

RESEARCH ARTICLE

Development of Molecular Markers for Determining Continental Origin of Wood from White Oaks (*Quercus* L. sect. *Quercus*)

Hilke Schroeder^{1*}, Richard Cronn², Yulai Yanbaev³, Tara Jennings⁴, Malte Mader¹, Bernd Degen¹, Birgit Kersten¹

1 Thuenen-Institute of Forest Genetics, Grossshansdorf, Germany, **2** US Forest Service, Pacific Northwest Research Station, Corvallis, Oregon, United States of America, **3** Bashkir State University, Ufa, Russia, **4** Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, United States of America

* hilke.schroeder@thuenen.de



OPEN ACCESS

Citation: Schroeder H, Cronn R, Yanbaev Y, Jennings T, Mader M, Degen B, et al. (2016) Development of Molecular Markers for Determining Continental Origin of Wood from White Oaks (*Quercus* L. sect. *Quercus*). PLoS ONE 11(6): e0158221. doi:10.1371/journal.pone.0158221

Editor: Dusan Gomory, Technical University in Zvolen, SLOVAKIA

Received: February 11, 2016

Accepted: June 13, 2016

Published: June 28, 2016

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: DNA sequences from Sanger sequencing of the indel containing fragments derived from this study have been deposited in GenBank (Accession numbers: KU201020-KU201034). Next-generation sequences supporting this study are available from the NCBI GenBank as BioProject PRJNA269970 (SRA Accessions SRS954648, SRS954649, and SRS954650, SRX1050444-1050539). SRA accession numbers of the next-generation sequencing data of the two DNA-pools are SRR3624658 (*Q. robur*) and SRR3624680 (*Q. mongolica*). Draft chloroplast genome references

Abstract

To detect and avoid illegal logging of valuable tree species, identification methods for the origin of timber are necessary. We used next-generation sequencing to identify chloroplast genome regions that differentiate the origin of white oaks from the three continents; Asia, Europe, and North America. By using the chloroplast genome of Asian *Q. mongolica* as a reference, we identified 861 variant sites (672 single nucleotide polymorphisms (SNPs); 189 insertion/deletion (indel) polymorphism) from representative species of three continents (*Q. mongolica* from Asia; *Q. petraea* and *Q. robur* from Europe; *Q. alba* from North America), and we identified additional chloroplast polymorphisms in pools of 20 individuals each from *Q. mongolica* (789 variant sites) and *Q. robur* (346 variant sites). Genome sequences were screened for indels to develop markers that identify continental origin of oak species, and that can be easily evaluated using a variety of detection methods. We identified five indels and one SNP that reliably identify continent-of-origin, based on evaluations of up to 1078 individuals representing 13 white oak species and three continents. Due to the size of length polymorphisms revealed, this marker set can be visualized using capillary electrophoresis or high resolution gel (acrylamide or agarose) electrophoresis. With these markers, we provide the wood trading market with an instrument to comply with the U.S. and European laws that require timber companies to avoid the trade of illegally harvested timber.

Introduction

Illegal logging is a serious issue not only for tropical rainforests and tropical trees, but it is also a concern for tree species in temperate latitude forests. White oaks from the genus *Quercus* sect. *Quercus* (Fagaceae) provide a relevant example of illegal logging in a temperate zone tree, and they highlight the challenge facing importers and regulatory agencies responsible for validating the taxonomic and geographic sources of timber products. White oaks account for a significant percentage of the hardwood flooring and furniture trade in Europe and the USA, and

will be submitted to Dryad soon. The open source software Variant Tools is available on <https://github.com/ThuenerFG/varianttools>.

Funding: This project has been financially supported by German Federal Ministry of Food and Agriculture (BMEL) (<http://www.bmel.de>), the Deutsche Bundesstiftung Umwelt (DBU) (<http://www.dbu.de>), the USDA Forest Service International Programs Office (<http://www.fs.fed.us/about-agency/international-programs>), the US Agency for International Development, and US Department of State (<http://www.state.gov/s/d/rm/rls/dosstrat/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

they represent one of the most important hardwoods in terms of logs and lumber exports from these regions. The most important trade woods of white oaks derive from the European species *Quercus robur* L. and *Q. petraea* (Mattuschka) Liebl., the CITES Appendix III-protected *Q. mongolica* Fisch. Ex Ledeb. native to East-Asia, and North American oaks, such as *Q. alba* L. and *Q. macrocarpa* Michx. [1]. Non-governmental organizations such as the Environmental Investigation Agency (<http://eia-global.org/news-media/liquidating-the-forests>) have documented increases in the rate of illegal logging for white oak wood, especially in the Russian Far East region. These activities increase the likelihood that international wood trading companies will market illegally harvested wood, an activity that is banned by the U.S. Lacey Act amendment of 2008 and the European Union timber regulation of 2010. Violation of these regulations can result in fines, forfeiture of wood, and additional payments, as was recently demonstrated with improperly documented shipments of white oak flooring in the United States [2]. Under these laws, timber companies are responsible for avoiding the trade of illegally harvested timber, and they are obligated to declare the species name and geographic origin of traded timber in order to reduce the risk that traded timber originated from illegal logging [3].

The increased attention to illegal logging has led to an increased demand for methods that can be used to provide precise species identification and geographic origin verification. Wood anatomical methods are widely used for tree species identification [3], but these methods cannot discriminate white oak species, nor identify geographic origin of oaks generally. Over the last decade, worldwide programs have been established using the potential of DNA as universal tool for identifying organisms (Barcode of Life www.barcodeoflife.org, [4]). In plants, the success of barcoding is highly dependent on several factors, including magnitude of primary divergence, frequency of secondary contact, and mutation rate of the DNA region [5], so the choice of suitable barcode regions in plants can be difficult [6–8]. Barcoding efforts in plants have focused on chloroplast genomes due their simple pattern of (typically) uniparental inheritance, low effective population size, and useful variation at the scale of geography and taxonomy across a wide range of species (e.g. [9–13]). With advances in next-generation sequencing, chloroplast genomes are affordable to sequence in their entirety by ‘skimming’ methods [14], and whole genome analysis can reveal substantial variation, even in unexpected genomic regions that are included in traditional barcoding efforts (e.g. [15–16]). DNA barcoding already has proven to be appropriate for revealing illegal trading (e.g. [17,18]), and it is increasingly used to identify plant species in commercial trade [19].

The aim of this study is to use chloroplast-genome scale information to develop a cost-efficient, easy-to-use assay that allows the identification of the geographic origin of white oak wood products to hemisphere (Old World vs. New World) and continent (Asia; Europe; North America) to support regulatory and commercial efforts to detect illegal logging of *Q. mongolica*.

Material and Methods

Plant material

For next-generation sequencing, we sequenced a single oak individual from four species and three continents to produce chloroplast genome references; included are *Q. mongolica* from Asia (sample QUMO5_CH_1; China), *Q. petraea* from Europe (sample QUPE2_PO_1; Poland), *Q. robur* from Europe (sample QURO2_SVT6; Germany), and *Q. alba* from North America (sample QUAL_VT_1; USA). To develop a panel of polymorphisms for Asia and Europe, we used next-generation sequencing to screen two pooled DNA samples that included 20 individual specimens of *Q. mongolica* or European *Q. robur*, respectively. Each of the *Q. robur*/*Q. mongolica* specimens were sampled from 10 geographically-widespread populations.

To develop a panel of polymorphisms for North American white oaks, we sequenced chloroplast genomes from additional specimens representing the following species: *Q. alba*, *Q. bicolor* Willd., *Q. garryana* Douglas ex. Hook., *Q. lyrata* Walter, *Q. macrocarpa* Michx., *Q. michauxii* Nutt., *Q. prinoides* Willd., and *Q. stellata* Wangenh.

For marker validation, DNAs from 13 *Quercus* species were screened. *Q. mongolica* and *Q. dentata* Thunb. represented Asian oaks (Far East Russia) with 200 and 10 specimens, respectively. *Q. robur*, *Q. petraea* and *Q. pubescens* Willd. represented European oaks, with 360, 210, and 200 specimens, respectively. Finally, eight white oak species were screened from North America (*Q. alba*, *Q. bicolor*, *Q. garryana*, *Q. lyrata*, *Q. macrocarpa*, *Q. michauxii*, *Q. prinoides*, *Q. stellata*), with between 5 and 25 specimens per species.

None of the oak specimens used in this study has been sampled in protected areas that require permission of any authority. Some of the samples were done on private land and all owners of the lands gave permission for sampling. The field studies did not involve endangered or protected species.

Next-generation sequencing analyses

Aliquots of DNA (~0.5–1 µg) were sheared to a median length of ~300 bp and converted into sequencing libraries using Illumina TruSeq v.2 kits at the USDA Forest Service (Corvallis, OR; individual samples, each indexed with dual-index adapters) or GATC Biotech AG (Konstanz, Germany; pooled samples). Sequencing was performed using three approaches: (A) using the Illumina MiSeq with 2x150 bp paired-end reads for individual *de novo* genome reference assemblies; (B) using the Illumina MiSeq with 2x300 bp paired-end reads for pooled samples and reference-guided mapping; and (C) using the Illumina HiSeq with 100 bp single-end reads for individual North American species and reference-guided mapping. All reactions used version 3 sequencing chemistry. Information on raw clusters, sequence yield, and approximate target sequence (chloroplast genome) coverage depth is provided in Table 1.

De novo reference construction. Raw read quality filtering was accomplished using Trimmomatic v0.30 [20], and we removed reads with a mean Phred score less than 33. Reads were digitally normalized to a coverage of 20 using khmer v0.7.1 [21], and kmer-filtered sequences were assembled using Velvet Optimizer v2.2.5 (<https://github.com/Victorian-Bioinformatics-Consortium/VelvetOptimiser.git>) and Velvet v1.2.10 [22]. K-mer lengths ranging from 21 to 121 were evaluated, with a final k-mer length of 121 selected for assembly. *De novo* contigs ≥ 100 bp in length were screened for homology to the *Quercus rubra* chloroplast

Table 1. Results of next-generation sequencing for oak reference assembly and polymorphism screening. Indexed individuals of oaks were sequenced using 150 bp paired-end reads and evaluated using *de novo* assembly (reference assembly). Pooled individuals were sequenced using 300 bp paired-end reads and evaluated using reference-guided assembly (polymorphism screen).

	Reference Assembly				Polymorphism Screen	
	QUM05	QUPE2	QURO2 ¹	QUAL	QUMON Pool	QUROB Pool
Sequences	6,183,339	5,281,517	6,676,925	4,280,049	17,370,324	28,438,226
Bases sequenced (Gbp)	1.747	1.502	0.977	1.207	10.42	17.06
Proportion of reads mapping to cp genome	3.3	1.9	7.6	0.4	0.95	2.25
<i>De novo</i> contigs (length sum, Mbp)	2,085	872	80	7,085	–	–
Longest <i>de novo</i> contig (bp)	135,603	60,701	27,043	57,523	–	–
Chloroplast contigs	4	12	19	20	–	–
Variants relative to QUM05	0	126	124	572	789	346

¹ *Q. robur* was sequenced using 100 bp single-end reads.

genome (NCBI NC_020152) using BLAT [23]. Contigs showing high similarity were retained for reference assembly and ordered against the *Q. rubra* chloroplast genome. Reference sequences were constructed to include the large single-copy (LSC) region, one of two inverted repeats (IR), and the small single-copy (SSC) region, equivalent to positions 1–135,502 from *Q. rubra* NC_020152.

Reference-guided identification of SNP and indel variation. Reference guided read mapping and polymorphism detection was performed using CLC Genomics Workbench version 7.5.1 (CLC-bio, a Qiagen company; Aarhus, Denmark). The reference chloroplast sequence of the *Q. mongolica* individual QUMO5_CH_1 generated by *de novo* reference construction (see above) was used as reference for read mapping. The trimmed Illumina data of the two pools (*Q. mongolica*, *Q. robur*) and the trimmed HiSeq data of representative North American individuals were mapped to the reference scaffold using a length fraction of 0.9 and a similarity fraction of 0.94. Variants detected by CLC Genomics Workbench included SNPs and small indels, and these were exported to tab-delimited files and processed using an in-house script (*Variant Tools*, see below) to identify species-specific polymorphism. Because the aim of this study was to develop markers that differentiate between the continents, the data set was reduced to these variants appearing with a frequency between 95 and 100%. For marker development, we focused in indels due to their simplicity of analysis and their easier handling using DNA extracted from timber.

Maximum likelihood (ML) analysis was performed so that variants could be understood within the phylogenetic context of New and Old World oak chloroplast genome evolution. We performed ML analyses using RAxML-HPC2 version 8.2.8 through the Cipres Science Gateway (<http://www.phylo.org/>) using the GTRGAMMA substitution model for determining the final best ML tree, and the GTRCAT model to conduct 1000 rapid bootstrap replicates [24]. For this analysis, *de novo* contigs were aligned to the *Q. rubra* chloroplast genome (NC_020152), *Q. rubra* was specified as the gougroup, and gaps were treated as missing information.

Post processing of identified SNPs and indels

To merge the SNP and indel tables and find common variants present in two or more individuals/pools, we developed *Variant Tools*, a command line program implemented in Ruby. This program merges individual sample SNP and indel tables (CSV format) produced by CLC GWB to create a multi-individual SNP and indel matrix. Required input options include the reference sequence (fasta format), an input directory containing the variant CSV tables, and an option specifying the input data type (SNP or indel). The reference fasta can contain one reference sequence or multiple reference contigs. Optionally, coverage tables (produced by CLC from read mappings to a reference) of every individual can be included in the analysis by specifying a directory containing coverage files (CSV format). Furthermore several filtering options are available to reduce the output according to user-provided thresholds. The output from *Variant Tools* is stored in a CSV file and contains several data columns: the reference sequence name, the reference position, the variant length, the calling type (SNP, MNP, deletion or insertion), the reference base(s), the alternative base(s) for every individual, the coverage for every individual at the reference position, different summary statistics, and sequences flanking the called variant.

The flanking sequences are calculated based on two given distance thresholds. An upper and a lower threshold define minimum distances in base pairs between two called variants on the genomic scale. If a variant occurs within the genomic range of the lower threshold no flanking sequence is created. If a variant resides within the genomic range of the upper threshold the length of the flanking sequence created is defined by the lower threshold. If no variant is found

within the range of both thresholds a flanking sequence with the length of the upper threshold is calculated. The default thresholds are 75 bp and 50 bp.

The *Variant Tools* create different summary statistics while the variant matrix is generated. *The number of individual alleles deviating from the reference* is a count for all found variants in all individuals at a specific genomic position. *The number of alleles matching the reference with minimal coverage* is a count for all positions in all individuals where no variant has been called and that are supported by a minimum coverage. The threshold for the minimum coverage is specified by the user. The default threshold is set to a minimum coverage of 8. *Critical forward reverse balance* is an indicator for systematic sequencing errors and describes how many forward and reverse reads are supporting the called variant. The value is averaged over all individuals showing the variant.

The *Variant Tools* are open source software under ongoing development. They are available under the terms of the ICS license and can be obtained from <https://github.com/ThuenenFG/varianttools>.

DNA extraction, PCR, restriction, and genotyping

Leaves. One cm² of a single leaf was ground to powder in liquid nitrogen. Total DNA was extracted, following a modified ATMAB protocol by [25]. PCR reactions for leaf-derived DNA contained ~30 ng template DNA, 10x PCR buffer, 1.5 or 1.75 mM MgCl₂, 200 μM dNTPs, 0.4 unit AmpliTaq Gold DNA polymerase (ThermoFisher Scientific, Darmstadt, Germany), and 0.05 to 0.13 μM of each primer in a total volume of 15 μl. PCR was carried out in a Sensoquest Thermocycler (Göttingen, Germany) with a pre-denaturation step at 94°C for 10 min, followed by 25 to 30 cycles of 94°C for 45 sec (30 sec *trnCD*), suitable annealing temperature for each primer combination (between 52°C and 57°C) (details of primer conditions are given in Table 2) for 45 sec (30 sec *trnCD*), 72°C for 45 sec (1 min *trnLF*) and a final elongation at 72°C for 10 min. PCR amplification products were checked relative to a 100 bp ladder (Life Technologies, Martinsried, Germany) on a 1% agarose gel stained with Roti-Safe GelStain (Carl Roth GmbH & Co. KG, Karlsruhe, Germany); afterwards, PCR products were run on an ABI3730 capillary sequencer. Fragment analysis was performed using GeneMarker™ software v. 2.4.0 (Softgenetics, State College, PA, USA).

Timber. For genotyping analysis of timber-derived DNA, mostly a special DNA extraction protocol has been developed and patented [26] based on the CTAB method. Exceptionally the innuPREP Plant DNA Kit from Analytik Jena (Germany) was used. Due to the small amount

Table 2. List of primers for the amplification and resequencing of the newly developed markers. Fluorescent-labeling of the primers is given in column “sequences”: FAM = blue, VIC = green, PET = red. In the last column, the accession numbers of the related markers for the three species *Q. robur*, *Q. mongolica* and *Q. alba* are given. “Length” means sequence length.

Marker-name	Primer	Sequence 5'-3'	length (bp)	Annea-ling (°C)	Acces. no.
<i>psal-ycf4</i>	psal	CGTGTGTAACATGATATATGAG_FAM	174-	55	KU2010
	ycf4	GAGTAATTCATCGAATTGGTTAG	178		20–22
<i>psbE-petL</i>	psbE	AAGGAATTGGTTAGTTGTCCAG_VIC	179-	55	KU2010
	petL	TTACATATCTTAAATTAGAGAGCC	185		23–25
<i>trnLF</i>	trnL1	CAATACATATCATTCTTGTACTG_PET	130-	53	KU2010
	trnF2	TAGATAACTTGAGTTTATGTCAATT	135		26–28
<i>trnCD</i>	trnC5	TTGGATAGACGAACGGGAAT_FAM	115-	57	KU2010
	trnD5	TATCATATTTAAATTGATTGCCGG	123		29–31
<i>trnDT</i>	trnD3	GGATAGGGATCAACAAGTTATTG_PET	187-	52	KU2010
	trnT4	CAAGACCGACCCTAATTGAAT_PET	189		32–34

doi:10.1371/journal.pone.0158221.t002

of total DNA extracted from timber, the DNA quantity wasn't measured; rather, a standard dilution of 1:10 (DNA:water) was used for all PCR reactions. PCR conditions were similar to those used for leaf material, but with slight modifications (described in [Results](#)).

For one marker (*trnDT*), a restriction enzyme digest was used to reveal a SNP polymorphism in the amplified fragment. The restriction digestion reaction contained 10 μ l PCR product, 2 μ l 10x CutSmart[®] buffer, 0.5 μ l enzyme (FastDigest[®] *Hinf*I, New England Biolabs, Ipswich, MA) in a final volume of 20 μ l. The reaction lasted 15 min at 37°C followed by an inactivation at 80°C for 20 min. Restriction products were either visualized relative to a 50 bp ladder (Life Technologies, Germany, Martinsried) using an 8% polyacrylamide gel stained with ethidium bromide, or using an ABI3730 capillary sequencer.

Probabilities for fixation of gene markers

The Thünen-Institute of Forest Genetics possesses a large collection of reference samples that contains oak species from Europe (*Quercus robur*, *Q. petraea*), North America (*Q. alba*, *Q. macrocarpa*, and others) and Asia (*Q. mongolica*, *Q. dentata*). Based on the numbers in this collection of white oaks from different continents, we computed the maximal potential frequency of variants that were not observed using 95% confidence intervals [27]. This can be described as a method to determine the risk (potential error rate) that a genetic variant (allele/haplotype) assumed to be exclusive to one continent is found in one or more individuals originating from another continent. The calculations were carried out using the online forma at <http://vassarstats.net/prop1.html> based on 962 European, 325 Asian and 61 American white oak individuals for the gene markers *psaI-ycf4*, *psbE-petL*, *trnLF*, and *trnCD*. For the gene marker *trnDT* the sample sizes were 115 European, 425 Asian and 19 American white oak individuals.

Results

Next-generation sequencing, reference genome assembly, and identification of cpDNA length variants in white oaks

Next-generation sequencing of four indexed oak exemplars (*Q. alba*, *Q. mongolica*, *Q. petraea*, *Q. robur*) yielded between nearly 1 and 1.75 Gbp per individual ([Table 1](#)). *De novo* contig assembly with Velvet produced between 872 and 7,085 contigs, and the longest contigs from each assembly were from the chloroplast genome. Our best *de novo* chloroplast assembly derived from *Q. mongolica* QUMO5_CHI_1, which was represented by a single large contig totalling 135.6 kb, and it spanned the three main chloroplast genome regions (large single copy, inverted repeat, small single copy) in their entirety. Alignment of these contigs against the published *Q. rubra* chloroplast genome yielded an alignment of 135,603 nucleotides (alignment excludes one inverted repeat).

Maximum likelihood analysis of these selected white oaks yielded a topology similar to the topology previously established with chloroplast DNA restriction site analysis [28] ([Fig 1](#)), with the Old World white oaks *Q. mongolica*, *Q. petraea* and *Q. robur* resolving as a sister group to the New World *Q. alba* ([Fig 1](#)). The best maximum likelihood tree resolved *Q. mongolica* and *Q. petraea* as sister, but bootstrap support for this resolution was low (40%) and supported by a small number of characters. Across white oak chloroplast genomes, we identified 672 single nucleotide variants and 189 indels, the vast majority of which discriminate white oaks from the outgroup *Q. rubra*, and New World from Old World oaks ([Fig 1](#)).

For production of a best maximum likelihood tree, we used all these variants ([Fig 1](#)) but indicated in [Fig 1](#) only the indels and one SNP that were selected for a marker set due to our

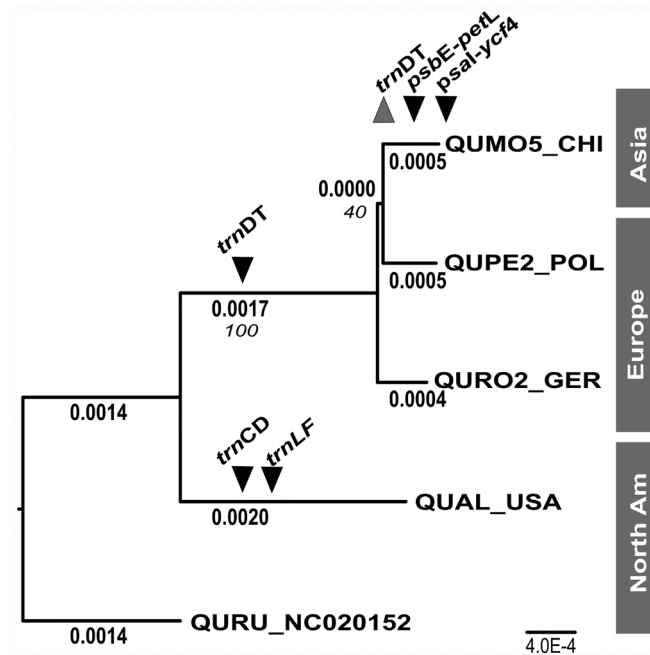


Fig 1. Phylogenetic relationship among chloroplast genomes of white oak species representing Old World and New World lineages. The best maximum likelihood tree is shown for four white oak chloroplast genomes (*Q. mongolica*; *Q. robur*; *Q. petraea*; *Q. alba*) and one outgroup genome (*Q. rubra*). Inferred branch lengths in maximum likelihood substitutions are shown in bold, and bootstrap support values are shown in italics. The phylogenetic resolution of informative indel markers are shown in black inverted triangles, and the resolution of the diagnostic PCR-RFLP marker is shown as a grey triangle.

doi:10.1371/journal.pone.0158221.g001

earlier described aim of this study. Nevertheless, the whole dataset is publically available via NCBI (Data Accessibility).

Next-generation sequencing of the *Q. mongolica* and *Q. robur* DNA pool produced 17.4 and 28.4 million paired-end 300 bp sequences (10.4 and 17.1 Gbp of sequence data respectively; Table 1). Mapping of paired-end reads from the *Q. mongolica* and *Q. robur* pools against the *Q. mongolica* QUMO5_CHI_1 chloroplast genome reference revealed 346 variant positions for the *Q. robur* pool, and 789 variant positions for the *Q. mongolica* pool (Table 1). After read mapping, the two variant tables were compared to filter those variants that showed fixed differences between the continents. The next step was to reduce the dataset to these variants appearing with a frequency between 95 and 100%. This analysis left five indels and 15 SNPs. For marker development, we focused in indels. Checking of these indels within the mapping revealed a (T)_N microsatellite, two indels with a difference in the fragment length of two bp, and one indel with one bp difference. These were removed from the further analyses and only the two longest indels in two spacer regions (*psaI-ycf4*, *psbE-petL*), one with four and one with six bp difference, remained.

Short read sequences from eight North American species were also mapped to the QUMO5 reference, and we specifically searched for indels differentiating North American species from the reference QUMO5 with a frequency of 95 to 100%. We identified three indels that consistently differentiated North America from Asia. These length polymorphisms were found in three spacer regions (*trnLF*, *trnCD*, and *trnDT*) and they ranged in length from 2 bp to 8 bp.

Primer design, marker validation and resequencing

For the five indel-including cpDNA regions, primers were designed using the reference QUMO5 to amplify fragments ranging from 110 bp to 190 bp. A preliminary validation performed with three individuals each of *Q. robur*, *Q. petraea*, *Q. mongolica*, *Q. alba*, and *Q. macrocarpa* revealed that all five cpDNA regions could successfully be amplified by PCR. Subsequent Sanger sequencing validated the sequence of the intervening region, the repeat type, and BLASTN analysis confirmed annotations (Table 2).

Two indels differentiate European (*Q. petraea*, *Q. robur*) and Asian (*Q. mongolica*) white oaks, and these are located in the *psaI-ycf4* linker (4 bp difference) and the *psbE-petL* linker (6 bp difference; Table 2), and these are inferred as mutations (deletions, specifically) are restricted to Asian white oaks. The further validation of these indels was conducted by screening the amplification products of 10 additional individuals of North American species, 50 individuals of two European species (*Q. robur*, *Q. petraea*), and two Asian species (*Q. mongolica*, *Q. dentata*). This validation revealed that white oaks from North America and Europe showed the same fragment length (Table 3, S1 Fig), and confirmed that the Asian species shared deletions that resulted in shorter, diagnostic fragment lengths.

Three indels differentiate Old World (*Q. petraea*, *Q. robur*, *Q. mongolica*) and North American white oaks, and these are located in the *trnL-trnF* (*trnLF*) linker (5 bp difference), the *trnC-petN* linker of the broader *trnC-trnD* (*trnCD*) linker region (8 bp) and in the *trnE-trnT* linker of the broader *trnD-trnT* (*trnDT*) linker region (2 bp; Table 2). These indels were validated with the same individuals of all above mentioned species. The validation revealed that white oaks from Asia and Europe showed identical fragment lengths (Table 3), and that the North American species shared mutations that resulted in diagnostic fragment lengths.

Resequencing (Sanger) of the *trnDT* region identified a single nucleotide polymorphism (SNP) differentiating Asia from Europe and North America. The SNP lies within a *HinfI* restriction site, and restriction digestion of this region was predicted to yield three fragments in Asian white oaks and two fragments in European and North American species (all oaks share

Table 3. Details for used species, individuals and markers. Given are number of individuals per species and continent tested with the five markers, and fragment length based on sequencing for each marker and species. Consensus sequences of the five markers are given in S1 Fig.

continent	species	¹ No. individuals	<i>psaI-ycf4</i>	<i>psbE-petL</i>	<i>trnLF</i>	<i>trnCD</i>	² <i>trnDT</i>
Europe	<i>Q. robur</i>	531 (103)	178	185	135	123	86/101
	<i>Q. petraea</i>	273 (12)	178	185	135	123	86/101
	<i>Q. pubescens</i>	158 (0)	178	185	135	123	
USA	<i>Q. alba</i>	15 (4)	178	185	130	115	88/101
	<i>Q. macrocarpa</i>	12 (7)	178	185	130	115	88/101
	<i>Q. bicolor</i>	7 (4)	178	185	130	115	88/101
	<i>Q. garryana</i>	4 (0)	178	185	130	115	
	<i>Q. lyrata</i>	4 (2)	178	185	130	115	88/101
	<i>Q. michauxii</i>	5 (1)	178	185	130	115	88/101
	<i>Q. stellata</i>	8 (1)	178	185	130	115	88/101
	<i>Q. prinoides</i>	6 (0)	178	185	130	115	
Asia	<i>Q. mongolica</i>	316 (420)	174	179	135	123	86/71 (30)
	<i>Q. dentata</i>	9 (5)	174	179	135	123	86/71 (30)

¹ The number of individuals tested includes numbers for all loci except *trnDT*, which is given in parentheses.

² Fragment lengths for *trnDT* show the lengths after restriction digestion with *HinfI*. Asian species contain an additional internal restriction site that yields one additional unlabeled fragment after *HinfI* digestion; these are shown in brackets.

one *Hin*I site). By labeling the forward and reverse amplification primer, restriction digestion of the PCR fragment with *Hin*I allows the visualization of two of the three fragments, one of which is diagnostic for Asian white oaks due to its truncated length. Since this single region offers the possibility to differentiate all three continents with one assay, we decided to include this SNP into the marker set. In total, the five markers have been evaluated with sample sizes ranging from 559 to 1241 (in detail: trnCD 1241, trnLF 1078, psaI_ycf4 1230, psbE_petL 1183, trnDT 559), with samples representing 13 oak species from the three continents (Table 3, genotype information per specimen is given in S1 File).

The nucleotide variations shown to be characteristic for Asian white oaks (*Q. mongolica* and *Q. dentata*) are also present in the complete chloroplast genome of one individual of *Q. aliena*, another Asian white oak species (GenBank accession KP301144.1) [29].

We computed the probabilities of fixation for the gene markers. Since we require a minimum of two independent markers for continent assignment, we performed the calculations for the risk of not identifying a rare variant in the reference samples based on a combination of two markers. By this means the risk of not identifying a rare variant in the reference samples was calculated to be less than 0.022% for Europe, 0.0051% for Asia and less than 0.098% for America. Thus, it is extremely unlikely that studied gene markers are not fixed to just one variant in the different groups.

Marker set design and optimization for timber

All above described analyses were performed using single PCR reactions and fragment analyses to optimize each marker separately. Subsequently, the markers were successfully multiplexed for fragment analysis using the fluorescence labeling as given in Table 2 (Fig 2).

For the development of the markers and multiplexing, DNA from fresh leaves was used. The protocol was later optimized for DNA from timber. From our experience, DNA from

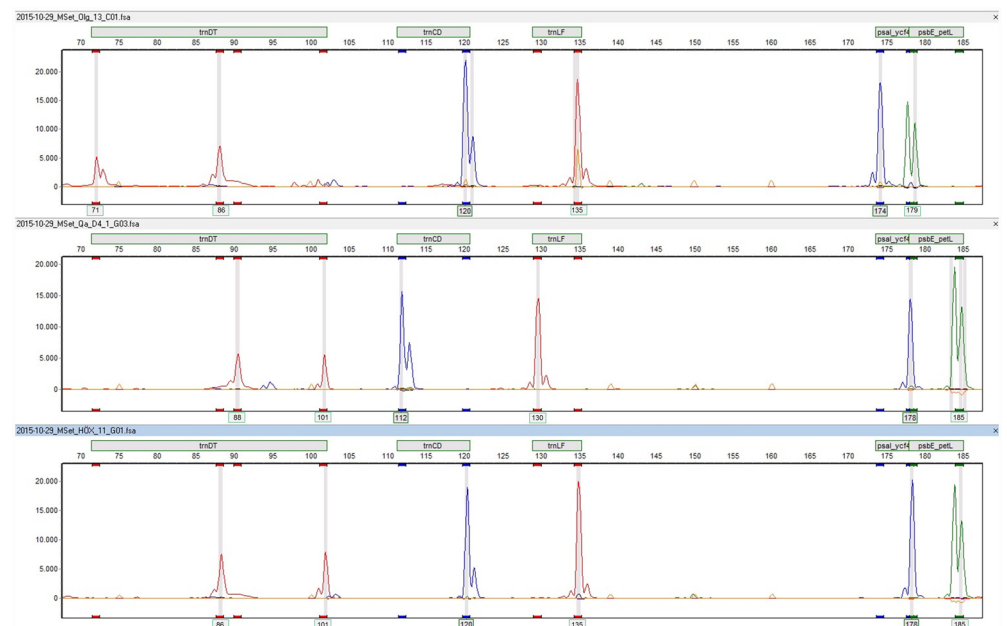


Fig 2. Fragment patterns of the five markers for individuals from Asia (top), North America (middle) and Europe (bottom). The sequence sizes for each peak as given in Table 3 are shown beneath the peaks. The first blue peaks appear smaller (112, 120) than the sequenced length (115, 123) given in Table 3. The color code of the peaks is as described in Table 2.

doi:10.1371/journal.pone.0158221.g002

Table 4. PCR conditions compared for leaf and timber. Only the differences are shown, all other parameters are as given in material and methods.

	Leaf				Timber			
	PCR cycles	Conc. MgCl ₂	Enhancer	Conc. Primer	PCR cycles	Conc. MgCl ₂	Enhancer	Conc. Primer
psal-ycf4	30	1.75 mM	no	0.07 μM	40	2.0 mM	no	0.1 μM
psbE-petL	30	1.75 mM	no	0.05 μM	40	2.0 mM	no	0.05 μM
trnLF	30	1.75 mM	yes	0.2 μM	40	2.0 mM	yes	0.2 μM
trnCD	25–30	1.5 mM	yes	0.13 μM	35	2.5 mM	yes	0.3 μM
trnDT	25	1.75 mM	yes	0.1 μM	40	2.0 mM	yes	0.1 μM

doi:10.1371/journal.pone.0158221.t004

timber is more sensitive to all PCR parameters, thus, all markers were singly tested with DNA from timber and the PCR was optimized for the DNA from timber. Differences in the PCR conditions used for the two different materials are given in Table 4. Due to the sensitivity of timber DNA in PCR, multiplexing of the PCRs of the five markers is not advisable, but multiplexing of the markers on the sequencer worked as well with timber DNA as with DNA from leaves. The only difference is that the PCR product from leaf DNA is diluted 1:50 and the PCR product from timber DNA 1:10 for use on the sequencer.

Our analyses used a capillary sequencer to visualize length polymorphisms of these fragments. However, due to the large size differences of these indels, all markers can be distinguished on a polyacrylamide gel, even for differences as small as two base pairs, as shown for the fragment *trnDT* (Fig 3). In this way, polymorphisms can be screened in laboratories where no sequencer is available.

The functionality of the optimized PCR protocols and the multiplexed sequencer runs has been tested by means of orders worked on in the “Thünen Centre of Competence on the Origin of Timber”. Timber from these orders included highly processed wood as flooring or parquet, as well as treated solid wood samples as different parts of furniture, barrels or boards, and unused solid wood as firewood. In total, over 80 processed timber samples and 130 treated solid wood samples have been evaluated (data not shown). Based on our experiences so far we have a success rate of sufficient DNA amplification for our gene markers for 58% for solid wood samples.

Discussion

A set of five chloroplast markers have been developed and optimized to analyze DNA from timber to identify the continental origin of white oak wood. Small fragment sizes (< 200 bp) were chosen because genotyping success with DNA from timber is highest when fragments under 200 bp are targeted. This has recently also been shown for DNA from old and dried insect specimens of museum collections when using mitochondrial barcoding regions [30]. For the identification of haplotypes within oak species from wood samples, Deguilloux et al. [31] similarly developed chloroplast microsatellite and SNP markers that targeted small DNA fragments.

The sequencing data revealed no specific indels to differentiate oak species within the classical barcoding regions *matK*, *rbcL* or the linker *trnH-psbA* [6,11,32–34]. Recently, the barcode regions *matK* and *trnH-psbA* were evaluated for their power to discriminate select species from oak sections *Cerris*, *Heterobalanus* (= “Group *Ilex*”), *Lobatae*, and *Quercus* [34]. In this study, the *matK* region proved to have too low resolution for the differentiation within the genus; interestingly, for *trnH-psbA*, the variability was too high to identify fixed interspecific differences [35].

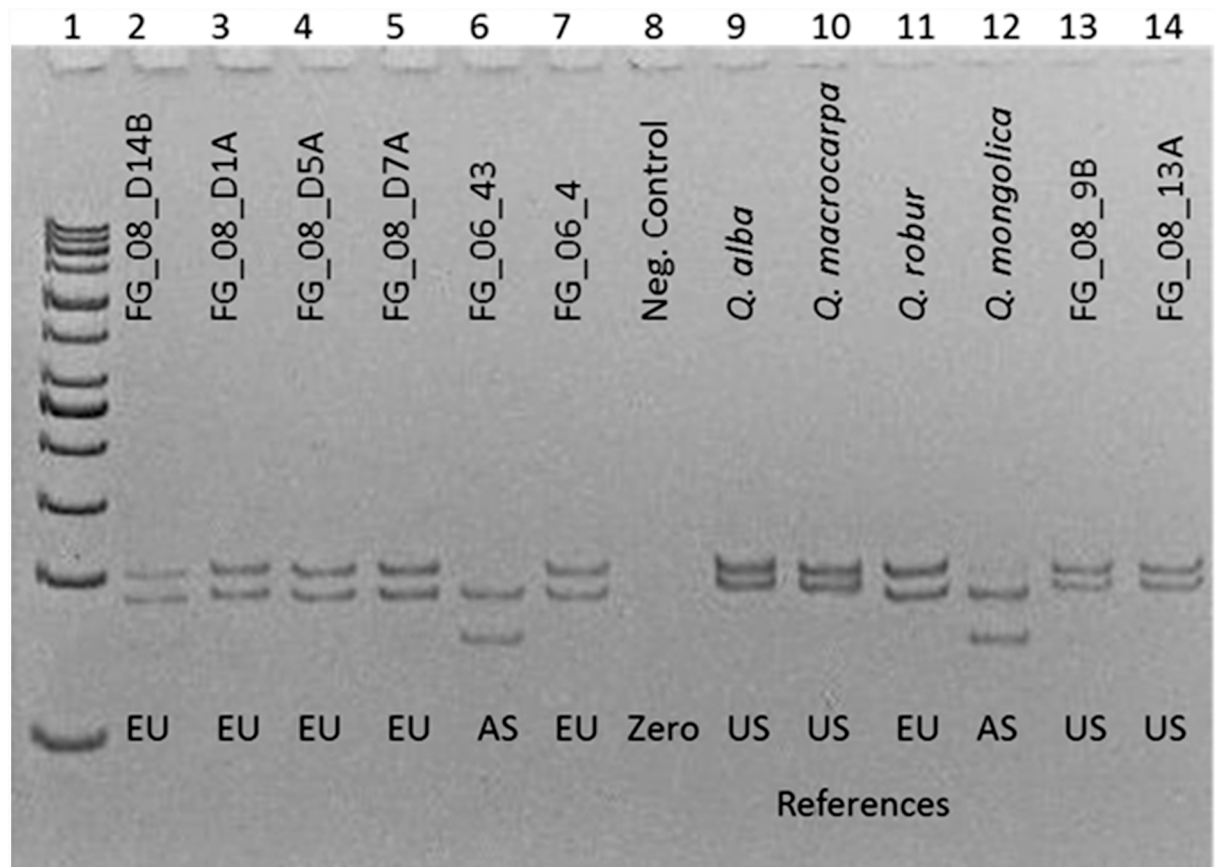


Fig 3. Marker *trnDT* visualized on a polyacrylamide gel. Lane 1: 50 bp ladder, lane 8: zero control, lane 2–7 and 13–14: analysis of wood-derived DNA, its location is inferred from genotypes, lane 9–12: references from North America (US), Europe (EU) or Asia (AS), respectively.

doi:10.1371/journal.pone.0158221.g003

The intergenic linkers *trnLF*, *trnCD* and *trnDT* we found valuable in this work have been widely tested in population and evolutionary genetic studies of plants, and they show wide variation in their ability to discriminate species and lineages [36]. For example, the intergenic linker *trnLF* proved to be not variable enough for overall barcoding approaches [6]. For the trial of phylogenetic reconstructions this *trnLF* linker lacked variation in closely related species [37]. Similarly, differentiation within the genus *Populus* failed using *trnLF* [38]. Nevertheless, there are other examples for successful use of this marker in molecular systematics [39,40] and citations therein) and for unraveling of the phylogeny of different plant species [41]. Similarly, *trnCD* and *trnDT* have been used in comprehensive studies of chloroplast DNA diversity in European white oaks [42–51]. In Japan, four oak species have been differentiated using *trnDT* among other chloroplast markers [52]. Hence, as for many regions within the chloroplast, the applicability of these spacers to questions of species identity depends on the specific phylogenetic and geographic context where they are used.

Forensic applications of molecular markers are already established with regard to illegal wildlife trade (parrots: [17]; sea turtles: [53]) or for identification of products made of endangered animal species ('whale meat': [54]; horn: [55]). The barcode of wildlife project (<http://www.barcodeofwildlife.org/>) has been originated especially for this purpose. Further on control of the seafood market in different countries is well supported by usage of barcoding markers

[18,56]. For identification of illegal logging of tropic tree species the use of molecular markers is already widespread [57–59], but the methods are less established for tree species from temperate zones. Thus, the presented markers should be applied to give commercial vendors of white oak wood the possibility to exercise ‘due diligence’ when placing timber on the European market and the public authorities to control timber imports should questions emerge on the correct declaration of wood.

Supporting Information

S1 Fig. Alignment for five cp regions with marked indels and SNPs used in the marker set.

For each of the described markers an alignment of a different number of individuals (between three and 50) of three species from the three continents has been made. Thus, the sequence for each species given in the figure is a consensus sequence. Further details of the markers and the related accession numbers of single individuals are given in [Table 2](#). QUALB: *Quercus alba* (USA), QUMON: *Q. mongolica* (Asia), QUROB: *Q. robur* (Europe). (TIF)

S1 File. Information details including genotypes of used *Quercus* individuals as references. (XLSX)

Acknowledgments

We thank Lasse Schindler for developing of the timber specific DNA extraction protocol, and Aki M. Höltken for initial discussions. We are grateful to Vivian Kuhlenkamp, Laura Schulz, Susanne Jelkmann and Ann-Christine Bergmann for technical assistance. We also thank collaborators and institutions that provided samples of North American oaks, including the Morton Arboretum, Lisle, IL (Andrew Hipp, Marlene Hahn), North Carolina State University, Raleigh, NC (Paul Manos), and the USDA Forest Service (Paul Berrang, Andy Bower; Britton Flash; Steven Forry; Ben Gombash; Mark Twery). Furthermore, we thank the Botanical Gardens of Duesseldorf, Bayreuth, Kiel, Goettingen, Giessen, Bonn, Bochum, Frankfurt, Ulm, Osnabrueck, and the Forest Botanical Gardens Tharandt and Eberswalde that provided us with some material of different oak species. Finally, we thank Mark Dasenko, Matthew Peterson, Chris Sullivan (Oregon State University Center for Genome Research and Biocomputing) for assistance with sequencing, and Brad Kinder, Alex Moad, and Shelley Gardner (US Forest Service) for project and planning assistance.

Author Contributions

Conceived and designed the experiments: HS RC YY BD. Performed the experiments: HS RC TJ BK. Analyzed the data: HS RC TJ BK. Contributed reagents/materials/analysis tools: MM BD. Wrote the paper: HS RC YY TJ MM BD BK. Overall coordination: HS RC BD BK.

References

1. Cassens D. White oak. In: Hardwood lumber and veneer series. Purdue Extension (2007); FNR-292-W.
2. US Department of Justice. Statement of facts, United States of American v. Lumber Liquidators. Case document 2:15-cr-00126-RAJ-LRL (2015). Available: <http://www.justice.gov/opa/file/787141/download>.
3. Dormontt EE, Boner M, Braun B, Breulmann G, Degen B, Espinoza E, et al. Forensic timber identification: It's time to integrate disciplines to combat illegal logging. *Biol Conserv* (2015); 191:790–798.
4. Hollingsworth PM, Forrest LL, Spouge JL. A DNA barcode for landplants. *Proc Natl Acad Sci U S A* (2009); 106:12794–12797. doi: [10.1073/pnas.0905845106](https://doi.org/10.1073/pnas.0905845106) PMID: [19666622](https://pubmed.ncbi.nlm.nih.gov/19666622/)

5. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. *PLoS One* (2011); 6(5):e19254. doi: [10.1371/journal.pone.0019254](https://doi.org/10.1371/journal.pone.0019254) PMID: [21637336](https://pubmed.ncbi.nlm.nih.gov/21637336/)
6. Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, et al. Land plants and DNA barcodes: short-term and long-term goals. *Philos Trans R Soc Lond B Biol Sci* (2005); 360:1889–1895. PMID: [16214746](https://pubmed.ncbi.nlm.nih.gov/16214746/)
7. Newmaster SG, Fazekas AJ, Ragupathy S. DNA barcoding in the land plants: valuation of *rbcl* in a multigene tiered approach. *Can J Bot* (2006); 84:335–341.
8. Kress WJ, Erickson DL. DNA barcodes: Genes, Genomics and bioinformatics. *Proc Natl Acad Sci U S A* (2008); 105(8):2761–2762. doi: [10.1073/pnas.0800476105](https://doi.org/10.1073/pnas.0800476105) PMID: [18287050](https://pubmed.ncbi.nlm.nih.gov/18287050/)
9. Kress WJ, Erickson DL. A two-locus DNA barcode for plants: the coding region *rbcl* gene complements the non-coding *trnH-psbA* spacer regions. *PLoS One* (2007); 6:e508.
10. Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valantini A, et al. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res* (2007); 35:e14. PMID: [17169982](https://pubmed.ncbi.nlm.nih.gov/17169982/)
11. Lahaye R, Van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, et al. DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci U S A* (2008); 105(8):2923–2928. doi: [10.1073/pnas.0709936105](https://doi.org/10.1073/pnas.0709936105) PMID: [18258745](https://pubmed.ncbi.nlm.nih.gov/18258745/)
12. Janzen DH, CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci U S A* (2009); 106(31):12794–12797. doi: [10.1073/pnas.0905845106](https://doi.org/10.1073/pnas.0905845106) PMID: [19666622](https://pubmed.ncbi.nlm.nih.gov/19666622/)
13. Huang X-C, Xi X-Q, Conran JG, Li J. Application of DNA barcodes in Asian tropical trees—a case study from Xishuangbanna nature reserve, Southwest China. *PLoS One* (2015); 10(6):e0129295. doi: [10.1371/journal.pone.0129295](https://doi.org/10.1371/journal.pone.0129295) PMID: [26121045](https://pubmed.ncbi.nlm.nih.gov/26121045/)
14. Straub SC, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *Am J Bot* (2012); 99(2):349–64. doi: [10.3732/ajb.1100335](https://doi.org/10.3732/ajb.1100335) PMID: [22174336](https://pubmed.ncbi.nlm.nih.gov/22174336/)
15. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol* (2009); 7:84. doi: [10.1186/1741-7007-7-84](https://doi.org/10.1186/1741-7007-7-84) PMID: [19954512](https://pubmed.ncbi.nlm.nih.gov/19954512/)
16. Parks M, Cronn R, Liston A. Separating the wheat from the chaff: mitigating the effects of noise in a plastome phylogenomic data set from *Pinus* L. (Pinaceae). *BMC Evol Biol* (2012); 12:100. doi: [10.1186/1471-2148-12-100](https://doi.org/10.1186/1471-2148-12-100) PMID: [22731878](https://pubmed.ncbi.nlm.nih.gov/22731878/)
17. Goncalves PFM, Oliveira-Marques AR, Matsumoto TE, Miyaki CY. DNA Barcoding identifies illegal parrot trade. *J Hered* (2015); 106:560–564. doi: [10.1093/jhered/esv035](https://doi.org/10.1093/jhered/esv035) PMID: [26245790](https://pubmed.ncbi.nlm.nih.gov/26245790/)
18. Pappalardo AM, Ferrito V. DNA barcoding species identification unveils mislabeling of processed flat-fish products in southern Italy markets. *Fish Res* (2015); 164:153–158.
19. Handy SM, Parks MB, Deeds JR, Liston A, de Jager LS, Luccioli S, et al. Use of the chloroplast gene *ycf1* for the genetic differentiation of pine nuts obtained from consumers experiencing dysgeusia. *J Agric Food Chem* (2011); 59:10995–11002. doi: [10.1021/jf203215v](https://doi.org/10.1021/jf203215v) PMID: [21932798](https://pubmed.ncbi.nlm.nih.gov/21932798/)
20. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* (2014); 30(15):2114–2120. doi: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170) PMID: [24695404](https://pubmed.ncbi.nlm.nih.gov/24695404/)
21. Crusoe MR, Alameldin HF, Awad S, Boucher E, Caldwell A, Cartwright R, et al. The khmer software package: enabling efficient nucleotide sequence analysis. *F1000 Research* (2015); 4:900. doi: [10.12688/f1000research.6924.1](https://doi.org/10.12688/f1000research.6924.1) PMID: [26535114](https://pubmed.ncbi.nlm.nih.gov/26535114/)
22. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* (2008); 18(5):821–829. doi: [10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107) PMID: [18349386](https://pubmed.ncbi.nlm.nih.gov/18349386/)
23. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Res* (2002); 12:656–664. PMID: [11932250](https://pubmed.ncbi.nlm.nih.gov/11932250/)
24. Stamatakis A: RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* (2014); 30(9):1312–1313. doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033) PMID: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)
25. Dumolin S, Demesure B, Petit RJ. Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theor Appl Genet* (1995); 91:1253–1256. doi: [10.1007/BF00220937](https://doi.org/10.1007/BF00220937) PMID: [24170054](https://pubmed.ncbi.nlm.nih.gov/24170054/)
26. Lowe AJ, Jardine DI, Cross HB, Degen B, Schindler L, Hoeltken AM. A method of extracting plant nucleic acids from lignified plant tissue. (2015) Patent WO/2015/070279 filed 2014-11-14 and issued 2015-05-21.
27. Newcombe RG. Two-Sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods. *Stat Med* (1998); 17:857–872. PMID: [9595616](https://pubmed.ncbi.nlm.nih.gov/9595616/)

28. Manos PS, Doyle JJ, Nixon KC. Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Mol Phylogenet Evol* (1999); 12(3):333–349. PMID: [10413627](#)
29. Lu S, Hou M, Du FK, Li J, Yin K. Complete chloroplast genome of the Oriental white oak: *Quercus aliena* Blume. *Mitochondrial DNA* (2015); 26:1–3.
30. Mitchell A. Collection in collections: a PCR strategy and primer set for DNA barcoding for decades-old dried museum specimens. *Mol Ecol Resour* (2015); 15:1102–1111. doi: [10.1111/1755-0998.12380](#) PMID: [25644663](#)
31. Deguilloux MF, Pemonge MH, Bertel L, Kremer A, Petit RJ. Checking the geographical origin of oak wood: molecular and statistical tools. *Mol Ecol* (2003); 12:1629–1636. PMID: [12755890](#)
32. Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G, et al. A proposal for a standardised protocol to barcode all land plants. *Taxon* (2007); 56(2):295–299.
33. Hollingsworth PM. Refining the DNA barcode for landplants. *Proc Natl Acad Sci U S A* (2011); 108(49):19451–19452. doi: [10.1073/pnas.1116812108](#) PMID: [22109553](#)
34. Tripathi AM, Tyagi A, Kumar A, Singh A, Singh S, Chaudhary LB, Roy S. The Internal Transcribed Spacer (ITS) region and *trnH-psbA* are suitable candidate loci for DNA barcoding of tropical trees species from India. *PLoS One* (2013); 8(2):e57934. doi: [10.1371/journal.pone.0057934](#) PMID: [23460915](#)
35. Simeone MC, Piredda R, Papini A, Vessella F, Schirone B. Application of plastid and nuclear markers to DNA barcoding of Euro-Mediterranean oaks (*Quercus*, Fagaceae): problems, prospects and phylogenetic implications. *Bot J Linn Soc* (2013); 172(4):478–499.
36. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, et al. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am J Bot* (2005); 92(1):142–166. doi: [10.3732/ajb.92.1.142](#) PMID: [21652394](#)
37. Dong W, Liu J, Yu J, Wang L, Zhou S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One* (2012); 7(4):e35071. doi: [10.1371/journal.pone.0035071](#) PMID: [22511980](#)
38. Schroeder H, Hoeltken AM, Fladung M. Differentiation of *Populus* species using chloroplast SNP-markers—essential for comprehensible and reliable poplar breeding. *Plant Biol* (2012); 14:374–381. doi: [10.1111/j.1438-8677.2011.00502.x](#) PMID: [21973311](#)
39. Martin-Bravo S, Meimberg M, Luceño M, Märkl W, Valcárcel V, Bräuchler C et al. Molecular systematics and biogeography of Resedaceae based on ITS and *trnL-F* sequences. *Mol Phylogenet Evol* (2007); 44:1105–1120. PMID: [17300965](#)
40. Pirie MD, Balca MP, Vargas Z, Botermans M, Bakker FT, Chatrou LW. Ancient paralogy in the cpDNA *trnL-F* region in Annonaceae: Implications for plant molecular systematics. *Am J Bot* (2007); 94(6):1003–1016. doi: [10.3732/ajb.94.6.1003](#) PMID: [21636470](#)
41. Drábková L, Kirschner J, Vlcek C, Paces V. *TrnL-trnF* intergenic spacer and *trnL* intron define major clades within *Luzula* and *Juncus* (Juncaceae): Importance of structural mutations. *J Mol Evol* (2004); 59:1–10. PMID: [15383902](#)
42. Bordács S, Popescu F, Slade D, Csaikl UM, Lesur I, Borovic A, et al. Chloroplast DNA variation of white oaks in northern Balkans and in the Carpathian Basin. *Forest Ecol Manag* (2002); 156:197–209.
43. Cottrell JE, Munro RC, Tabbener HE, Gillies ACM, Forrest GI, Deans JD, et al. Distribution of chloroplast DNA variation in British oaks (*Quercus robur* and *Q. petraea*): the influence of postglacial colonisation and human management. *Forest Ecol Manag* (2002); 156:181–195.
44. Csaikl UM, Burg K, Fineschi S, König AO, Mátyás C, Petit RJ. Chloroplast DNA variation of white oaks in the alpine region. *Forest Ecol Manag* (2002); 156:131–145.
45. Fineschi S, Turchini D, Grossoni P, Petit RJ, Vendramin GG. Chloroplast DNA variation of white oaks in Italy. *Forest Ecol Manag* (2002); 156:103–114.
46. Jensen JS, Gillies A, Csaikl U, Munro R, Madsen SF, Roulund H, et al. Chloroplast DNA variation within the Nordic countries. *Forest Ecol Manag* (2002); 156:167–180.
47. König AO, Ziegenhagen B, van Damb BC, Csaikl UM, Coart E, Degen B, et al. Chloroplast DNA variation of oaks in western Central Europe and genetic consequences of human influences. *Forest Ecol Manag* (2002); 156:147–166.
48. Olalde M, Herrán A, Espinel S, Goicoechea PG. White oaks phylogeography in the Iberian Peninsula. *Forest Ecol Manag* (2002); 156:89–102.
49. Petit RJ, Csaikl UM, Bordács S, Burg K, Coart E, Cottrell J, et al. Chloroplast DNA variation in European white oaks phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecol Manag* (2002); 156:5–26.

50. Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, et al. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecol Manag* (2002); 156:49–74.
51. Petit RJ, Latouche-Hallé C, Pemonge MH, Kremer A. Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. *Forest Ecol Manag* (2002); 156:115–129.
52. Kanno M, Yokoyama J, Suyama Y, Ohyama M, Itoh T, Suzuki M. Geographical distribution of two haplotypes of chloroplast DNA in four oak species (*Quercus*) in Japan. *J Plant Res* (2004); 117:311–317. PMID: [15232717](#)
53. Rehman A, Jafar S, Raja NA, Mahar J. Use of DNA Barcoding to control the Illegal Wildlife Trade: A CITES case report from Pakistan. *J Bioresour Manag* (2015); 2(2):19–22.
54. Baker CS, Steel D, Choi Y, Lee H, Kim KS, Choi SK, et al. Genetic evidence of illegal trade in protected whales links Japan with the US and South Korea. *Biol Lett* (2010); 6:647–650. doi: [10.1098/rsbl.2010.0239](#) PMID: [20392716](#)
55. Yan D, Luo JY, Han YM, Peng C, Dong XP, Chen SL, et al. Forensic DNA barcoding and Bio-Response studies of animal horn products used in traditional medicine. *PLoS One* (2013); 8(2):e55854. doi: [10.1371/journal.pone.0055854](#) PMID: [23409064](#)
56. Carvalho DC, Palhares RM, Goncalves Drummond M, Frigo TB. DNA Barcoding identification of commercialized seafood in South Brazil: A governmental regulatory forensic program. *Food Control* (2015); 50:784–788.
57. Höltnen AM, Schröder H, Wischniewski N, Degen B, Magel E, Fladung M. Development of DNA-based methods to identify CITES-protected timber species: A case study in the Meliaceae family. *Holzforchung* (2012); 66:97–104.
58. Nithaniyal S, Newmaster SG, Ragupathy S, Krishnamoorthy D, Vassou SL, Parani M. DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of India. *PLoS One* (2014); 9(9):e107669. doi: [10.1371/journal.pone.0107669](#) PMID: [25259794](#)
59. Hartvig I, Czako M, Kjær ED, Nielsen LR, Theilade I. The use of DNA barcoding in identification and conservation of rosewood (*Dalbergia* spp.). *PLoS One* (2015); 10(9):e01382