© Indian Academy of Sciences

ONLINE RESOURCES

Fingerprinting of Fagaceae individuals using intermicrosatellite markers

J. P. COUTINHO, A. CARVALHO and J. LIMA-BRITO*

Institute of Biotechnology and Bioengineering (IBB), Centre of Genomics and Biotechnology (CGB), University of Tras-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

[Coutinho J. P., Carvalho A. and Lima-Brito J. 2014 Fingerprinting of Fagaceae individuals using intermicrosatellite markers. *J. Genet.* **93**, e132–e140. Online only: http://www.ias.ac.in/jgenet/OnlineResources/93/e132.pdf]

Introduction

The genera *Castanea* (chestnuts), *Fagus* (beechs) and *Quercus* (oaks) are the only genera of the family Fagaceae distributed throughout the temperate forests of the northern hemisphere (Kremer *et al.* 2007). Comprising more than 900 species and nine genera (Manos *et al.* 2001), this family is considered highly complex at the phylogenetic level. However, its taxonomic classification is presently being widely revised due to the availability of new molecular data.

Inter-simple sequence repeats (ISSRs) have been used in different Fagaceae species for genetic diversity analyses, DNA fingerprinting and estimation of interspecific and intraspecific relationships (Goulão *et al.* 2001; Carvalho *et al.* 2009; Coutinho *et al.* 2014). ISSRs could be used alone or in combination with other markers such as the simple sequence repeats (SSRs) (López-Aljorna *et al.* 2007). Recent technological progresses made in single-locus markers reduced the importance of multi-locus markers. SSRs are the marker system of choice for genetic diversity evaluation in species with reduced polymorphism, otherwise, both SSRs and ISSRs could be used for the same (López-Aljorna *et al.* 2007; Gomes *et al.* 2008; Kumar and Sharma 2011; Jia *et al.* 2011; Noormohammadi *et al.* 2013). Many expressed sequence tags (EST) were developed for Fagaceae species such as *Quercus petraea*, *Q. robur*, *C. sativa* and *F. sylvatica* (NCBI dbEST, http://www.ncbi.nlm.nih.gov/dbest/index.html). The ESTs are useful for comparative mapping studies but their transferability is dependent on the occurrence of point mutations in the primer annealing sites, and could lead to some uncertainty in phylogenetic inferences. On the other hand, ISSRs are ubiquitous, simple, affordable and produce a large number of polymorphic fragments (increasing their resolving power and polymorphic information content), thus being suitable for genetic diversity assessment of distant and closely related species and for the estimation of phylogenies.

Due to their economic and ecological importance, seven Fagaceae species were selected for this study: *Castanea sativa* Mill., *Fagus sylvatica* L., *Quercus ilex* L., *Q. rubra* L., *Q. pyrenaica* L., *Q. robur* L. and *Q. petraea* (Mattuschka) Liebl. According to Denk and Grimm (2010), these oak species belong to the infrageneric groups *Ilex* (Ilex oaks), *Lobatae* (red oaks) and *Quercus* (white oaks). Except for *Q. robur* (Eurasian oak) and *Q. rubra* (American oak), the remaining oaks have European ecological origins.

The ISSR amplified products of closely related individuals usually generate a similar pattern across an entire species. In this context, we aim to evaluate the potential of the cost-effective ISSR markers for the assessment of intraspecific genetic diversity among individuals belonging to these seven species of Fagaceae with high economic and ecological importance, as well as examining their phylogenetic relationships.

Materials and methods

Plant material and genomic DNA extraction

To avoid misclassification and hypothetical hybridizations between the progenitors of these individuals, the seeds were collected from 13 different botanical gardens to ensure that all the parents were nonrelated trees (table 1). After seed viability evaluation, dormancy break and germination, the young plantlets were grown in peat moss under greenhouse conditions at UTAD (Vila Real, Portugal). Genomic DNA was extracted from young leaves following a CTAB based protocol with some modifications.

ISSRs amplification

Fifty ISSR primers from the set 100/9 of the University British Columbia (UBC) were tested. The amplifications were performed in a final volume of 20 μL containing 25 ng of genomic DNA, 8 μL of distilled water (Qiagen, Venlo,

*For correspondence. E-mail: jbrito@utad.pt.

Keywords. intraspecific genetic diversity; intermicrosatellite (ISSR) markers; Fagaceae; *Quercus*; *Castanea*; *Fagus*.

Table 1. Taxonomical classification, ecological distribution, seeds origin and voucher numbers of the Fagaceae individuals studied here. The HVR (herbarium vila real) designation is in accordance to Index Herbariorum codes (Thiers 2010).

Genus	Infrageneric group	Species	Ecological distribution	Seed origins[a]	Voucher number
Quercus	*Ilex*	*Q. ilex* L. (holm oak)	Mediterranean	04	HVR20141
				09	HVR20159
				10	HVR20160
	Lobatae	*Q. rubra* L. (northern red oak)	NE of America	02	HVR20179
				06	HVR20176
				10	HVR20177
				11	HVR20175
				12	HVR20178
	Quercus	*Q. petraea* (Mattuschka) Liebl. (sessile oak)	Europe, SW Asia	06	HVR20161
				07	HVR20146
				10	HVR20162
				12	HVR20163
		Q. pyrenaica L. (pyrenean oak)	SW Europe, N. Africa	02	HVR20164
				10	HVR20150
		Q. robur L. (pedunculate oak)	Europe, SW Asia	01	HVR20170
				02	HVR20169
				04	HVR20168
				06	HVR20165
				09	HVR20166
				12	HVR20167
Castanea		*C. sativa* Mill. (sweet chestnut)	Europe, SW Asia	08	HVR20158
				10	HVR20157
				11	HVR20130
Fagus		*F. sylvatica* L. (european beech)	Europe	03	HVR20174
				05	HVR20131
				06	HVR20173
				11	HVR20171
				13	HVR20172

NE, north east; SW, south west; N, north. [a]Seed origins: 01, Arboretumul Simeria, Hunedoara, Romania; 02, Botanical Garden of the University of Tras-os-Montes and Alto Douro, Vila Real, Portugal; 03, Botanischen Garten der Johannes Gutenberg-Universität, Mainz, Germany; 04, Botanischer Garten der Westfälische Wilhelms, Universität Münster, Germany; 05, Dubrava Arboretum, Kaunas, Lithuania; 06, Herbarium, University of Aarhus, Aarhus, Denmark; 07, Hortus Botanicus Tallinnensis, Tallinn, Estonia; 08, Hortus Botanicus, Universitatis Latviensis, Riga, Latvia; 09, Jardin Botanique de la Ville de Lyon, Lyon, France; 10, Ökologisch Botanischer Garten, Universität Bayreuth, Germany; 11, Rivierenh of Park, Deurne, Belgium; 12, Späth-Arboretum, Humboldt-Universität zu Berlin, Berlin, Germany; 13, Stad Antwerpen Plantentuin, Antwerp, Belgium.

The Netherlands), 10 μL of Taq Master Mix (Qiagen) and 5 pmol of primer. The amplification conditions followed those reported by Carvalho *et al.* (2009) and the amplification reactions were repeated twice.The ISSR amplified products were visualized after electrophoresis on 1.8% w/v agarose gels stained with ethidium bromide and analysed for the presence (1) or absence (0) of bands to construct binary matrices for statistical analyses. Each band was considered an ISSR locus and only reproducible bands were used.

Statistical analyses

To measure the informative and discriminatory capacity of each primer, the polymorphic information content (PIC) (Botstein *et al.* 1980) and the resolving power (Rp) (Prevost and Wilkinson 1999) were calculated. The Nei's gene diversity (h) and the average gene diversity were calculated for each primer as $h = 2n\ (1 - \sum x_i^2)/(2n - 1)$, where h is the gene diversity for a single locus; x_i, the frequency of the ith marker; n, the number of bands (Nei 1978).

The genetic relationships were estimated with the software PopGene v.1.32 (Yeh *et al.* 1999) by calculating each taxonomic level, the percentage of polymorphic bands, Shannon's index of phenotypic diversity (I) and the expected heterozygosity (H_{E}).

With the pool of ISSR data, a genetic similarity UPGMA (unweighted pair group method with arithmetic mean) dendrogram was generated using the simple matching coefficient and sequential agglomerative hierarchical and nested (SAHN) algorithm with the software NTSYS-PC 2.02 (Rohlf 1998). The clustering confidence was tested by bootstrap

analysis using the software Winboot (Yap and Nelson 1996) based on 1000 replications. The modules COPH and MXCOMP of the NTSYS-PC 2.02 calculated the cophenetic correlation coefficient to test the goodness of fit of the UPGMA clustering to the data matrices.

The STRUCTURE 2.3.4 software (Pritchard *et al.* 2000) was used to confirm the NTSYS UPGMA-clustering and to evaluate the genetic structure of the plant material, with the analysis being performed under the 'no admixture model' suitable for dominant markers (Falush *et al.* 2007) for different population numbers (K) with 50,000 generations of burn-in period and 100,000 MCMC (Markov chain Monte Carlo) iterations.

Results and discussion

Among the 50 primers tested for ISSR amplification, 17 were selected for having produced the largest polymorphic ISSR patterns between individuals of the same species. These primers presented eight repetition motifs, redundant anchoring at the 3′ end and were asymmetric to avoid primer slippage and redundancy or duplication of the amplified loci.

Within each species, we detected some monomorphic bands among individuals from different provenances that generated a similar pattern (figure 1). Also, several polymorphic ISSR markers were found among the individuals of the same species. On pooling all the ISSR data obtained from the 17 primers, high intraspecific ISSR polymorphism was detected in most of the studied species (figure 1; table 2).

The size of the ISSR bands ranged from 150 to 2900 bp and an average of 9.2 bands per primer were amplified. The primers 826, 835 and 850 produced the lowest number of bands (16), whereas the primer 841 presented the highest value (31). Primers 835, 841 and 880 were also successfully used for the amplification of ISSR markers in *Q. suber* individuals from different provenances (López-Aljorna *et al.* 2007). The oligonucleotide 841 produced the highest number of fragments in this study, similar to that reported for *Q. suber* (López-Aljorna *et al.* 2007) and *Q. coccifera* (Carvalho *et al.* 2009). These results suggest that this primer is highly discriminative in different Fagaceae species and may be useful for DNA fingerprinting and assessment of intraspecific genetic diversity. Further, the capacity of each primer to differentiate individuals belonging to the same species was evaluated by calculating the Rp, and the highest Rp values also corresponded to the oligonucleotide 841 (table 2). The Rp values ranged from 0.67 in *C. sativa* with the primers 826, 835 and 842 to 8.80 in *Q. rubra* with the primer 841 (table 2). The highest Rp average was 4.16 in *Q. ilex* and the lowest 2.86 in *C. sativa*, while the overall Rp average value produced by these 17 primers in the seven species was 3.75 (table 2). These values are lower than the ones reported in other studies involving the same ISSR primers in barley (Fernández *et al.* 2002) and coffee (Mishra *et al.* 2011).

The minimum PIC values were obtained in *C. sativa* (0.178) with primer 842, and *F. sylvatica* (0.315) with primer 836 (table 2), suggesting that these primers are more discriminative in species of the genus *Quercus*. However, with the exception of *Q. robur*, the primer 810 showed the highest PIC value (0.375) for the remaining species of the three genera (table 2). Considering that dominant markers may provide a maximum PIC value of 0.5 (Riek *et al.* 2001), it reveals that primer 810 was the more discriminative for the assessment of intraspecific genetic diversity. The maximum value (0.5) of Nei's gene diversity (h, Nei 1978 was detected with several primers in different species (table 2), while the lowest value (0.198) was found in *C. sativa* with primer 842. Additionally, all PIC values were slightly lower than the corresponding h, revealing that the 17 primers did not overestimate the heterozygosity, proving the ISSRs ability to correctly discriminate individuals per species and genus.

The ISSR pool data showed high average percentages of polymorphism that ranged from 76.5% in *Q. pyrenaica* to 99.2% in *Q. robur* (table 3). The highest values for Shannon's index of phenotypic diversity (I; 0.518) and expected heterozygosity (H_E; 0.3412) were found in *Q. ilex*.

Previous genetic studies performed with ISSRs have reported high interspecific and intraspecific genetic variability in the Fagaceae family, particularly, within the genus *Quercus* (López-Aljorna *et al.* 2007; Carvalho *et al.* 2009; Coutinho *et al.* 2014). The high phenotypic plasticity and genetic variation of this genus could explain its ecological

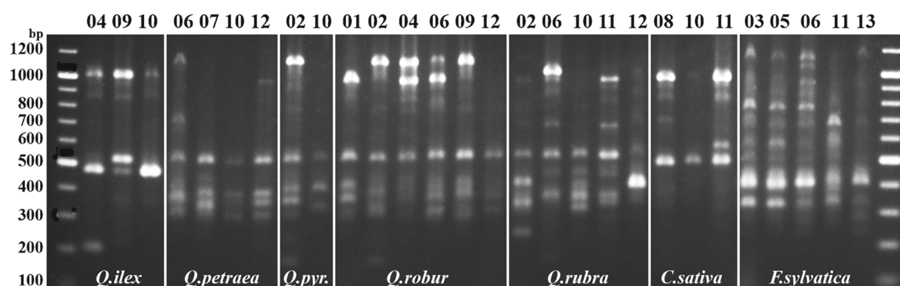

Figure 1. ISSR profiles amplified with primer 841. Above each lane is indicated the provenance of each individual (according to table 1). M, molecular weight marker, Gene Ruler DNA Ladder Mix (Fermentas).

Table 2. ISSR results achieved per primer among the studied Fagaceae species.

Primer (5′ → 3′)	Species	*T*	*M*	%P	PIC	*h*	Rp
807	*C. sativa*	8	0	100	0.368	0.486	5.33
$(AG)_8T$	*F. sylvatica*	9	1 (1150 bp)	88.9	0.369	0.488	5.20
	Q. ilex	8	1 (875 bp)	87.5	0.368	0.486	4.67
	Q. petraea	7	0	100	0.370	0.490	4.00
	Q. pyrenaica	5	3 (850, 625, 550 bp)	40.0	0.269	0.320	2.00
	Q. robur	10	0	100	0.339	0.433	5.00
	Q. rubra	10	0	100	0.360	0.471	4.80
810	*C. sativa*	8	0	100	0.375	0.500	5.33
$(GA)_8$ T	*F. sylvatica*	10	2 (650, 550 bp)	80.0	0.375	0.499	4.80
	Q. ilex	6	0	100	0.372	0.494	4.00
	Q. petraea	7	0	100	0.374	0.497	4.50
	Q. pyrenaica	3	0	100	0.375	0.500	3.00
	Q. robur	8	0	100	0.359	0.469	4.67
	Q. rubra	11	0	100	0.327	0.413	6.00
817	*C. sativa*	9	2 (925, 725 bp)	77.7	0.375	0.499	4.67
$(CA)_8A$	*F. sylvatica*	10	1 (750 bp)	90.0	0.371	0.493	4.80
	Q. ilex	7	2 (950, 750 bp)	71.4	0.370	0.490	3.33
	Q. petraea	7	2 (925, 750 bp)	71.4	0.375	0.500	5.00
	Q. pyrenaica	8	1 (625 bp)	87.5	0.371	0.492	7.00
	Q. robur	8	0	100	0.353	0.457	5.00
	Q. rubra	10	1 (950 bp)	90.0	0.375	0.500	6.00
818	*C. sativa*	9	2 (1150, 750 bp)	77.7	0.372	0.494	4.67
$(CA)_8G$	*F. sylvatica*	6	3 (1600, 1150, 1050 bp)	50.0	0.346	0.444	1.60
	Q. ilex	8	2 (1600, 1050 bp)	75.0	0.346	0.444	4.00
	Q. petraea	7	0	71.4	0.370	0.490	3.00
	Q. pyrenaica	4	2 (925, 550 bp)	50.0	0.305	0.375	2.00
	Q. robur	9	0	100	0.366	0.483	5.33
	Q. rubra	9	0	100	0.353	0.458	4.40
823	*C. sativa*	5	1 (925, 525 bp)	80.0	0.346	0.444	2.00
$(TC)_8C$	*F. sylvatica*	5	1 (750 bp)	80.0	0.375	0.499	2.80
	Q. ilex	7	0	100	0.370	0.490	4.67
	Q. petraea	11	0	100	0.356	0.463	6.00
	Q. pyrenaica	5	2 (750 bp)	80.0	0.365	0.480	4.00
	Q. robur	7	0	100	0.360	0.472	5.33
	Q. rubra	6	1 (750 bp)	83.3	0.374	0.498	3.20
825	*C. sativa*	5	2 (950, 750 bp)	60.0	0.346	0.444	2.00
$(AC)_8T$	*F. sylvatica*	8	1 (725 bp)	87.5	0.365	0.480	3.60
	Q. ilex	10	0	100	0.365	0.480	6.67
	Q. petraea	8	1 (1600 bp)	87.5	0.375	0.500	4.00
	Q. pyrenaica	6	0	100	0.375	0.500	6.00
	Q. robur	10	0	100	0.339	0.433	5.00
	Q. rubra	9	0	100	0.375	0.500	3.60
826	*C. sativa*	3	2 (1100, 750 bp)	33.3	0.286	0.346	0.67
$(AC)_8C$	*F. sylvatica*	9	1 (1900 bp)	88.9	0.369	0.488	4.40
	Q. ilex	6	0	100	0.346	0.444	4.00
	Q. petraea	6	0	100	0.359	0.469	4.50
	Q. pyrenaica	3	0	100	0.375	0.500	3.00
	Q. robur	8	0	100	0.328	0.413	4.00
	Q. rubra	10	2 (1600, 1050 bp)	80.0	0.298	0.365	4.80
827	*C. sativa*	4	1 (925 bp)	75.0	0.346	0.444	2.00
$(AC)_8G$	*F. sylvatica*	11	1 (1300 bp)	90.9	0.365	0.480	5.20
	Q. ilex	9	0	100	0.358	0.466	6.00
	Q. petraea	6	1 (925 bp)	83.3	0.346	0.444	3.00
	Q. pyrenaica	5	3 (1050, 925, 575 bp)	40.0	0.269	0.320	2.00
	Q. robur	8	0	100	0.337	0.430	3.67
	Q. rubra	10	0	100	0.348	0.449	6.40

Table 2 *(contd)*

Primer (5′ → 3′)	Species	*T*	*M*	%P	PIC	*h*	Rp
834 $(AG)_8YT$*	*C. sativa*	4	0	100	0.375	0.500	2.67
	F. sylvatica	6	0	100	0.365	0.480	3.20
	Q. ilex	6	1 (850 bp)	83.3	0.372	0.494	3.33
	Q. petraea	3	1 (550 bp)	66.7	0.346	0.444	1.00
	Q. pyrenaica	2	0	100	0.375	0.500	2.00
	Q. robur	3	0	100	0.321	0.401	1.67
	Q. rubra	4	0	100	0.372	0.495	2.00
835 $(AG)_8YC$*	*C. sativa*	1	0	100	0.346	0.444	0.67
	F. sylvatica	4	1 (1400 bp)	75.0	0.332	0.420	1.20
	Q. ilex	5	0	100	0.374	0.498	3.33
	Q. petraea	4	1 (450 bp)	75.0	0.375	0.500	2.00
	Q. pyrenaica	3	2 (925, 450 bp)	33.3	0.239	0.278	1.00
	Q. robur	4	0	100	0.305	0.375	2.00
	Q. rubra	4	0	100	0.332	0.420	2.00
836 $(AG)_8YA$*	*C. sativa*	6	2 (825, 650 bp)	66.6	0.372	0.494	2.67
	F. sylvatica	6	0	100	0.315	0.391	2.80
	Q. ilex	8	0	100	0.373	0.497	5.33
	Q. petraea	7	2 (850, 350 bp)	71.4	0.370	0.490	3.00
	Q. pyrenaica	6	3 (850, 525, 350 bp)	50.0	0.305	0.375	3.00
	Q. robur	4	0	100	0.368	0.486	2.00
	Q. rubra	10	0	100	0.341	0.435	5.60
841 $(GA)_8YC$*	*C. sativa*	6	2 (500, 975 bp)	66.6	0.346	0.444	2.67
	F. sylvatica	12	3 (340, 425, 675 bp)	75.0	0.375	0.500	4.80
	Q. ilex	6	3 (475, 875, 1000 bp)	50.0	0.321	0.401	2.00
	Q. petraea	9	2 (375, 525 bp)	77.7	0.374	0.499	4.50
	Q. pyrenaica	9	2 (425, 525 bp)	77.7	0.362	0.475	7.00
	Q. robur	14	1 (525 bp)	92.9	0.363	0.477	8.33
	Q. rubra	15	1 (525 bp)	93.3	0.365	0.480	8.80
842 $(GA)_8YG$*	*C. sativa*	3	2 (825, 275 bp)	33.3	0.178	0.198	0.67
	F. sylvatica	8	1 (725 bp)	87.5	0.372	0.495	3.20
	Q. ilex	7	0	100	0.370	0.490	4.67
	Q. petraea	8	0	100	0.371	0.492	6.00
	Q. pyrenaica	5	1 (275 bp)	80.0	0.365	0.480	4.00
	Q. robur	7	0	100	0.312	0.387	3.67
	Q. rubra	8	0	100	0.332	0.420	3.60
846 $(CA)_8RT$*	*C. sativa*	8	0	100	0.346	0.444	5.33
	F. sylvatica	9	1 (1300 bp)	88.9	0.374	0.498	4.40
	Q. ilex	8	0	100	0.373	0.497	5.33
	Q. petraea	9	0	100	0.346	0.444	6.00
	Q. pyrenaica	3	1 (1400 bp)	66.7	0.346	0.444	2.00
	Q. robur	7	0	100	0.280	0.337	3.00
	Q. rubra	7	0	100	0.338	0.431	4.00
850 $(GT)_8YC$*	*C. sativa*	4	0	100	0.346	0.444	2.67
	F. sylvatica	8	0	100	0.351	0.455	4.40
	Q. ilex	7	0	100	0.374	0.499	4.67
	Q. petraea	6	0	100	0.346	0.444	4.00
	Q. pyrenaica	6	0	100	0.375	0.500	6.00
	Q. robur	3	0	100	0.321	0.401	1.67
	Q. rubra	3	0	100	0.269	0.320	1.20
856 $(AC)_8YA$*	*C. sativa*	4	0	100	0.368	0.486	2.67
	F. sylvatica	9	1 (1700 bp)	88.9	0.365	0.480	5.20
	Q. ilex	6	1 (450 bp)	83.3	0.362	0.475	3.33
	Q. petraea	8	2 (750, 450 bp)	75.0	0.374	0.498	3.50
	Q. pyrenaica	7	0	100	0.375	0.500	7.00
	Q. robur	4	0	100	0.305	0.375	2.00
	Q. rubra	5	0	100	0.298	0.365	2.40

Table 2 *(contd)*

Primer (5′ → 3′)	Species	*T*	*M*	%P	PIC	*h*	Rp
880	*C. sativa*	4	1 (1400 bp)	75.0	0.346	0.444	2.00
$(GGAGA)_3$	*F. sylvatica*	8	1 (1900 bp)	87.5	0.365	0.480	4.00
	Q. ilex	3	1 (750 bp)	66.7	0.372	0.494	1.33
	Q. petraea	4	1 (650 bp)	75.0	0.371	0.492	1.50
	Q. pyrenaica	3	0	100	0.375	0.500	3.00
	Q. robur	4	0	100	0.275	0.330	1.67
	Q. rubra	3	1 (750 bp)	66.6	0.374	0.498	0.80

T, total number of amplified ISSR bands; *M*, number of monomorphic bands; %P, percentage of ISSR polymorphism; PIC, polymorphic information content; *h*, (Nei 1978), Nei's gene diversity; Rp, resolving power produced by each primer within each Fagaceae specie.
*Single letter abbreviations for mixed positions R, (purine); Y, (pyrimidine)

and adaptive success (Dodd and Kashani 2003). The high intraspecific and intrastand polymorphisms reported for the *Quercus* species have been explained based on their mating system, the low distance among stands or small size of the stands (López-Aljorna *et al.* 2007). In fact, we detected monomorphic ISSR markers among individuals from different species, such as those common to *Q. pyrenaica* and *Q. petraea* amplified with primers 827 (925 bp), 835 (450 bp); 836 (350 bp and 850 bp) and 841 (525 bp). However, despite their provenance, several monomorphic ISSR markers were also detected among individuals of the same species. Thus, the existence of monomorphic markers support the amplification of the observed monomorphic bands (table 3), and might explain some of the lowest Rp values presented in table 2.

Some of the monomorphic ISSR bands found in *C. sativa*, *F. sylvatica*, *Q. ilex* and *Q. petraea* were considered as species-specific bands (table 3). These monomorphic bands may be useful in different approaches such as quantitative trait loci (QTL) association studies, gene-tagging, discrimination between species with similar morphological traits, detection of natural hybrids or development of sequenced characterized amplified regions (SCARs). A similar approach was performed by (López-Aljorna *et al.* 2007) regarding the DNA fingerprinting of cork oak elite trees from different provenances based on ISSRs and simple sequence repeats (SSRs). According to them, the molecular data derived from the characterization of cork oak elite trees could be applied to the management of genetic resources, being useful for the definition of strategies for reforestation, maintenance of rural areas and optimization of cork production (López-Aljorna *et al.* 2007). Moreover, the high number of unique ISSR fragments amplified per oak species in this study and in Carvalho *et al.* (2009) also reinforced the potential of these molecular markers for the DNA fingerprinting and intraspecific genetic variability assessment in species of the family Fagaceae.

The pool of the ISSR data was used to construct an UPGMA dendrogram of genetic similarity to estimate the genetic relationships among the 28 Fagaceae individuals under study (figure 2). Despite having different origins and belonging to different species, the UPGMA dendrogram reflected only 17.5% of genetic diversity among the individuals studied (figure 2). Except for *C. sativa*, the ISSR data revealed higher interspecific rather than intraspecific genetic variability. The high cophenetic correlation coefficient ($r = 0.84$) indicated that the ISSR data matrix was properly represented by the UPGMA clustering. A consensus tree constructed with the NTSYS software was compared with the one obtained with the Winboot software. Despite only bootstrap values higher than 50% and common to both consensus trees are indicated (Felsenstein 1985) in figure 2, they support several branches. Despite the relatively high genetic identity, all the 28 individuals were adequately clustered in

Table 3. Global results achieved per Fagaceae species with the 17 ISSR primers.

Species	*T*	*M*	Np	% P	M (species specific)	H_E	*I*
C. sativa	91	18	10	80.2	$880_{1400\ bp}$	0.2934 (0.1624)	0.4415 (0.2298)
F. sylvatica	138	19	13	86.2	$835_{1400\ bp}$; $880_{1900\ bp}$	0.2705 (0.1592)	0.4174 (0.2151)
Q. ilex	117	11	7	90.6	$841_{1000\ bp}$	0.3472 (0.1301)	0.5178 (0.1714)
Q. petraea	117	13	9	88.9	$834_{550\ bp}$	0.2892 (0.1525)	0.4425 (0.2062)
Q. pyrenaica	81	19	10	76.5	0	0.3093 (0.1813)	0.4516 (0.2646)
Q. robur	118	1	1	99.2	0	0.2626 (0.1256)	0.4204 (0.1550)
Q. rubra	134	6	5	95.5	0	0.2636 (0.1232)	0.4214 (0.1550)

T, total number of amplified ISSR bands per species; *M*, number of monomorphic bands per species; Np, number of primers that amplified the M bands; %P, average percentage of ISSR polymorphism per species; number of monomorphic species-specific bands; average values (and standard deviation within brackets) of expected heterozygosity (H_E); I, Shannon's index of phenotypic diversity.

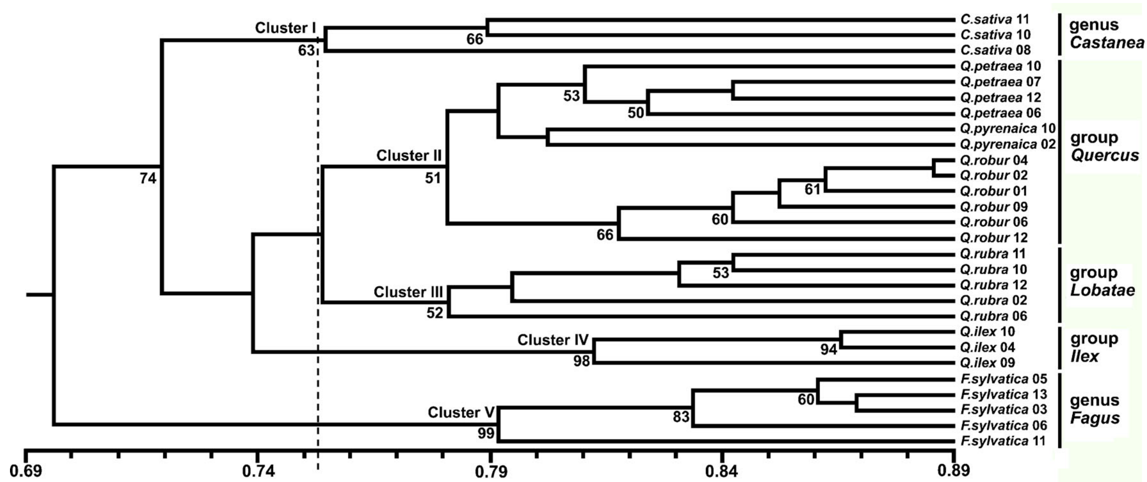

Figure 2. UPGMA dendrogram of genetic similarity among the 28 Fagaceae individuals based on the ISSR data pool obtained from 17 primers and calculated using the simple matching (SM) coefficient. The bootstrap percentage values are indicated which represent 1000 bootstrap cycles (99% confidence).

the dendrogram and the relationships were generally supported by the bootstrap values. All the Fagaceae individuals from different provenances were clustered per species and genus. In this study and in a previous work (Coutinho *et al.* 2014) performed on Fagaceae species based on ISSRs, we found high genetic similarity between the genera *Castanea* and *Quercus*. The similarity at the intermicrosatellite regions of these genera constitutes no surprise due to the high transferability of SSR loci between chestnut and oak species, as previously demonstrated by Barreneche *et al.* (2004).

For a cut-off value of 0.7525, five clusters could be considered, which included: cluster I, all *C. sativa* individuals; cluster II, oaks from group *Quercus*; cluster III, oaks from group *Lobatae*; cluster IV, oaks from group *Cerris*; and cluster V, the *F. sylvatica* individuals. The ISSRs have been proved to be specific for the establishment of phylogenetic relationships and discrimination of plants per species, genus or other taxa, ecological origins or distribution areas (Chokchaichamnankit *et al.* 2008; Carvalho *et al.* 2009; Myking *et al.* 2011; Cipriano *et al.* 2013; Coutinho *et al.* 2014). Nevertheless, after the estimation of the genetic relationships among cork oak elite trees from different provenances based on ISSRs, López-Aljorna *et al.* (2007) reported that these markers did not show any clear relationship with provenance regions. In our opinion, the discriminatory ability of ISSRs increases as much as the number of genotypes, number of primers and ISSR markers amplified per study. This study is based on 28 individuals and 17 primers that produced a total number of 334 ISSR amplified bands. Also, the molecular data were validated by a Bayesian analysis that reflected the genetic substructure of the plant material studied using the STRUCTURE 2.3 software (figure 3) and corroborated its UPGMA clustering (figure 2).

The genetic substructure analysis revealed five clusters ($K = 5$) that, as mentioned, correspond to the discrimination of the Fagaceae individuals per genus and *Quercus* infrageneric groups, which was also been revealed by the UPGMA dendrogram. The F_{ST} indexes obtained revealed moderate to high genetic differentiation in four clusters: genus *Fagus* ($F_{ST} = 0.2242$); genus *Castanea* ($F_{ST} = 0.2231$); group *Ilex* ($F_{ST} = 0.4281$) and group *Lobatae* ($F_{ST} = 0.4251$). The remaining cluster corresponded to the infrageneric group *Quercus*, whose F_{ST} value of 0.1628 indicates a reduced genetic differentiation. The cluster II, in the UPGMA clustering, included all individuals from the three oak species that belong to this infrageneric

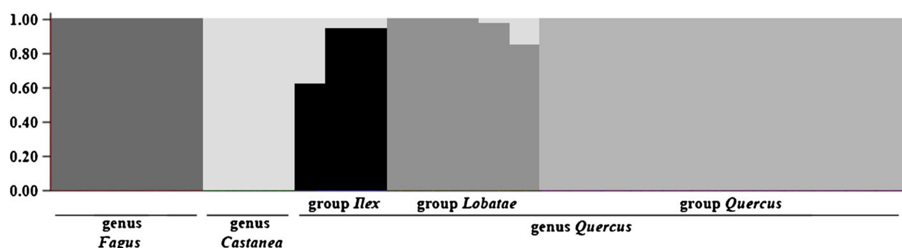

Figure 3. Bar plot that represents the genetic structure for $K = 5$, based on the ISSR data pool.

group. Despite their different provenances and species, these individuals shared monomorphic bands (probably conserved ISSR regions in these oak genomes) contributing for the reduced genetic differentiation of this cluster as extrapolated by the STRUCTURE analysis. ISSRs constitute rapid evolving regions but the maintenance of conserved regions among oaks and other Fagaceae species might have an important evolutionary role. The detection of species-specific ISSR markers in individuals from very distant provenances should be characterized in the future, once they could correspond to coding regions for interesting traits useful for improvement, gene-tagging or markers-assisted selection. Specific ISSR markers were previously used for gene tagging of interesting agronomic traits (Ammiraju *et al.* 2001) and for marker-assisted selection (Dayteg *et al.* 2008) in other plant species. Even if some degree of genetic similarity among individuals of the genera *Castanea* and *Quercus* was revealed (particularly with oaks from the groups *Ilex* and *Lobatae*), globally the proposed five clusters were adequately discriminative (figure 3). The occurrence of high gene flow and natural hybridization among oak species greatly contribute to taxonomical problems such as the alternative consideration of two distinct species or two different subspecies at the level of infrageneric groups or sections (Aykut *et al.* 2013). The potential of the ISSR markers to discriminate individuals per infrageneric group, as demonstrated in this study, is adequate to overcome such taxonomical problems by revealing real phenetic and phylogenetic relationships that could help to define the limits for each taxon.

Regarding the number of ESTs developed till July 2014, publicly available in the NCBI dbEST, a large number of ESTs were developed in *Q. petraea* (more than 58,000), *Q. robur* (more than 80,000), *C. sativa* (around 1,000) and *F. sylvatica* (over 31,000). On the contrary, only two ESTs can be found for *Q. rubra*, three for *Q. ilex* and none for *Q. pyrenaica*. Subjected to the level of known genomic sequences of a species, the development of new SSR markers may be achieved either by the *in vivo* standard method (construction of small-insert library, hybridization and sequencing of candidate clones; Liu *et al.* 1996), by using the ISSR-suppression PCR technique (Lian *et al.* 2001), or *in silico* by sequence database mining (Kantety *et al.* 2002). Thus, several monomorphic bands found in the complete genome outlook performed in this study, may promote the identification of new SSRs and ESTs, particularly, for the less studied species. In fact, SSRs have been used for the construction of the genetic maps of *F. sylvatica* (Scalfi *et al.* 2004) or *Q. robur* (Barreneche *et al.* 1998), but ISSRs have also revealed their potential for mapping purposes in *C. sativa* (Casasoli *et al.* 2001).

Globally, the ISSRs have proved to be discriminative enough for the clustering of Fagaceae individuals per genus, species and infrageneric groups of the genus *Quercus*, revealing their ability to estimate phylogenics and taxonomic inferences.

Acknowledgements

This study was supported by the Ph.D. grant SFRH/BD/42837/2008, attributed by FCT (Portuguese Foundation for Science and Technology), and by the IBB-CGB/UTAD. The authors also thank the Botanical Gardens listed in table 1 for the collection of seeds used in this work.

References

Ammiraju J. S. S., Dholakia B. B., Santra D. K., Singh H., Lagu M. D., Tamhankar S. A. *et al.* 2001 Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. *Theor. Appl. Genet.* **102**, 726–732.

Aykut Y., Emel U. and Tekin B. M. 2013 Molecular diversity among Turkish oaks (*Quercus*) using random amplified polymorphic DNA (RAPD) analysis. *Afr. J. Biotechnol.* **12**, 6358–6365.

Barreneche T., Bodenes C., Lexer C., Trontin J. F., Fluch S., Streiff R. *et al.* 1998 A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme, and 5S rDNA markers. *Theor. Appl. Genet.* **97**, 1090–1103.

Barreneche T., Casasoli M., Russell K., Akkak A., Meddour H., Plomion C. *et al.* 2004 Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theor. Appl. Genet.* **108**, 558–566.

Botstein D., White R. L., Skolnick M. and Davis R. W. 1980 Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**, 314–331.

Carvalho A., Paula A., Guedes-Pinto H., Martins L., Carvalho J. and Lima-Brito J. 2009 Preliminary genetic approach based on both cytogenetic and molecular characterisations of nine oak species. *Plant Biosyst.* **143**, 25–33.

Casasoli M., Mattioni C., Cherubini M. and Villani F. 2001 A genetic linkage map of European chestnut (*Castanea sativa* Mill.) based on RAPD, ISSR, and isozyme markers. *Theor. Appl. Genet.* **102**, 1190–1199.

Chokchaichamnankit P., Chulalaksananukul W., Phengklai C. and Anamthawat-Jónsson K. 2008 Species and genetic diversity of Fagaceae in northern Thailand based on ISSR markers. *J. Trop. For. Sci.* **20**, 8–18.

Cipriano J., Carvalho A., Fernandes C., Gaspar M. J., Pires J., Bento J. *et al.* 2013 Evaluation of genetic diversity of Portuguese Pinus sylvestris L. populations based on molecular data and inferences about the future use of this germplasm. *J. Genet.* **92**, e41–e48.

Coutinho J., Carvalho A. and Lima-Brito J. 2014 Genetic diversity assessment and estimation of phylogenetic relationships among 26 Fagaceae species using ISSRs. *Biochem. Syst. Ecol.* **54**, 247–256.

Dayteg C., Rasmussen M., Tuvesson S., Merker A. and Jahoor A. 2008 Development of an ISSR-derived PCR marker linked to nematode resistance (Ha2) in spring barley. *Plant Breed.* **127**, 24–27.

Denk T. and Grimm G. W. 2010 The oaks of western Eurasia: Traditional classifications and evidence from two nuclear markers. *Taxon* **59**, 351–366.

Dodd R. S. and Kashani N. 2003 Molecular differentiation and diversity among the California red oaks (Fagaceae; *Quercus* section *Lobatae*). *Theor. Appl. Genet.* **107**, 884–892.

Falush D., Stephens M. and Pritchard J. K. 2007 Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* **7**, 574–578.

Felsenstein J. 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Fernández M. E., Figueiras A. M. and Benito C. 2002 The use of ISSR and RAPD markers for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. *Theor. Appl. Genet.* **104**, 845–851.

Gomes S., Martins-Lopes P., Lima-Brito J., Meirinhos J., Lopes J., Martin A. *et al.* 2008 Evidence for clonal variation in 'Verdeal-Transmontana' olive using RAPD, ISSR and SSR markers. *J. Hortic. Sci. Biotech.* **83**, 395–400.

Goulão L., Valdiviesso T., Santana C. and Oliveira C. M. 2001 Comparison between phenetic characterisation using RAPD and ISSR markers and phenotypic data of cultivated chestnut (*Castanea sativa* Mill.) *Genet. Resour. Crop Ev.* **48**, 329–338.

Jia X. D., Wang T., Zhai M., Li Y. R. and Guo Z. R. 2011 Genetic diversity and identification of Chinese-grown pecan using ISSR and SSR markers. *Molecules* **16**, 10078–10092.

Kantety R. V., Rota M. L., Matthews D. E. and Sorrells M. E. 2002 Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* **48**, 501–510.

Kremer A., Casasoli M. and Barreneche T. [and others] 2007 Chapter 5 Fagaceae Trees. In *Genome mapping and molecular breeding in plants*, vol. 7 (ed. C. Kole), pp. 161–187. Springer, Leipzig, Germany.

Kumar V. and Sharma S. N. 2011 Comparative potential of phenotypic, ISSR and SSR markers for Characterization of sesame (Sesamum indicum L.) varieties from India. *J. Crop Sci. Biotech.* **14**, 163–171.

Lian C., Zhou Z. and Hogetsu T. 2001 A simple method for developing microsatellite markers using amplified fragments of inter-simple sequence repeat (ISSR). *J. Plant Res.* **114**, 381–385.

Liu Z. W., Biyashev R. M and Saghai-Maroof M. A. 1996 Development of simple sequence repeat DNA markers and their integration into a barley linkage map. *Theor. Appl. Genet.* **93**, 869–876.

López-Aljorna A., Bueno M. A., Aguinagalde I. and Martín J. P. 2007 Fingerprinting and genetic variability in cork oak (*Quercus suber* L.) elite trees using ISSR and SSR markers. *Ann. For. Sci.* **64**, 773–779.

Manos P. S., Zhou Z. K. and Cannon C. H. 2001 Systematics of Fagaceae: phylogenetic tests of reproductive trait evolution. *Int. J. Plant Sci.* **162**, 1361–1379.

Mishra M. K., Nishani S. and Jayarama 2011 Molecular identification and genetic relationships among coffee species (coffeal) inferred from ISSR and SRAP marker analyses. *Arch. Biol. Sci.* **63**, 667–679.

Myking T., Yakovlev I. and Ersland G. A. 2011 Nuclear genetic markers indicate Danish origin of the Norwegian beech (*Fagus sylvatica* L.) populations established in 500–1,000 AD. *Tree Genet. Genomes* **7**, 587–596.

Nei M. 1978 Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583–590.

Noormohammadi Z., Hasheminejad-Ahangarani Farahani Y., Sheidai M., Ghasemzadeh-Baraki S. and Alishah O. 2013 Genetic diversity analysis in Opal cotton hybrids based on SSR, ISSR, and RAPD markers. *Genet. Mol. Res.* **12**, 256–269.

Prevost A. and Wilkinson M. J. 1999 A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor. Appl. Genet.* **98**, 107–112.

Pritchard J. K., Stephens M. and Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Riek J. D., Calsyn E., Everaert I., Bockstaele E. V. and Loose M. D. 2001 AFLP based alternatives for the assessment of distinctness, uniformity and stability of sugar beet varieties. *Theor. Appl. Genet.* **103**, 1254–1265.

Rohlf F. J. 1998 NTSYS-pc ver. 2.02. Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket, New York, USA.

Scalfi M., Troggio M., Piovani P., Leonardi S., Magnaschi G., Vendramin G. G. *et al.* 2004 A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.) *Theor. Appl. Genet.* **108**, 433–441.

Thiers B. 2010 Index herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. Available at: http://sweetgum.nybg.org/ih.

Yap I. V. and Nelson R. J. 1996 WinBoot: A program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrogram, discussion paper series 14. International Rice Research Institute, Manila, Philippines.

Yeh F. C., Yang R. C. and Boyle T. 1999 POPGENE Microsoft Windows-based freeware for population genetic analysis. Release 1.32. University of Alberta, Edmonton.

Received 21 May 2014, in revised form 23 July 2014; accepted 28 July 2014
Unedited version published online: 7 August 2014
Final version published online: 9 December 2014