Chloroplast \textit{trnD-trnT} Region Sequencing for Quick Haplotypeing of Oak Populations

Josef Vlasak*, Helena Cvrckova*, Pavlína Machova* and Jana Malá*

Abstract

Method of DNA isolation from young oak leaves is described that enables efficient amplification of \textit{trnD-trnT} region of chloroplast DNA and direct sequencing of crude PCR reaction mixture. Sequence data of \textit{trnD-trnT} region have been used to examine \textit{Quercus robur} and \textit{Q. petraea} population’s variability. High discrimination power of this method is demonstrated, comparable with previously used restriction analysis of four similar chloroplast DNA regions. Variable positions in \textit{trnT-trnD} chloroplast DNA fragment are mapped and their use for oak haplotyping indicated.

Keywords

CpDNA; Genetic diversity; Phylogeny; Phylogeography; \textit{Quercus} spp

Introduction

Chloroplast DNA (cpDNA) markers have often been used to assess genetic variability among European oak populations [1,2] and to examine their postglacial migration history [3]. Fossil oak pollen maps [4] have shown that in times of major Pleistocene glaciations, European oak species \textit{Quercus robur} L. and \textit{Q. petraea} (Matt.) Liebl were restricted to a few refugial populations in southern Europe. Molecular studies indicated that most of cpDNA differentiation took place in these refugia during prolonged periods of isolation [5]. Because of slow rate of cpDNA mutation, northward expansion of oak at the end of the last ice age retained many refugial cpDNA characteristics in recently established populations. Cp DNA analysis of native European oak populations thus enables to identify the original refugium in Southern Europe as well as probable migration routes, because some changes occurred also during and after migration and these can be exploited for tracing of migration history of local oak populations [3].

In these oak migration studies, thousands of oak DNA samples have been prepared in many laboratories over the whole Europe and a very simple method of cpDNA sequence variation analysis has been adopted. Generally, noncoding cpDNA sequence flanked by strongly conserved regions mostly of rRNA genes has been amplified using primers derived from the conserved regions. The resulting PCR product was digested with restriction enzymes in an effort to find restriction sites typical for some populations. Ferris [5,6,7] have mostly used about 500 bp long intron in \textit{trnL} gene, or short intergeneric region (380 bp) between \textit{trnL} and \textit{trnF} genes [8] exploiting one variable site in both regions. Petit [9] have used about 1700 bp long intergeneric region DT between \textit{trnD} and \textit{trnT} genes where mutations generating new \textit{TaqI} or \textit{AluI} sites and also insertions/deletions in \textit{TaqI} fragments were discovered in some populations, which allowed defining 11 haplotypes of European oaks. Exploiting in a similar way of three other variable regions, \textit{trnT / trnF} (TF, about 1800 bp), \textit{psaA/trnS} (AS, about 3700 bp), and \textit{psbC/trnD} (CD, about 3100 bp) enabled to split the original DT haplotypes in several subtypes so that the number of European oak haplotypes could be increased to 32 [9].

Even if successful at last, this method includes two obvious defects. First is the rather awkward scoring of the presence/absence of restriction fragments after electrophoresis in PAGE gels. Fragments have been sorted as DT1, DT2, DT3 etc. by their rather different sizes, and particular fragment variants caused by small insertions/deletions have been ranked by the order in which the variants move through the gel. Their exact size in base pairs was not determined. This may cause problems in variant identification in the laboratory where not all fragment variants are available. Also, long insertions which are not really present in cpDNA may change the migration rate so much that the shorter fragment (e.g. DT3) mimics longer fragment variant (DT2). Second and major defect is that a lot of sequence information remained undisclosed by this restriction procedure. Our preliminary results with population sequencing of these regions have shown that especially DT region of cpDNA contains several yet undiscovered variable sites which can define many more haplotypes than original 11, and perhaps all or nearly all haplotypes identified by Petit et al. (2002a) [9] using four variable chloroplast regions. In this paper, we describe in detail a quick procedure for sequencing and sequence analysis of cp DNA DT region aimed at disclosing genetic structure of \textit{Q. petraea} and \textit{Q. robur} populations. Preliminary results with some Czech Republic oak populations are also presented.

Materials and Methods

Plant materials

Oak twigs with young leaves were sampled in April and immediately processed or lyophilized. 10 populations of \textit{Q. robur} and 10 of \textit{Q. petraea} were selected for sampling with about 20 randomly selected trees in a minimal distance of 50 m sampled from each population. 100 mg of fresh or 20 mg of lyophilized young leaves were homogenized under liquid nitrogen by mortar and pestle and transferred to 1.5 ml Eppendorf. Total genomic DNA was extracted using DN easy Plant Mini Kit (QIAGEN, Venlo, Netherlands) following the Quick-Start protocol.

PCR amplification and sequencing

PCR of cpDNA DT region was performed using the universal chloroplast primers \textit{trnD} (5´-ACC AAT TGA ACT ACA ATC CC) and \textit{trnT} (5´-CTA CCA CTG AGT TAA AAG GG) of Demesure [10] which match the intergenic region between chloroplast \textit{trnD} and \textit{trnT} genes. For 40 DNA samples, 600 µl of PCR master mix was prepared containing 6 µl (30 U) of Ex Taq DNA polymerase

http://dx.doi.org/10.4172/2229-4473.1000124

Received: June 09, 2016 Accepted: June 20, 2016 Published: June 27, 2016
Direct sequencing of crude PCR products can reduce the cost and time substantially, but it can be performed only after perfect amplification. CpDNA DT region was well amplified in parallel experiments with beech (Fagus sylvatica L.), fir (Abies alba Mill.) and pine (Pinus sylvestris L.), but DNA from older oak leaves performed poorly, giving no or contaminated amplified DNA. Thereafter we developed a method of DNA isolation from very young leaves (see Materials and Methods) which yields amplifiable DNA in most cases. In some preparations, however, contaminating band of 180 bp was observed in the trace that could be corrected, however, after manual checking of the original chromatogram. When processing larger sets of sequences, we took advantage of the fact that the central 600 bp of the DT fragment does not contain any informative sites and this region could be deleted from the procedure of Petit et al. (Petit et al. 2002a)[9] and they probably represent up to now unrecognized splitting of 5a haplotype.

Individual trees in some populations yielded three other very divergent haplotypes (V, VI, VII) which belong to west-European lineage B, not present in the Czech Republic native oak forests [9].

### Discussion

The arguments that sequencing of cp DNA amplified fragments is costly and time-consuming [6] were relevant twenty years ago when sequencing of crude PCR reaction mixture produced in most cases clean chromatograms readable to about 900 bps. Accordingly sequencing from both ends enabled us to read the whole 1750 bp long sequence. Sequencing errors sometimes occurred in the last 200 bp of the trace that could be corrected, however, after manual checking of the original chromatogram. When processing larger sets of sequences, we took advantage of the fact that the central 600 bp of the DT fragment does not contain any informative sites and this region could be deleted from the alignment, losing no informative data. Clean sequence results have been obtained by sequencing, suitable for widely used programs for sequence relationships computing (Figure 2).
the use of molecular methods for characterization of oak populations has been tested. At present, however, the opposite is true. Standard methods of oak molecular haplotyping as RFLP (restriction fragment length polymorphism) or CAPS (cleaved amplified polymorphic sequences) require enzymatic reactions, PAGE gels, difficult scoring of positive bands, which is time-consuming compared to commercial
European oak populations.
to be useful for the assessment of the origin and genetic structure of
DNA restriction analysis used in most studies. The method proved
quicker, less expensive and more informative than chloroplast
amplified chloroplast DNA was developed and tested, which is
3.

References
Agriculture of the Czech Republic no. QJ1230334.

Acknowledgement
This research was supported by the project of the Ministry of Agriculture of the Czech Republic no. QJ1230334.

Conclusions
Method of oak haplotyping based on direct sequencing of amplified chloroplast DNA was developed and tested, which is quicker, less expensive and more informative than chloroplast DNA restriction analysis used in most studies. The method proved to be useful for the assessment of the origin and genetic structure of European oak populations.

Author Affiliations
1) Biology Centre of the Academy of Sciences of the Czech Republic, Brno, Czech Republic
2) Forestry and Game managements Research Institute, Jílovíště-Strnady, Czech Republic

Submit your next manuscript and get advantages of SciTechnol submissions
● 50 Journals
● 21 Day rapid review process
● 1000 Editorial team
● 2 Million readers
● Publication immediately after acceptance
● Quality and quick editorial review processing

Submit your next manuscript at www.sciotechnol.com/submission

Figure 3: Map of the Q. robur DT region of chloroplast DNA (in bp). Constant Taq I sites are shown which define DT1, DT2, DT3 and DT4 regions scored in [3,9]. Occasional Taq I sites are in blue; 85 bp Taq I indicates our haplotype I [3], 1137 bp Taq I our haplotype IV (26 of Petit et al 2002a). Violet bars indicate newly identified variable sites. Of them, mutation 252 bp creates Alu I site typical for our haplotypes VI and VII (24a in [Petit [3,9]). Green bars indicate positions of poly A resp. poly T (dark green) stretches.

sequencing and sequence analysis. The costs are particularly low when we take advantage of the optimized method described above that enables to sequence unpurified PCR products.

DT region of cpDNA of oaks is optimal for sequence-based haplotyping because of proper size, good amplification, relatively high variability, and relevance to standard restriction haplotyping that has been also mostly based also on this region. 8 variable positions of which 5 were identified in this paper enable to define tens of oak haplotypes and have the same discrimination power as previously used restriction analysis of four cp DNA fragments.

20 Czech Republic oak populations were sampled and 8 different haplotypes identified. Of six cp DNA haplotypes previously reported (Petit et al., 2002a) two (2 and 15) could not be found and four were newly discovered. The same haplotypes were often found in both Quercus robur and Q. petraea has already stated by (Petit et al., 2002a), which could be explained by occasional hybridization between these species. Preliminary phylogenetic sequence analysis (not shown) clustered haplotypes in four lineages A, B, C, E (Table 1) defined by Petit et al. (2002a). Trees of B lineage do not make populations in the Czech Republic and represent very probably man-introduced individuals.