Vlasak et al., Vegetos 2016, 29:2
http://dx.doi.org/10.4172/2229-4473.1000124

Vegetos- An International Journal of Plant Research

Research Article A SCITECHNOL JOURNAL

Chloroplast *trnD-trnT* Region Sequencing for Quick Haplotyping of Oak Populations

Josef Vlasak[1*], Helena Cvrckova[2], Pavlína Machova[2] and Jana Mala[2]

Abstract

Method of DNA isolation from young oak leaves is described that enables efficient amplification of *trnD-trnT* region of chloroplast DNA and direct sequencing of crude PCR reaction mixture. Sequence data of *trnD-trnT* region have been used to examine *Quercus robur* and *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 *trnT -trnD* chloroplast DNA fragment are mapped and their use for oak haplotyping indicated.

Keywords

CpDNA; Genetic diversity; Phylogeny; Phylogeography; *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 *Quercus robur* L. and *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 tRNA 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 *trnL* gene, or short intergeneric region (380 bp) between *trnL* and *trnF* genes [8] exploiting one variable site in both regions. Petit [9] have used about 1700 bp long intergeneric region DT between *trnD* and *trnT* genes where mutations generating new *TaqI* or *AluI* sites and also insertions/deletions in *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, *trnT* / *trnF* (TF, about 1800 bp), psaA/trnS (AS, about 3700 bp), and *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 rare 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 *Q. petraea* and *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 *Q. robur* and 10 of *Q. petrea* 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 *trnD* (5´-ACC AAT TGA ACT ACA ATC CC) and *trnT* (5´-CTA CCA CTG AGT TAA AAG GG) of Demesure [10] which match the intergeneric region between chloroplast *trnD* and *trnT* genes. For 40 DNA samples, 600 µl of PCR master mix was prepared containing 6 µl (30 U) of Ex Taq DNA polymerase

***Corresponding author:** Josef Vlasak, Biology Centre of the Academy of Sciences of the Czech Republic, Branišovská 31, 370 05 České Budějovice, Czech Republic, Tel: +420 387 775 514; Fax +420 385 310 356; E-mail vlasak@umbr.cas.cz

Received: June 09, 2016 **Accepted:** June 20, 2016 **Published:** June 27, 2016

TAKARA, Ex*Taq* buffer TAKARA, 25 µM of each deoxynucleotide triphosphate and 0.12 µM of each primer. 15 µl f master mix was added to ~ 5 ng (0.5 µl) of DNA in 0.2 ml PCR tube and amplified in BIOMETRA T3 thermal cycler using initial denaturation for 3 min. at 94oC followed by 35 cycles of denaturation (5 s at 94°C), annealing (15 s at 57°C) and extension (1 min. at 72°C) with a final extension 5 min. at 72°C. Amplified DNAs were visualized by electrophoresis of 4 µl of PCR product on a horizontal 1.5% agarose gel with 100 bp DNA ladder New England Biolabs. To 5 µl of unpurified PCR product, 6 µl of either *trnD* or *trnT* primer (4 µM) was added and the mixture was commercially sequenced by GATC Biotech AG, Germany. Sequences representing important haplotypes were submitted to GenBank (see Table 1).

DNA sequence analysis

The sequences were aligned with Clustal X and manually pruned. After removing unalignable regions, the dataset contained 1671 characters of which 19 were variable and 17 parsimony informative.

Results

DNA isolation, amplification and sequencing

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 b appeared after PCR, which disturbs sequencing of the 5′ and 3′ ends of PCR-DT product. This band could be eliminated by using very low primer concentration in the PCR master mix (see Materials and Methods). Only in rare cases of recalcitrant samples we have used gel isolation of positive bands from agarose gels (Figure 1).

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).

Variable positions in CpDNA DT region

Sequencing of about 400 individuals from 20 Czech Republic populations of *Q. robur* and *Q. petraea* revealed 5 newly recognized positions where point mutation occur in some populations and also four regions of 4–11 bp long, variable polyA resp. polyT stretches (Figure 3). The homobasic stretches add, however, no additional differentiation as the number of repeating bases corresponds to particular haplotype in all individuals. Similarly, 5 bp deletion at 1481 is always associated with haplotypes I and VIII and 19 bé insertion after 1010 (not shown in Figure 3) with haplotype V. In rare cases, other deletions and/or insertions (up to 209 bp long!) occur in some individuals, but never in the whole population and these stretches were removed from the final alignment.

These newly fixed variable sites together with those already known (Table 1) enable to define many tens of haplotypes, thus providing enough data for identification of oak populations.

Preliminary results obtained with Czech Republic oak populations

Petit have defined six main lineages A, B, C, D, E, F of European oaks from the haplotyping of 2600 European oak populations using restriction analysis of four cp DNA regions. More detailed analyses have shown that only 6 haplotypes have been determined in our region [3] 1 and 2 (lineage C), 5 and 26 (A), and perhaps 15 and 17 (E) around the Czech Republic southern border. This we could confirm by sequencing of a DT region only. Nearly all sequenced oak populations belonged to our haplotypes I, II and IV that could be identified as haplotypes 1, 5a and 26 in (Petit et al., 2002a, 2002b) (Table 1). One population of *Q. petraea* from the border with Austria contained haplotype 17 (=VIII). The match between published and our haplotypes was supported by additional sequencing of TF region, which is of similar size (see Table 1, not shown in detail). TF region sequencing confirmed the identity of haplotypes II resp. IV with 5a resp. 26 (Petit et al. 2002a) because of presence of *CfoI* restriction site.

Clearly different haplotypes II and III could not be separated by 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

Table 1: *Oak* DT haplotypes and their variable positions. Unique mutations identifying respective haplotypes are **bold** underlined. TF* column refers to additional sequencing of TF region, where the presence of *Cfo I* site is indicated by 9 and absence by 1, following notation in (Petit [3]).

Haplotype		*Lineage*	85 bp	98 bp	149 bp	252 bp	363 bp	1137 bp	1543 bp	1544 bp	TF* CfoI	GenBank
Our	Petit											
I	1	***C***	**C**	C	A	T	T	C	T	G	1	KP256524
II	5a	***A***	A	C	C	T	T	C	**A**	G	9	KP256523
III	5a	***A***	A	C	A	T	T	C	T	G	9	KP256525
IV	26	***A***	A	C	C	T	T	**G**	T	T	9	KP256526
V	?	***B***	A	C	A	T	G	C	T	G	1	KP256528
VI	24a	***B***	A	T	A	**C**	G	C	T	T	1	KP256529
VII	24a	***B***	A	T	A	**C**	G	C	T	G	1	
VIII	17	***E***	A	C	C	T	T	C	T	G	1	

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

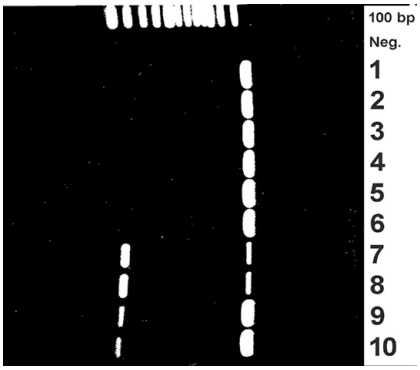

Figure 1: Typical elfo of Q. *robur* amplified chloroplast DT region.1-6: samples from individual trees yielding good amplified DNA for direct sequencing,7, 8: unsuccessful samples, 9,10: samples requiring gel-isolation of positive bands, Neg.: negative control.

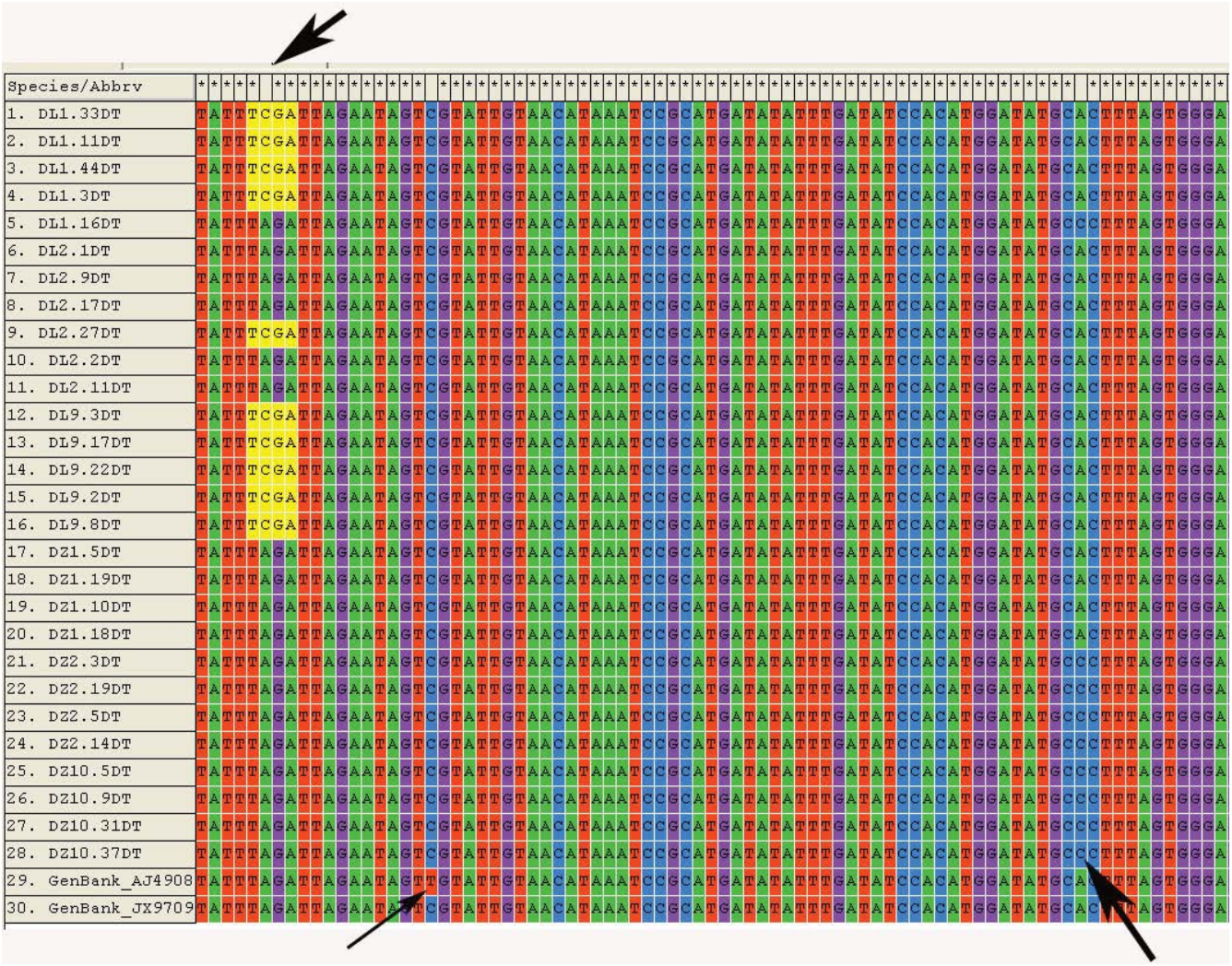

Figure 2: Alignment of cp DNA *trnD/trnT* sequences of several oaks from 5 Czech oak populations (example).Variable position 85bp with occasional TaqI site, and newly discovered variable positions 98 bp and 149 bp are marked by arrows.

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

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

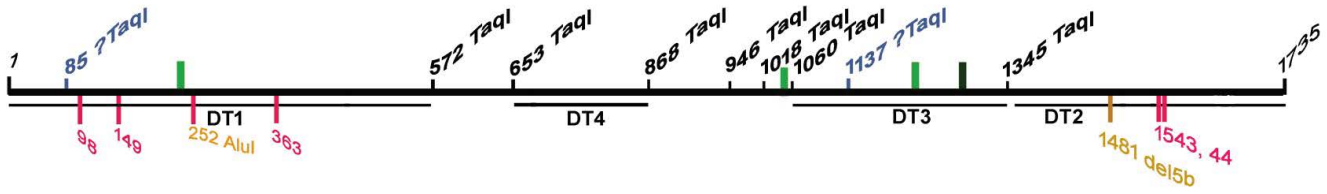

Figure 3: Map of the *Q. robur* DT region of chloroplast DNA (in bp). Constant *Taq I* sites are shown which define DT1, DT2, DT3 and DT4 regions scored in [3,9]. Occasional *Taq I* sites are in blue; 85 bp Taq I indicates our haplotype I [3], 1137 bp *Taq I* our haplotype IV (26 of Petit et al 2002a). 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 Pettit et al. (2002a). Trees of B lineage do not make populations in the Czech Republic and represent very probably man-introduced individuals.

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.

Acknowledgement

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

References

1. Kremer A, Petit R, Zanetto A, et al. (1991) Nuclear and organelle gene diversity in quercus robur and Q petraea. Genetic Variation of Forest Tree Populations in Europe, Sauerlanders Verlag, Frankfurt, Germany.
2. Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic structure of white oaks throughout the european continent. Genetics 146: 1475-1487.
3. Petit R J, Brewer S, Bordacs S, Burg K, Cheddadi R, et al. (2002) Identification of refugia and post-glacial colonisation routes of european white oaks based on chloroplast dna and fossil pollen evidence. Forest Ecol Manag 156: 49-74.
4. Huntley B, Birks HJB (1983) An atlas of past and present pollen maps for europe: 0-13000 years ago. Cambridge University Press, Cambridge, United Kingdom.
5. Ferris C, Oliver RP, Davy AJ, Hewitt GM (1993) Native oak chloroplasts reveal an ancient divide across europe. Mol Ecol 2: 337-344.
6. Ferris C, Oliver RP, Davy AJ, Hewitt GM (1995) Using chloroplast DNA to trace postglacial migration routes of oaks into britain. Mol Ecol 4: 731-738.
7. Ferris C, King RA, Väinölä R, Hewitt GM (1998) Chloroplast DNA recognizes three refugial sources of european oaks and suggests independent eastern and western immigrations to finland. Heredity (Edinb) 80 : 584-593.
8. Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17: 1105-1109.
9. Petit RJ, Brewer S, Bordacs S, Burg K, Cheddadi R, et al.(2002) Chloroplast DNA variation in european white oaks. Phylogeography and patterns of diversity based on data from over 2600 populations. Forest Ecol Manag 156: 5-26.
10. Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4: 129-131.

Author Affiliations

Top

[1]*Biology Centre of the Academy of Sciences of the Czech Republic, Branišovská, Czech Republic*

[2]*Forestry and Game managements Research Institute, Jíloviště-Strnady, Czech Republic*