PROGRESS IN MORPHOLOGICAL AND MOLECULAR SYSTEMATICS

Leaf morphological analyses in four European oak species (Quercus) and their hybrids: A comparison of traditional and geometric morphometric methods

V. VISCOSI¹, O. LEPAIS²,³, S. GERBER²,³, & P. FORTINI¹

¹Dipartimento di Scienze e Tecnologie per l’Ambiente e il Territorio, Università del Molise, Pesche, Italy, ²INRA, Cestas, France, and ³Université de Bordeaux 1, Cestas, France

Abstract

In this study, leaf morphology was assessed in a mixed oak stand (western France) using two geometric morphometric (landmark and outline) datasets and one dataset of 19 leaf measures. Adult oaks (817 oaks), comprising four white oak species (Quercus petraea, Q. robur, Q. pubescens and Q. pyrenaica), were sampled for DNA extraction and genetic analysis (nuclear microsatellites). Leaf morphology was assessed on 336 oaks, comprising pure species and hybrids as determined by genetic assignment. This comparative study of oak leaf morphology, based on the use of two free size geometric morphometric methods and a set of leaf measurements, combined with the genetic assignment of individuals to pure species or hybrids, provided information about the differences among species and the intermediate leaf morphology of their hybrids.

Keywords: Genetic assignment, hybridization, traditional morphometrics, geometric morphometrics, leaf morphology, nuclear microsatellites, Quercus

Introduction

The genus Quercus (Fagaceae) is widespread in the northern hemisphere. Deciduous oaks are significant components of the temperate forests of North America and Eurasia (Camus 1936–1954; Manos et al. 1999). In Europe, the subgenus Quercus spread quickly during the second half of the Tertiary era (Axelrod 1983) and during the ice ages of the Quaternary period (120,000–18,000 years ago). Several studies have investigated post-glacial routes of oaks using the maternally inherited chloroplast genome (cpDNA), while palynological techniques (Brewer et al. 2002; Petit, Brewer, et al. 2002) have established that the dynamics involved movement from refuge areas towards the north. In France, oak populations originated from all three primary refuge areas, located in the Italian, Balkan and Iberian peninsulas (Petit, Latouche-Hallé, et al. 2002). Oak species were not correlated with a particular refugium, however, spatial analysis of cpDNA at a regional scale indicated that species occurring in sympathy tend to share the same maternal genome indicating a potential role of hybridization during postglacial migration (Petit et al. 1997, 2003).

Interspecific gene flow is a common phenomenon in the genus Quercus, owing to the absence of real reproductive barriers within the genus (Burger 1975; Spellenberg 1995; Gonzàles-Rodrìguez et al. 2004); moreover, natural hybridization was observed by mating system analysis with gene markers in natural mixed populations (Bacilieri et al. 1996; Streiff et al. 1999). As a consequence, introgression has led to a proliferation of fertile hybrids displaying a wide range of intermediate morphological features (Gellini & Grossoni 1997).

Hybridization has been attributed to both genetic and ecological effects (Belahbib et al. 2001; Boavida et al. 2001; Petit et al. 2003). When the range of distribution of two or more species overlap, hybrids occur more frequently at their geographical or ecological margins (Valbuena-Carabaña et al. 2005) where environmental conditions might be more
suitable for hybrids with intermediate characters (Williams et al. 2001).

Despite this, the existence of interspecific differences has been shown in some studies based on genetic markers. Bordacs and Burg (1997) demonstrated the existence of two species-specific RAPD markers differentiating *Q. robur* from *Q. petraea*, while more recently Scotti-Saintagne et al. (2004) have shown that interspecific differentiation was heterogeneous and spatially autocorrelated across the genomes using a variety of molecular markers (isozymes, AFLPs, SCARs, microsatellites and SNPs). Independently, natural oak population studies based on microsatellites and individual-based genetic assignment methods (Muir et al. 2000; Muir & Schlötterer 2005; Valbuena-Carabaña et al. 2005, 2007; Curtu et al. 2007; Gugerli et al. 2007) provided multiple evidences that oak species represent separate taxonomical units.

Morphological analyses have also been performed for species differentiation in European white oaks (Dupouey & Badeau 1993; Bacilieri et al. 1995; Kleinschmit et al. 1995; Bruschi et al. 2000; Kremer et al. 2002; Borazan & Babac 2003) but, as a hybrid identification method was not available, information on hybrid morphology remained poor. However, two recent studies combining genetic assignment and morphological analysis (Curtu et al. 2007; Gugerli et al. 2007) demonstrated a general congruence of both species determination. Hybrids, as defined by genetic assignment, did not necessarily present intermediate morphologies, so that intermediate morphology does not systematically mean hybridization. These results confirm the polygenic nature of morphological species differences in oak species as demonstrated by quantitative traits loci analyses (QTL; Saintagne et al. 2004). In addition, complexity, morphological variability and plasticity of the phenotype could render morphological features inadequate for quantifying the degree of hybridization.

In this study, the morphological analysis of leaves was undertaken to investigate the interspecific relationships among four sympatric oak species (*Quercus pubescens* Willd., *Q. robur* L., *Q. petraea* (Matt.) Liebl. and *Q. pyrenaica* Willd.) and their potential hybrids, in a mixed stand of western France. Using a Bayesian approach, genetic markers (nuclear microsatellites) were used to assign individuals to a species or hybrid class without a priori classification. In the present paper we will refer to species determination as a result of genetic assignment only, following the blind approach (Duminil et al. 2006). This method provides an independent classification of individual trees that can be compared to morphological analysis. Leaf morphology was assessed by means of morphological measures (Kremer et al. 2002), well suited to discriminate two additional species and two geometric morphometric methods (landmark and outline) described in Viscosi et al. (2009).

The aims of the present study were, firstly, to investigate the morphological relationships among the four oak species and their genetic hybrids in a mixed stand, secondly to analyse the correlation among the three morphological datasets used in relation to the genetic assignment of individuals to species or hybrid groups. Finally, the functional capacity of geometric morphometric methods to analyse the leaf shape morphology of European white oaks, when molecular species assignment of individuals or more generally when molecular markers are available, is discussed. In addition to population genomics tools, this new method could be useful to provide new insight into morphological species’ differences with a more functional perspective.

**Materials and methods**

All adult oaks (817 oaks) were sampled in a mixed oak stand (Briouant) located in south-western France and comprising four white oak species (*Quercus petraea* (Matt.) Liebl., *Q. robur* L., *Q. pubescens* Willd. and *Q. pyrenaica* Willd.) for DNA extraction and genetic analysis. Leaf morphology was assessed on a subset of 336 oaks comprising pure species and hybrids as determined by genetic assignment. For all selected trees, 10 leaves were sampled in the highest part of the tree for morphological analyses.

**Genetic analysis**

Genotyping of 10 microsatellite loci (Steinkellner et al. 1997; Kampfer et al. 1998) was performed on all adult trees according to a multiplex protocol described in Lepais et al. (2006). A Bayesian assignment method implemented in the software Structure version 2.1 (Pritchard et al. 2000; Falush et al. 2003) was carried out to assign individuals to species without any other information than multilocus genetic data (field species determination was not used for a priori classification). Since such a method requires that all parental populations should be sampled to recover the genetic structure, and given the very small number of oaks that resemble *Q. petraea*, genetic assignment analyses were performed with several others oak populations with a total of 2163 individuals (Lepais et al. 2009), providing a sufficient number of individuals per species.

Briefly, the Structure was run with the admixture model, 100,000 burn-in periods and 1,000,000 MCMC repeats after burn-in. The number of clusters was set to *K* = 4, each cluster corresponding to
one oak species, providing the best partition of the genetic data. Genetic assignment of individuals to species was obtained based on the probability of belonging to one of the four clusters computed by Structure. A specimen was assigned to a species when its probability for one cluster was >0.90. Individuals that could not be assigned to a pure species were considered as hybrids originated from a cross between the two species showing the higher probabilities. This assignment procedure, tested by simulations, was found to (1) slightly underestimate hybrid proportions (Lepais et al. 2009), and (2) assign correctly up to the second hybrid generation. So, we defined hybrids as groups of first (F1) and second generation hybrids (back-crosses and F2).

All specimens were assigned to four species and six hybrid classes. The 336 trees analysed for leaf morphology were assigned as follows: Q. robur (77), Q. petraea (6), Q. pubescens (54) and Q. pyrenaica (67) for pure species; Q. robur × petraea (10), Q. robur × pubescens (21), Q. robur × pyrenaica (29), Q. pubescens × pyrenaica (41), Q. petraea × pubescens (8), Q. petraea × pyrenaica (23) for hybrids.

Morphological analysis

Leaf morphological analyses were performed by means of three different methods: a morphological analysis of leaf variables as described in Kremer et al. (2002) and two geometric mophometric analyses of leaf shape (landmark and outline) performed following Viscosi et al. (2009).

Leaf morphology was assessed by means of 19 variables (Kremer et al. 2002). Petiole length (pl), lamina length (ll), lobe width (lw), sinus width (sw) and length of the lamina at largest width (wp) were measured on leaf pictures and converted in millimetres according to the scale of each picture. The numbers of lobes (nl) and veins (nv) were counted for each leaf. Basal shape of the lamina (bs) was scored from 1 to 9 and pubescence density was estimated from 0 (no pubescence) to 5 (dense pubescence) for abaxial lamina (hr), adaxial lamina (sr), central nervation (nr) and petiole (ptr). We also used five transformed variables: lamina shape (ob), petiole ratio (pr), lobe depth ration (ldr), percentage of venation (pv) and lobe width ratio (lwr). For overall lamina morphology, we used the software ImageJ (Rasband 1997–2007) to measure leaf perimeter and circularity calculated as $c = 4\pi r^2$ which gives a value of 0 for a perfect circle and 1 for an increasingly elongated polygon.

For geometric morphometric analyses, scanned images were used to record 11 landmarks on the right half of each leaf; the outlines of leaves were automatically recorded and reduced to Fourier coefficients for 50 harmonics.

Statistical analysis

For morphological measurements, the mean of 19 leaf variables for all trees was computed. Principal Component Analysis (PCA) was performed on this dataset shifting variables to the zero centre and scaling them to have a unit variance. For landmark data, Relative Warp Analysis (Rohlf & Slice 1990) was performed on a standardized matrix of 336 consensus extracted from all individual trees, after multivariate analysis of outliers; all configurations were optimally rotated to minimize the squared differences between the corresponding landmarks by using the Generalized Procrustes Analysis (Rohlf & Slice 1990). For outline data, PCA was performed on a variance-covariance matrix extracted from the coefficients of 50 harmonics; the size and orientation of each contour was standardized by the longest radius method (Kuhl & Giardina 1982; Iwata & Ukai 2002).

Both landmark and outline datasets were implemented by means of multivariate analysis of the outliers (Gschwandtner & Filzmoser 2004; Filzmoser 2005). All the outliers identified through this procedure were eliminated from the data matrices of landmarks and outlines for each individual tree. A consensus configuration (landmark data) and the mean of EFDs (outline data) were extracted for each tree.

Afterwards, to realize a model for each dataset, all pure species trees were subjected to canonical discriminant analysis (CVA) using principal component scores (for morphological and outline data) and relative warp scores (for landmark data). The discriminant functions extracted were used to compute the pairwise group comparisons (step-by-step method) to verify discrimination degree among species. Moreover, because in discriminant function analysis the solutions are often unreliable, cross-validation (the process of testing a model on more than one sample) was performed to assess the reliability and general ability of the findings for each dataset.

Then, to verify if hybrids had an intermediate leaf morphology between the two parental species, three new matrices containing hybrids were obtained and analysed by means of PCA. For each pure species and hybrid group we computed the mean of leaf variables (morphological measures), the consensus configuration (landmarks) and the mean EFDs (outlines).

Finally, three Mantel tests (Mantel 1967) were run among the Euclidean distance matrices of the three datasets which were computed on the average canonical scores of species without hybrids (using NTSYSpc 2.2).
Results

Morphological leaf variables

In the PCA, the two first axes explained 58.50% of the variance, with 33.89% for PC1 and 24.61% for PC2, respectively. The ordering of individuals and variable vectors on the first two PCs (Figure 1) revealed that Q. robur individuals were clearly separated from the others mainly due to the effect of basal shape, circularity and other variables (pl, hr, sr, nr, nl, pr and ptr). Q. pyrenaica and Q. pubescens partially overlapped, but were separated by perimeter and circularity variables. For Q. petraea clear results were not found, because of the low number of individuals in the dataset.

In the CVA, the greater part of total variance (99.39%) was explained along the first two canonical variates (CVs): CV1 (Wilks' $\lambda = 0.036$; df = 15; $p < 0.001$) and CV2 (Wilks' $\lambda = 0.375$; df = 8; $p < 0.001$) explained 85.58% and 13.81% of the total variance, respectively; while CV3 (Wilks' $\lambda = 0.939$; df = 3; $p = 0.006$) explained only 0.61% of the total variance.

The scatterplot of CV1 and CV2 (Figure 2) showed that, along CV1, Q. robur was separated from the other three species, while CV2 explained differences between Q. pyrenaica and Q. pubescens.

From pairwise group comparisons, the differences among the four species appeared significant, the greater differences being those between Q. robur, Q. pubescens and Q. pyrenaica; also Q. petraea was significantly discriminated from all other species (Table I). According to the cross-validation test, 90.7% of oaks where correctly classified: All specimens for Q. robur, the greater part for Q. pubescens and Q. pyrenaica and none for Q. petraea (Table II).

Landmark data

RWA showed that the greater part of the total variance (83.72%) was explained along the first four RWs: The first two RWs explained only 55.31% of the total variance, with 32.85% and 22.49%, respectively for RW1 and RW2; RW3 and RW4 explained 16.75% and 11.66% of the total variance, respectively.

The pattern of the shape variation was clarified in the canonical variate space (Figure 3), where species were significantly discriminated along CV1 (Wilks’ $\lambda = 0.092$; df = 18; $p < 0.001$) and CV2 (Wilks’ $\lambda = 0.411$; df = 10; $p < 0.001$) that explained 99.20% of the total variance. CV3 was not significant (Wilks’ $\lambda = 0.961$; df = 4; $p = 0.095$). Specifically, CV1 (71.60% of the total variance) explained differences in the length of the petiole and the shape of the basal and apical regions, and clearly distinguished Q. robur from Q. pubescens and Q. pyrenaica; CV2 (27.50%) explained leaf variability with respect to depth of lobes and shape of the apical region, and separated Q. pubescens from Q. pyrenaica. Pairwise group comparisons were computed among species and the differences among Q. robur, Q. pubescens and Q. pyrenaica were meaningful; Q. petraea was also discriminated from other species (Table I).

The test of cross-validation gave a high percentage of correctly classified cases (87.7%). For Q. robur, Q. pubescens and Q. pyrenaica, the greater part of specimens were correctly classified (Table II).

Outline data

From the PCA it resulted that PC1 and PC2 explained 67.15% of the total variance, i.e., 55.70% and 11.45%, respectively. The discriminant analysis yielded similar results to other datasets: Three
species were significantly discriminated along the
first two CVs (Figure 4); from pairwise group
comparisons it resulted that only the differences
between Q. petraea and Q. pubescens were not signif-
icant (Table I). The test of cross-validation (leave-
one-out method) showed a high percentage of
correct classification (Table II).

### Analyses of hybrids

For each dataset, the mean leaves of the 10 groups,
comprising four pure species and six hybrid groups,
were analysed by means of PCA.

For morphological leaf variables, PC1 and
PC2 explained 80.52% of the total variance. It

### Table I. Pairwise group comparisons.

<table>
<thead>
<tr>
<th>Species</th>
<th>Q. petraea</th>
<th>Q. pubescens</th>
<th>Q. pyrenaica</th>
<th>Q. robur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q. petraea</td>
<td>—</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. pubescens</td>
<td>11.62</td>
<td>—</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. pyrenaica</td>
<td>10.85</td>
<td>59.93</td>
<td>—</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. robur</td>
<td>19.17</td>
<td>248.89</td>
<td>306.75</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Q. petraea</th>
<th>Q. pubescens</th>
<th>Q. pyrenaica</th>
<th>Q. robur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q. petraea</td>
<td>—</td>
<td>0.002</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. pubescens</td>
<td>3.61</td>
<td>—</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. pyrenaica</td>
<td>4.65</td>
<td>43.59</td>
<td>—</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. robur</td>
<td>8.88</td>
<td>94.30</td>
<td>91.94</td>
<td>—</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Q. petraea</th>
<th>Q. pubescens</th>
<th>Q. pyrenaica</th>
<th>Q. robur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q. petraea</td>
<td>—</td>
<td>0.064</td>
<td>&lt;0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>Q. pubescens</td>
<td>1.79</td>
<td>—</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. pyrenaica</td>
<td>3.74</td>
<td>30.64</td>
<td>—</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q. robur</td>
<td>2.91</td>
<td>35.08</td>
<td>52.65</td>
<td>—</td>
</tr>
</tbody>
</table>

F-values (below) and p-values (above) are indicated for each
species comparison at Step 5. Degree of freedom (df) were 5.196,
6.195 and 10.191 for 19 leaf variables, landmark data and outline
data, respectively. F-values significant at $p < 0.001$ are in bold type.

### Table II. Results of the cross-validation test of the three datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Q. petraea</th>
<th>Q. pubescens</th>
<th>Q. pyrenaica</th>
<th>Q. robur</th>
</tr>
</thead>
<tbody>
<tr>
<td>19 Leaf variables</td>
<td>0.0</td>
<td>16.7</td>
<td>50</td>
<td>33.3</td>
</tr>
<tr>
<td>Q. petraea</td>
<td>1.9</td>
<td>87.0</td>
<td>11.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Q. pubescens</td>
<td>0.0</td>
<td>7.5</td>
<td>91.0</td>
<td>1.5</td>
</tr>
<tr>
<td>Q. pyrenaica</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Q. robur</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>98.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Q. petraea</th>
<th>Q. pubescens</th>
<th>Q. pyrenaica</th>
<th>Q. robur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landmark data</td>
<td>0.0</td>
<td>50.0</td>
<td>33.3</td>
<td>16.7</td>
</tr>
<tr>
<td>Q. petraea</td>
<td>1.9</td>
<td>79.5</td>
<td>13.0</td>
<td>5.6</td>
</tr>
<tr>
<td>Q. pubescens</td>
<td>0.0</td>
<td>9.0</td>
<td>89.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Q. pyrenaica</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>98.7</td>
</tr>
<tr>
<td>Q. robur</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>98.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Q. petraea</th>
<th>Q. pubescens</th>
<th>Q. pyrenaica</th>
<th>Q. robur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outline data</td>
<td>16.7</td>
<td>33.3</td>
<td>16.7</td>
<td>33.3</td>
</tr>
<tr>
<td>Q. petraea</td>
<td>9.3</td>
<td>74.1</td>
<td>9.3</td>
<td>7.3</td>
</tr>
<tr>
<td>Q. pubescens</td>
<td>3.0</td>
<td>6.0</td>
<td>86.5</td>
<td>4.5</td>
</tr>
<tr>
<td>Q. pyrenaica</td>
<td>7.8</td>
<td>3.9</td>
<td>0.0</td>
<td>88.3</td>
</tr>
</tbody>
</table>

F-values (below) and p-values (above) are indicated for each
dataset at Step 5. Degree of freedom (df) were 5.196,
6.195 and 10.191 for 19 leaf variables, landmark and outline
data, respectively. F-values significant at $p < 0.001$ are in bold type.
Figure 3. Scatterplot of the first and the second canonical variate scores derived from landmark data. Deformation grids showing major shape features correlated with the extremes of variation along the axes, and were obtained through regression of shape variables on discriminant scores. Filled circles = *Q. robur*; filled diamonds = *Q. pubescens*; open triangles = *Q. pyrenaica*; open upturned triangles = *Q. petraea*; hybrids not projected.

Figure 4. Scatterplot of the first and the second canonical variate scores derived from outline data. Filled circles = *Q. robur*; filled diamonds = *Q. pubescens*; open triangles = *Q. pyrenaica*; open upturned triangles = *Q. petraea*; hybrids not projected.
Leaf morphological analyses in European oaks

appeared that the projections of the hybrid groups Q. robur × pubescens, Q. robur × pyrenaica, Q. petraea × pubescens and Q. pubescens × pyrenaica were located between the parental species, for each PC (Figure 5).

RWA was performed for landmark data: The first two RWs explained 84.09% of the total variance, resulting in an intermediate leaf shape for the hybrid groups relative to parental species, except for Q. robur × petraea (Figure 6).
For outline data, PC1 and PC2 explained 90.43% of the total variance, and all hybrid groups, relative to parental species, occupied intermediate positions on the first or on both PCs (Figure 7).

**Correlation between matrices**

From the three Mantel tests performed, all correlation results were significant and the datasets were highly correlated. In particular, landmark data were more strictly related to morphological measurements than outline data (Table III).

**Discussion**

In this paper, an integrated analysis of leaf morphology of some European white oaks species was performed. Species identification was performed by means of genetic markers and assignment of individuals to species or hybrids using unsupervised Bayesian analyses. This approach assumes that specimens from the same species will have a more similar genotype than the individuals from different species (Congdon 2001; Baudouin et al. 2004).

Many reports about intermediate phenotypes have received special attention, first by delineating a specific class of intermediate phenotypes in comparison to reference populations (Rushton 1983; Ietswaart & Feij 1989; Elsner 1993) and then by designating these trees as introgressed forms. Other papers indicate that the within-species variation encompassed these intermediate forms, despite the interfertility between the two species (Kremer et al. 2002). Natural hybridization has been shown by mating system analysis with gene markers in natural mixed populations (Bacilieri et al. 1996; Sreiff et al. 1998).
Leaf morphological analyses in European oaks

1999) and interspecific hybridization was also demonstrated by controlled crossings (Kleinschmit & Kleinschmit 2000).

This comparative study of oak leaf morphology, based on the use of two free size geometric morphometric methods (landmark and outline) and of a set of leaf measurements, combined with the genetic assignment of individuals to pure species or hybrids, provided information about the differences among species and the intermediate leaf morphology of their hybrids.

Genetic assignment of individuals clearly demonstrated that the analysed mixed oak stand consisted of four species (Q. robur, Q. pubescens, Q. petraea and Q. pyrenaica) and their genetic hybrids.

Multivariate statistical analyses (PCA, CVA) were performed on all three morphological datasets, and similar results were obtained for leaf morphology of species and the relationships with their hybrids.

For all morphological datasets, the original variables were summarized in several synthetic variables (PCs) that were subjected to discriminant analysis (CVA). The result was that all species were discriminated by means of all three morphological methods: The greater part of differences was among Q. robur, Q. pubescens and Q. pyrenaica, while Q. petraea was less clearly differentiated, because of the low number of individuals sampled.

For each dataset a mean was computed for all definite groups: Four pure species and six hybrid groups. The three new matrices obtained were subjected to PCA, and the mean of each group was projected on PC1 and PC 2 axes to visualize the location of hybrid groups relative to their parental species. For all three datasets, hybrids were located between the two parental species on the PC1 and PC2 axes, indicating that hybrids are globally characterized by an intermediate, albeit highly variable, leaf morphology.

Like other papers that reported the use of the same methods on European white oaks (Dupouey & Badeau 1993; Kremer et al. 2002; Curtu et al. 2007; Viscosi et al. 2009), our results confirmed that oak species were distinguishable by means of leaf morphology.

Morphological data analyses reported by Kremer et al. (2002) were able to discriminate Q. petraea from Q. robur. In the present study, three additional variables (pubescence of the adaxial area, perimeter and roundness) were added to improve the discrimination power for Q. pyrenaica and Q. pubescens. To discriminate Q. petraea from Q. robur, Kremer et al. (2002) reported that the petiole length (pl and pr), intercalary venation (nv and pv), pubescence (pu) and sinus width (sw and ldr) were the original variables that exhibited the highest weight in species differentiation, while the basal shape of the lamina (bs) and the number of lobes (nl) contributed moderately. Our results confirmed these discriminant variables and identified others that were discriminating for other species. While pubescence variables differentiated Q. robur from Q. pubescens and Q. pyrenaica, the leaf circularity and some dimensional variables, such as lobe depth ratio (ldr), lamina length (ll), perimeter, length of the lamina at largest width (wp) and lobe width (lw), distinguished Q. pyrenaica from other species. Q. pubescens was principally distinguished by means of percentage venation (pv) and the number of veins (nv).

Leaf shape analyses (landmark and outline methods) confirmed the observations by Viscosi et al. (2009): Q. robur was characterized by a short petiole and obovate leaf blade with a narrow and auriculate basal shape; Q. petraea was distinguished by an elliptical leaf blade with a sub-acute apical lobe, and for Q. pubescens the leaf blade was elliptic-obovate with cordate basal region and obtuse apex. Finally, Q. pyrenaica was characterized by an elliptical and deeply lobate leaf blade.

In this paper, we tested the discrimination power of the geometric morphometric approaches in comparison with other morphological measurements and analysed morphological relationships among species and their putative hybrids, identified using molecular markers (microsatellites). We discovered that genetic assignment using 10 microsatellites is able to discriminate hybrids up to the second generation. This is clearly a limitation since the species group will contain some third or later generation hybrids which can partially explain the high morphological variability observed in pure species. A more precise vision of hybrid morphology will be gained by adding additional genetic markers to obtain a deeper and more complete view of the morphological consequences of hybridization. Due to the increasing availability of Expressed Sequence Tags (EST) in many species including oaks (e.g. European “Network of Excellence” EVOLTREE: Evolution of Trees as Drivers of Terrestrial Biodiversity), rapid development of numerous new molecular markers, in particular microsatellites located in transcribed regions of the genome (EST-SSR), will lead to an increasing number of loci and to fast genotyping as well. These kinds of loci could show a much higher genetic differentiation between species because they are located in transcribed regions (Pashley et al. 2006; Ellis & Burke 2007). As a result, it will soon be possible to genotype oaks with a much higher number of loci showing a high genetic differentiation, a technological advancement that will increase resolution for the identification of later generation hybrids. In addition, this will offer a more functional and evolutionary perspective in the field...
of morphological differences between species (see Kashi & King 2006 for a recent review), especially when coupled with population genomics approaches (i.e. using a high number of genetic markers to saturate the whole genome; Scotti-Saintagne et al. 2004). The geometric morphometric analysis, compared in the present study with a classical morphological analysis, provides a fast and efficient tool for phenotyping oak leaf morphology, in conjunction with the upcoming new powerful genotyping techniques. Finally, we observed that hybrids are a mosaic of phenotypes with parental and intermediate characters. This result could be in part due to the limitation of hybrid identification by our genetic markers as discussed above, but it also highlights the fact that hybrid groups are not homogeneous, instead they are a mixture of heterogeneous genotypes (Anderson & Stebbins 1954; Arnold 1997). Nevertheless, when mean leaves were analysed in hybrid groups, they resulted in intermediate leaf morphology, relative to parental species. This confirms the difficulties of using morphological features for hybrid identification. This issue will probably be clarified to some extent by analysing a larger sample of species and locations in order to understand the taxonomic and geographical distributions of species hybrids in European white oaks.

Acknowledgements
We thank Jérôme Willm and Guy Roussel for help in collecting leaves and Florian Alberto for leaf measurements. Genotyping was performed at the Genotyping and Sequencing facility of Bordeaux (grants from the Conseil Régional d’Aquitaine no. 20030304002FA and 20040305003 FA and from the European Union, FEDER no. 2003227).

References


