G}$ matrix with the adaptive landscape (Arnold 1992). While commonality in the alignment between $\mathbf{G}$ matrices across related species defines the multi-species primary optima, the actual optimum for each related species is primarily driven by selection, environmental factors, and background perturbations in traits inheritance (Hansen 1997).

While BS, NV, and NL represent the most discriminant morphology characteristics of fully expanded leaves (see Porth et al. 2016, in agreement with previous analyses by Kremer et al. 2002), at oak plants' ontogenetically young stages, especially, NV provides only a poor discriminative index (Ponton et al. 2004). Only the pilosity density on the lamina and the average angle of auricles at the lamina base (equivalent to the BS characteristic described here) were overall the most powerful species discriminant characters (Ponton et al. 2004). Therefore, one can assume that, for phenotypic differentiation between species, the part of the $\mathbf{G}$ matrix, especially for NV, is unstable and varying with plants' ontogenetic development. As mentioned above, the unstable part of the $\mathbf{G}$ matrix across species can result from shifting species primary optima to species-specific actual optima (Hansen 1997). Thus, when species enter new environments, the adaptive optimum peak can be moved causing the observed deviation from species primary optima (Arnold et al. 2001).

Trait phenotypic integration in the context of sympatric species differentiation

Species' phenotypic differentiation is initiated by directional selection, and also, a certain level of stabilizing selection is required to maintain detectable differentiation (Weissing et al. 2011). Stabilizing selection should theoretically produce additive genetic variance reduction (Roff 2012), however, there is no clear explanation of such pattern regarding the genetic variation that is commonly present in populations under stabilizing selection (Johnson and Barton 2005). Hunt et al. (2007) found genetic variation reduction when stabilizing selection acts on a small number of traits. Our analyses showed that the main traits responsible for interspecific differentiation (BS, NV, and NL) between oaks did not suffer reduced genetic variance, as this would be reflected in their heritability estimates but levels were similar across the investigated traits (Tables S1a and S2a and Fig. 4). These traits exhibited considerable

change only in their means (results not provided). Interestingly, BS, NV, and NL were the only traits showing low levels of integration (Table 1); thus, indeed, one requirement (Hunt et al. 2007) for hypothetical genetic variation reduction for traits under stabilizing selection was uncovered. Hence, such traits would lead to poor genetic covariances/correlations estimation under multivariate evaluation and a decreased precision in the estimation of evolutionary response to selection would be the consequence. Evolutionary integration represents the ability of two or more traits to evolve jointly through species' differentiation processes (Armbruster et al. 2014), but we did not find any evidence in oak species' diagnostic leaf shape/structural traits (i.e., BS, NV, and NL). These traits showed generally low levels of genetic integration, as opposed to the leaf dimensional traits LL, LW, and WP, see above. Moreover, the genetic relationships for BS, NV, and NL within each investigated species were weak and even weaker in *Q. robur* than *Q. petraea* (Fig. 5). Such results support the lack of evolutionary integration for certain species' diagnostic traits and suggest that they rather evolved as independent traits. This supports the assumption of a more efficient response to selection for traits that show higher modularity and provides the argument that these traits diverged separately. Consequently, less integrated traits can effectively drive *Quercus* species divergence. These traits may also be plastic based on the weak genetic integration uncovered in our study (see also Ponton et al. 2004 discussed above). Additionally, our study found $\mathbf{G} \times \mathbf{E}$ interactions across the studied traits which were strongly driven by the level of introgression between species across the tested forest stands (Fig. S4). Therefore, the stability of the $\mathbf{G}$ matrix decreases with an increasing level of genetic introgression in sympatric species. Leroy et al. (2020) found historical introgression of *Q. robur* to *Q. petraea* with the highest extent of such introgression within cooler and wetter environments. Additionally, the geographical pattern in the level of introgression was corroborated by the clinal variation in genes contributing to phenological divergence.

We feel that the evolutionary response to selection for the NV trait (Fig. 6) might be directly related to leaf veins' function (Sack and Scoffoni 2013), and therefore, to the known difference in water use efficiency and gas exchange between both oaks (Ponton et al. 2002). Certainly, selection for differential gas exchange properties would affect more characters in addition to NV, because efficiency increase in only one trait would cause limitations to all other implicated leaf traits (McKown et al. 2010). Thus, there is strong evidence that vein architecture traits are well integrated with other important structural and gas exchange-related functional traits (Sack and Scoffoni 2013), yet not considered in the present study. Thus, results from our analyses indicate that such traits would indeed belong collectively to a

different module leading to the correlated selection of involved characters (Sinervo and Svensson 2002). However, such selection can develop independently of the above-mentioned global regulation in leaf dimension development.

As stated above, the studied leaf morphology characters are genetically complex traits (Saintagne et al. 2004; Gailing 2008). For example, the development of lobed leaves is driven by spatially specific expression of a *KNOX* gene (Pires and Dolan 2012). Porth et al. (2016) found genetic associations between β -tubulin gene sequence variability and the species discriminative NL. Additionally, in *Q. robur* × *Q. petraea* interspecific crosses, the same study found an underrepresentation of the β -tubulin allele highly associated with NL, an observation, likely established through possible post-mating species barriers (Abadie et al. 2012).

Conclusions

While we implemented SSRs as molecular markers for our study, the ongoing progress in next-generation sequencing technologies has resulted in the development of new genomic resources for our study species (Ueno et al. 2010; Lepoittevin et al. 2015; Lesur et al. 2015; Bodénés et al. 2016; Lang et al. 2018) which enabled detailed insights into the genetic similarity within populations (Lesur et al. 2018) and further allowed the discovery of fine-scale introgression patterns between oaks (Lang et al. 2018). Finally, all these efforts cumulated in the *Q. robur* reference genome (Ploimion et al. 2018).

The European white oak complex *Q. petraea/robur* represents the best-studied system to date for temperate tree speciation and population dynamics (Kremer and Hipp 2019). Generally known for its extremely high genetic diversity among forest tree species, oak exhibits an exceptional ecological and morphological diversification. It is widely accepted that leaf morphology is adapted to environmental conditions; for example, leaf sizes and shapes correlate with temperature and moisture (Fu et al. 2017). Leaf lobation may enable faster gas exchange and heat dissipation (Rupp and Gruber 2019). In our study, we further elucidated the importance of genetic, epistatic, and plastic effects associated with leaf morphology in this important genus. By considering leaf morphology and genetic data, our study provided important insights into the presence/extent of leaf traits integration within two frequently hybridizing oak species and these oak traits' evolutionary responses to selection. The highest level of traits' genetic integration was found for leaf dimensions (termed "leaf size gradient" cluster) supporting evidence of conserved processes taking place during vegetative bud establishment and development (global regulation). By contrast,

lamina basal shape and intercalary veins' number showed little within-species genetic integration while allowing for the highest difference in evolutionary response to selection between species. Consequently, we propose that organismal modularity allowed for the independent evolution in oak species' morphology that relates to ecological adaptation, thus further driving *Quercus* species divergence.

Data availability

All data used in this analysis are included as a supplement.

Acknowledgements We thank Drs. Ulrike Csaikl, Alexis Ducousso, Bernd Degen, Pablo Goicoechea, Santiago Espinel, Joukje Buiteveld, and Joan Cottrell for the provision of oak genotypic and leaf morphological trait data. We acknowledge funding through the Commission of the European Communities, Community research program "Quality of Life and Management of Living Resources" (Project OAKFLOW QLK5-2000-00960) for this study.

Compliance with ethical standards

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Abadie P, Roussel G, Dencausse B, Bonnet C, Bertocchi E, Louvet JM et al. (2012) Strength, diversity and plasticity of postmating reproductive barriers between two hybridizing oak species (*Quercus robur* L. and *Quercus petraea* (Matt) Liebl.). *J Evol Biol* 25(1):157–173
- Alberch P (1982) Developmental constraints in evolutionary processes. In: Bonner JT (eds) *Evolution and development*. Dahlem workshop reports (Life science research report), vol 22. Springer, Berlin, Heidelberg.
- Armbruster WS, Pélabon C, Bolstad GH, Hansen TF (2014) Integrated phenotypes: understanding trait covariation in plants and animals. *Philos Trans R Soc B* 369(1649):20130245
- Arnold SJ (1992) Constraints on phenotypic evolution. *Am Nat* 140: S85–S107
- Arnold SJ (2005) The ultimate causes of phenotypic integration: lost in translation. *Evolution* 59(9):2059–2061
- Arnold SJ, Pfender ME, Jones AG (2001) The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112–113:9–32
- Barnola P, Alatou D, Lacointe A, Lavarenne S (1990) Etude biologique et biochimique du déterminisme de la croissance rythmique du chène pédoncule (*Quercus robur* L.). Effets de l'ablation des feuilles. *Ann For Sci* 47(6):619–631
- Berg R (1960) The ecological significance of correlation pleiades. *Evolution* 14(2):171–180
- Bijarpari MM, Shahraji TR, Lahiji HS (2019) Genetic variability and heritability of some morphological and physiological traits in *Fagus orientalis* Lipsky along an elevation gradient in Hyrcanian forests. *Folia Oecol* 46(1):45–53
- Bijma P, Bastiaansen JW (2014) Standard error of the genetic correlation: how much data do we need to estimate purebred-crossbred genetic correlations? *Genet Sel Evol* 46:1–6

- Bissell E, Diggle P (2010) Modular genetic architecture of floral morphology in *Nicotiana*: quantitative genetic and comparative phenotypic approaches to floral integration. *J Evol Biol* 23(8):1744–1758
- Bodénès C, Chancerel E, Ehremann F, Kremer A, Plomion C (2016) High-density linkage mapping and distribution of segregation distortion regions in the oak genomes. *DNA Res* 23:115–124
- Brock MT, Weinig C (2007) Plasticity and environment-specific covariances: an investigation of floral – vegetative and within flower correlations. *Evolution* 61(12):2913–2924
- Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual. The State of Queensland, Department of Primary Industries and Fisheries, Brisbane.
- Cheverud JM (1982) Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution* 36(3):499–516
- Cheverud JM (1996) Quantitative fanatic analysis of cranial morphology in the cotton-top (*Saguinus oedipus*) and saddle-back (*S. fuscicollis*) tamaris. *J Evol Biol* 9:5–42
- Conner JK, Via S (1993) Patterns of phenotypic and genetic correlation among morphological and life-history traits in wild radish, *Raphanus raphanistrum*. *Evolution* 47:704–711
- Couturier E, Du Pont SC, Douady S (2009) A global regulation inducing the shape of growing folded leaves. *PLoS ONE* 4(11):e7968
- de los Campos G, Pérez-Rodríguez P (2013) BGLR: Bayesian generalized linear regression. R package version 1(3). Accessed 24 Apr 2016
- de los Campos G, Sorensen D, Gianola D (2015) Genomic heritability: what it is? *PLoS Genet* 11:e1005048
- de los Campos G, Grüneberg A (2016) MTM (multiple-trait model) package. <http://quantgen.github.io/MTM/vignette.html>. Accessed 10 May 2018
- De Micco V, Aronne G (2012) Morpho-anatomical traits for plant adaptation to drought. In: Aroca R (ed.) Plant responses to drought stress. Springer, Berlin, Heidelberg
- Dow DB, Ashley MV, Howe HF (1995) Characterization of highly variable (GA/CT)_n microsatellites in the bur oak, *Quercus macrocarpa*. *Theor Appl Genet* 91:137–141
- Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* 6:e19379
- Esteve-Altava B (2017) In search of morphological modules: a systematic review. *Biol Rev* 92:1332–1347
- Feng C, Feng C, Yang L, Kang M, Rausher MD (2019) Genetic architecture of quantitative flower and leaf traits in a pair of sympatric sister species of *Primulina*. *Heredity* 122(6):864–876
- Fu G, Dai X, Symanzik J, Bushman S (2017) Quantitative gene-gene and gene-environment mapping for leaf shape variation using tree-based models. *N Phytol* 213:455–469
- Gailing O (2008) QTL analysis of leaf morphological characters in *Quercus robur* full-sib family (*Q. robur* x *Q. robur* ssp. *slavonica*). *Plant Biol* 10(5):624–634
- Gianola D, Norton H (1981) Scaling threshold characters. *Genetics* 99(2):357–364
- Hadfield JD (2010) MCMC methods for multi-response generalized linear models: the MCMCglmm R package. *J Stat Softw* 33(2):1–22
- Hammond D (1941) The expression of genes for leaf shape in *Gossypium hirsutum* L. *Am J Bot* 28(2):124–138
- Hansen TF (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51:1341–1351
- Hansen TF (2006) The evolution of genetic architecture. *Annu Rev Ecol Evol Syst* 37:123–157
- Hansen TF, Houle D (2008) Measuring and comparing evolvability and constraint in multivariate characters. *J Evol Biol* 21(5):1201–1219
- Healy TM, Brennan RS, Whitehead A, Schulte PM (2018) Tolerance traits related to climate change resilience are independent and polygenic. *Glob Change Biol* 24:5348–5360
- Hermant M, Prinzing A, Vernon P, Convey P, Hennion F (2013) Endemic species have highly integrated phenotypes, environmental distributions and phenotype-environment relationship. *J Biogeogr* 40(8):1583–1594
- Hill WG, Zhang J (2012) On the pleiotropic structure of the genotype-phenotype map and the evolvability of complex organisms. *Genetics* 190(3):1131–1137
- Hodel RG, Segovia-Salcedo MC, Landis JB, Crowl AA, Sun M, Liu X et al. (2016) The report of my death was an exaggeration: a review for researchers using microsatellites in 21st century. *Appl Plant Sci* 4(6):1600025
- Hubert F, Grimm GW, Jousset E, Berry V, Franc A, Kremer A (2014) Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus*. *Syst Biodivers* 12:405–423
- Hunt J, Blows MW, Zajitschek F, Jennions MD, Brooks R (2007) Reconciling strong stabilizing selection with the maintenance of genetic variation in a natural population of black field crickets (*Teleogryllus commodus*). *Genetics* 177(2):875–880
- Johnson T, Barton N (2005) Theoretical models of selection and mutation on quantitative traits. *Philos Transl R Soc B* 360(1459):1411–1425
- Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11):1403–1405
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11(1):1
- Jones AG, Arnold SJ, Bürger R (2003) Stability of the G-matrix in a population experiencing stabilizing selection, pleiotropic mutation, and genetic drift. *Evolution* 57:1747–1760
- Jones AG, Arnold SJ, Bürger R (2004) Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* 58:1639–1654
- Kanage MK, Ryel RJ, Mock KE, Pfreder ME (2008) Quantitative-genetic variation in morphological and physiological traits within a quaking aspen (*Populus tremuloides*) population. *Can J For Res* 38:1690–1694
- Karaman E, Lund MS, Anche MT, Janss L, Su G (2018) Genomic predictions using multitrait weighted GBLUP accounting for heterogeneous variances and covariances across genome. *Genes Genomes Genet* 8:3549–3558
- Kashtan N, Alon U (2005) Spontaneous evolution of modularity and network motifs. *Proc Natl Acad Sci USA* 102(39):13773–13778
- Klápště J, Lstibůrek M, El-Kassaby YA (2014) Estimates of genetic parameters and breeding values from western larch open-pollinated families using marker-based relationship. *Tree Genet Genomes* 10(2):241–249
- Klingenberg CP (2008) Morphological integration and developmental modularity. *Annu Rev Ecol Evol Syst* 39:115–132
- Kremer A, Dupouey JL, Deans JD, Cottrell J, Csaikl U, Finkeldey R et al. (2002) Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across Western European mixed oak stands. *Ann For Sci* 59(7):777–787
- Kremer A, Hipp AL (2019) Oak: an evolutionary success story. *N Phytol* 226:987–1011
- Kremer A, Lowe AJ, Munro RC, Dick J, Cavers S et al. (2008) Intra and interspecific gene flows in oaks as mechanisms promoting genetic diversity and adaptive potential. OAKFLOW. Final report (2001–2005). European Commission, 73 pp (Quality of life and management of living resources (1.1.1.), CEH Project Number: C01615)
- Lande R (1980) The genetic covariances between characters maintained by pleiotropic mutation. *Genetics* 94:203–215

- Lang T, Abadie P, Leger V, Decourcelle T, Frigerio J-M, Burban C et al. (2018) High-quality SNPs from genic regions highlight introgression patterns among European white oaks (*Quercus petraea* and *Q. robur*). bioRxiv <https://doi.org/10.1101/388447>
- Laughlin DC, Messier J (2015) Fitness of multidimensional phenotypes in dynamic adaptive landscape. *Trends Ecol Evol* 30 (8):487–496
- Lepoittevin C, Bodénès E, Chancerel E, Villate L, Lang T, Lesur I et al. (2015) Single-nucleotide polymorphism discovery and validation in high-density SNP array for genetic analysis in European white oaks. *Mol Ecol Resour* 15(6):1446–1459
- Leroy T, Louvet J-M, Lalanne C, Le Provost G, Labadie K, Aury J-M et al. (2020) Adaptive introgression as a driver of local adaptation to climate in European white oaks. *N Phytol* 226:943–946
- Lesur I, Alexandre H, Bourry C, Chancerel E, Plomion C, Kremer A (2018) Development of target sequence capture and estimation of genomic relatedness in a mixed oak stands. *Front Plant Sci* 9:996
- Lesur I, Le Provost G, Bento P, Da Silva C, Leplé J-C, Murat F et al. (2015) The oak gene expression atlas: insights into Fagaceae genome evolution and the discovery of genes regulated during bud dormancy release. *BMC Genomics* 16:112
- Lévy G, Becker M, Duhamel D (1992) A comparison of the ecology of pedunculate and sessile oaks: radial growth in the Centre and NW of France. *For Ecol Manag* 55:51–63
- Lewis MC (1972) The physiological significance of variation in leaf structure. *Sci Prog* 60(237):25–51
- Lewontin RC (1978) Adaptation. *Sci Am* 239:212–231.
- Lickliter R (2014) Developmental evolution and the origins of phenotypic variation. *Biomol Concepts* 5(4):343–352
- Lippert C, Quon G, Kang EY, Kadie CM, Listgarten J, Heckerman D (2013) The benefits of selecting phenotype-specific variants for applications of mixed models in genomics. *Sci Rep* 3:1815
- Lipson H, Pollack JB, Suh NP (2002) On the origin of modular variation. *Evolution* 56(8):1549–1556
- Lo Gullo MA, Salleo S (1988) Different strategies of drought resistance in three Mediterranean sclerophyllous trees growing in the same environmental conditions. *N Phytol* 108:267–276
- Lucas LK, Nice CC, Gompert Z (2018) Genetic constraints on wing pattern variation in *Lycaeides* butterflies: a case study on mapping complex, multifaceted traits in structured populations. *Mol Ecol Resour* 18:892–907
- Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics* 152(4):1753–1766
- Mariette S, Cottrell J, Csaikl UM, Goikoechea P, König A, Lowe AJ et al. (2002) Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed *Q. petraea* (Matt.) Liebl. and *Q. robur* L. stands. *Silvae Genet* 51:72–79
- Marroig G, Melo D, Porto A, Sebastiao H, Garcia G (2011) Selection Response Decomposition (SRD): a new tool for dissecting differences and similarities between matrices. *Evol Biol* 38 (2):225–241
- McKown AD, Cochard H, Sack L (2010) Decoding leaf hydraulics with a spatially explicit model: principles of venation architecture and implications for its evolution. *Am Nat* 175:447–460
- Melo D, Garcia G, Hubbe A, Assis AP, Marroig G (2015) EvolQG—an R package for evolutionary quantitative genetics. *F1000Research* 4:925
- Muñoz F, Sanchez L (2017) breedR: statistical methods for forest genetic resources analysis. R package version 0.12-2. Accessed 12 Feb 2018
- Murren CJ (2002) Phenotypic integration in plants. *Plant Species Biol* 17(2-3):89–99
- Murren CJ (2012) The integrated phenotype. *Integr Comp Biol* 52 (1):64–76
- Olson EC, Miller RL (1958) Morphological integration. University of Chicago Press, Chicago
- Ordano M, Fomoni J, Boege K, Domínguez CA (2008) The adaptive value of phenotypic floral integration. *N Phytol* 179(4):1183–1192
- Pepper JW (2003) The evolution of evolvability in genetic linkage patterns. *Biosystems* 69(2-3):115–126
- Pew J, Muir PH, Wang J, Frasier TR (2015) Related: an R package for analysing pairwise relatedness from codominant molecular markers. *Mol Ecol Resour* 15(3):557–561
- Phillips PC, Arnold SJ (1989) Visualizing multivariate selection. *Evolution* 43(6):1209–1222
- Pigliucci M (2003) Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecol Lett* 6(3):265–272
- Pires ND, Dolan L (2012) Morphological evolution in land plants: new designs with old genes. *Philos Trans R Soc B* 367(1588):508–518
- Plomion C, Aury J-M, Amselem J, Leroy T, Murat F, Duplessis S et al. (2018) Oak genome reveals facets of long lifespan. *Nat Plants* 4:440–452
- Ponton S, Dupouey J-L, Bréda N, Dreyer E (2002) Comparison of water-use efficiency of seedlings from two sympatric oak species: genotype x environment interactions. *Tree Physiol* 22:413–422
- Ponton S, Dupouey J-L, Dreyer E (2004) Leaf morphology as species indicator in seedling of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl.: modulation by irradiance and growth flush. *Ann For Sci* 61(1):73–80
- Porth I, Koch M, Berenyi M, Burg A, Burg K (2005a) Identification of adaptation-specific differences in mRNA expression of sessile and pedunculate oak based on osmotic stress-induced genes. *Tree Physiol* 25(10):1317–1329
- Porth I, Klápště J, McKown AD, La Mantia J, Hamelin R, Skyba O et al. (2014) Extensive functional pleiotropy of REVOLUTA substantiated through forward genetics. *Plant Physiol* 164 (2):548–554
- Porth I, Garnier-Géré P, Klápště J, Scotti-Saintagne C, El-Kassaby YA, Burg K et al. (2016) Species-specific alleles at a β -tubulin gene show significant associations with leaf morphological variation within *Quercus petraea* and *Q. robur* populations. *Tree Genet Genomes* 12(4):81
- Porth I, Scotti-Saintagne C, Barreneche T, Kremer A, Burg K (2005b) Linkage mapping of osmotic stress induced genes of oak. *Tree Genet Genomes* 1(1):31–40
- Press MC (1999) The functional significance of leaf structure: a search for generalizations. *N Phytol* 143(1):213–219
- Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution* 43(2):258–275
- Ramírez-Valiente JA, Lorenzo Z, Soto A, Valladares F, Gil L, Aranda I (2009) Elucidating the role of genetic drift and natural selection in cork oak differentiation regarding drought tolerance. *Mol Ecol* 18(18):3803–3815
- Roff DA (2012) Evolutionary quantitative genetics. Springer Science & Business Media, Dordrecht
- Royer DL, Wilf P, Janesko DA, Kowalski EA, Dilcher DL (2005) Correlations of climate and plant ecology to leaf size and shape: potential proxies for the fossil record. *Am J Bot* 92:1141–1151
- Rupp AIKS, Gruber P (2019) Biomimetic groundwork for thermal exchange structures inspired by plant leaf design. *Biomimetics* 4:75
- Sack L, Scoffoni C (2013) Leaf venation: structure, function, development, evolution, ecology and adaptation in the past, present and future. *N Phytol* 198:983–1000
- Saintagne C, Bodenes C, Barreneche T, Pot D, Plomion C, Kremer A (2004) Distribution of genomic regions differentiating oak species assessed by QTL detection. *Heredity* 92(1):20–30
- Savell KR, Auerbach BM, Roseman CC (2016) Constraint, natural selection, and the evolution of human body form. *Proc Natl Acad Sci USA* 113(34):9492–9497

- Schluter D (1996) Adaptive radiation along genetic lines of least resistance. *Evolution* 50:1766–1774
- Scotti-Saintagne C, Mariette S, Porth I, Goicoechea PG, Barreneche T, Bodénès C et al. (2004) Genome scanning for interspecific differentiation between two closely related oak species (*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.). *Genetics* 168:1615–1626
- Sinervo B, Svensson E (2002) Correlational selection and the evolution of genomic architecture. *Heredity* 89:329–338
- Smith SD (2016) Pleiotropy and the evolution of floral integration. *N Phytol* 209:80–85
- Steinkellner H, Fluch S, Turetschek E, Lexer C, Streiff R, Kremer A et al. (1997) Identification and characterization of (GA/CT)_n microsatellite loci from *Quercus petraea*. *Plant Mol Biol* 33:1093–1096
- Streiff R, Labbe T, Bacilieri R, Steinkellner H, Glossl J, Kremer A (1998) Within-population genetic structure in *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. assessed with isozymes and microsatellites. *Mol Ecol* 7(3):317–328
- Thomas SC (2005) The estimation of genetic relationships using molecular markers and their efficiency in estimating heritability in natural populations. *Philos Trans R Soc B* 360:1457–1467
- Ueno S, Le Provost G, Leger V, Klopp C, Noirot C, Frigerio J-M et al. (2010) Bioinformatics analysis of ESTs collected by Sanger and pyrosequencing methods for a keystone forest tree species: oak. *BMC Genomics* 11:650
- Van de Casteele T, Galbusera P, Matthysen E (2001) A comparison of microsatellite-based pairwise relatedness estimators. *Mol Ecol* 10 (6):1539–1549
- VanRaden PM (2008) Efficient methods to compute genomic prediction. *J Dairy Sci* 91:4414–4423
- Viscosi V, Lepais O, Gerber S, Fortini P (2009) Leaf morphological analyses in four European oak species (*Quercus*) and their hybrids: a comparison of traditional and geometric morphometric methods. *Plant Biosyst* 143(3):564–574
- Wagner GP, Pavlicev M, Cheverud JM (2007) The road to modularity. *Nat Rev Genet* 8(12):921–931
- Wagner GP, Zhang J (2011) The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet* 12(3):204–213
- Wang Z, Liao B-Y, Zhang J (2010) Genomic patterns of pleiotropy and the evolution of complexity. *Proc Natl Acad Sci USA* 107:18034–18039
- Weissing FJ, Edelaar P, Sander van Doorn G (2011) Adaptive speciation theory: a conceptual review. *Behav Ecol Sociobiol* 65(3):461–480
- Young RL, Badyaev AV (2006) Evolutionary persistence of phenotypic integration: influence of developmental and functional relationships on complex trait evolution. *Evolution* 60:1291–1299
- Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF et al. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* 38:203–208