RESEARCH PAPER

CrossMark

Geometric morphometric analyses of leaf shapes in two sympatric Chinese oaks: *Quercus dentata* Thunberg *and Quercus aliena* Blume (Fagaceae)

Yuan Liu[1] · Yuejuan Li[1] · Jialin Song[2] · Ruipu Zhang[3] · Yu Yan[1] · Yuyao Wang[1] · Fang K. Du[1]

Abstract

· *Key message* Geometric morphometric analyses (GMMs) of the leaf shape can distinguish two congeneric oak species *Quercus dentata* Thunberg and *Quercus aliena* Blume in sympatric areas.

· *Contexts* High genetic and morphological variation in different *Quercus* species hinder efforts to distinguish them. In China, *Q. dentata* and *Q. aliena* are generally sympatrically distributed in warm temperate forests, and share some leaf morphological characteristics.

· *Aims* The aim of this study was to use the morphometric methods to discriminate these sympatric Chinese oaks preliminarily identified from molecular markers.

· *Methods* Three hundred sixty-seven trees of seven sympatric *Q. dentata* and *Q. aliena* populations were genetically assigned to one of the two species or hybrids using Bayesian clustering analysis based on nSSR. This grouping served as a priori classification of the trees. Shapes of 1835 leaves from the 367 trees were analyzed in terms of 13 characters (landmarks) by GMMs. Correlations between environmental and leaf morphology parameters were studied using linear regression analyses.

· *Results* The two species were efficiently discriminated by the leaf morphology analyses (96.9 and 95.9% of sampled *Q. aliena* trees and *Q. dentata* trees were correctly identified), while putative hybrids between the two species were found to be morphologically intermediate. Moreover, we demonstrated that the leaf morphological variations of *Q. aliena*, *Q. dentata*, and their putative hybrids are correlated with environmental factors, possibly because the variation of leaf morphology is part of the response to different habitats and environmental disturbances.

· *Conclusion* GMMs were able to correctly classify individuals from the two species preliminary identified as *Q. dentata* or *Q. aliena* by nSSR. The high degree of classification accuracy provided by this approach may be exploited to discriminate other problematic species and highlight its utility in plant ecology and evolution studies.

Keywords Geometric morphometrics · Genetic assignment · Leaf morphology · *Quercus* · Sympatric distribution

---

✉ Fang K. Du
  dufang325@bjfu.edu.cn

[1] College of Forestry, Beijing Forestry University, 100083 Beijing, People's Republic of China

[2] School of Biological Science and Technology, University of Jinan, 250022 Jinan, People's Republic of China

[3] Kunyushan National Forest Park, 264000 Yantai, People's Republic of China

INRA
SCIENCE & IMPACT

Springer

1 Introduction

There are many phenotypic traits which can be used for the differentiation of taxa, such as leaf morphology, anatomical descriptors, and other structural characters. Among them, leaf morphology is important for differentiating plant taxa (Costa et al. 2009), especially congeneric woody species, which tend to have little variation in floral features and short flowering seasons (Elias 1980). This is particularly true for oaks (*Quercus* L., Fagaceae), in which leaf features are "the most important" characters for discrimination (Stace 1982). *Quercus* is a genus of evergreen or deciduous shrubs and trees, and the largest genus in the Fagaceae family. About 450 species are widely distributed in temperate and subtropical regions of the northern hemisphere (Nixon 1993, 1997; Huang et al. 1999a, b; Manos et al. 1999), and 35 species are widely distributed in China (Huang et al. 1999a, b). *Quercus* taxonomy is extremely complex due to wide geographical distribution and species overlap, high levels of variability and phenotypic plasticity, and widespread hybridization where species ranges overlap (Rushton 1993; Bacilieri et al. 1996; Manos et al. 1999; Petit et al. 2005; Gerber et al. 2014).

Thus, discriminating oak species in mixed stands is not straightforward. However, progress has been made in the development of leaf morphological methods capable of resolving these problems in recent years. Traditional morphological measurements typically apply statistical techniques to a wide range of measurements, such as distances and distance ratios, angles, areas, and volumes (Mitteroecker and Gunz 2009). It can be efficient for discriminating the species. For example, Kremer et al. (2002) distinguished two oak species in nine mixed oak stands using 14 shape-related traits. However, traditional morphological methods are not effective for shape analysis, and do not allow interpretable graphic representation. Modern geometric morphometric methods (GMMs) using landmark and outline data are suitable for differentiating complex taxa because they can partition total coordinate variation into shape and non-shape (translation, rotation, and size) components, regardless of location, orientation, and scale factors (Slice 1996; MacLeod and Forey 2002; Adams et al. 2004; Mitteroecker and Gunz 2009; Viscosi et al. 2009). GMMs are also helpful for studying bilateral asymmetry and allometric patterns. Separate analysis of symmetric and asymmetric components of leaf variation in a principal component analysis (PCA) and two-block partial least-square (2B-PLS) analysis enables leaf shape variability to be recovered, thereby reducing the variation caused by developmental noise and instability (Viscosi 2015). Using GMMs, geometric information on morphological differences can be displayed in various images and diagrams, thereby highlighting differences in patterns and increasing the interpretability of results (Rohlf 2010). Due to these advantages, GMMs are increasingly used in morphological analysis (e.g., Viscosi and Cardini 2011;

Viscosi and Fortini 2011; Souza et al. 2012; Viscosi et al. 2009; Tucić et al. 2018). Moreover, recent advances have shown that combinations of leaf morphological characters and molecular markers can significantly improve taxonomic delimitations (Viscosi et al. 2012; Rellstab et al. 2016).

In recent years, molecular markers were widely used in species discrimination due to the rapid development of DNA technology (e.g., Jaramillo-Correa et al. 2008; Du et al. 2011 for gymnosperm-conifers; Arnold et al. 2012; Eaton et al. 2015 for angiosperms). Various molecular markers have been used for species classification. Among them, nuclear microsatellite markers or simple sequence repeats (nSSRs) are biparentally inherited and experience extensive gene flow through both pollen and seeds (e.g., Petit et al. 2005; Du et al. 2013). Due to their co-dominant inheritance, high degree of polymorphism and relative ease of transfer between closely related species, nSSRs have become one of the most popular sources of molecular markers (Guichoux et al. 2011). Leaf shape is mainly genetically fixed but also affected by environmental factors (Graham et al. 2010). The environmental heterogeneity such as elevation, temperature, water, etc. can also affect the shape and size of leaves, e.g., leaves must be as wide as possible to absorb sufficient light and as flat and thin as possible to facilitate gas exchange, but will be dried out if too wide and thin (Tsukaia 2005). Thus, it is important to take environment factors into account when trying to explain the morphological differences.

Quercus dentata Thunberg and *Quercus aliena* Blume are generally sympatrically distributed in warm temperate forests, and they are important components of both deciduous broad-leaved and mixed broad-leaved forests (Ren and Wang 1985; Peng et al. 2007). The two species are found, and play major ecological roles, on southern slopes of mountains in northern China, and the neighboring countries Japan and Korea (Kazue and Atsuko 2015). In addition, like other oaks, the two species have great economic values, e.g., their leaves are used for sericulture, fruits for starch extraction, and bark for tannin extraction (Peng et al. 2007). Genetically, the two species are members of the *Quercus* (white oak) section, closely related, and regarded as sister species (Hubert et al. 2014). There are several major morphological differences between leaves of *Q. dentata* and *Q. aliena*: the former have shorter petioles, more densely stellate tomentose abaxial surfaces, more rounded bases, larger apices with shorter, blunter tips, and less secondary veins on each side of midvein compared with *Q. aliena* (Huang et al. 1999a, b). However, intermediate types of leaves have been observed in sympatric forests of *Q. dentata* and *Q. aliena*, which hinders discrimination of the two species (Huang et al. 1999a, b). Many previous studies on these two closely related species have mainly focused on their physiology and ecology (e.g., Fujiwara and Harada 2015; Box and Fujiwara 2015; Nugroho et al. 2016), while relatively few articles have addressed their morphological characteristics.

In this study, we collected numerous leaves from *Q. dentata* and *Q. aliena* trees and then investigated their morphological characteristics using selected leaf characters in a multivariate setting based on an a priori classification, to address the following three questions. First, can the two species be distinguished using selected leaf characters? Second, are symmetric and asymmetric patterns of variation helpful for distinguishing the two species and their hybrids? Finally, are leaf shape and size influenced by their environments, and if so which environmental factors are involved?

2 Materials and methods

2.1 Sampling strategy

We sampled trees in seven sympatric populations of *Q. dentata* and *Q. aliena,* covering most of their sympatric distribution in northern China (Fig. 1). Details of each population are listed in Table 1. Five intact, mature leaves were collected from each of 367 selected trees (which were spaced at least 5 m apart to minimize risks of selecting clones) for morphological analyses and DNA extraction. Thus, 1835 leaves were collected in total, all of which were taken to the laboratory for analysis, and voucher specimens were preserved at the Molecular Ecology Laboratory of Beijing Forestry University, China.

2.2 Genetic classification of the specimens

Genomic DNA was extracted from leaves of the 367 oak trees using a plant genomic DNA kit (Tiangen Biotech, China) following the manufacturer's instructions. DNA concentrations in the extracts were measured by an ultramicro-spectrophotometer (Thermo Fisher, USA), and adjusted to a final concentration of 20–30 ng/μl for subsequent use. Twelve pairs of polymorphic nuclear simple sequence repeats (nSSRs) primers were selected for genotyping. These markers were originally developed in other oak species (Dow et al. 1995; Ueno et al. 2008, 2009; Durand et al. 2010) and the
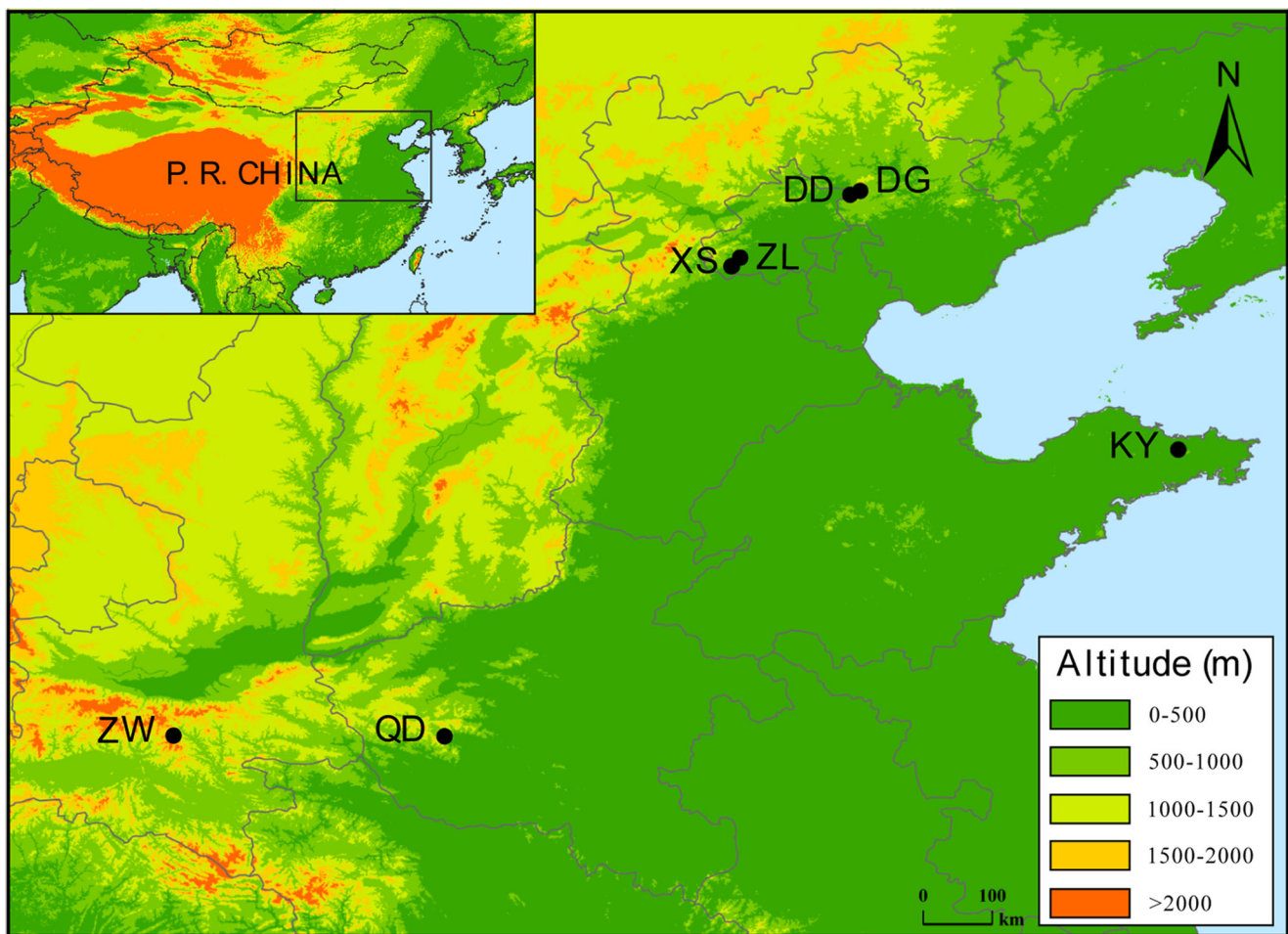

Fig. 1 Map showing the *Quercus dentata* and *Quercus aliena* sampling locations in this study, XS (XiangShulin), ZL (ZhongLou), DG (DongGou), DD (DaDianzi), QD (QiaoDuan), KY (KunYushan), and ZW (ZiWuyu)

Table 1 Locations and numbers of individuals of *Quercus dentata* and *Quercus aliena* based on genetic assignment

Pop	*Q. dentata*	*Q. aliena*	Hybrids	Total	Latitude (º)	Longitude (º)
XS	19	92	7	118	39.67	115.49
ZL	16	54	10	80	39.41	115.50
QD	23	7	8	38	33.55	112.08
DD	21	10	4	35	40.36	117.22
DG	6	10	4	20	40.39	117.30
ZW	23	2	9	34	33.56	108.55
KY	21	19	2	42	37.28	121.64
Total	129	194	44	367		

existence of repeats and polymorphisms were confirmed in the species we studied by Lyu et al. (2018). Detailed information about the primers, amplification, and genotyping procedures is presented in Supplementary Table S1. Bayesian cluster analysis was used to assign individuals to K clusters (1–10) on the basis of genotypes by STRUCTURE V2.3 (Pritchard et al., 2000). In each case, the program was set to run 200,000 Markov Chain Monte Carlo cycles (MCMC), following 100,000 burn-in cycles without any species identification information (USEPOPINFO = 0). Twenty iterations were performed for each K to estimate the best K according to Evanno et al. (2005). The most likely number of clusters was defined by computing ΔK and mean LnP(K) according Evanno et al. 2005 and Janes et al. 2017 (Fig. S1) using the structure harvester program (Earl and Vonholdt 2012). In order to validate the number of genetic clusters, we have provided two STRUCTURE plots (K = 2 and 3) for a visual comparison in Fig. S2. Admixture coefficient (Q) values were used to determine whether there were pure species or hybrids. The choice of the threshold of the Q value was dependent on efficiency and accuracy of assigning individuals to hybrid and purebred groups (Vähä and Primmer 2006). The reported threshold Q value for oaks is $Q \leq 0.9$ (Lepais et al. 2009; Peñaloza-Ramírez et al. 2010) or $Q \leq 0.8$ (Zeng et al. 2010). In this study, a strict threshold of 0.9 was chosen for a priori classification of the two species based on previous work in oaks (e.g., Lepais et al. 2009; Viscosi et al. 2012; Lyu et al. 2018). Individuals with $Q \geq 0.9$ were considered to be purebreds, while individuals with $Q < 0.9$ from two genetic groups were considered to be hybrids individuals between two oak species.

2.3 Landmark configuration

The leaves were pressed, dried, and then scanned using a CanoScan 5600 F scanner (Canon Inc. Japan) at 600 dpi resolution. Scanned images were digitized to landmarks by Image J (Abràmoff et al. 2005). Thirteen landmarks were

selected on each leaf, including the primary veins, symmetric structures, and homologous part (Fig. 2) (Jensen et al. 2002; Savriama and Klingenberg 2011; Viscosi et al. 2009). The landmarks consisted of three distributed along the middle axis of leaves (LM 1–LM 3) and nine others symmetrically distributed in other parts of the leaves (LM 4–LM 13). These landmarks were converted to a configuration of 13 pairs of Cartesian coordinates (x, y) for each leaf and the data were stored in ".txt" file format using the Image J program.

2.4 Statistical analysis

Cartesian x and y coordinates of the landmarks were used for the following analysis using the MorphoJ program (Klingenberg 2011). All the configurations were firstly rotated to generate a mean leaf configuration with minimal differences between corresponding landmarks by generalized Procrustes analysis (GPA) and were generated to a mirror image (Gower 1975; Rohlf and Slice 1990; Klingenberg et al. 2002; Klingenberg 2003; Zelditch et al. 2012). Secondly, leaf-level shape information was extracted for deviation detection by using the Procrustes distance to measure the absolute magnitude of the shape deviation and the Mahalanobis distance to provide an indication of the individual abnormality relative to others in the sample and a new dataset with no outliers (specimens that deviate strongly from the average) was constructed (Klingenberg and Monteirob 2005). Then, subsets of landmarks for the symmetric and asymmetric components of leaf shape were generated by GPA of the original configurations and their mirror image. Here, symmetrical component is composed of median axis and joint mirror image shifts of pairs of corresponding landmarks in any direction, while asymmetric component quantifying differences between the original configurations and mirrored images as described in the manual of MorphoJ (Klingenberg et al. 2002). Fourthly, the variation of leaf shape hierarchically distributed among leaves, trees, and species was used to create mean configurations at each of these levels. Leaf-level data were used for leaf shape variation analysis, tree-level data for species discrimination, and species-level data (including the hybrid individuals) for multivariate analysis. Centroid size, symmetric, and asymmetry components of leaves were used for further analysis. Finally, matrix correlations of each hierarchical level were used as datasets for subsequent multivariate statistical analyses.

The normalized matrices were then subjected to PCA to identify latent (principal) components capturing most of the variation in the dataset (Benzécri 1992), i.e., global variation in leaf landmark configurations, and identify the leaf shape variability at leaf level. To investigate the leaf morphology variations of the entire samples, the symmetric and asymmetric components were separately subjected to leaf-level PCA. Because leaf shape variability (symmetric and asymmetric
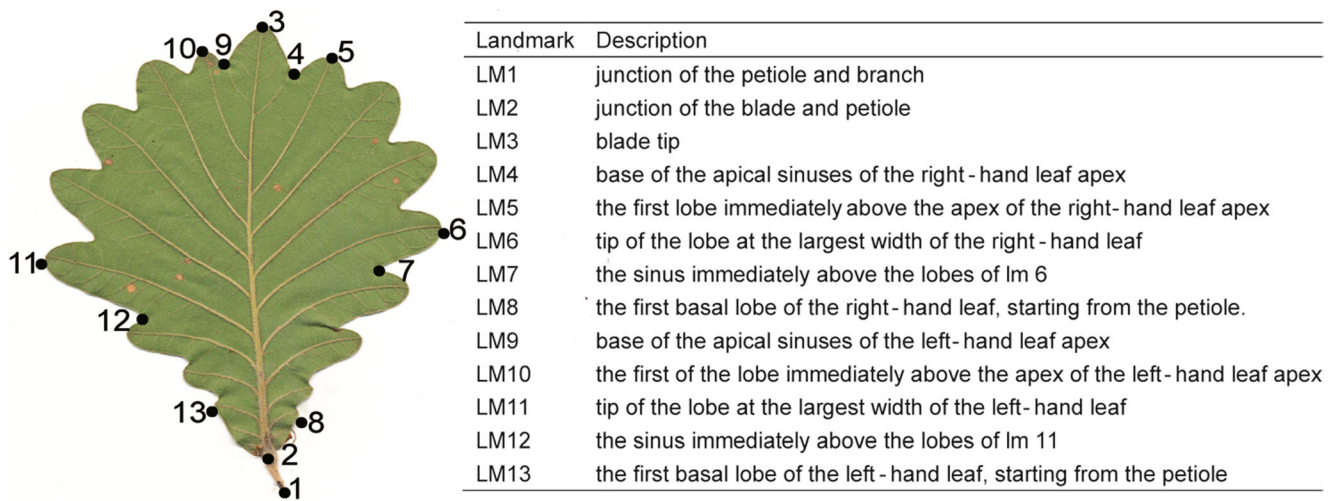
Landmark	Description
LM1	junction of the petiole and branch
LM2	junction of the blade and petiole
LM3	blade tip
LM4	base of the apical sinuses of the right-hand leaf apex
LM5	the first lobe immediately above the apex of the right-hand leaf apex
LM6	tip of the lobe at the largest width of the right-hand leaf
LM7	the sinus immediately above the lobes of lm 6
LM8	the first basal lobe of the right-hand leaf, starting from the petiole.
LM9	base of the apical sinuses of the left-hand leaf apex
LM10	the first of the lobe immediately above the apex of the left-hand leaf apex
LM11	tip of the lobe at the largest width of the left-hand leaf
LM12	the sinus immediately above the lobes of lm 11
LM13	the first basal lobe of the left-hand leaf, starting from the petiole

Fig. 2 Configuration of *Quercus dentata* and *Quercus aliena* leaves, showing locations of the 13 features used as landmarks in this study, with descriptions of the landmarks on the right

components) might be associated with leaf size (i.e., allometry) (Klingenberg 1998), 2B-PLS analysis was performed to detect covariation of leaf shape and size with a 10,000 permutation test using MorphoJ.

Discriminant analysis (DA) and canonical variate analysis (CVA) were used for tree-level species discrimination. Generally, DA is the most widely used statistical method for investigating taxonomic differences, focusing on two groups, while CVA is used for investigating three or more groups (Klingenberg 2011). In this study, DA was performed using cross-validated scores classification tables with T^2 statistics (P value for tests with 1000 permutations < 0.0001) while CVA was performed with permutation tests using Mahalanobis distances (T^2 statistics; 10,000 permutations per test). Results of the DA (which compares each pair of groups separately) were visualized in a frequency bar chart and the CVA results were presented in a scatter plot of the significant canonical variate (CV) scores.

GMMs allow shape differences between specimens to be recognized, using diagrams to visualize variability along the axes (i.e., PC/CV/PLS). In this study, an overlapping of transformation grid and wireframe graph was used to reconstruct the leaf shape by thin-plate spline analysis (Bookstein 1996; Klingenberg 2011; Viscosi and Cardini 2011; Viscosi 2015).

Linear regression analysis was used to study the relationships between environmental variables and both leaf shape (symmetric and asymmetric components) and size parameters. The leaf shape and size dataset was generated by MorphoJ, using symmetric and asymmetric components that contributed most strongly to PC1 and thus explained most of the variation in leaf shape (61.7 and 61.6%, respectively). The climate data were generated using bioclimatic parameters from the WORLDCLIM database (http://www.worldclim.org) for the current climate across the collection localities at 30 s resolution by ArcMAP 10.2 (http://www.esri.com/software/

arcgis/arcgis-for-desktop). Because collinearity can lead to instability of parameter estimates in regression models (Chatterjee et al. 2006; Dormann et al. 2013), environmental factors that were strongly correlated with other environmental factors were removed by variance inflation factor (VIF) analysis with the threshold values of 0.7 implemented in the "usdm" package, R. Finally, mean temperature of the warmest quarter (MTWQ), precipitation in the wettest month (PWM), and altitude were selected from 19 climate factors and three geographical factors (longitude, latitude, and altitude) for the analysis of correlations between leaf parameters and environmental variables. Values of the environmental factors at each population's location are listed in Table S2.

3 Results

3.1 A priori assignment of the individuals to the two species

STRUCTURE analysis clearly divided the 367 individuals into two main clusters with different genotypes, with indications of wide and asymmetric gene flow between them (Fig. S1). The genetic analyses assigned 129 individuals to pure *Q. dentata*, 194 individuals to pure *Q. aliena*, and 44 individuals to hybrids, with a 0.9 admixture coefficient (Q) threshold (Table 1).

3.2 Leaf morphology analysis

3.2.1 Multivariate analysis of raw data

The raw coordinate matrix obtained from image analysis of the 1835 leaves was translated into a normalized matrix by Procrustes fitting, and the configurations were rotated to the

concentrated distribution around 13 leaf landmarks based on the raw coordinate matrix (Fig. 3a), symmetric components (Fig. 3b), and asymmetric components (Fig. 3c). After removing outliers from the landmark data matrices, data for 1815 of the 1835 leaves were retained and used for subsequent multivariate analyses.

3.2.2 Interspecific leaf shape variation

In the scatter plot of PC1 and PC2 scores for symmetric components generated by PCA, leaves of the two species formed distinct groups with some overlap, whereas leaves of their hybrids were scattered (upper panel of Fig. 4a). In detail, the first and second PCs accounted for more than two thirds of the total leaf shape variation (61.7 and 14.6%, respectively) and the first five of PCs accounted for 96.6% of the total leaf shape variation (Table S3). Symmetric components with extreme PC1 and PC2 scores (the maximum and minimum values of the scatters projected on both ends of the PC axis) revealed the main differences between leaves of the two species (lower panel of Fig. 4a). The PC1 results showed that *Q. aliena* leaves had a longer petiole (distance between LM 1 and 2); narrower blade tip (manifested by contraction of the grid and squeezing between the leaf tip at LM 3 and widest point, at LM 6 and 11, relative to the maximal width), and relatively wider basal region (distance between LM 8 and 13) compared to *Q. dentata* leaves. The distance between LM 7 and 12 was related to the degree of leaf blade lobes division. PC2 was mainly related to the locations of lobes at the maximal width (LM 11 and 6) and their sinuses (LM 12 and 7), the negative values were associated with deep lobes, and positive values were associated with shallow lobes.

In the PC1 versus PC2 scatter plot for asymmetric components, the specimens were rather fully and densely overlapping than scattered (Fig. 4b). In detail, the first and second PCs accounted for 76.1% of the total leaf shape variation (61.6 and 14.5%, respectively): PC1 reflected the asymmetry of both lobes at the greatest width of the leaf while PC2 reflected the asymmetry due to bending of the leaf blade.

3.3 Allometry

2B-PLS analyses of relations between allometry and symmetric components at the tree level resulted in a RV coefficient (squared Pearson correlation coefficient) of 0.1089, $P < 0.001$ (Fig. 5a), indicating significant covariation between these components and leaf size. Graphical reconstruction of the symmetric components along PLS2 showed that negative values are associated with longer petioles and narrow to obovate leaf shape, while positive values are associated with short petioles and narrow to inverted water drop leaf shape. The graphical reconstruction along PLS1 (log centroid size axis) indicated that *Q. dentata* has larger leaves and positive leaf shape values while *Q. aliena* has smaller leaves with negative leaf shape values. However, no significant allometric patterns of asymmetric variation were detected for leaves of *Q. aliena*, *Q. dentate*, or their hybrids (RV coefficient, 0.0068, $P < 0.1673$) (Fig. 5b).

3.4 Species discrimination

The cross-validated DA revealed significant differences in leaf shape between *Q. dentata* vs. *Q. aliena* (Fig. 6a), large overlap in shapes of leaves of *Q. aliena* vs. the hybrids (Fig. 6b), and small overlap in shapes of leaves of *Q. dentata* vs. the hybrids (Fig. 6c). For the pairwise comparisons, T^2 values (all significant, $P < 0.0001$) were 1659, 220, and 84 for *Q. aliena* vs. *Q. dentata*, hybrids vs. *Q. dentata* and hybrids vs. *Q. aliena*, respectively. In these comparisons, 95.9 and 96.9% of *Q. aliena* and *Q. dentata*, 64.4 and 81.5% of the hybrids and *Q. aliena*, and 71.1 and 97.7% of the hybrids and *Q. dentata* trees were correctly classified, respectively.

Scatter plots of CVA scores showed that leaves of *Q. dentata* and *Q. aliena* trees mainly distributed along the CV1 axis (61.7%) with scores of − 4 to 0 and − 2 to 4,

Fig. 3 Results of generalized Procrustes analysis of the leaf shape of *Quercus dentata*, *Quercus aliena*, and hybrid based on the full raw coordinate matrix (**a**), symmetric components (**b**), and asymmetric components (**c**)
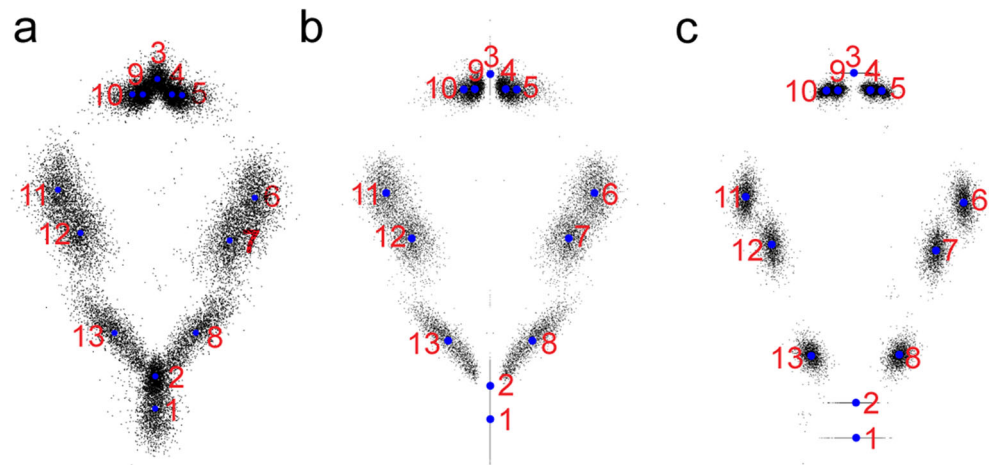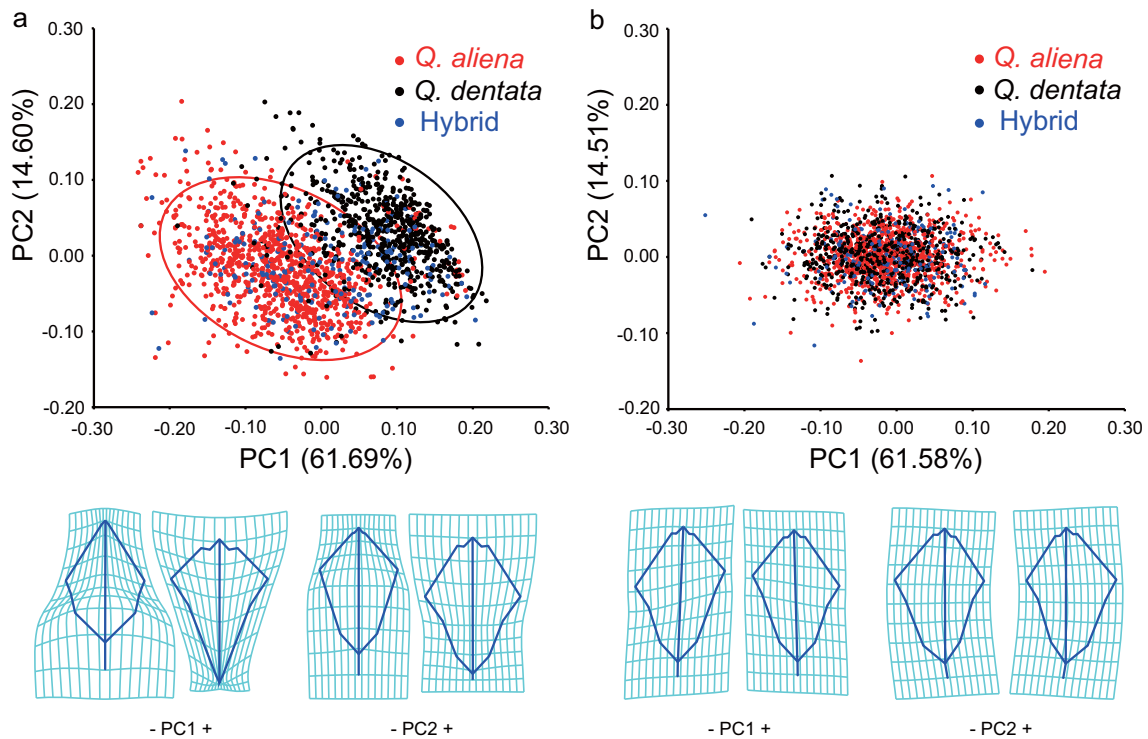
Fig. 4 Results of leaf-level principal component analysis (PCA) of *Quercus dentata*, *Quercus aliena*, and hybrid leaves' symmetric components (**a**) and asymmetric components (**b**). Scatter plots of PC1 and PC2 scores, with 95% confidence ellipses in **a**. Transformation Grids: the left and right graphs represent shapes corresponding to extreme positive (+) and negative (−) PC scores

respectively (Fig. 7). In addition, the leaf morphology of the hybrids is intermediate to some extent but also highly variable with a certain overlap with *Q. aliena*. Overlapping transformation grid and wireframe graphs showed that the main differences between leaves of *Q. dentata* and *Q. aliena* lay in the patterns of expansion and contraction from the leaves' bases to their apexes (Fig. 7). In contrast, *Q. dentata*, *Q. aliena*, and hybrid leaves largely overlapped along CV2 (38.2%).

3.5 Regression analysis of relations between environmental variables and leaf shape and size parameters

The linear regression analyses showed clear correlations between the leaf morphology and environmental factors (Table 2). Symmetric components of *Q. aliena* leaves were significantly correlated with MTWQ, PWM, and altitude,
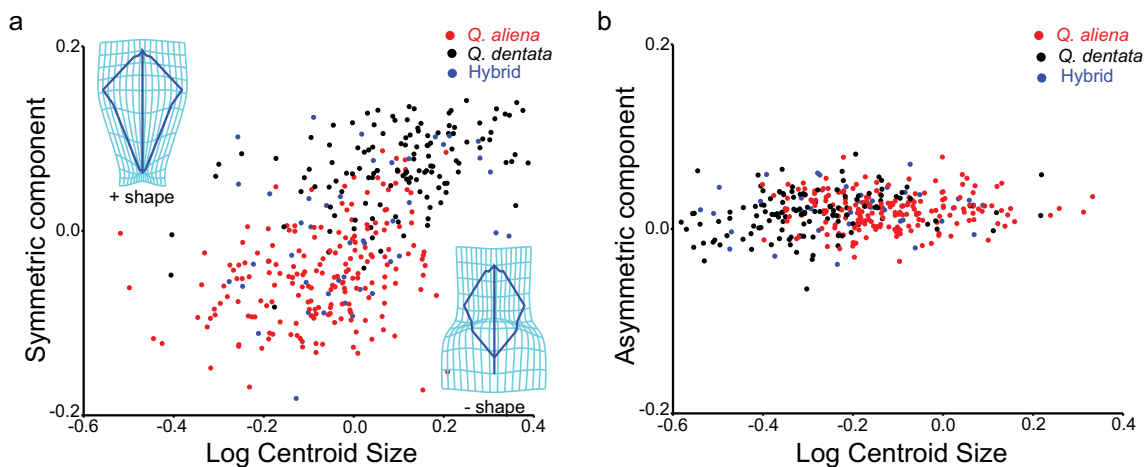

Fig. 5 Scatterplots obtained from 2B-PLS analysis of the relationship between size (log centroid size) and both symmetric components (**a**) and asymmetric components (**b**) of the shape of leaves of *Quercus dentata*, *Quercus aliena*, and their hybrids. The thin-plate spline deformation grids in **a** represent leaf shapes reflecting the negative (− shape) and positive (+ shape) extremes of the PLS axis for symmetric components
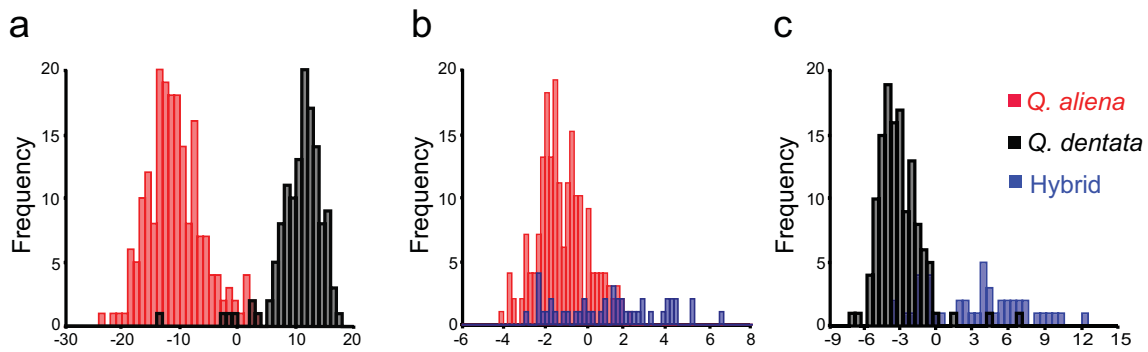
Fig. 6 Results of discriminant analysis (DA) of the shapes of leaves of *Quercus dentata* vs. *Quercus aliena* (**a**), *Q. aliena* vs. hybrids (**b**), and *Q. aliena* vs. hybrids (**c**). Black bars, *Q. dentata*; red bars, *Q. aliena*; blue bars, hybrids

those of *Q. dentata* leaves were correlated with MTWQ and altitude, while those of hybrids were significantly correlated with PWM. In contrast, asymmetric components of both sympatric oak species and their hybrids were only correlated with MTWQ. In addition, leaf size of *Q. aliena* was significantly correlated with altitude, leaf size of *Q. dentata* was significantly correlated with PWM, and leaf size of the hybrids was correlated with MTWQ and PWM.

4 Discussion

GMMs based on a priori classification of genetic analysis can provide reliable approaches for quantitative and graphical resolution of variation in leaf shape, with several major advantages over traditional morphological analysis

(Viscosi et al. 2012). Based on the a priori classification by nSSR and GMMs, we were able to develop a robust methodology for discriminating the sympatric oak species and explore the relationships between leaf allometry, shape, size, and environmental factors, which might be useful for the morphological studies of closely related oaks in China.

The results showed that analysis of symmetric components of leaf shape can successfully distinguish the two species if we presume that nSSR assignment well reflects taxonomic separation. In PCA score plots for symmetric components, *Q. dentata* and *Q. aliena* leaves formed two clearly separated groups, while leaves of their hybrids were scattered (Fig. 4a). Deformation grids showed that *Q. dentata* leaves have a shorter petiole, narrower basal region, deeper clefts, larger apical region, less lobed blade, and broader blade tip than

Fig. 7 Results of tree-level canonical variate analysis (CVA) of *Quercus dentata*, *Quercus aliena*, and hybrids' leaf shapes (symmetric components) showing 90% confidence ellipses for the species. Transformation grids represent leaf shapes reflecting the negative (− shape) and positive (+ shape) extremes of the CV1 axis
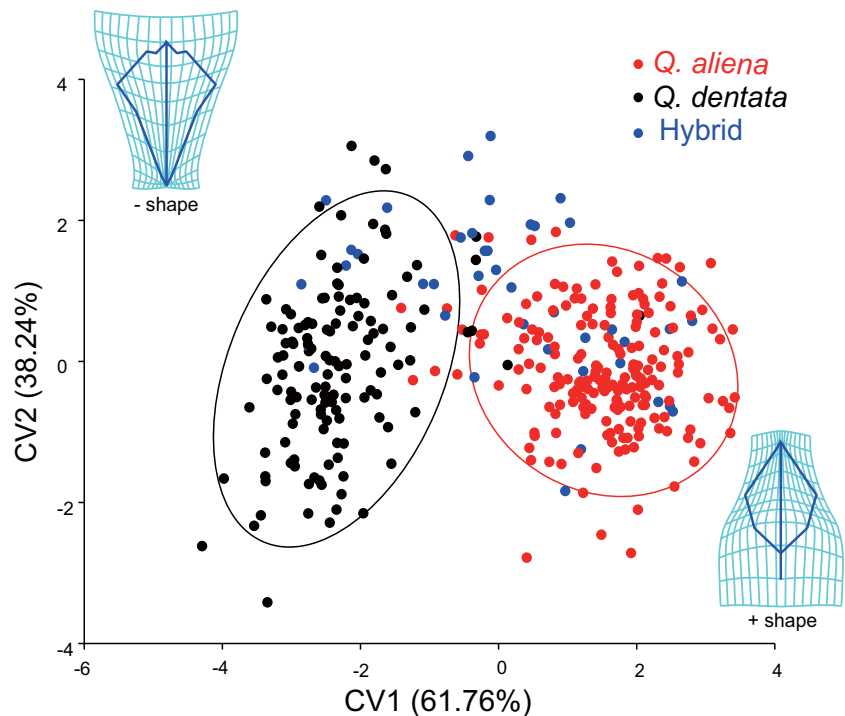
Table 2 Results of linear regression analysis of relationships between environmental factors

	MTWQ (°C)	PWM (mm)	Altitude (m)
Q. aliena			
Symmetric components	**	***	***
Asymmetric components	*	0.16	0.09
Leaf size	0.23	0.95	***
Q. dentata			
Symmetric components	*	0.28	*
Asymmetric components	**	0.64	0.07
Leaf size	0.18	***	0.40
Hybrid			
Symmetric components	0.50	**	0.37
Asymmetric components	*	0.86	0.09
Leaf size	*	*	0.88

Q. aliena leaves. Moreover, the morphological discrimination analysis after a priori classification at the molecular level proved to be useful to distinguish the two species and their hybrids. The DA analysis highlighted significant differences in leaf morphology variables between *Q. dentata* and *Q. aliena* (Fig. 6). Leaf shape DA could correctly classify genetically pure individuals with > 95% probability, but only correctly classify hybrids with much lower probability (ca. 64–71%). Interestingly, DA provided higher probability to correctly distinguish *Q. dentata* from hybrids than *Q. aliena* from hybrids. In addition, in CVA score plots, the two species were significantly separated along CV1, and the hybrids were mainly located between them, albeit closer to the *Q. aliena* cluster (Fig. 7). The CVA and DA results suggest that the hybrids' leaf shapes were affected by both parental species, but most strongly by the dominant species, *Q. aliena*. Interestingly, a corresponding pattern of variation in hybrids and parental species has been observed in populations of European white oaks (*Quercus frainetto* and *Quercus petraea*) (Viscosi et al. 2012; Stephan et al. 2018). This phenomenon may be caused by extensive and imbalanced gene flow between oak species (Gerber et al. 2014) but needs further investigation.

Several studies have found that symmetric components of leaf shape are related to genetic factors, while asymmetric components are generally related to environmental perturbations (Yoshioka et al. 2004; Bresson et al. 2011; Viscosi 2015). However, both the symmetric and asymmetric components of the shape of leaves studied here were significantly related to environmental factors. Moreover, responses of the symmetric components of leaves of the two species and their hybrids to several environmental factors differed, while the only tested environmental factor that influenced the asymmetric components was the mean temperature of the warmest quarter, MTWQ. It is reported that leaf phenology and ecophysiological properties are more environmentally influenced than genetically controlled in highly adaptive and plastic genotypes as described in Vitasse et al. (2010) and Bresson et al. (2011). However, further investigations such as controlled experiments are needed for the validation of leaf shape adaptation to different environments.

The size of *Q. aliena* leaves was significantly related to elevation (Table 2) as reported for other oaks (Viscosi 2015; Bresson et al. 2011). However, the leaf sizes of *Q. dentata* and hybrids were related to precipitation in the wettest month or mean temperature of the warmest quarter. Thus, leaf size can clearly be influenced by different environmental factors and genotypes, and their effects warrant further attention.

In conclusion, the GMM analysis of the leaf morphology of the two sympatrically distributed oak species suggests that we can clearly separate leaves of *Q. dentata* and *Q. aliena* morphologically. Moreover, individuals classified as putative hybrids through genetic analysis were consistent with morphologically intermediate individuals. In field observations of *Q. dentata* and *Q. aliena* at the XS and ZL sites in 2017, we found that their flowering times were asynchronous but the male florescences of *Q. dentata* and female florescences of *Q. aliena* overlap phenologically to a limited extent. The overlapping of the flowering time of the species suggests the possibility of introgression or hybridization between the two species. Introgression as an adaptive process is a common phenomenon in some tree species, such as oak and poplar (e.g., Lepais et al. 2009; Suarez - Gonzalez et al. 2018). But carefully controlled pollination experiments are required to confirm this possibility especially on the non-sympatric areas of the two species.

5 Conclusion

Thorough understanding of variation in leaf shape is highly important for taxonomists and ecologists, and morphological variation of *Quercus* components in mixed forest stands has interested biologists and botanists for decades (Kremer et al. 2002; Gugerli et al. 2007). GMMs based on genetic classification provided detailed, accurate information on the morphological variation in two sympatrically distributed Chinese oaks, and revealed correlations between leaf morphology and environmental factors. The data we have acquired could facilitate research into the taxonomy, phylogeny, and ecology of oaks. Moreover, extending the applied approach to studies of other plants will likely provide new insights into the biology and evolution of both oaks and other taxa.

INRA
SCIENCE & IMPACT

Springer

Data availability The datasets generated and/or analyzed during the current study are available from the authors on reasonable request.

Compliance with ethical standards

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

References

Abràmoff MD, Magalhães PJ, Ram SJ (2005) Image processing with Image J Part II. Biophoton Int 11:36–43. https://doi.org/10.1117/1.3589100

Adams DC, Slice DE, Rohlf FJ (2004) Geometric morphometrics: ten years of progress following the 'revolution'. Ital J Zool 71:5–16. https://doi.org/10.1080/11250000409356545

Arnold ML, Ballerini ES, Brothers AN (2012) Hybrid fitness, adaptation and evolutionary diversification: lessons learned from Louisiana irises. Heredity 108:159–166. https://doi.org/10.1038/hdy.2011.65

Bacilieri R, Ducousso A, Kremer A (1996) Comparison of morphological characters and molecular markers for the analysis of hybridization in sessile and pedunculate oak. Ann Sci For 53:79–91. https://doi.org/10.1051/forest:19960106

Benzécri JP (1992) Correspondence analysis handbook. Biometrics 49:672. https://doi.org/10.2307/2532585

Bookstein FL (1996) Combining the tools of geometric morphometrics. In: Advances in Morphometrics. Springer, Boston, pp 131–151

Box EO, Fujiwara K (2015). Warm-temperate deciduous forests: concept and global overview. In: Warm-temperate deciduous forests around the Northern hemisphere. Geobotany studies. Springer Cham, pp 7–26

Bresson CC, Vitasse Y, Kremer A, Delzon S (2011) To what extent is altitudinal variation of functional traits driven by genetic adaptation in European oak and beech? Tree Physiol 31:1164–1174. https://doi.org/10.1093/treephys/tpr084

Chatterjee S, Hadi AS, Price B (2006) Simple linear regression. Regression analysis by example. John Wiley and Sons, New York, pp 21–51

Costa C, Paglia G, Salvador FR, Lolletti D, Rimatori V, Menesatti P (2009) Hazelnut cultivar identification with leaf morphometric analysis: preliminary results. Acta Hortic (845):245–248. https://doi.org/10.17660/ActaHortic.2009.845.34

Dormann CF, Elith J, Bacher S, Buchmann C, Carl G, Carré G (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dow BD, 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. https://doi.org/10.1007/BF00220870

Du FK, Peng XL, Liu JQ, Lascoux M, Hu FS, Petit RJ (2011) Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. New Phytol 192:1024–1033. https://doi.org/10.1111/j.1469-8137.2011.03853.x

Du FK, Xu F, Qu H, Feng S, Tang J, Wu R (2013) Exploiting the transcriptome of euphrates poplar, *Populus euphratica* (Salicaceae) to develop and characterize new EST-SSR markers and construct an EST-SSR database. PLoS One 8:e61337. https://doi.org/10.1371/journal.pone.0061337

Durand J, Bodénès C, Chancerel E, Frigerio JM, Vendramin G, Sebastiani F (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11:570. https://doi.org/10.1186/1471-2164-11-570

Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing structure output and implementing the Evanno method. Conserv Genet Resour 4:359–361 http://taylor0.biology.ucla.edu/structureHarvester

Eaton DA, Hipp AL, González-Rodríguez A, Cavender-Bares J (2015) Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution 69:2587–2601. https://doi.org/10.1111/evo.12758

Elias TS (1980) The complete trees of North America. Van Nostrand Reinhold Co., New York

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fujiwara K, Harada A (2015) Character of warm-temperate *Quercus* forests in Asia. Warm-temperate deciduous forests around the northern hemisphere. Springer, Cham, pp 27–80

Gerber S, Chadoeuf J, Gugerli F, Lascoux M, Buiteveld J, Cottrell J, Goicoechea PG (2014) High rates of gene flow by pollen and seed in oak populations across Europe. PLoS One 9:e85130. https://doi.org/10.1371/journal.pone.0085130

Gower JC (1975) Generalised Procrustes analysis. Psychometrika 40:33–51

Graham JH, Raz S, HelOr H, Nevo E (2010) Fluctuating asymmetry: methods, theory, and applications. Symmetry 2:466–540. https://doi.org/10.3390/sym2020466

Gugerli F, Walser JC, Dounavi K, Holderegger R, Finkeldey R (2007) Coincidence of small-scale spatial discontinuities in leaf morphology and nuclear microsatellite variation of *Quercus petraea* and *Q. robur* in a mixed forest. Ann Bot 99:713–722. https://doi.org/10.1093/aob/mcm006

Guichoux E, Lagache L, Wagner S, Petit R (2011) Current trends in microsatellites genotyping. Mol Ecol Resour 11:591–611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Huang CJ, Zhang YT, Bruce B (1999a) Fagaceae. Flora of China, vol 4. Science Press, Beijing, pp 314–400

Huang CJ, Zhang YT, Bartholomew B (1999b) Fagaceae. In: Wu ZY, Raven PH (eds) Flora of China Vol. 4 (in English). Science Press, Beijing, pp 370–380

Hubert F, Grimm GW, Jousselin E, Berry V, Franc V, Kremer A (2014) Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus*. Syst Biodivers 12:405–423. https://doi.org/10.1080/14772000.2014.941037

Janes JK, Miller JM, Dupuis JR, Malenfant RM, Gorrell JC, Cullingham CI, Andrew RL (2017) The K=2 conundrum. Mol Ecol 26:3594–3602. https://doi.org/10.1111/mec.14187

Jaramillo-Correa JP, Aguirre-Planter E, Khasa DP, Eguiarte LE, Piñero D, Furnier GR (2008) Ancestry and divergence of subtropical montane forest isolates: molecular biogeography the genus *Abies* (Pinaceae) in southern Mexico and Guatemala. Mol Ecol 17:2476–2490. https://doi.org/10.1111/j.1365-294X.2008.03762.x

Jensen RJ, Ciofani KM, Miramontes LC (2002) Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. Taxon 51:475–492. https://doi.org/10.2307/1554860

Klingenberg CP (1998) Heterochrony and allometry: the analysis of evolutionary change in ontogeny. Biol Rev 73:79–123. https://doi.org/10.1017/S000632319800512X

Klingenberg CP (2003) A developmental perspective on developmental instability: theory, models and mechanisms. In: Polak M, ed. Developmental instability: causes and consequences. Camb Law J, New York pp 14–34

Klingenberg CP (2011) MorphoJ: an integrated software package for geometric morphometrics. Mol Ecol Resour 11:353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x

Klingenberg CP, Monteirob LR (2005) Distances and directions in multidimensional shape spaces: implications for morphometric applications. Syst Biol 54:678–688. https://doi.org/10.1080/10635150590947258

Klingenberg CP, Barluenga M, Meyer A (2002) Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. Evolution 56:1909–1920. https://doi.org/10.1111/j.0014-3820.2002.tb00117.x

Kremer A, Dupouey LJ, Deans JD, Cottrell J, Csaikl U, Finkeldey R, Ducousso A (2002) Leaf morphological variation in mixed oak stands (*Quercus robur* and *Quercus petraea*) in stable western European population. Ann For Sci 59:777–787. https://doi.org/10.1051/forest:2002065

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A (2009) Species relative abundance and direction of introgression in oaks. Mol Ecol 18:2228–2242. https://doi.org/10.1111/j.1365-294X.2009.04137.x

Lyu J, Song J, Liu Y, Wang Y, Li J, Du FK (2018) Species boundaries between three sympatric oak species: *Quercus aliena*, *Q. dentata*, and *Q. variabilis* at the northern edge of their distribution in China. Front Plant Sci 9:414. https://doi.org/10.3389/fpls.2018.00414

MacLeod N, Forey PL (2002). Introduction: morphology, shape, and phylogenetics. In: Morphology, shape, and phylogeny. Taylor & Francis, London, pp 1–7

Manos PS, Doyle JJ, Nixon KC (1999) Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). Mol Phylogenet Evol 12:333–349. https://doi.org/10.1006/mpev.1999.0614

Mitteroecker P, Gunz P (2009) Advances in geometric morphometrics. Evol Biol 36:235–247. https://doi.org/10.1007/s11692-009-9055-x

Nixon KC (1993) Infrageneric classification of *Quercus* (Fagaceae) and typification of sectional names. Ann For Sci 50:25s–34s. https://doi.org/10.1051/forest:19930701

Nixon KC (1997) *Quercus*. In: Flora of North America editorial committee (eds) vol. 3. Oxford University Press, New York, pp 445–447

Nugroho A, Song BM, Su HS, Choi JS, Choi J, Choi JY (2016) HPLC analysis of phenolic substances and anti-Alzheimer's activity of Korean *Quercus* species. Nat Prod Sci 22:299. https://doi.org/10.20307/nps.2016.22.4.299

Peñaloza-Ramírez JM, González-Rodríguez A, Mendoza-Cuenca L, Caron H, Kremer A, Oyama K (2010) Interspecific gene flow in a multispecies oak hybrid zone in the Sierra Tarahumara of Mexico. Ann Bot 105:389–399. https://doi.org/10.1093/aob/mcp301

Peng YS, Chen L, Li JQ (2007) Study on numerical taxonomy of *Quercus* L. (Fagaceae) in China. J Wuhan Bot Res 25:149–157. https://doi.org/10.1016/j.jcis.2009.12.058

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol 14:689–701. https://doi.org/10.1111/j.1365-294X.2004.02410.x

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Rellstab C, Bühler A, Graf R, Folly C, Gugerli F (2016) Using joint multivariate analyses of leaf morphology and molecular-genetic markers for taxon identification in three hybridizing European white oak species (*Quercus*, spp.). Ann For Sci 73:1–11. https://doi.org/10.1007/s13595-016-0552-7

Ren XW, Wang LM (1985) Geographic distribution of deciduous oaks in China. Journal of Beijing Forestry University 2:57–69

Rohlf FJ (2010) tpsDig, digitize landmarks and outlines, ver. 2.16. Department of Ecology and Evolution, State University of New York at Stony Brook, NJ. life. http://bio.sunysb.edu/morph

Rohlf FJ, Slice DE (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Zool 39:40–59. https://doi.org/10.2307/2992207

Rushton BS (1993) Natural hybridization within the genus *Quercus* L. Ann For Sci 50:73–90. https://doi.org/10.1051/forest:19930707

Savriama Y, Klingenberg CP (2011) Beyond bilateral symmetry: geometric morphometric methods for any type of symmetry. BMC Evol Biol 11:280. https://doi.org/10.1186/1471-2148-11-280

Slice DE (1996) A glossary for geometric morphometrics. Advances in morphometrics, pp 531–551

Souza SMF, Moreira DAI, Joseph MS (2012) Geometric morphometrics of leaf blade shape in *Montrichardia linifera* (Araceae) populations from the Rio Parnaíba delta, north-east Brazil. Bot J Linn Soc 170: 554–572. https://doi.org/10.1111/j.1095-8339.2012.01309.x

Stace CA (1982) Plant taxonomy and biosystematics. Brittonia 34:80–80

Stephan JM, Teeny PW, Vessella F, Schirone B (2018) Oak morphological traits: between taxa and environmental variability. Flora 243: 32–44. https://doi.org/10.1016/j.flora.2018.04.001

Suarez - Gonzalez A, Hefer CA, Lexer C, Douglas CJ, Cronk Q (2018) Introgression from *Populus balsamifera* underlies adaptively significant variation and range boundaries in *P. trichocarpa*. New Phytol 217:416–427. https://doi.org/10.1111/nph.14779

Tsukaia H (2005) Leaf shape: genetic controls and environmental factors. Int J Dev Biol 49:547–555. https://doi.org/10.1387/ijdb.041921ht

Tucić B, Budečević S, Manitašević SJ, Vuleta A, Klingenberg CP (2018) Phenotypic plasticity in response to environmental heterogeneity contributes to fluctuating asymmetry in plants: first empirical evidence. J Evol Biol 31:197–210. https://doi.org/10.5061/dryad.8th5m

Ueno S, Taguchi Y, Tsumura Y (2008) Microsatellite markers derived from *Quercus mongolica* var. *crispula* (Fagaceae) inner bark expressed sequence tags. Genes Genet Syst 83:179–187. https://doi.org/10.1266/ggs.83.179

Ueno S, Aoki K, Tsumura Y (2009) Generation of expressed sequence tags and development of microsatellite markers for *Castanopsis sieboldii* var. *sieboldii* (Fagaceae). Ann For Sci 66:509. https://doi.org/10.1051/forest/2009037

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol Ecol 15:63–72. https://doi.org/10.1111/j.1365-294X.2005.02773.x

Viscosi V (2015) Geometric morphometrics and leaf phenotypic plasticity: assessing fluctuating asymmetry and allometry in European white oaks (*Quercus*). Bot J Linn Soc 179:335–348. https://doi.org/10.1111/boj.12323

Viscosi V, Cardini A (2011) Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. PLoS ONE 6: e25630. https://doi.org/10.1371/annotation/bc347abe-8d03-4553-8754-83f41a9d51ae.t0023.cn

Viscosi V, Fortini P (2011) Leaf shape variation and differentiation in three sympatric white oak species revealed by elliptic Fourier analysis. Nord J Bot 29: 632-640. doi: 10.1111/j.1756-1051.2011.01098.x

Viscosi V, Fortini P, Slice DE, Loy A, Blasi C (2009) Geometric morphometric analyses of leaf variation in four oak species of the subgenus

Quercus (Fagaceae). Plant Biosyst 143:575–587. https://doi.org/10.1080/11263500902775277

Viscosi V, Antonecchia G, Lepais O, Fortini P, Gerber S, Loy A (2012) Leaf shape and size differentiation in white oaks: assessment of allometric relationships among three sympatric species and their hybrids. Int J Plant Sci 173:875–884. https://doi.org/10.1086/667234

Vitasse Y, Bresson CC, Kremer A (2010) Quantifying phenological plasticity to temperature in two temperate tree species. Funct Ecol 24: 1211–1218. https://doi.org/10.1111/j.1365-2435.2010.01748.x

Yoshioka Y, Iwata H, Ohsawa R, Ninomiya S (2004) Analysis of petal shape variation of *Primula sieboldii* by elliptic Fourier descriptors and principal component analysis. Ann Bot 94:657–664. https://doi.org/10.1093/aob/mch190Zelditch

Zelditch ML, Swiderski DL, Sheets HD (2012) Glossary-geometric morphometrics for biologists, (2nd edn). Geometric morphometrics for biologists, pp 455–469

Zeng YF, Liao WJ, Petit RJ (2010) Exploring species limits in two closely related Chinese oaks. PLoS One 5:e15529. https://doi.org/10.1371/journal.pone.0015529