Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (Quercus spp.)

E. GUICHOUX,*†‡ L. LAGACHE,*† S. WAGNER,*†§ P. LÉGER*† and R.J. PETIT*†

*INRA, UMR Biodiversité Gènes & Communautés 1202, F-33610 Cestas, France, †University of Bordeaux, UMR Biodiversité Gènes & Communautés 1202, F-33610 Cestas, France, ‡Centre de Recherche Pernod-Ricard, F-94000 Créteil, France, §University of Bonn, Steinmann Institut, D-53115 Bonn, Germany

Abstract

Multiplex PCR is a fast and cost-effective technique allowing increased genotyping throughput of microsatellites. We developed two multiplexes for Quercus petraea and Q. robur, a 12-plex of EST-SSRs (eSSRs) and an 8-plex of genomic SSRs (gSSRs). We studied the origin of allele calling errors at the human reader and software levels. We showed that the robustness of allele identification can be improved by binning on raw peak sizes prior to genetic data analysis. We checked through simulation the power of these markers for species delimitation and hybrid detection. The resolution achieved with all 20 markers was greatly improved compared to that of previous studies based on a subset of the markers. Preliminary PCR tests suggest that these multiplexes might be useful to study other oak species as well. The strategy used for multiplex microsatellite development (from PCR conditions to the definition of allele calling rules) should be broadly applicable.

Keywords: manual binning, microsatellites, molecular identification, Quercus spp.

Received 5 October 2010; revision received 2 December 2010; accepted 10 December 2010

Introduction

Oaks (Quercus spp.) are widely distributed across the Northern Hemisphere. They are often dominant forest tree species and play therefore key ecological and economical roles. For instance, in France, they represent 40% of the forests and almost 60% of wood lumber production. The two major temperate European species (Quercus petraea and Q. robur) have become important models for population genetic and speciation studies (Streiff et al. 1998, 1999; Muir et al. 2000; Petit et al. 2002, 2004; Barreneche et al. 2004; Scotti-Saintagne et al. 2004; Prida et al. 2007; Lepais et al. 2009; Morin et al. 2010). Studying the evolutionary dynamics of such closely related species requires suitable genetic markers (Váhá & Primmer 2006). In recent studies, simple sequence repeats (SSRs) have been the markers of choice to study hybridization (Burgarella et al. 2009; Viscosi et al. 2009; Ortego & Bonal 2010; Penalosa-Ramirez et al. 2010) and population genetic structure (Neophytou et al. 2010). At the same time, single-nucleotide polymorphisms (SNP) genotyping is emerging as a possible alternative in oaks as in other tree species (Namroud et al. 2008; Eckert et al. 2009; Lascoux & Petit 2010). Nevertheless, many basic or applied questions in population genetics only require a small number of highly polymorphic markers on large sample numbers. High-density SNP genotyping is not suitable in such cases. Instead, multiplexing SSRs can improve genotyping throughput as well as cost-effectiveness. Multiplexing is the amplification of several markers in a single PCR (polymerase chain reaction) and must be distinguished from pool plexing, where pooling takes place after PCR. Multiplex PCR is increasingly used (Hayden et al. 2008; Kawalko et al. 2009). However, large multiplexes involving eight or more markers are still uncommon (Hill et al. 2009), because of long development procedures and complex reaction interactions. Since a few years, new tools for multiplex development have appeared, including software for primer design to limit interactions between primers during PCR and for selecting the best combinations of loci (Holleley & Geerts 2009). Moreover, the generalization of second generation sequencing techniques now allows fast and affordable SSR identification (Abdelkrim et al. 2009; Santana et al. 2010).
2009). In oaks, although microsatellites have been available for many years (Dow et al. 1995; Steinkellner et al. 1997; Kampfer et al. 1998), multiplexing efforts were limited, with only two studies reporting multiplexing at no more than five loci (Dzialuk et al. 2005; Lepais et al. 2006). Thus, analysing large oak populations at multiple markers remains expensive and time-consuming. In this study, we developed two multiplex kits, a 12-plex of expressed sequence tag-SSRs (eSSRs) and an 8-plex of genomic SSRs (gSSRs), paying particular attention to genotyping accuracy and cost-effectiveness. We describe the whole procedure, with a focus on the binning phase (i.e. the identification of peaks corresponding to the different alleles) by comparing the performance of two genotyping software. Finally, we test the assignment power of both multiplex kits using simulated oak genotypes and study their transferability on congeneric species and on species belonging to other genera within the Fagaceae family.

Material and methods

Material

Part of the material used is coming from a 5-ha mixed oak stand comprising both Q. petraea and Q. robur located in the western part of France (Petite Charnie State Forest, Sarthe, latitude: 48.08° N, longitude: 0.17° W). This stand has been intensively studied for many years for gene flow, species differentiation, phenology and wood characteristics (Bacilieri et al. 1993, 1994, 1995; Streiff et al. 1998, 1999; Prida et al. 2006, 2007; Lepais et al. 2009). In 2000, 273 adult trees from this stand were grafted in a nursery (Guémené-Penfao, Loire-Atlantique, France). Each genotype was cloned eight times. A total of 898 surviving ramets were sampled (number of ramet per genotype: 1–8, mean: 2.2). In addition, 3780 trees belonging to 51 half-sib families (originating from seeds collected on 28 Q. robur and 23 Q. petraea adult trees from the Petite Charnie stand) were planted in 1998 and 2001, close to the adult stand. In 2009, we sampled 1257 trees from 35 half-sib families (18 Q. robur and 17 Q. petraea). For each tree, one leaf or several buds were stored in sealed plastic bags with 10 g of silica gel. The taxonomic status of the adult trees had previously been characterized using 19 leaf measures. Trees were classified into three categories: Q. petraea, Q. robur or intermediate (Kremer et al. 2002). The two multiplex kits were further tested on Q. pubescens, Q. pyrenaica, Q. alba, Q. rubra, Q. faginea, Q. suber, Q. ilex, Castanea sativa and Fagus sylvatica (number of samples per species: 5–48) and sampled in southwest of France in natural populations or in an arboretum (for Q. alba, Q. rubra, Q. faginea, Q. suber and Q. ilex).

DNA isolation

Five leaf discs (5 mm diameter) or two buds for each tree to standardize the starting quantity of tissue were collected in 96-well plates. DNA was isolated with Invisorb DNA plant HTS 96 kit (Invitek, Germany), following the manufacturer instructions, except for the lysis step (1 h at 65°C). Disruption of plant material was carried out using a Mixer Mill MM300 (Retsch, Germany). In each well of the 96-well plates, a 3-mm tungsten bead was added and the plates were frozen in liquid nitrogen for 2 min before a 1-min disruption step at 30 Hz. DNA quality was estimated on a 1% (w/v) agarose gel stained with GelRed (Biotium, USA). DNA concentration was evaluated on an eight channel Nanodrop spectrophotometer, and concentration of each sample was adjusted to 10 ng/μL on a STARlet 8-channel robot (Hamilton, USA).

Multiplex PCR optimization

Kit-1. Sixty-four eSSRs (Durand et al. 2010, Table S2, Supporting Information) derived from expressed sequence tags (ESTs) were first tested on 24 samples from across the European range (12 Q. petraea and 12 Q. robur trees). They were analysed on a 4000L automatic DNA sequencer (LI-COR Biosciences, USA). Criteria for SSR selection were as follows: good amplification quality, no slippage and high number of alleles (>5). We then determined which specific combination of loci provides the highest species assignment power with the software WHICHLOC (Banks et al. 2003). A subset of 17 loci was selected for further evaluation.

Kit-2. In the second kit, we included highly validated genomic SSRs (gSSRs) (Dow et al. 1995; Steinkellner et al. 1997; Kampfer et al. 1998), some of which had already been multiplexed (Lepais et al. 2006). We selected 10 loci suitable for species differentiation to develop a second multiplex (8-plex) and to increase taxonomic resolution in combination with kit-1.

We first validated all SSRs in simplex using the M13-tail technique (Schnuelle 2000), which allows direct visualization of the PCR product on capillary sequencer. Hence, SSRs presenting low-quality profiles, i.e. excessive stuttering, weak alleles, triple bands, unspecific products or heterogeneous profiles (more than 50% of difference in fluorescence intensity between the two alleles of a heterozygote), were excluded or redesigned from original sequences (Dow et al. 1995; Steinkellner et al. 1997; Kampfer et al. 1998; Durand et al. 2010) using Primer3Plus (Untergasser et al. 2007). To help null-allele detection, 12 families (composed of the female parent and seven offspring) were genotyped at all loci. We also tested microsatellite loci for null alleles, large allele
dropout and scoring errors because of stutter peaks with MICRO-CHECKER v.2.2.0.3 (Van Oosterhout et al. 2004). Further validations (microsatellites scoring and error rate measurement) were only performed on kit-1 because gSSRs (kit-2) are already highly validated (Dow et al. 1995; Steinkellner et al. 1997; Kampfer et al. 1998). Once validated in simplex, and prior to multiplexing, primers were examined for possible interactions using a local BLAST. The complementary threshold (the maximum number of AT or CG matches for any two primers within a multiplex reaction) was set to seven (Holleley & Geerts 2009). The multiplex reactions were then carried out with the Qiagen Multiplex PCR kit (Qiagen, Germany), following the manufacturer instructions. Final volume was optimized (10 μL) as well as final concentration of Mastermix (0.6x), reducing eight times the final cost.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequences (5’–3’)</th>
<th>Reference</th>
<th>LG</th>
<th>Dye</th>
<th>[C]</th>
<th>Motif Size (bp)</th>
<th>A</th>
<th>H_o</th>
<th>F_Is</th>
<th>F_ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIE20</td>
<td>CGAAGGGCTCTCTTAAATACAGAACT</td>
<td>Durand et al. 2010</td>
<td>1</td>
<td>FAM</td>
<td>1.00</td>
<td>AG</td>
<td>97–119</td>
<td>11</td>
<td>0.668</td>
<td>-0.002</td>
</tr>
<tr>
<td>PIE22</td>
<td>TAGAAGCCCAACAGGCCTAC</td>
<td>Durand et al. 2010</td>
<td>2</td>
<td>FAM</td>
<td>1.00</td>
<td>GGT</td>
<td>197–221</td>
<td>9</td>
<td>0.749</td>
<td>-0.057</td>
</tr>
<tr>
<td>PIE152</td>
<td>TGCTACCTGTTTCTCTCTTTAAACT</td>
<td>Durand et al. 2010</td>
<td>2</td>
<td>FAM</td>
<td>3.75</td>
<td>TA</td>
<td>230–260</td>
<td>15</td>
<td>0.842</td>
<td>-0.024</td>
</tr>
<tr>
<td>PIE242</td>
<td>TGAGGGAAAAAGAACAAATGC</td>
<td>Durand et al. 2010</td>
<td>3</td>
<td>VIC</td>
<td>1.00</td>
<td>TA</td>
<td>102–128</td>
<td>12</td>
<td>0.803</td>
<td>0.045</td>
</tr>
<tr>
<td>PIE102</td>
<td>ACCTTCCATGGCTCAAAAGATG</td>
<td>Durand et al. 2010</td>
<td>11</td>
<td>VIC</td>
<td>0.50</td>
<td>CT</td>
<td>131–161</td>
<td>9</td>
<td>0.722</td>
<td>-0.047</td>
</tr>
<tr>
<td>PIE243</td>
<td>GGGTGCAATGAGCGACTTCT</td>
<td>Durand et al. 2010</td>
<td>10</td>
<td>VIC</td>
<td>0.25</td>
<td>AG</td>
<td>208–222</td>
<td>6</td>
<td>0.151</td>
<td>0.677</td>
</tr>
<tr>
<td>PIE239</td>
<td>TCAAACAATGGCTCAACAGTG</td>
<td>Durand et al. 2010</td>
<td>NA</td>
<td>PET</td>
<td>0.63</td>
<td>AT</td>
<td>70–83</td>
<td>11</td>
<td>0.590</td>
<td>-0.082</td>
</tr>
<tr>
<td>PIE227</td>
<td>TACCATGATGCTGGGGAAGCAAC</td>
<td>Durand et al. 2010</td>
<td>NA</td>
<td>PET</td>
<td>0.38</td>
<td>TGG</td>
<td>156–177</td>
<td>5</td>
<td>0.546</td>
<td>-0.064</td>
</tr>
<tr>
<td>PIE271</td>
<td>AACATCACCAACCACCTACCC</td>
<td>Durand et al. 2010</td>
<td>2</td>
<td>PET</td>
<td>0.50</td>
<td>TC</td>
<td>180–197</td>
<td>10</td>
<td>0.759</td>
<td>0.019</td>
</tr>
<tr>
<td>PIE267</td>
<td>TCCAACCATGAGCCATCCCTCC</td>
<td>Durand et al. 2010</td>
<td>3</td>
<td>NED</td>
<td>0.25</td>
<td>AG</td>
<td>80–105</td>
<td>10</td>
<td>0.824</td>
<td>-0.038</td>
</tr>
<tr>
<td>PIE258</td>
<td>TTTCGACTCAAAAAAACAAATGAC</td>
<td>Durand et al. 2010</td>
<td>2</td>
<td>NED</td>
<td>0.75</td>
<td>TC</td>
<td>128–159</td>
<td>19</td>
<td>0.880</td>
<td>0.005</td>
</tr>
<tr>
<td>PIE215</td>
<td>TACGGATATGCGCTGTTGACC</td>
<td>Durand et al. 2010</td>
<td>12</td>
<td>NED</td>
<td>0.30</td>
<td>GAG</td>
<td>188–206</td>
<td>6</td>
<td>0.553</td>
<td>0.036</td>
</tr>
</tbody>
</table>

QrZAG7  | CAACCTGACCTCTGAGTGAAC                     | Kampfer et al. 1998 | 2 | FAM | 0.50 | TC | 115–153 | 19 | 0.874 | -0.015 | 0.025 |
MsQ13   | ACACCACTGACACCACTACATCCCC                | Dow et al. 1995 | 6 | FAM | 0.50 | GA | 191–221 | 16 | 0.785 | 0.055 | 0.052 |
QrZAG112 | TCTTCTCCTCTCTGAGGGCG                      | Kampfer et al. 1998 | 12 | VIC | 0.40 | GA | 85–96 | 12 | 0.579 | -0.005 | 0.128 |
QrZAG20  | CCCAAACCACTGACCTAACCTTGTGCTAGC          | Kampfer et al. 1998 | 1 | VIC | 0.15 | TA | 160–200 | 19 | 0.874 | -0.015 | 0.025 |
QpZAG15  | CGATTTTGATATGACAGCTATG                   | Steinkellner et al. 1997 | 9 | PET | 0.50 | AG | 108–152 | 14 | 0.764 | -0.026 | 0.024 |
†QpZAG110 | GAGCTTGGCTCTCTGACCTCTATTTG               | Steinkellner et al. 1997 | 8 | PET | 0.50 | AG | 206–262 | 16 | 0.765 | 0.009 | 0.024 |
QrZAG96  | CCCAGACATCATATCGATCTGCTCC              | Kampfer et al. 1998 | 10 | NED | 0.15 | TC | 135–194 | 18 | 0.628 | 0.015 | 0.149 |
†QrZAG111 | CTTGGAACATCGAGGATGCTCC            | Kampfer et al. 1998 | 10 | NED | 0.40 | AG | 238–267 | 21 | 0.828 | -0.031 | 0.075 |

NA, not available.
*LG, linkage group (Catherine Bodénes, personal communication), [C]: final concentration in each primer premix (μM), A: allelic richness, H_o: observed heterozygosity.

redesigned.
PCR mix was composed of 3.5 μL of sterile water, 3 μL of Qiagen Multiplex Buffer (2×), 1 μL of primer premix and 2.5 μL of DNA (10 ng/μL). Concentrations for each primer pair in the primer premix are shown in Table 1. The cycling conditions were as follows: an initial step at 95 °C for 15 min; followed by 30 cycles at 94 °C for 30 s, 56 °C for 1 min and 72 °C for 45 s; and a final incubation at 60 °C for 10 min. PCR products were separated on 3% agarose gel stained with GelRED (Biotium, USA), diluted 20 times in pure water and run on ABI-3730 (Applied Biosystems, USA), with LIZ600 as internal lane size standard. Similarity between profiles from simplex and multiplex was also checked.

Diversity analyses and assignment power

Allelic richness (A), observed heterozygosity (H_o), Fis and FST were estimated on 273 adult trees of both species using GenAlEx 6 (Peakall & Smouse 2006). We used simulated data, generated from allele frequencies of purebred individuals with HYBRIDLAB v.1.0 (Nielsen et al. 2006), to test the assignment power of the two multiplexes alone and in combination (Burgarella et al. 2009; Lepais et al. 2009). Allele frequencies for Q. robur and Q. petraea were first estimated on a subset of 88 purebred samples per species (based on their genotype at 20 SSRs), identified with STRUCTURE v.2.3.3 (Pritchard et al. 2000; Falush et al. 2003), with a burn-in of 50 000 steps followed by 50 000 Markov chain Monte Carlo repetitions. We calculated the average result over 10 runs with K (number of groups) set to two, corresponding to the two species, and used a threshold of 0.9 to identify pure individuals from each species. Assignment of simulated genotypes (10 000 purebreds and 10 000 F1 hybrids) relied on the same method, except that we used theoretical intervals of 0–0.25 and 0.75–1 for purebreds and 0.25–0.75 for F1 hybrids (only F1 were generated, not backcrosses, so these thresholds should be optimal to distinguish between parental species and hybrids in the simulations).

Microsatellites scoring (kit-1 only)

Individual genotypes were determined using both Genemapper (Applied Biosystems, USA) and STRand (http://www.vgl.ucdavis.edu/STRand). Alleles were sorted by raw size to detect discrete size variants, with an Excel macro inspired from FLEXIBIN (Amos et al. 2007). The results were used to assign each allele to a bin. We also compared raw sizes between software to test the reproducibility of data obtained with two different algorithms (Advanced Peak Detection Algorithm implemented in Genemapper and Local Southern Algorithm implemented in STRand) on a subset of 490 samples.

Error rate measurement (kit-1 only)

A first error rate was estimated using 80 duplicated samples (6% of the complete dataset) that had been randomly selected, by counting mismatches (Johnson & Haydon 2007). A second error rate, called ‘disagreement rate’ between human readers, was measured on all 490 samples. Incoherencies were classified as follows: Type A is when one genotype is classified as heterozygous for one reader and as homozygous for the other reader, and Type B is when different alleles are selected by both readers. When two different genotypes were obtained for the same sample, we tried to identify a consensus genotype. In a few cases, no consensus genotype could be determined and corresponding data was considered as missing.

Results

Multiplex PCR optimization

Among the 27 preselected SSRs (17 eSSRs and 10 gSSRs), seven were excluded (five with null alleles, one with triple bands and one with low signal once multiplexed). Three primer pairs were redesigned: one locus having a weak allele and two showing overlapping sizes in our first tests. The final profiles obtained for each kit were sharp with homogeneous amplification of the loci (Figs S1 and S2, Supporting Information). Moreover, the analysis of the 35 half-sib families did not reveal a single case of null allele at any of the 20 SSR markers. Four of the 20 SSR markers, all with di-nucleotide repeat (PIE152, PIE239, PIE258 and PIE271), had interfering from the expected periodicity of two base pairs). These alleles were shown to segregate in progenies and are therefore not amplification artefacts. Interestingly, initial analysis with classical automatic-binning mode (implemented in most commercial software and widely used by many researchers) failed to identify these alleles, resulting in incoherencies when checking for Mendelian segregation (data not shown). With binning based on raw allele size, these alleles are easily identified, increasing the total number of alleles for the corresponding markers. These results confirm the necessity to analyse samples using raw sizes and to bin the alleles afterwards.

SSR properties

We found that gSSRs are more polymorphic than eSSRs (mean allelic richness: 16.9 for gSSRs and 10.3 for eSSRs). This difference is partly because of the presence of SSRs with tri-nucleotide repeats in kit-1, as loci with longer
repeats are known to be less variable (Kelkar et al. 2008). The loci that best differentiate Q. robur from Q. petraea are distributed on the two kits (Table 1), with interspecific $F_{ST}$ reaching 0.20 (mean: 0.06, Table 1).

Assignment power

Results of assignment tests on 20 000 simulated genotypes are shown in Fig. 1. The three classes (Q. robur, Q. petraea and F1 hybrids) are well delimited, resulting in low assignment error rates, even though Q. petraea and Q. robur are closely related species. Assignment with all 20 SSRs is much more effective than when using only 8 or 12 loci; the proportion of incorrect assignments is divided by four or five when the two kits are combined, compared to the proportion observed with only one of the two kits (with thresholds of 0.25 and 0.75, see Table 2). Note that the thresholds chosen are considered as optimal. If they had been set to other values, incorrect assignments would have increased for one category (purebreds or F1 hybrids) and decreased for the other one, but the overall error rate would have been increased (Fig. S3, Supporting Information).

SSR transferability

All 20 loci amplified in the other oak species tested (Q. pubescens, Q. pyrenaica, Q. alba, Q. rubra, Q. faginea, Q. suber and Q. ilex). Our first tests on more distant species showed that all 20 SSRs amplified in C. sativa. In

![Fig. 1](image-url) Assignment of 20 000 simulated genotypes (purebred for both parental species and F1 hybrids). (a) kit-1 (12-plex). (b) kit-2 (8-plex). (c) kit-1 + kit-2 (12-plex + 8-plex).
F. sylvatica, three loci from kit-1 (PIE020, PIE152 and PIE271) and four from kit-2 (MsQ13, QpZAG15, QrZAG20 and QrZAG96) failed to amplify with our conditions, even though transferability of gSSRs from kit-1 has been previously validated in simplex (Barreneche et al. 2004). Depending on the species, we noticed highly heterogeneous profiles and amplification was not successful on all samples, perhaps because of low DNA quality or technical difficulties. The Mendelian segregation analysis and further amplification tests on large populations remain necessary before concluding that these markers can be successfully transferred to these species. Still, it appears that eSSRs (kit-1) have a better transferability than gSSRs (kit-2), as found in previous studies on other species (Varshney et al. 2005).

Microsatellites scoring and binning (kit-1)

True allele sizes recovered with Genemapper and STRand were similar (mean deviation: 0.03 bp). However, moderate deviation (>0.1 bp) was observed between sizes measured with each software in 7.8% of genotypes and large deviation (>0.25 bp) was observed in 2.9% of genotypes (maximum deviation: 0.48 bp). These deviations are directly induced by the algorithm used to relate internal size marker and allele sizes. This result indicates that even if raw sizes are used for analysis, problems might still occur when samples from different data sets scored with different methods are integrated (Morin et al. 2009).

Error rate measurement (kit-1)

Disagreement rates between both human readers ranged from 0 to 3.6% across all loci (mean 1.1%). Most differences (78%) were because of calling a heterozygous genotype as homozygous by one of the two readers (type A error). Wrong allele calling (type B error) represented only 22% of incoherencies. Type A errors are easily avoidable as they result most of the time in careless mistakes. Type B errors can be decreased by defining clearer reading rules across readers. While corrections involving only 1% of the samples might seem costly in view of the extra-work involved, it can be critical in studies that are very sensitive to genotyping errors such as parentage analysis (Kalinowski et al. 2007). After establishing consensus genotypes between the two readers, error rates measured by checking the conformity of blindly repeated genotypes ranged from 0% to 1.6%, with a mean of only 0.26% across loci, illustrating the high robustness of markers (Table S1, Supporting Information).

Conclusion

Multiplex PCR allows fast, accurate and cost-effective genotyping but requires significant efforts for its development. Primer validation in simplex is the key step of the overall process. If carried out carefully, subsequent multiplexing becomes much easier. Furthermore, if automatic binning seems to save time, genotyping errors appear to be more frequent. As a consequence, we recommend to analyse samples in raw sizes and to bin the data afterwards, which allows accurate analysis of off-ladder microvariants. We believe that these two highly validated multiplexes will be helpful for future studies on oaks by providing powerful and accurate genotyping tools. In particular, our results confirm the power of microsatellites for hybrid identification. With a larger reference database, assignment rates should be further improved. In combination with additional markers, these two multiplexes should be useful in more complex situations involving more than two species or later-generation hybrids. More generally, this development strategy for medium-throughput genotyping assay (presented here from multiplex PCR development to the definition of allele calling rules) could be efficiently transferred to other species.

Authors’ contributions

EG, LL, SW and PL performed the experiments and produced the data. EG analysed the data and performed the simulations. EG wrote the paper with the help of RP. All authors have checked and approved the final version of the manuscript.

Acknowledgements

We thank Christophe Boury for developing robotic applications used in this project. We thank Alexis Doucouso, François Hubert, Catherine Bodénès, Emilie Chancerel and Jérôme
Durand for their assistance during various steps of the project. We thank Nicolas Langlade and two anonymous reviewers for their suggestions that greatly improved the manuscript. Part of the sampling was performed at the State Forest Nursery of Guémené-Penfao with the assistance of Jean-Pierre Huveline. Genotyping was performed in the Genome-Transcriptome facility of the Functional Genomic Center of Bordeaux with the help of Franck Salin and Sarah Monilier. Experiments were funded by the Research Center of Pernod-Ricard (CRPR) as part of Erwan Guichoux PhD, by the EVOLTREE Network of Excellence supported by the 6th Framework Programme of the European Commission no. 016322 and by the LINKTREE project from the Eranet Biodiversa programme (ANR-08-BDVA-006).

References


© 2011 Blackwell Publishing Ltd


**Supporting Information**

Additional supporting information may be found in the online version of this article.

**Table S1** Disagreement rate measured on 490 samples with *kit-1*. Error rate was measured on 80 samples (6% of the complete dataset).

**Table S2** List of 64 EST-SSRs tested to develop *kit-1*. The 12 selected loci are in red.

**Fig S1.** Multiplex profile with *kit-1*.

**Fig S2.** Multiplex profile with *kit-2*.

**Fig S3.** Incorrect assignments for simulated *Q. robur*, *Q. petraea* and F1 hybrids with different intervals used for hybrid assignment.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.