

## Supporting material

Using joint multivariate analyses of leaf morphology and molecular-genetic markers for taxon identification in three hybridizing European white oak species (*Quercus* spp.)

Christian Rellstab, Andreas Bühler, René Graf, Catherine Folly & Felix Gugerli

Annals of Forest Science 2016, doi:10.1007/s13595-016-0552-7

### **Appendix - PCR reactions and fragment length analysis**

Multiplex PCR reactions were conducted with the Qiagen Multiplex PCR Master Mix (Qiagen, Hilden, Germany) using labeled forward (Sigma-Aldrich, St. Louis, USA) and unlabeled reverse primers (Microsynth, Balgach, Switzerland). We used a total reaction volume of 11  $\mu$ l containing 5.5  $\mu$ l of 2x Multiplex PCR Master Mix, 1.1  $\mu$ l of primer premix (0.2-0.5  $\mu$ M per primer), 3.4  $\mu$ l of RNase-free water, and 1  $\mu$ l of 1:30 diluted DNA (originally 61.8 ng/ $\mu$ l  $\pm$  42.8 SD). The cycling protocol consisted of 15 min initial denaturation and polymerase activation at 95°C, 33 cycles of 30 s denaturation at 94°C, 60 s primer annealing at 56°C, 45 s extension at 72°C, and a final extension step of 15 min at 60°C. The PCR products were diluted with 20  $\mu$ l water, and 1  $\mu$ l of the diluted PCR products was mixed with 9.3  $\mu$ l of 1:100 mixture of size standard (GeneScan™-500 LIZ®) and Hi-Di™ formamide (Applied Biosystems, Foster City, USA) for fragment length analysis on a 3730 DNA Analyzer (Applied Biosystems).

**Table S1** Overview of the *Quercus* populations from Switzerland sampled in this study. Given are population number, coordinates (latitude/longitude in decimal degrees), community name, elevation, and number of trees included in the analyses.

Population	Latitude °N	Longitude °E	Community	Elevation (m.a.s.l.)	No. of trees
101	46.3683	6.1797	Arnex-sur-Nyon	465	20
102	46.4412	6.2309	Genolier	564	19
103	46.4805	6.3315	Mont-sur-Rolle	712	20
104	46.6964	6.4718	Romainmôtier-Envy	714	20
106	46.8408	6.6431	Champagne	603	20
107	46.9214	6.8035	Bevaix	521	20
109	47.4632	7.1478	Bonfol	451	20
110	47.4905	7.1945	Bonfol	441	20
111	47.3283	7.6371	Laupersdorf	903	20
112	47.4395	7.5512	Brislach	425	20
113	47.4831	7.5033	Hofstetten-Flüh	446	19
114	47.4602	7.6774	Büren (SO)	641	20
118	47.5370	7.7850	Rheinfelden	389	20
120	47.5359	7.9934	Eiken	340	20
122	47.7607	8.6160	Merishausen	638	20
123	47.6811	8.5266	Neunkirch	491	20
124	47.5186	8.1627	Remigen	555	20
125	47.3923	7.4641	Bärschwil	691	20
202	46.8137	6.7927	Cheyres	546	20
205	46.9247	7.2060	Ulmiz	505	17
209	47.1239	7.3563	Büren an der Aare	558	20
210	47.0923	7.3520	Diessbach bei Büren	506	20
211	46.8872	7.5319	Belp	516	15
212	47.0967	7.5393	Fraubrunnen	490	20
217	47.3267	8.0710	Muhen	519	18
218	47.2492	8.2138	Beromünster	574	19
219	47.4360	8.2553	Birmenstorf	380	20
223	47.5356	8.6171	Buch am Irchel	680	20
225	47.6735	8.6371	Laufen-Uhwiesen	552	16
229	47.5625	9.3399	Romanshorn	438	19
302	46.6189	7.0731	Bulle	770	19
307	47.0678	7.6953	Heimiswil	834	20
310	47.2392	8.8394	Rapperswil-Jona	437	20
313	47.2030	9.0174	Kaltbrunn	420	20
317	47.4381	9.6267	Au (SG)	563	20
319	46.8764	8.2509	Sachslen	590	20
401	46.3039	6.8810	Vionnaz	1000	17
404	46.0873	6.9868	Finhaut	1364	17
406	46.0866	7.0924	Bovernier	992	20
412	46.2663	7.3980	Ayent	881	20
414	46.2384	7.5706	Anniviers	986	16
419	46.1706	7.1666	Saillon	795	20
421	46.2212	7.2552	Vétroz	811	20
424	46.1763	7.0123	Evionnaz	798	20
430	46.3203	7.5810	Salgesch	827	20
433	46.3140	7.7174	Gampel-Bratsch	649	20
440	46.2879	7.8890	Visp	768	20
442	46.3018	7.7157	Turtmann	667	18
501	46.3480	8.6071	Cevio	617	20
502	46.1794	8.6811	Centovalli	750	20
503	46.2338	8.7097	Maggia	338	20
504	46.1607	8.7576	Losone	331	20
505	46.1954	8.8632	Gordola	1016	20
508	46.2505	9.1359	Grono	636	19
509	46.3797	9.2265	Mesocco	901	20
510	46.0861	8.9107	Mezzovico-Vira	568	20
511	46.0420	8.9498	Cureglia	494	18
512	46.0114	9.0091	Lugano	414	18
513	45.9708	8.9215	Collina d'Oro	357	17
514	45.9623	8.8866	Caslano	341	19
515	45.9439	8.9069	Morcote	446	20
516	45.8576	9.0307	Breggia	703	17
518	46.2906	9.1770	Verdabbio	420	19
521	46.2598	8.6880	Maggia	332	20
522	46.1759	8.7463	Losone	230	19
601	47.0425	9.4909	Fläsch	831	20
606	46.7696	9.4378	Tomils	1031	19
608	46.9787	9.5920	Malans	872	20
609	46.8190	9.4321	Domat/Ems	599	19
610	46.8280	9.3666	Trin	934	20
615	46.7853	9.1424	Rueun	1009	16

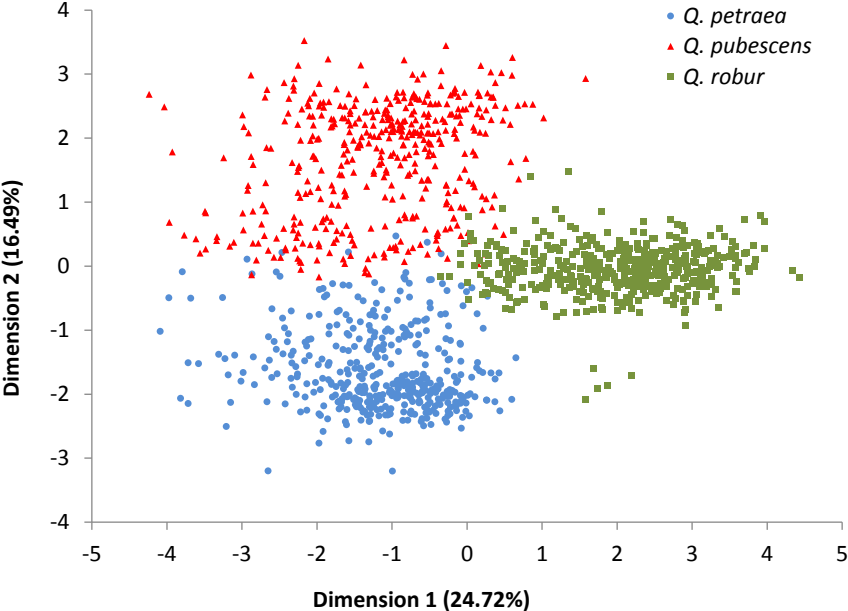
**Table S2** Ranking of the five categorical and nine continuous variables in the hierarchical clustering of principle components (HCPC). Based on the first ten dimensions of the factor analysis of mixed data (FAMD). Ranking according to p values of the  $\chi^2$ -tests (categorical variables) and one-way analysis of variance (ANOVA). Abbreviations of the variables are explained in Table 3 of the manuscript. df=degrees of freedom, Eta<sup>2</sup>=effect size.

Categorical variables ( $\chi^2$ -test)		
	p value	df
LC	2.39 E-214	2
LS	1.26 E-201	2
VC	2.66 E-171	2
LI	3.83 E-162	2
VS	1.09 E-03	2

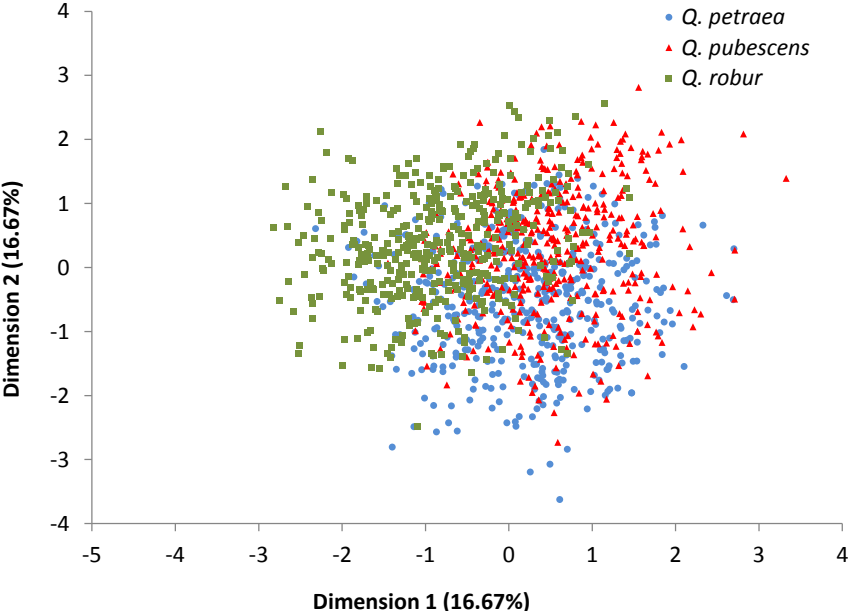
Continuous variables (one-way ANOVA)		
	Eta <sup>2</sup>	p value
PCo1	0.738	<1.01 E-220
PR	0.524	1.01 E-220
RPV	0.441	3.77 E-173
BS	0.177	2.15 E-58
PCo2	0.162	3.36 E-53
LDR	0.120	1.68 E-38
OB	0.107	3.10 E-34
LLNR	0.092	1.53 E-29
LDRW	0.076	3.94 E-24
PCo4	0.058	1.95 E-18
PCo3	0.056	1.03 E-17
PCo6	0.027	5.50 E-09
LWR	0.019	1.78 E-06
PCo5	0.015	3.63 E-05

**Figure S1** Species assignment of 1369 *Quercus* individuals with a factor analysis of mixed data (FAMD) and subsequent hierarchical clustering of principal components (HCPC) using (A) 13 morphological and (B) six synthetic genetic variables. Shown is the position of each tree along the first two dimensions of the FAMD, with the clustering of the HCPC in different colors and symbols. Species assignment of the three proposed clusters was done according to hair types.

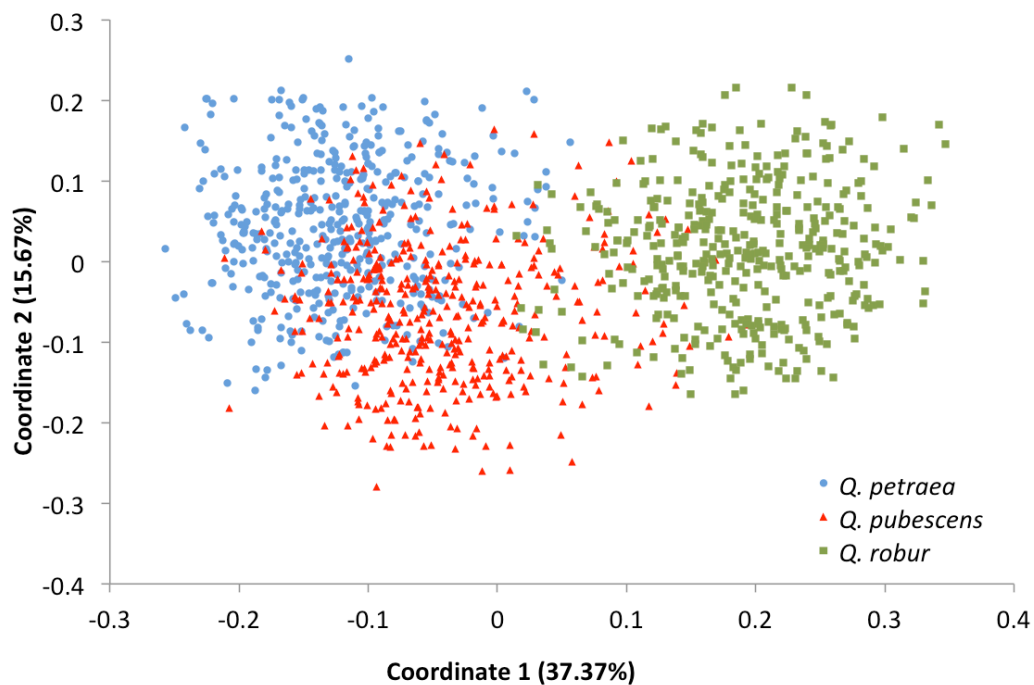
(A)



(B)

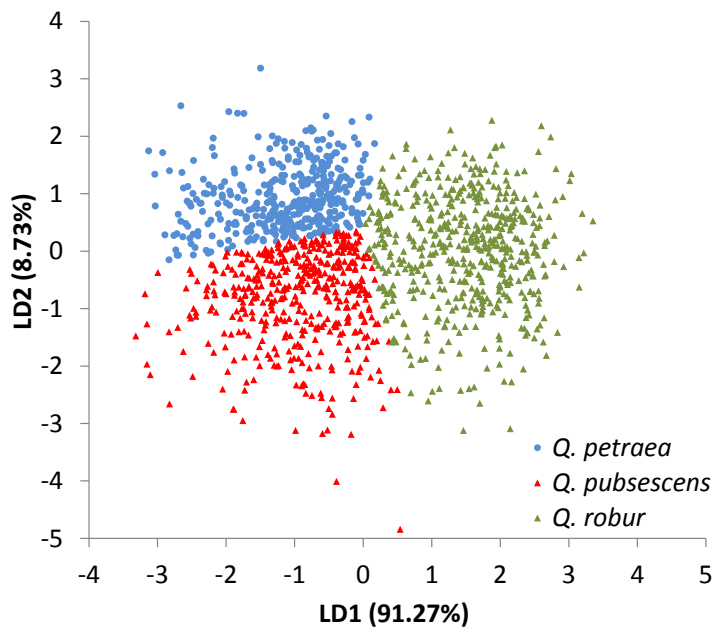


**Figure S2** Results of the principal coordinates analysis (PCoA) of 1369 *Quercus* individuals based on eight microsatellite markers. Shown is the position of each tree along the first two axes of the PCoA. The clustering, shown in different colors and symbols, is based on subsequent factor analysis of mixed data (FAMD) and hierarchical clustering of principal components (HCPC) using the six synthetic variables of the PCoA. Species assignment of the three proposed clusters was done according to hair types. The separation/cohesion ratio (Table 1) of this analysis is 2.46.

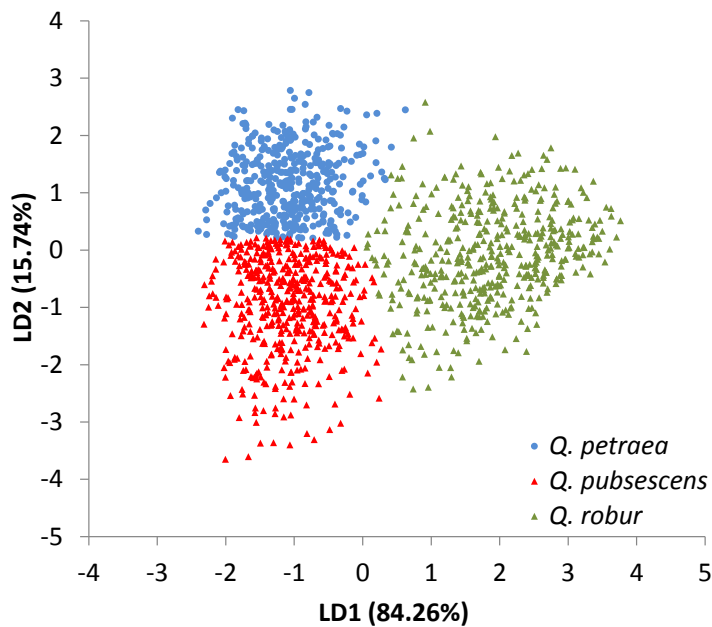


**Figure S3** Species assignment of 1369 *Quercus* individuals with a linear discriminant analysis (LDA) using (A) eight morphological and (B) six synthetic genetic variables. Unambiguous species-specific hair types were used to pre-define the groups. Shown are the position and classification of each tree along the two discriminant functions using highest assignment probability (majority rule) for species assignment.

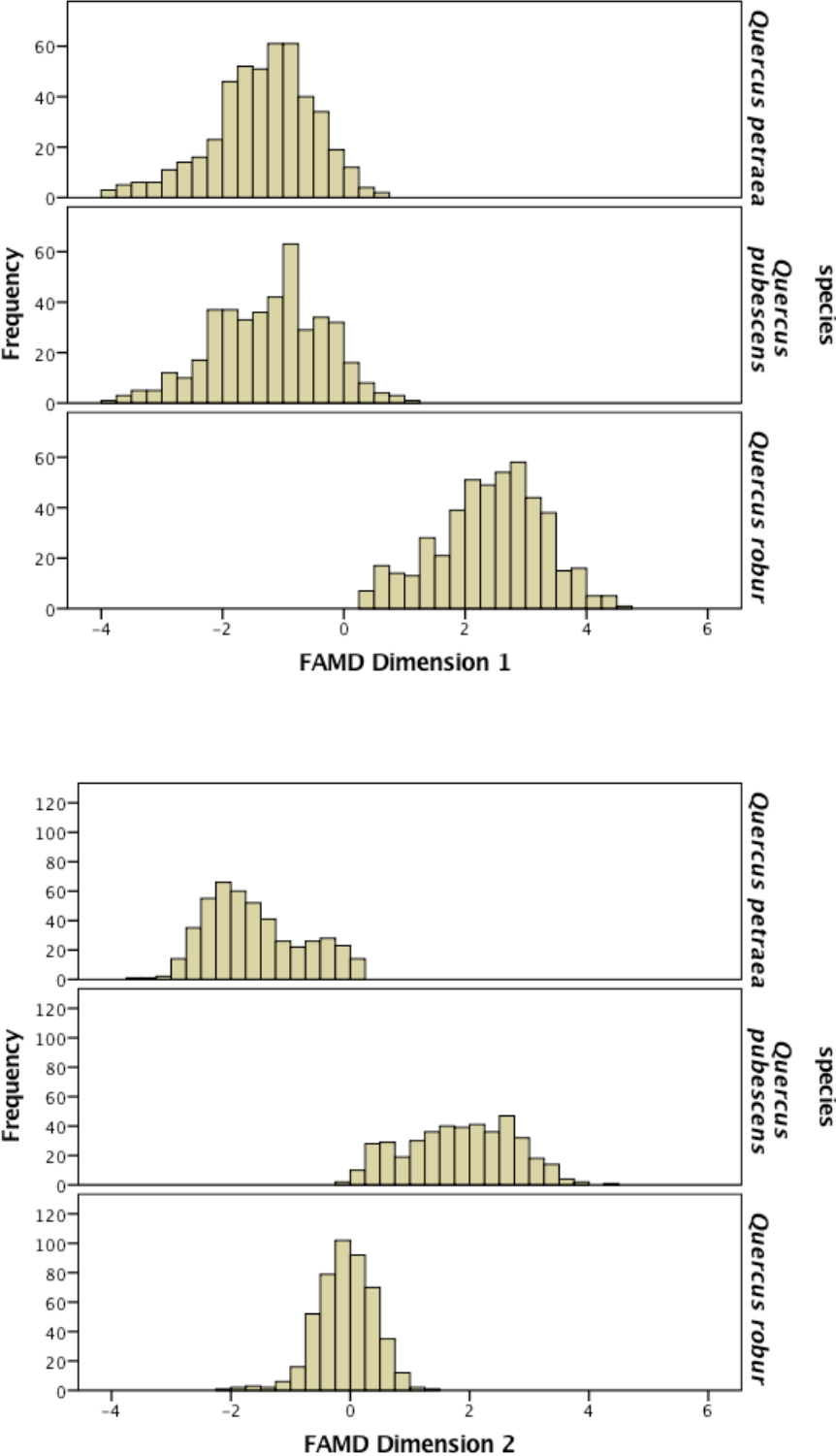
(A)



(B)

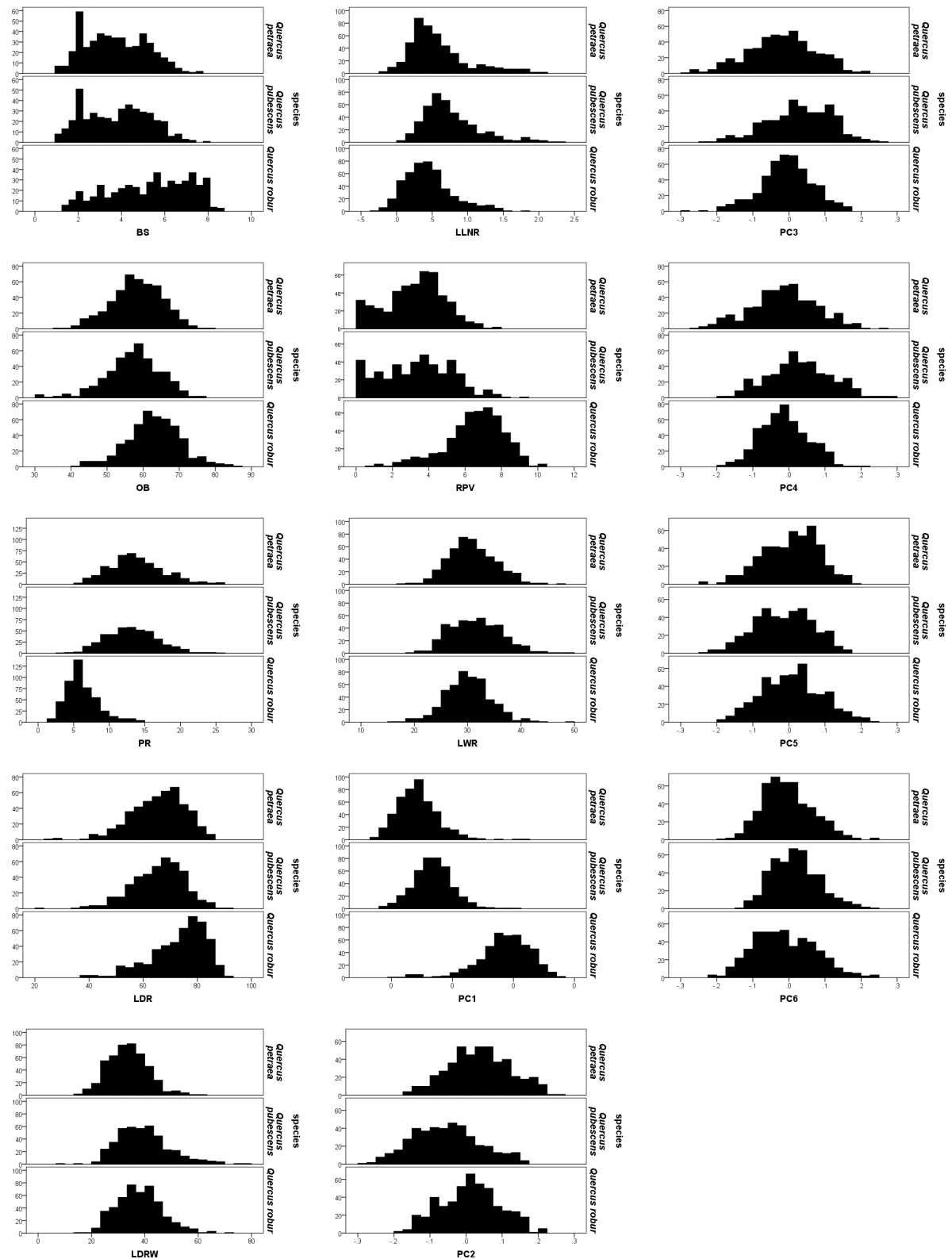


**Figure S4** Histograms (no. of trees) of the two first dimensions of the factor analysis of mixed data (FAMD) using all variables. The species assignment of the groups proposed by hierarchical clustering of principle components (HPCP) was done according to hair types.

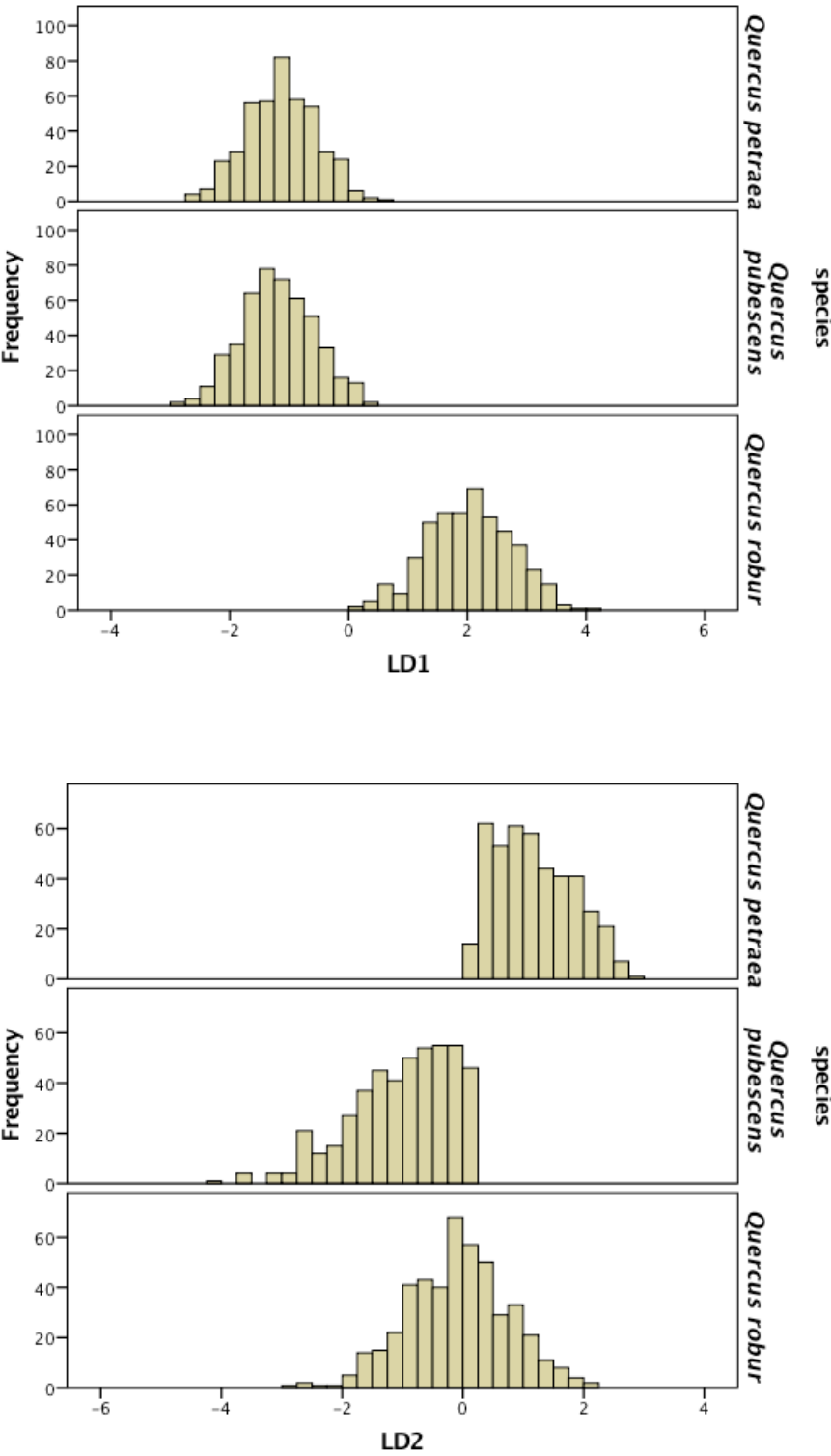




**Figure S5** Histograms (no. of trees) of all 14 continuous variables used in the factor analysis of mixed data (FAMD) for each species according to the clustering using hierarchical clustering of principle components (HCPC). Species assignment of the groups proposed by HPCP was done according to hair types. Abbreviations are explained in Table 3 of the manuscript.

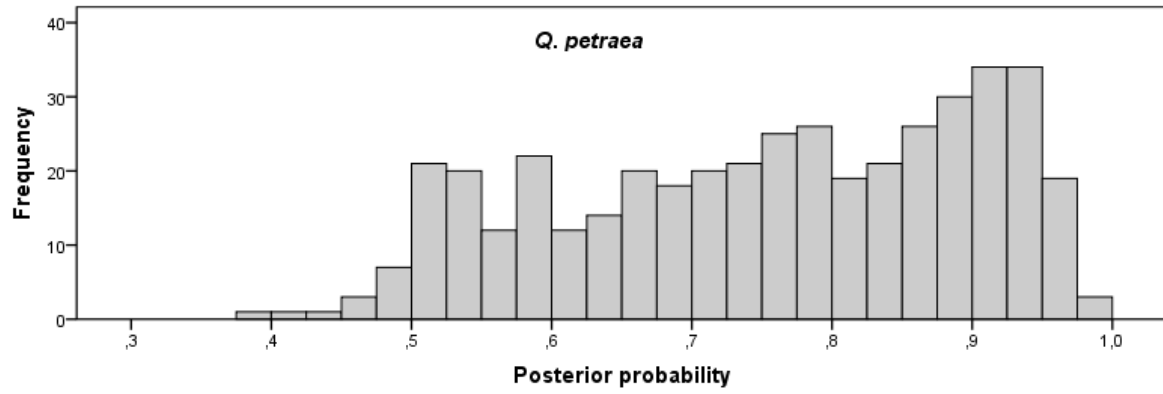


**Figure S6** Histograms (no. of trees) of the two linear discriminants (LDs) obtained from linear discriminant analysis (LDA) using all variables. Species assignment according to majority rule.

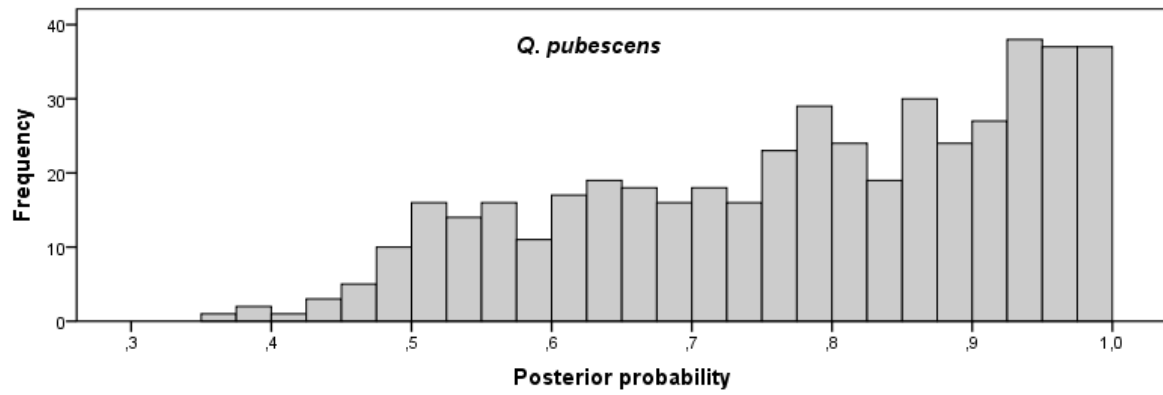


**Figure S7** Posterior probability distribution of the 1369 *Quercus* individuals to belong to a species in the linear discriminant analysis (LDA).

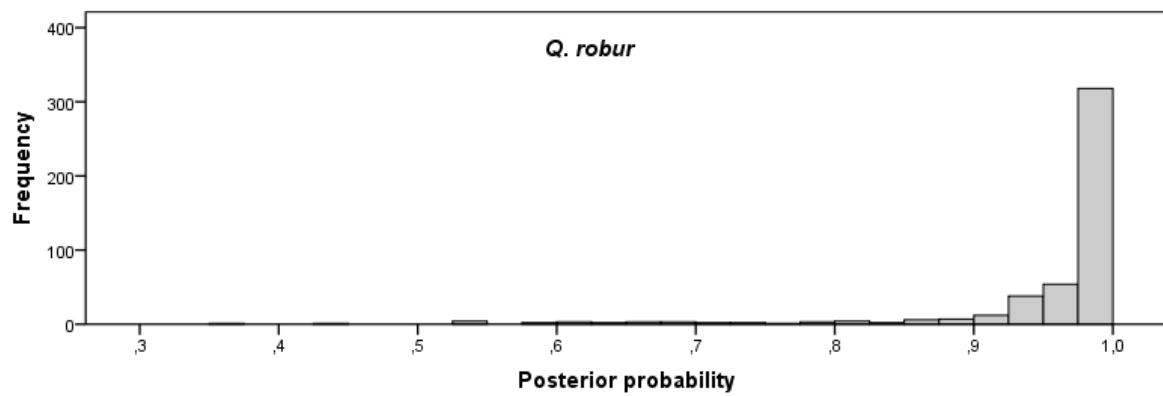
(A)



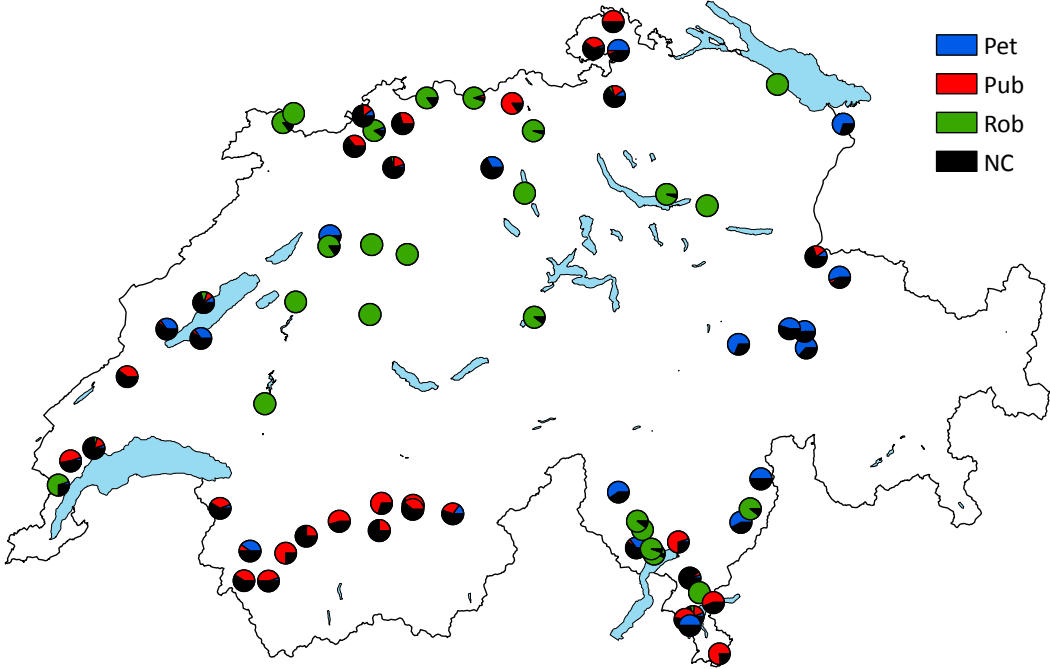
(B)



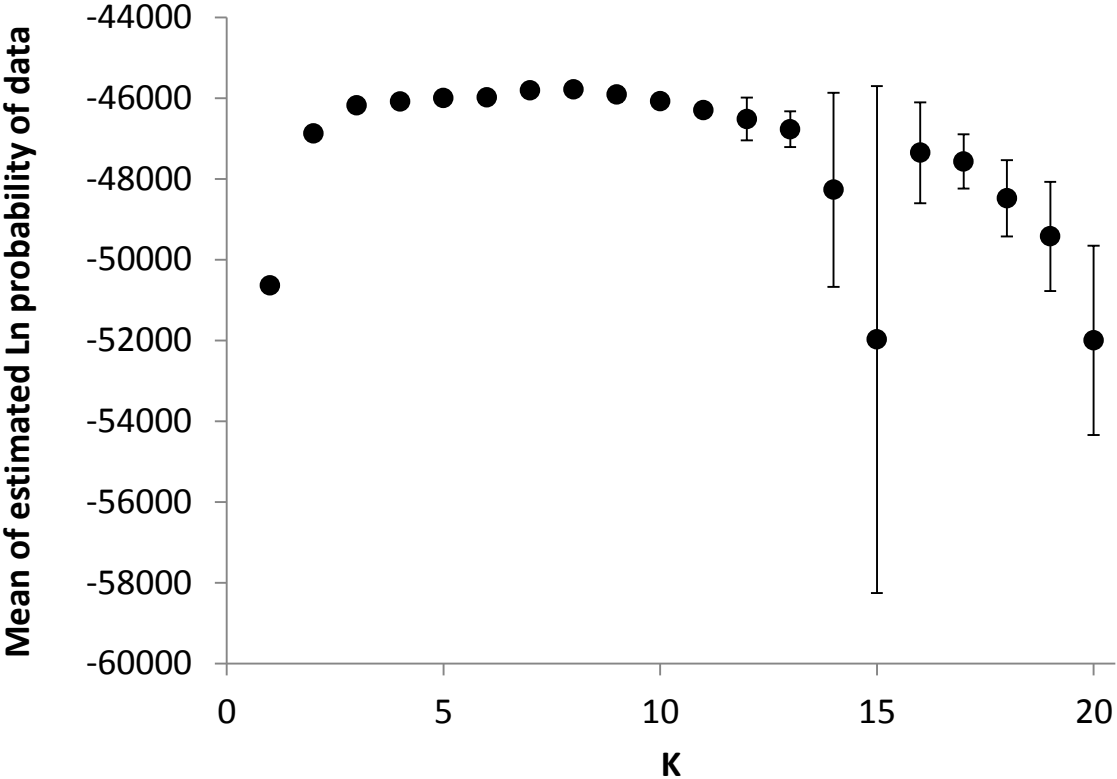
(C)



**Figure S8** Species composition of the 71 investigated *Quercus* populations in Switzerland using linear discriminant analysis (LDA) with an assignment probability threshold of 80%. See also Fig. 2b of the manuscript. Pet=*Q. petraea*, Pub=*Q. pubescens*, Rob=*Q. robur*, NC=not classified.

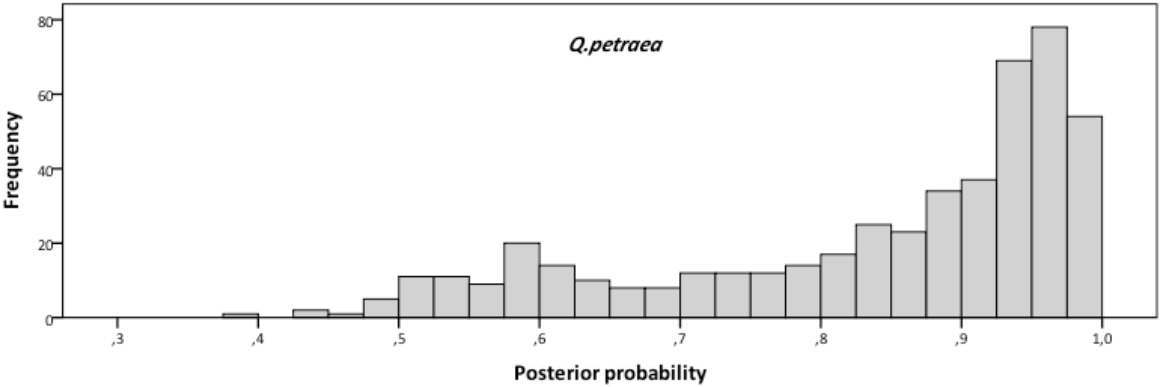


**Figure S9** Results of the STRUCTURE analysis (Pritchard et al. 2000). Shown is the mean likelihood ( $\pm$ SD) of the data at various numbers of  $K$  (1 to 20).

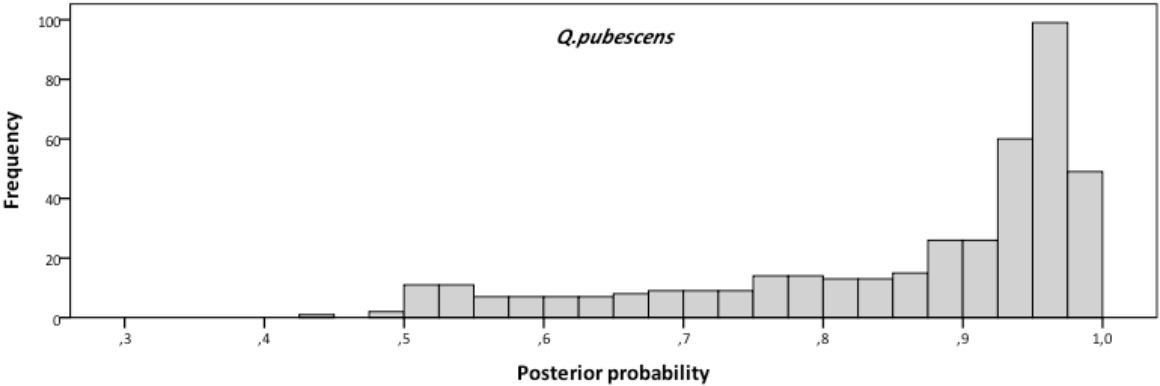


**Figure S10** Posterior probability distribution of the 1369 *Quercus* individuals to belong to a species using STRUCTURE (Pritchard et al. 2000) with  $K=3$ . Species assignment of the three clusters was done according to hair types

(A)



(B)



(C)

