

# **Abschlussbericht**

## **DNA-basierte Arten- und Herkunfts- bestimmung für Holz und Holzprodukte der Weißeichen (Sektion *Quercus*)**

Aktenzeichen DBU: 31837/01-31

**Hilke Schröder, Birgit Kersten, Bernd Degen**

Thünen-Institut für Forstgenetik  
Sieker Landstrasse 2  
22927 Großhansdorf

Großhansdorf, Juni 2017

|  |                  |  |                   |   |                     |
|--|------------------|--|-------------------|---|---------------------|
| 06/02  |                  | <b>Projektkennblatt</b><br>der<br><b>Deutschen Bundesstiftung Umwelt</b>   |                   |  |                     |
| Az   | <b>31837</b>     | Referat  | <b>31</b>         | Fördersumme   | <b>259.000 Euro</b> |
| <b>Antragstitel</b>  |                  | <b>DNA-basierte Arten- und Herkunftsbestimmung für Holz und Holzprodukte der Weißeichen (Sektion Quercus)</b>  |                   |   |                     |
| <b>Stichworte</b>  |                  | Weißeichen, DNA-Marker, Holzherkunft, Artbestimmung, illegaler Holzeinschlag   |                   |   |                     |
|  | Laufzeit         | Projektbeginn  | Projektende       | Projektphase(n)   |                     |
|  | <b>2,5 Jahre</b> | <b>01.10.2014</b>  | <b>31.03.2017</b> | <b>1</b>  |                     |
|  | Zwischenberichte | <b>30.09.2015</b>  |                   |   |                     |
| <b>Bewilligungsempfänger</b>   |                  | Thünen-Institut für Forstgenetik<br>Sieker Landstraße 2, 22927 Großhansdorf  |                   | Tel   | 04102-696101        |
|  |                  |  |                   | Fax   | 04102-696200        |
|  |                  |  |                   | Projektleitung  |                     |
|  |                  |  |                   | Dr. Bernd Degen   |                     |
|  |                  |  |                   | Bearbeiter  |                     |
|  |                  |  |                   | Dr. Hilke Schröder  |                     |
| <b>Kooperationspartner</b>   |                  | Prof. Dr. Yulay Yanbaev, Bashkir State University und Russ. Akademie der Wiss. in Ufa, Russland<br>Brad Kinder und Shelley Gardener, US Forest Service International, Washington DC, USA<br>Prof. Dr. Richard Cronn, US Forest Service, Pacific NW Res Stn, USDA, Corvallis, OR 97331 USA<br>Prof. Dr. Meng-Zhu Lu, Chinese Academy of Forestry, Peking, China<br>Dr. Tae-Su Kim, Korean Forest Research Institute (KFRI), Seoul, Süd-Korea<br>Prof. Dr. Andrew Lowe, The University of Adelaide, Australien |                   |   |                     |
| <b>Zielsetzung und Anlass des Vorhabens</b>  |                  |  |                   |   |                     |
| Moderne DNA-basierte Methoden sind ein wichtiges Instrument im Kampf gegen den illegalen Holzhandel, von dem auch in den gemäßigten Breiten diverse Baumarten stark betroffen sind. Dazu zählt insbesondere das sehr wertvolle Holzsortiment der Weißeiche (Sektion Quercus). Die neue Gesetzeslage mit dem Inkrafttreten der EU-Holzhandelsverordnung im März 2013 und der hohe Marktanteil der Weißeichen sind Gründe für eine ständig steigende Nachfrage an eindeutigen Art- und Herkunftsnachweisen. Dieser Problematik soll in diesem Projekt mit der Entwicklung von molekularen Markern zur Identifizierung sowohl der Eichenarten als auch der geographischen Herkunft innerhalb einzelner Eichenarten begegnet werden. Insgesamt sollen folgende Ziele verfolgt werden: (i) Entwicklung von Genmarkern zur Unterscheidung von Weißeichen-Arten und –herkünften, (ii) Aufbau von genetischen Referenzdaten zur Herkunftskontrolle ausgewählter Weißeichen-Arten, (iii) Technologie-Transfer, Gentests als Audits zur Absicherung von Handelsketten. |                  |  |                   |   |                     |
| Deutsche Bundesstiftung Umwelt • An der Bornau 2 • 49090 Osnabrück • Tel 0541/9633-0 • Fax 0541/9633-190 • <a href="http://www.dbu.de">http://www.dbu.de</a>   |                  |  |                   |   |                     |

### **Darstellung der Arbeitsschritte und der angewandten Methoden**

Es wurden sowohl Genmarker entwickelt, mit denen die Artzugehörigkeit von Holz innerhalb der Sektion *Quercus* ermittelt werden kann als auch Genmarker, die erste Hinweise zur Herkunft des Holzes liefern (Ostasien, Amerika oder Europa sowie Herkunftsidentifizierung innerhalb der Kontinente). Die Entwicklung der Marker basiert auf der Illumina MiSeq „Next Generation Sequencing“-Methode. Es wurden zwei Pools mit je 20 Individuen aus *Q. robur* bzw. *Q. mongolica*, ein Pool von fünf *Q. petraea* Individuen sowie ein Referenzindividuum von *Q. mongolica* mit dieser Methode sequenziert. Eine fast vollständig zusammengesetzte Chloroplastensequenz des Referenzindividuums wurde als Referenz zur Anordnung (Mapping) der gewonnenen Sequenzen aus den *Q. robur*- bzw. *Q. mongolica*-Pools genutzt, um DNA Variationen (SNPs und InDels) in den jeweiligen Pools zu identifizieren. Des Weiteren wurden bereits SNPs und InDels im Mitochondriengenom identifiziert. Zusätzlich wurden Daten aus einer RAD-Sequenzierung von zwei verschiedenen *Q. robur* Individuen für die Identifizierung von ersten Kern-Variationen genutzt. Weitere Kernvariationen, die aus parallelen Mappings der *Q. petraea*- und der *Q. robur*-Pooldaten auf Kern-Referenzsequenzen gewonnen wurden, werden derzeit validiert. Eine große Anzahl an neu entwickelten Markern wurde über einen MassArray-Ansatz analysiert. Zuordnungen zu Arten oder Herkunftten wurden statistisch abgesichert durch sog. „Self-Assignment-Tests“.

Zum Aufbau von genetischen Referenzdaten zur Herkunftskontrolle von ausgewählten Weißeichen-Arten wurden Referenzproben aus dem Verbreitungsgebiet der Stieleiche in Europa, der Mongolischen Eiche im Osten Russlands sowie der US-amerikanischen Eichenarten mit den entwickelten Markersets analysiert.

### **Ergebnisse und Diskussion**

Es sind molekulare Marker aus dem Plastiden- und Mitochondriengenom zu einem Markersets zusammengefasst worden, um den kontinentalen Ursprung von *Quercus*-Arten aus den USA, Europa und Asien differenzieren zu können. Dieses Markersets ist umfangreich validiert und für die Anwendung auf Holzproben optimiert worden. Anhand einer MassArray Analyse mit einer großen Anzahl an Kern-Markern konnte eine weitere Auftrennung der Arten über Häufigkeitsverteilungen und Berechnung von Zuordnungs-Wahrscheinlichkeiten erreicht werden. Außerdem ist innerhalb von Europa eine Herkunftsbestimmung mittels eines Markersets aus einer Kombination von etablierten und neu entwickelten Markern möglich. Die Marker wurden mit Hilfe von leicht umzusetzenden und damit überall anwendbaren Methoden in die Praxis überführt und sind ebenfalls bereits für die Nutzung bei Holz und Holzprodukten optimiert. Populationen der mongolischen Eiche (*Q. mongolica*) lassen sich mit einem Set aus fünf plastidären und mitochondrialen Genmarkern großräumig (westliche, südliche und nördlich gelegene Populationen) unterscheiden. Die Variation innerhalb der mongolischen Eiche aus den Stichprobengebieten im Fernen Russland und Chinas ist deutlich geringer als in den europäischen oder amerikanischen Arten sowohl in den haploiden Genome (Chloroplasten und Mitochondrien) als auch im diploiden Kerngenom. Auf der Konzessionsebene können mit diesem Markersets Individuen den jeweiligen Populationen mit hoher Wahrscheinlichkeit zugeordnet werden. Die Kernvariationen aus den *Q. petraea* Pool-Daten werden zurzeit noch validiert, um Marker für die Differenzierung der europäischen Arten zu erhalten. Ein umfangreicher Flyer zur Information von potentiellen Kunden wird derzeit im Rahmen des Holz-Kompetenzzentrums erstellt.

## Öffentlichkeitsarbeit und Präsentation

**Tab. 1:** Besuche von Medienvertretern zur Erstellung von Fernsehbeiträgen

| Datum         | Mediananstalt                   | Titel Beitrag   |
|---------------|---------------------------------|---|
| April 2015    | ZDF zoom                        | Ikea, Höffner und Co.<br>Woher kommen unsere Billigmöbel? |
| Dezember 2016 | NDR, Schleswig-Holstein Magazin | Forscher arbeiten an weltgrößter<br>Baumdatenbank         |
| März 2017     | Deutsche Welle                  | Germany's wood detectives                                 |

**Tab. 2:** Besuchte Veranstaltungen mit Posterbeiträgen oder Vorträgen zu dem Thema und zu den Arbeiten des Projekts

| Datum                    | Veranstaltung   | Ort                      | Thema  | Person          |
|--------------------------|---|--------------------------|--|-----------------|
| 30.11. bis<br>02.12.2014 | Treffen Global Timber<br>Tracking Network   | Rom,<br>Italien          | Application of DNA fingerprints to<br>control tree species and<br>geographic origin of timber        | Bernd Degen     |
| 04.12.2014               | Meeting of The British<br>Timber Trade<br>Federation                                      | London,<br>UK            | Application of DNA fingerprints to<br>control tree species and<br>geographic origin of timber        | Bernd Degen     |
| 10.12. bis<br>11.12.2014 | Conference of The<br>International<br>Consortium on<br>Combating Wildlife<br>Crime (ICWC) | Wien,<br>Austria         | Application of DNA fingerprints to<br>control tree species and<br>geographic origin of timber        | Bernd Degen     |
| 12.06.2015               | Jahrestagung<br>Gesamtverband<br>Deutscher Holzhandel<br>(GD-Holz)                        | Düsseldorf               | Genetische Holzarten- und<br>Herkunftsbestimmung   | Bernd Degen     |
| 09.05. bis<br>11.05.2016 | 4th Plant Genomics<br>Congress  | London,<br>UK            | NGS-based development of<br>molecular markers for<br>determining continental origin of<br>white oaks | Hilke Schroeder |
| 15.06. bis<br>16.06.2016 | Tagung der Sektion<br>Forstgenetik  | Chorin                   | Anwenderfreundliche DNA-<br>Marker zur Herkunftsidenti-<br>fizierung von Eichenholz                  | Hilke Schroeder |
| 21.06.2016               | Deutsch-Chinesischer<br>Workshop zur<br>Holzforschung                                     | Hamburg                  | Application of DNA fingerprints to<br>control tree species and<br>geographic origin of timber        | Bernd Degen     |
| 15.12.2016               | Plant genetics<br>colloquium and forest<br>genetics seminar                               | Universität<br>Göttingen | GeoAssign: A new genetic<br>approach to assign individuals to<br>species and populations             | Bernd Degen     |
| 08.06.2017               | CITES workshop  | Bergedorf                | Application of DNA fingerprints to<br>control tree species and<br>geographic origin of timber        | Hilke Schroeder |

### Fazit

Die im Projekt etablierten Markersets erlauben eine kontinentale Zuordnung von Weißeichen, eine Artdifferenzierung (für einige Arten), eine erweiterte Herkunftsbestimmung innerhalb Europas und – für die mongolische Eiche - sowohl eine großräumige Unterscheidung von Populationen als auch eine kleinräumige Zuordnung auf der Konzessionsebene. Alle entwickelten Markersets sind in umfangreichen MassArray Analysen validiert worden. Außerdem sind die Optimierung der Marker für Holz abgeschlossen und die Markersets in der Praxis bereits etabliert. Die im Projekt gestellten Aufgaben konnten damit sehr erfolgreich fertig gestellt werden.

## Inhaltsverzeichnis

|   |    |
|---|----|
| <b>Projektkennblatt</b> .....   | 2  |
| <b>Abbildungsverzeichnis</b> .....                                    | 6  |
| <b>Tabellenverzeichnis</b> .....                                      | 6  |
| <b>Abkürzungsverzeichnis</b> .....                                    | 6  |
| <b>Zusammenfassung</b> .....  | 7  |
| <b>1. Einführung / Motivation</b> .....                               | 8  |
| <b>2. Methodik / Vorgehensweise / Projektablauf</b> .....             | 8  |
| 2.1 Sequenzierungen und SNP-Detektionen .....                         | 8  |
| 2.2 Markerentwicklung .....   | 10 |
| 2.2.1 Vorauswahl von SNPs (Pre-Screening) .....                       | 10 |
| 2.2.2 Statistische Methoden .....                                     | 11 |
| 2.3 Validierung .....   | 11 |
| 2.4 Projektablauf .....   | 11 |
| <b>3. Projektergebnisse</b> .....                                     | 12 |
| 3.1 Materialbeschaffung .....   | 12 |
| 3.2 Kontinent- und Artdifferenzierung .....                           | 14 |
| 3.2.1 Kontinentdifferenzierung .....                                  | 14 |
| 3.2.2 Artdifferenzierung über alle drei Kontinente .....              | 16 |
| 3.3 Artdifferenzierung innerhalb Europas .....                        | 16 |
| 3.3.1 RAD-Marker .....  | 16 |
| 3.3.2 MiSeq Sequenzierung von <i>Q. petraea</i> .....                 | 17 |
| 3.4 Herkunftsidentifizierung .....                                    | 18 |
| 3.4.1 Herkunftsidentifizierung innerhalb von Europa .....             | 18 |
| 3.4.2 Herkunftsidentifizierung innerhalb von Asien .....              | 20 |
| 3.5 Praxistest .....  | 22 |
| <b>4. Öffentlichkeitsarbeit / Veröffentlichungen / Vorträge</b> ..... | 23 |
| 4.1 Öffentlichkeitsarbeit .....                                       | 23 |
| 4.2 Vorträge / Poster .....   | 24 |
| 4.3 Publikationen .....   | 24 |
| <b>5. Fazit / Ausblick</b> .....                                      | 25 |
| <b>6. Literaturangaben</b> .....                                      | 26 |

### Abbildungsverzeichnis

|  |    |
|--|----|
| Abbildung 1: Markerset zur Kontinentdifferenzierung .....                      | 15 |
| Abbildung 2: Clusteranalyse zur Artdifferenzierung .....                       | 16 |
| Abbildung 3: SNAPShot Methode zur Differenzierung der europäischen Arten ..... | 17 |
| Abbildung 4: Verbreitungskarten der häufigsten europäischen Haplotypen .....   | 19 |
| Abbildung 5: Haplotypenverteilung innerhalb von <i>Quercus mongolica</i> ..... | 21 |
| Abbildung 6: Karte für die Analyse auf Konzessionsebene .....                  | 22 |

### Tabellenverzeichnis

|   |    |
|---|----|
| Tabelle 1: Zeitplan .....   | 12 |
| Tabelle 2: Referenzproben .....   | 13 |
| Tabelle 3: Marker-Zusammenstellung zur Differenzierung der europäischen Haplotypen ...  | 20 |
| Tabelle 4: Marker-Zusammenstellung zur Differenzierung der asiatischen Haplotypen ..... | 20 |
| Tabelle 5: Besuche von Medienvertretern .....   | 23 |
| Tabelle 6: Veranstaltungsbesuche mit Posterbeiträgen und Vorträgen .....                | 24 |

### Abkürzungsverzeichnis

|          |  |
|----------|--|
| BLAST    | „Basic Local Alignment Search Tool“: Methode zum gleichzeitigen Vergleich vieler Sequenzen   |
| CITES    | “Convention on International Trade in Endangered Species of Wild Fauna and Flora”: Washingtoner Artenschutzübereinkommen   |
| cp       | Chloroplast  |
| FLEGT    | „Forest Law Enforcement, Governance and Trade“: EU-Aktionsplan zur Bekämpfung des illegalen Holzeinschlags   |
| InDel    | “Insertion/Deletion”: Längenpolymorphismus, der auf einer kombinierten Insertions/Deletions- Mutation beruht   |
| mt       | Mitochondrium  |
| NGS      | „Next Generation Sequencing“: Methoden zur Hochdurchsatz-Sequenzierung   |
| PCR      | „Polymerase Chain Reaction“: Methode zur Vervielfältigung einzelner DNA-Abschnitte   |
| PCR-RFLP | „Restriction Fragment Length Polymorphism“: Auf einer PCR und einer Restriktion basierende Methode zur Analyse von Fragmentlängenunterschieden   |
| RAD      | „Restriction site associated DNA“: Methode zur DNA-Marker Identifizierung basierend auf NGS-Sequenzierung von DNA-Bereichen, die Restriktionsschnittstellen („restriction sites“) flankieren |
| SNP      | „Single Nucleotide Polymorphism“: Vererbare Variation eines einzelnen Nukleotides im DNA-Strang  |

## Zusammenfassung

Illegaler Holzeinschlag betrifft nicht nur tropische Baumarten, sondern auch in den gemäßigten Zonen vorkommende Arten. Insbesondere für die CITES-geschützte mongolische Eiche (*Quercus mongolica*) sind besonders wertvolle Ökosysteme durch illegalen Holzeinschlag gefährdet. Verschiedene Übereinkünfte in Europa und den USA (FLEGT, Zertifizierungssysteme, Lacey Act) haben bereits zu ersten Erfolgen im Kampf gegen den illegalen Holzeinschlag geführt. Die EU-Holzhandelsverordnung hat den internationalen Holzhandel zusätzlich unter Druck gesetzt, gehandeltes Holz auf botanische Art und Herkunft überprüfen zu lassen.

Zu diesem Zweck sind anhand von umfangreichen Sequenzierungen der nächsten Generation (Next Generation Sequencing = NGS) erste molekulare Marker aus dem Plastiden- und Mitochondriengenom entwickelt worden, die eine Differenzierung von Eichenarten aus Asien, Europa und den USA erlauben. Außerdem ist bereits innerhalb von Europa eine Herkunftsbestimmung anhand von etablierten Haplotypen möglich. Diese haben wir im Projekt für die Anwendung bei Holz und Holzprodukten optimiert und die Auflösung innerhalb Europas mit weiteren im Projekt entwickelten Markern erhöht. Außerdem lassen sich Populationen der mongolischen Eiche anhand von sechs Haplotypen geographisch differenzieren. Die entwickelten molekularen Marker sind alle auf Holzproben optimiert worden und in die praktische Anwendung bei der Bearbeitung von Prüfaufträgen eingeflossen. Ein weiterer umfangreicher Datensatz befindet sich zurzeit noch in der Analyse, um Marker zur Differenzierung der beiden europäischen Eichenarten *Q. robur* und *Q. petraea* zu entwickeln.

## 1. Einführung / Motivation

Der illegale Holzhandel ist ein weltweit schwerwiegendes Problem, von dem auch in den gemäßigten Breiten diverse Baumarten stark betroffen sind. Obwohl der Fokus häufig auf tropischen Baumarten liegt, ist z.B. das sehr wertvolle Holzsortiment der Weißeiche (Sektion *Quercus*) ebenfalls vom illegalen Handel betroffen. Wichtige Handelshölzer aus diesem Sortiment sind die europäischen Arten *Quercus robur* und *Q. petraea* sowie mehrere amerikanische Arten, insbesondere *Q. alba* und *Q. macrocarpa* und die CITES-geschützte in Asien verbreitete Art *Q. mongolica*.

International agierende Holzhandelsunternehmen laufen Gefahr gegen internationale Gesetze (FLEGT Abkommen, Lacey Act) zu verstoßen, die sie zur genauen Deklaration der botanischen Art und Herkunft der gehandelten Hölzer verpflichtet (Degen & Höltken 2011). Spätestens mit dem Inkrafttreten der EU-Holzhandelsverordnung im März 2013 wuchs daher die Nachfrage nach Methoden zum Art- und Herkunftsnachweis.

Für die Gattung *Quercus* ist eine Bestimmung anhand von holzanatomischen Methoden nur insoweit möglich, als dass die Sektionen Rot- und Weißeichen unterschieden werden können. Hier bieten sich moderne molekulare Methoden zur Artunterscheidung an (Barcode of Life, [www.barcodeoflife.org](http://www.barcodeoflife.org), Hollingsworth et al. 2009). Zur Identifizierung von Pflanzenarten werden häufig verschiedene Regionen der Chloroplasten-DNA genutzt (z. B. Kress & Erickson 2007, Taberlet et al. 2007, Lahaye et al. 2008). Speziell zur Differenzierung der *Quercus*-Arten gibt es erste Ansätze zur Markerentwicklung in klassischen Barcoding-Regionen im Chloroplasten, wie *trnH-psbA* und *matK* (Hollingsworth 2011, Simeone et al. 2013). Bisher war mit diesen Markern lediglich eine Differenzierung von einigen Artengruppen (entsprechen Sektionen) möglich. Ein Verfahren zur Bestimmung der geographischen Herkunft innerhalb mehrerer europäischer Arten ist anhand von Haplotypen in einem europaweiten Kooperationsprojekt unter anderem für die drei Arten *Q. robur*, *Q. petraea* und *Q. pubescens* entwickelt worden (Petit et al. 2002a, b). Allerdings sind nicht diese Arten anhand der Haplotypen unterscheidbar, sondern die Herkunft innerhalb von Europa ist unabhängig von der Art bestimmbar. Ebenso wenig war die gleichzeitige Bestimmung von Herkunft und Art machbar. Damit fehlte die Möglichkeit der Überprüfung der Holzhandelsketten.

Damit also das Holzhandels-Sicherungs-Gesetz als deutsche Umsetzung der EU-Holzhandelsverordnung in der Praxis greift, müssen fälschungssichere Methoden zur Überprüfung der Art und Herkunft zur Verfügung gestellt werden. Das vorliegende Projekt hatte deshalb zum Ziel, mit der Entwicklung von molekularen Markern zur Identifizierung sowohl der Eichenarten als auch der geographischen Herkunft von Individuen innerhalb einzelner Eichenarten praxistaugliche Testverfahren bereitzustellen.

## 2. Methodik / Vorgehensweise / Projektablauf

### 2.1 Sequenzierungen und SNP-Detektionen

Für die Entwicklung von Markern sowohl zur interspezifischen Differenzierung als auch zur Identifizierung von Populationen innerhalb einzelner Arten wurden zwei Pools aus jeweils 20 *Q. mongolica* und 20 *Q. robur* Individuen (mit je zwei Individuen aus 10 verschiedenen Populationen) gebildet. Die DNA dieser zwei Pools und die DNA eines einzelnen Individuums von *Q. mongolica* (Referenzindividuum) wurden mit einer 10-fachen „Coverage“ (Abdeckung) mit der Illumina MiSeq „Next Generation Sequencing“-Methode sequenziert. Mit

dieser Methode wurden 2x300 Basen lange „paired-end reads“ (gepaarte Sequenzstücke) erstellt, mit einem Gesamtsequenzdatenvolumen von etwa 15 GB. Diese Sequenzierungen wurden aus dem im Antrag erwähnten Co-Projekt „Russia Far East“ finanziert.

Die nachfolgenden bioinformatischen Datenanalysen erfolgten mit Hilfe der Software CLC Genomics-Workbench (CLC-GWB; CLC bio, a QIAGEN company, Aarhus, Dänemark). Im ersten Schritt wurden aus den Reads des *Q. mongolica* Referenzindividuums Contigs (größere Sequenzstücke) zusammengesetzt (assembliert). Anhand von BLAST-Analysen wurden artfremde Contigs (Kontaminationen von Algen, Bakterien oder Pilzen) entfernt und die verbleibenden Contigs in plastidäre, mitochondriale und Kern-spezifische Contigs unterteilt. Diese Contigs wurden im Weiteren als Referenzcontigs verwendet. Zusätzlich wurde uns von Kooperationspartnern (R. Cronn, USA) eine fast vollständig zusammengesetzte Chloroplastensequenz des o.g. *Q. mongolica* Referenzindividuums bereitgestellt, die direkt für eine möglichst schnelle Entwicklung von Chloroplasten-Markern eingesetzt werden konnte. Auf diese *Q. mongolica* Referenzsequenz wurden die NGS-Daten, die aus den beiden Pools generiert wurden (Pool-Daten), parallel „gemappt“ (anhand der Referenzsequenz angeordnet), um anschließend SNPs und InDels in beiden „Mappings“ zu detektieren (CLC-GWB). Später wurden die genannten Analysen auch unter Verwendung der Referenzcontigs von Mitochondrien und Kern durchgeführt, um auch in der mitochondrialen bzw. nukleären genomischen DNA SNPs und InDels zu identifizieren.

Zur potentiellen Artdifferenzierung wurden in erster Linie SNPs und InDels ausgewählt, die in den Pools einer (oder mehrerer Arten) auf ein Allel fixiert waren (Allelfrequenz = 100%). Bei derartigen SNPs und InDels ist davon auszugehen, dass sie in allen Individuen des Pools vorliegen, also potentiell artspezifisch sind. Für die Identifizierung von SNPs und InDels, die potentiell geographische oder Populations-Unterschiede innerhalb der Arten abbilden, wurden SNPs und InDels selektiert, bei denen mindestens zwei Allele im Pool vorhanden waren, wobei die Frequenz eines Alleles im Pool jeweils im Bereich von 10-90% liegen sollte und an deren Position in der Referenzsequenz nur ein Allel vorlag. Die selektierten SNPs und InDels wurden dann nachfolgend für die Markerentwicklung verwendet.

Mit dieser Methode (Illumina MiSeq und nachfolgende Bioinformatik) können SNPs und InDels im Chloroplasten und Mitochondrien in großer Anzahl und in relativ kurzer Zeit detektiert werden. Da das Genom des Kerns sehr viel größer ist und dementsprechend sehr viele Contigs bei einer Gesamtsequenzierung des Genoms und der Assemblierung der gewonnenen Sequenzen gebildet werden, ist die Auswahl von DNA-Variationen (SNPs und InDels) hier bioinformatisch deutlich aufwendiger. Daher wurde zur Gewinnung potentieller Kernmarker zusätzlich auf Daten aus einer RAD-Sequenzierung von zwei einzelnen *Q. robur* Individuen zugegriffen, die im Institut bereits vorlagen. Bei der RAD-Sequenzierung wird das gesamte Genom mit spezifischen Restriktionsenzymen (hier ein „Achter-cutter“, d.h. ein Restriktionsenzym, das eine Erkennungssequenz von acht Basen hat) geschnitten. Über spezifische Adapter, die an diese Schnittstellen binden, werden DNA-Bereichen, welche die Restriktionsschnittstellen flankieren, selektiert und sequenziert. Die resultierenden Sequenzabschnitte beider Individuen werden miteinander verglichen und ebenfalls auf SNPs und kurze InDels (ein oder zwei Basen) hin untersucht. Diese Analyse ist bioinformatisch deutlich weniger aufwendig, resultiert aber in einer sehr viel geringeren Anzahl an möglichen DNA-Variationen; diese sind dafür aber sehr schnell verfügbar.

Zur weiteren Erhöhung der Chancen, Kernmarker zur Differenzierung zu erhalten, wurden die Kern-Contigs aus der MiSeq-Sequenzierung ausgewertet und in die Markerentwicklung mit einbezogen.

Im Zuge der kostenneutralen Verlängerung ist außerdem für eine detailliertere Differenzierung der europäischen Arten eine MiSeq Sequenzierung von einem Pool aus fünf Individuen von *Q. petraea* durchgeführt worden. Die Vorgehensweise entspricht weitestgehend der oben für die Pools aus *Q. mongolica* und *Q. robur* beschriebenen MiSeq Sequenzierung, allerdings wurde für *Q. petraea* mit einer 30-fachen coverage gearbeitet, um auf jeden Fall Kern-Marker identifizieren zu können. Die resultierenden NGS-Daten wurden parallel zu den Daten des *Q. robur* Pools auf inzwischen publizierte „Scaffolds“ (assemblierte lange Abschnitte des Kern-Genoms) von *Q. robur* (Plomion et al. 2016) gemappt. Mit Hilfe einer im Hause entwickelten Software (Variant Tools, Schroeder et al. 2016a) wurden SNPs und InDels in beiden Pools detektiert und anschließend miteinander verglichen. Auch hier wurde für die Auswahl potentiell differenzierender Marker zunächst wieder auf die auf ein Allel fixierten (100% Allelfrequenz) SNPs bzw. InDels zugegriffen.

## 2.2 Markerentwicklung

### 2.2.1 Vorauswahl von SNPs (Pre-Screening)

Die identifizierten DNA-Variationen (SNPs und InDels) aus 2.1 sowohl von der MiSeq-Sequenzierung als auch der RAD-Sequenzierung wurden für die weitere Entwicklung der Marker verwendet. Zunächst wurden für die Entwicklung von artspezifischen Markern aus den Sequenzumgebungen der Kandidaten-SNPs/InDels (Allelfrequenz von 100%) des Chloroplastengenoms flankierende Primer abgeleitet und für PCRs mit einer Auswahl von zunächst je drei Individuen je Art verwendet. Die erhaltenen PCR-Produkte wurden sequenziert (Sanger) und mit den ursprünglichen Sequenzen verglichen. Für alle potentiellen Marker, bei denen die PCR-Reaktion funktioniert hat und in der Sequenz die ausgewählten SNPs bzw. InDels vorhanden waren, wurden die PCRs mit einer größeren Anzahl von Individuen aus einem repräsentativen Bereich des Verbreitungsgebiets der zu differenzierenden Arten wiederholt.

Da sehr viel mehr SNPs bzw. InDels eine Allelfrequenz von unter 100% (nicht auf eine einzige Variante fixiert waren) aufwiesen, also für eine Differenzierung von Arten sowie eine Unterscheidung von Populationen innerhalb von Arten mit statistischen Methoden in Frage kamen, wurde hier auf ein Hochdurchsatzverfahren zugegriffen. Für ein „Pre-Screening“ der potentiellen Marker wurde die PCR- und Massenspektrometrie- basierte Methode des „MassArray“ verwendet. Hierbei werden in einem Multiplexverfahren bis zu 30 SNPs gleichzeitig an einer großen Anzahl an Individuen getestet. Dieses Verfahren wurde zunächst für eine Auswahl von 120 plastidären und mitochondrialen SNPs angewendet. Hierbei wurden insgesamt 95 Individuen verteilt auf die drei europäischen Arten *Q. robur*, *Q. petraea* und *Q. pubescens*, sieben amerikanische Arten *Q. alba*, *Q. macrocarpa*, *Q. stellata*, *Q. lyrata*, *Q. garryana*, *Q. michauxii* und *Q. bicolor* sowie vier asiatische Arten *Q. mongolica*, *Q. dentata*, *Q. aliena* und *Q. serrata* in die Analyse einbezogen. Außerdem wurde eine MassArray Analyse mit 233 aus der RAD-Sequenzierung selektierten Kern-SNPs mit insgesamt 190 Individuen der o.a. Arten durchgeführt. Beide MassArray Analysen dienten der Suche nach weiteren potentiellen SNPs zur Art- und Populationsdifferenzierung.

### 2.2.2 Statistische Methoden

Die Qualität der einzelnen SNP Genmarker wurde statistisch mit Hilfe des Programms „GDA-NT“ (Degen 2008) analysiert. Als Auswahlkriterien für SNPs berechnet das Programm für jeden Genort a) die Vollständigkeit der Daten, b) den Anteil Heterozygoter im Vergleich zu einer Zufallskombination der Allele, c) die genetische Diversität und d) die genetische Differenzierung.

Ein übliches statistisches Verfahren, um die Eignung von Genmarkern zur Herkunftsunterscheidung zu bestimmen, sind die sog. „Self-Assignment-Tests“. Hierbei werden der Reihe nach die einzelnen Individuen des Referenzdatensatzes herausgenommen und über den jeweiligen Zuordnungsalgorithmus einer Herkunft zugeordnet (Degen et al. 2013). Der Anteil hierbei richtig zugeordneter Individuen ist ein Maß für die Genauigkeit der Herkunftsbestimmung. Diese Algorithmen sind z.B. in der frei verfügbaren Software GeneClass 2 (Piry et al. 2004) und in der im Hause programmierten Software GeoAssign (Degen et al. 2017) enthalten.

### 2.3 Validierung

Bei Genmarkern zur Artunterscheidung suchen wir – wie oben beschrieben – solche, die bei verschiedenen Arten auf unterschiedliche genetische Varianten (Allele, Haplotypen) fixiert sind. Das heißt, idealerweise haben dann alle Individuen der einen Art nur die eine Variante (ein Allel) und alle anderen Individuen haben bei dem Genmarker eine oder mehrere andere genetische Varianten (Allele). Wir müssen ausreichend große Stichproben von Individuen der einzelnen Arten analysieren, um abzuschließen, dass die Genmarker tatsächlich fixiert sind. Nach Krusche & Geburek (1991) benötigt man, um mit 95%iger Wahrscheinlichkeit abzuschließen, dass die alternative genetische Variante mit höchstens 5% in der Grundgesamtheit vorkommt, 30 Individuen bei diploiden Genmarkern und 60 Individuen für haploide Genmarker. Diese Individuen sollten idealerweise eine Stichprobe darstellen, die repräsentativ für das geographische Verbreitungsgebiet der jeweiligen Art ist.

Die Validierung der Genmarker für die Herkunftsbestimmung erfolgte mit Hilfe der oben angesprochenen „Self-Assignment-Tests“ und später dann anhand von Blindtests.

Einfach anzuwendende Methoden zur experimentellen Testung einzelner Individuen mit den Markern (Genotypisierung an der jeweiligen SNP/InDel-Position) umfassen PCR-Reaktionen mit einer anschließenden Größenauftrennung (bei InDels) oder nachfolgender Restriktion und Größenauftrennung (PCR-RFLP) auf einem ABI-Sequenzierer. Diese Methoden wurden ebenfalls sowohl für art- als auch für herkunftsspezifische Marker getestet.

### 2.4 Projektablauf

Das Projekt konnte weitestgehend im ursprünglich geplanten Zeitrahmen abgeschlossen werden. Es stehen nur noch wenige Schritte der innerhalb der kostenneutralen Verlängerung begonnenen Arbeiten aus. Die bereits abgeschlossenen Arbeitsschritte sind in Tabelle 1 in grün dargestellt, die kurz vor dem Abschluss befindlichen Schritte in gelb. Die Bezeichnungen der Arbeitsschritte entsprechen denen im Antrag (3.1 bis 3.4) bzw. denen im Antrag für die kostenneutrale Verlängerung (1-5) (Tab. 1).

**Tab. 1:** Übersicht über die zeitliche Aufeinanderfolge der Arbeitspakete.

|                                    | 2014  | 2015  |       |       |       | 2016  |       |       |       |       | 2017 |
|------------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|
|                                    | 10-12 | 01-03 | 04-06 | 07-09 | 10-12 | 01-03 | 04-06 | 07-09 | 10-12 | 01-03 |      |
| 3.1 Materialbeschaffung            |       |       |       |       |       |       |       |       |       |       |      |
| 3.2a NGS                           |       |       |       |       |       |       |       |       |       |       |      |
| 3.2b NGS Analyse                   |       |       |       |       |       |       |       |       |       |       |      |
| 3.2c Markerentwicklung             |       |       |       |       |       |       |       |       |       |       |      |
| Validierung                        |       |       |       |       |       |       |       |       |       |       |      |
| 3.3a Marker-Transfer auf Holz-DNA  |       |       |       |       |       |       |       |       |       |       |      |
| 3.3b kostengünstige Methodik       |       |       |       |       |       |       |       |       |       |       |      |
| 3.4a Holz für Methoden-Entwicklung |       |       |       |       |       |       |       |       |       |       |      |
| 3.4b Überführung in die Praxis     |       |       |       |       |       |       |       |       |       |       |      |
| 1. MassArray <i>Q. robur</i>       |       |       |       |       |       |       |       |       |       |       |      |
| 2. MassArray <i>Q. mongolica</i>   |       |       |       |       |       |       |       |       |       |       |      |
| 3. <i>Q. petraea</i> NGS           |       |       |       |       |       |       |       |       |       |       |      |
| Auswertung                         |       |       |       |       |       |       |       |       |       |       |      |
| 4. Praxistest                      |       |       |       |       |       |       |       |       |       |       |      |
| 5. Infoblatt                       |       |       |       |       |       |       |       |       |       |       |      |

Die MiSeq Sequenzierung von *Q. petraea* sollte gemäß Antrag für eine kostenneutrale Verlängerung bereits im Oktober 2016 erfolgen. Jedoch gab es technische Probleme aufgrund einer schlechten DNA-Qualität der zu sequenzierenden Proben, so dass mehrfach neu extrahiert werden musste. Die Proben konnten daher erst im Dezember 2016 zur Sequenzierung geschickt werden. Die NGS-Daten lagen Ende Januar 2017 vor. Durch diese Verzögerung ist die Analyse der Daten noch nicht abgeschlossen.

Im Rahmen des Thünen-Kompetenzzentrums Holzherkünfte erstellen wir zurzeit einen Flyer für potentielle Kunden, der neben der Eiche auch andere Baumarten mit enthält. Dieser Übersichts-Prospekt soll etwas umfangreicher werden, daher wird daran im Moment noch gearbeitet. Sobald dieser fertig ist, wird er der DBU zur Verfügung gestellt (Infoblatt unter 5. in Tab. 1).

### 3. Projektergebnisse

#### 3.1 Materialbeschaffung

Im Antrag war bereits aufgeführt, dass umfangreiche Referenzproben diverser *Quercus*-Arten im Thünen-Institut für Forstgenetik verfügbar sind. Die Anzahl an Referenzproben konnte deutlich erweitert werden (Tab. 2). Dies erfolgte entweder durch Aufträge an Dritte oder im Austausch mit Kooperationspartnern. Insbesondere die Einsammlung der europäischen Arten *Q. robur*, *Q. petraea*, *Q. pubescens* und der asiatischen Art *Q. mongolica* konnte auf ein großes Verbreitungsgebiet ausgeweitet werden. Die

amerikanischen Arten wurden von Partnern des im Antrag als Co-Projekt aufgeführten Russia-Far-East Projektes zur Verfügung gestellt.

Die Proben wurden mit einer Thünen-ID versehen und sind in unserer internen Probandenbank mit entsprechenden Angaben, u.a. zu Art und Herkunft erfasst.

**Tab. 2:** Übersicht über die im Thünen-Institut für Forstgenetik verfügbaren Referenzproben. Grün dargestellt sind die Arten bzw. Regionen, die im Laufe des Projekts neu hinzugekommen sind und Zahlen, die im Laufe des Projekts erhöht wurden

| Art                  | Herkunft       | [N]<br>Populationen | [N] Individuen je Pop. | [N] gesamt |
|----------------------|----------------|---------------------|------------------------|------------|
| <i>Q. mongolica</i>  | Ost-Russland   | 48                  | 10 bis 37              | 1520       |
|                      | China          | 3                   | 3 bis 10               | 23         |
|                      | Korea          | 2                   | 6-8                    | 14         |
| <i>Q. dentata</i>    | Russland       | 1                   | 10                     | 10         |
|                      | Korea          | 1                   | 6                      | 6          |
| <i>Q. robur</i>      | Deutschland    | > 30                | 10 bis 30              | > 600      |
|                      | Baschkortostan | 6                   | 9 bis 40               | 140        |
|                      | West-Russland  | 21                  | 10 bis 35              | 571        |
|                      | Ukraine/Krim   | 8                   | 28 bis 34              | 252        |
|                      | Weißrussland   | 8                   | 18 bis 35              | 249        |
|                      | Lettland       | 1                   | 30                     | 30         |
|                      | Finnland       | 4                   | 4 bis 24               | 38         |
|                      | Polen          | 1                   | 10                     | 10         |
|                      | Schweiz        | 1                   | 10                     | 10         |
|                      | Ungarn         | 2                   | 10                     | 20         |
| <i>Q. petraea</i>    | Frankreich     | 6                   | 10 bis 30              | 120        |
|                      | Deutschland    | 7                   | 5 bis 27               | 69         |
|                      | West-Russland  | 2                   | 25 bis 28              | 53         |
|                      | Ukraine/Krim   | 8                   | 30 bis 32              | 242        |
| <i>Q. pubescens</i>  | Deutschland    | 15                  | 10                     | 150        |
|                      | Polen          | 1                   | 20                     | 20         |
|                      | Kroatien       | 1                   | 20                     | 20         |
|                      | Frankreich     | 3                   | 30 bis 58              | 118        |
|                      | West-Russland  | 1                   | 31                     | 31         |
|                      | Ukraine/Krim   | 6                   | 20 bis 58              | 203        |
| <i>Q. macrocarpa</i> | Bot. Gärten    |                     | 1 bis 2                | 12         |
|                      | USA            | 7                   | 1 bis 4                | 14         |
| <i>Q. alba</i>       | Bot. Gärten    |                     | 1 bis 2                | 10         |
|                      | USA            | 7                   | 1 bis 5                | 16         |
| <i>Q. bicolor</i>    | Bot. Gärten    |                     | 1 bis 2                | 7          |
|                      | USA            | 4                   | 1 bis 2                | 6          |
| <i>Q. garryana</i>   | Bot. Gärten    |                     |                        | 1          |
|                      | USA            | 2                   | 1 bis 3                | 4          |
| <i>Q. lyrata</i>     | Bot. Gärten    |                     |                        | 2          |
|                      | USA            | 4                   | 1                      | 4          |
| <i>Q. stellata</i>   | Bot. Gärten    |                     |                        | 2          |
|                      | USA            | 6                   | 1 bis 2                | 8          |
| <i>Q. michauxii</i>  | Bot. Gärten    |                     |                        | 2          |
|                      | USA            | 4                   | 1 bis 2                | 5          |
| <i>Q. prinoides</i>  | USA            | 4                   | 1 bis 3                | 7          |
| <i>Q. acutissima</i> | Korea          | 1                   | 5                      | 5          |
| <i>Q. serrata</i>    | Korea          | 1                   | 5                      | 5          |
| <i>Q. aliena</i>     | Korea          | 1                   | 5                      | 5          |

Für die Beschaffung der asiatischen Arten aus China wurden Kontakte zu Wissenschaftlern aufgenommen. Eine Kooperation, die einen Austausch von Probenmaterial und wissenschaftlichem Know-how zur Grundlage hatte, war bis zum Vertragsentwurf gediehen.

Diese Kooperation beinhaltete neben dem Forschungs-Aufenthalt einer chinesischen Wissenschaftlerin in unserem Institut die Einsammlung von über 1000 Proben von *Q. mongolica* und *Q. liaotungensis*. Leider wurde diese Kooperation seitens der chinesischen Partner vor Vertragsabschluss storniert. Weitere Bemühungen, Probenmaterial aus China zu erhalten, waren leider nicht erfolgreich.

## 3.2 Kontinent- und Artdifferenzierung

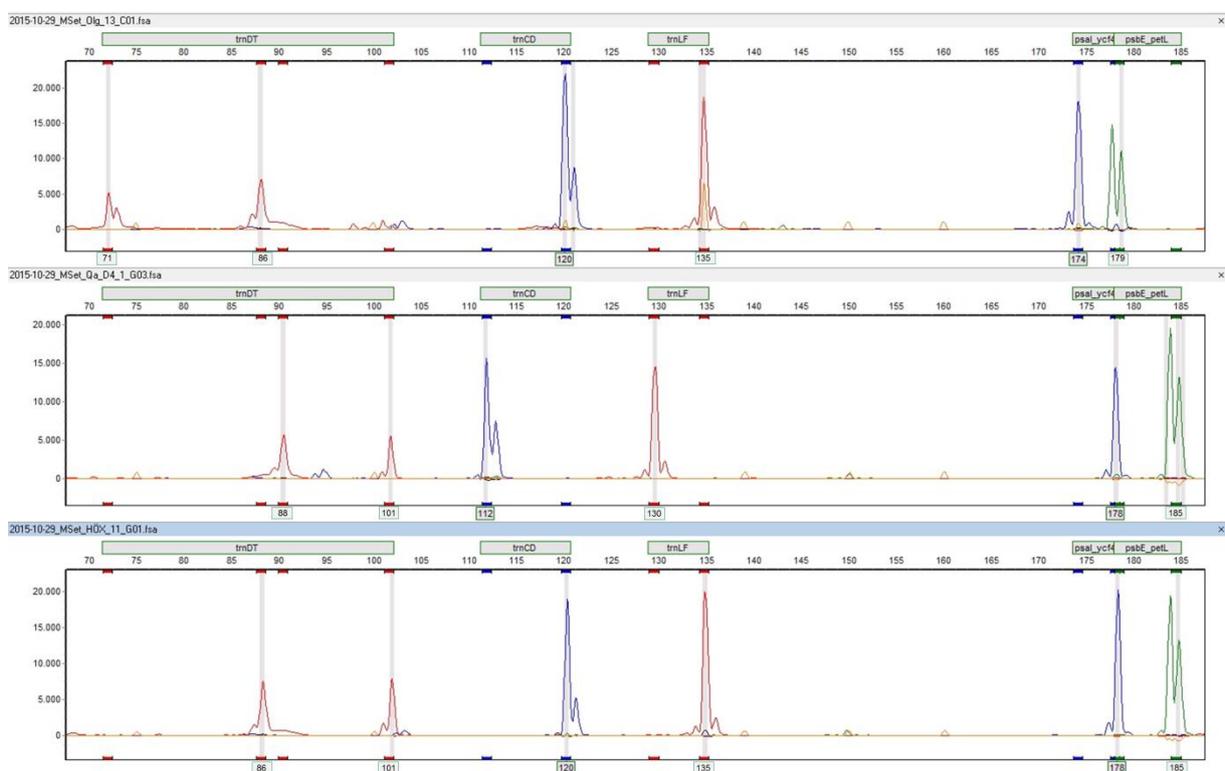
### 3.2.1 Kontinentdifferenzierung

Aus der in 2.1 beschriebenen MiSeq Sequenzierung und SNP-Detektion mit anschließender MassArray Analyse von 120 SNPs (2.2) aus dem Chloroplasten-Genom, die einen 100%igen Unterschied in der Allelfrequenz aufwiesen, ergaben sich 15 SNPs und sechs InDels, die potentiell artspezifisch waren. Die SNPs wurden auf Restriktionsstellen überprüft und es verblieben neun SNPs, die über PCR-RFLP Analysen mit mindestens 10 Individuen je Art getestet wurden. Für fünf dieser SNPs entsprachen die Ergebnisse der PCR-RFLPs den erwarteten Resultaten basierend auf den Sequenzen. Diese SNPs stehen für weitere Tests zur Verfügung und können später in ein Markerset überführt werden.

Die sechs InDels wurden ebenfalls an einer größeren Anzahl Individuen getestet. Hierbei wurden neben den ursprünglich sequenzierten Arten *Q. mongolica* und *Q. robur* auch *Q. dentata* und die amerikanischen Arten *Q. macrocarpa* und *Q. alba* einbezogen. Mit den beiden größten InDels mit vier bzw. fünf Basenpaaren Unterschied können *Quercus*-Arten aus Asien von denen aus Europa bzw. Amerika unterschieden werden. Diese beiden InDels (psal\_ycf4 und psbE\_petL; Abb. 1) wurden mit früheren Ergebnissen (InDels in trnLF und trnCD sowie je ein SNP und ein InDel in trnDT; Abb. 1) kombiniert und in ein Markerset überführt, das für ein Multiplex-Verfahren auf dem ABI-Sequenzierer optimiert wurde. Dieses Markerset enthält fünf InDels und einen SNP und ermöglicht eine kontinentale Differenzierung von Asien, Europa und Amerika (Abb. 1, Schroeder et al. 2016a-c).

Die DNA-Marker „trnCD“ und „trnLF“ unterscheiden zwischen Weißeichen aus entweder Nordamerika oder Eurasien. Bei „trnCD“ basiert diese Unterscheidung auf einem Größenunterschied („InDel“) von 8 Basenpaaren. Markerfragmente, die per PCR aus amerikanischen Eichenproben amplifiziert werden, sind kürzer als die von europäischen oder asiatischen Bäumen. Der Marker „trnLF“ basiert auf einem Größenunterschied der amplifizierten Fragmente von 5 Basenpaaren (die amplifizierten Markerfragmente amerikanischer Weißeichen sind kürzer). Die DNA-Marker „psal-ycf4“ und „psbE-petL“ unterscheiden zwischen Weißeichen, die aus Asien oder aber aus Europa bzw. Nordamerika stammen. Der Marker „psal-ycf4“ basiert auf einem Längenunterschied von 4 Basenpaaren (hier sind amplifizierte Markerfragmente asiatischer Weißeichen kürzer). Der Marker „psbE-petL“ basiert auf einem Größenunterschied von 6 Basenpaaren (amplifizierte Markerfragmente asiatischer Weißeichen sind kürzer). Der Marker „trnDT“ unterscheidet zwischen allen 3 zuvor aufgeführten Kontinenten. Nordamerikanische Weißeichenarten weisen hierbei einen Längenpolymorphismus von 2 zusätzlichen Basenpaaren auf, der sie von eurasischen Arten unterscheidet. Asiatische und Europäische Arten können durch die Analyse eines Einzelnukleotidunterschieds (SNP) unterschieden werden, indem eine sequenzspezifische Endonuclease eingesetzt wird (Schroeder et al. 2016a). Im Thünen-Institut für Forstgenetik ist dieses Set (Abb. 1) bereits an zahlreichen Holzproben getestet und in die Praxis überführt worden (Degen 2017).

In Abb. 1 sind die sechs Marker dargestellt. Die ersten beiden roten Marker entsprechen trnDT, der Marker mit einem InDel und einem SNP, der zwischen allen drei Kontinenten unterscheidet. Hier ist ein typischer Peak bei 71 für Asien plus ein Peak bei 86, der die gleiche Größe hat wie für Europa. Der erste rote Peak für USA ist bei 88, also eindeutig von Asien und Europa zu unterscheiden, gefolgt von einem Peak bei 101 für USA und Europa. Der erste blaue Marker (trnCD) zeigt einen Peak bei 120 bp für Eichen aus Eurasien oder bei 112 bp für Eichen aus Amerika. Der folgende rote Marker (trnLF) differenziert ebenfalls Amerika (Peak bei 130 bp) von Europa und Asien (Peak bei 135 bp). Der dann folgende blaue Marker (psaI\_ycf4) differenziert Asien (Peak bei 174 bp) von Europa und Amerika (Peak bei 178 bp) ebenso wie der letzte, grüne, Marker (psbE\_petL) mit einem Peak bei 179 bp für Asien und bei 184 bp für Europa und Amerika (Abb. 1).

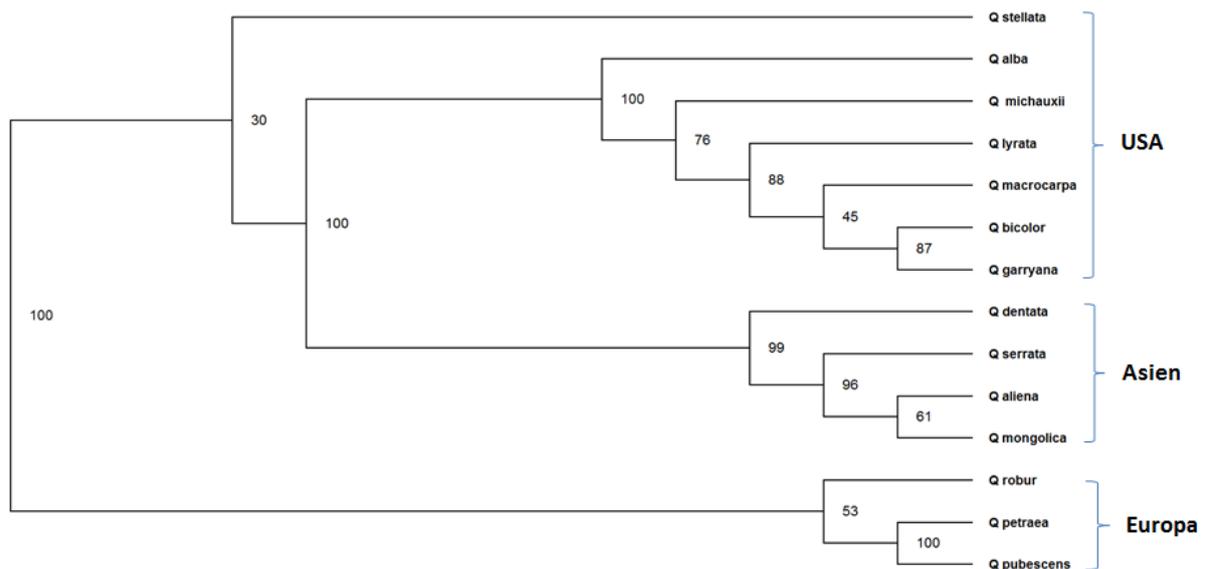


**Abb. 1:** Darstellung der Peaks des Markersets zur Differenzierung der Kontinente Asien (oben), Amerika (Mitte) und Europa (unten). Nähere Erläuterung im Text und in den Publikationen zum Markersset (Schroeder et al. 2016a-c).

In einem Multiplex-Verfahren kann somit mit sechs Markern der kontinentale Ursprung von Proben überprüft werden. Nicht immer gelingt es, wegen der degradierten DNA im Holz, für alle Genmarker ein PCR-Produkt zu bekommen. Daher ist es sinnvoll, mehrere Marker mit derselben Aussage für Holzuntersuchungen zu entwickeln, was in diesem Markersset verwirklicht wurde.

### 3.2.2 Artdifferenzierung über alle drei Kontinente

Für eine größere Auflösung auf Art-Ebene wurden die 233 Kern-SNP-Marker aus der RAD-Sequenzierung nach verschiedenen Qualitätsparametern (Kapitel 2.2.2) selektiert und in dem unter 2.2 beschriebenen MassArray Ansatz an 190 Individuen überprüft. Hierbei wurden die Individuen so ausgewählt, dass sowohl eine Auswahl an SNPs für Artdifferenzierungen als auch eine Herkunftsidentifizierung innerhalb von Arten ermöglicht werden sollte. Von den 233 untersuchten SNPs konnten für 179 SNPs ausreichende Daten für die GDA-NT-Analyse gewonnen werden. Von den 179 SNPs hatten 147 SNPs eine Vollständigkeit der Daten von mindestens 80%. Diese 147 SNPs waren alle polymorph, wobei die genetische Diversität (effektive Anzahl an Allelen) zwischen 1,001 und 1,57 variierte. Die genetische Differenzierung (Delta, Gregorius 1996) lag zwischen 0,001 und 0,460. Alle 147 polymorphen SNPs wurden genutzt, um genetische Abstände zwischen den Arten zu berechnen. Diese genetischen Abstände gingen dann in eine Clusteranalyse mit dem Programm PAST ein (Hammer et al. 2001). Das resultierende Dendrogramm zeigt eine Einteilung der verwendeten Arten nach Kontinenten in drei große Cluster (Abb. 2). Je kürzer die Äste des in Abb. 2 dargestellten Dendrogramms sind, desto ähnlicher sind die Arten einander.



**Abb. 2:** Clusteranalyse anhand von 147 Kern-SNPs zur Differenzierung von 14 Arten aus den USA, Europa und Asien. Die Werte an den Cluster-Übergängen entsprechen „bootstrapping values“, d.h. basierend auf einem Signifikanztest der Gruppierung der einzelnen Taxa zueinander, wobei 100 = 100 %, also sehr zuverlässig bedeutet (z.B. Hillis & Bull 1993).

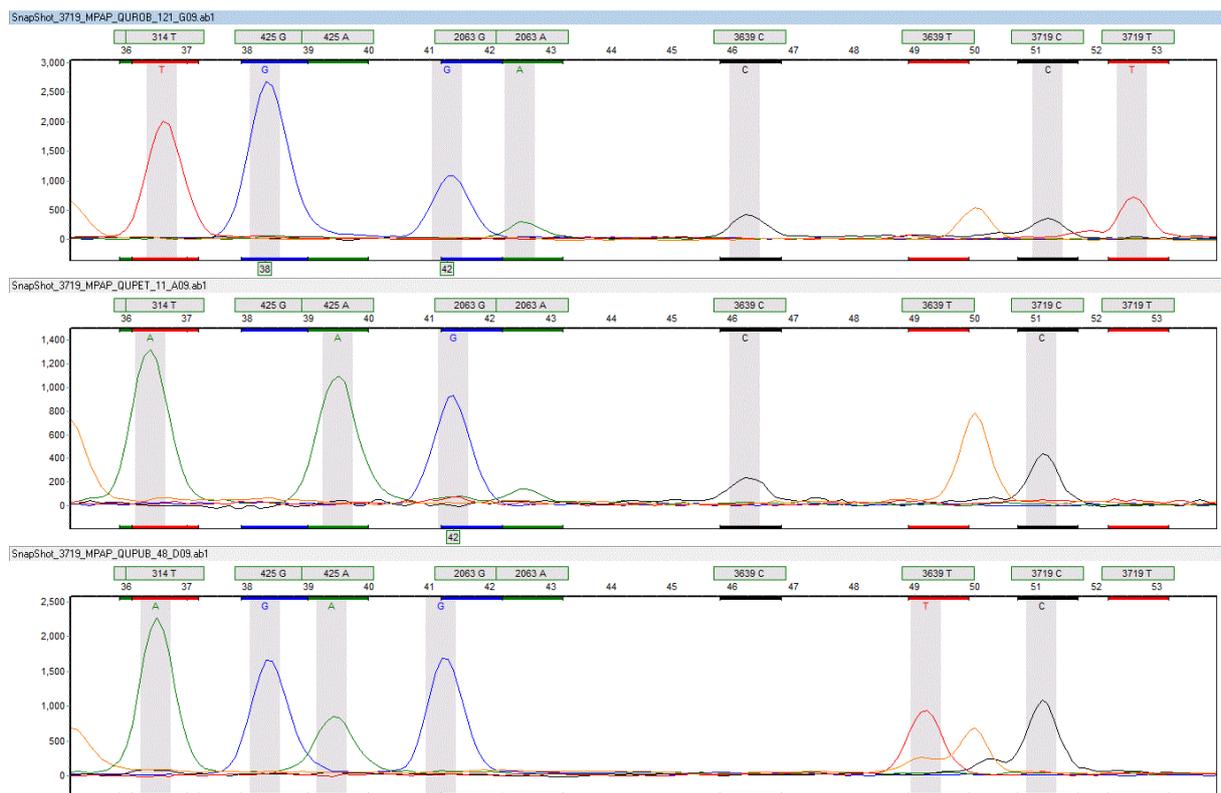
### 3.3 Artdifferenzierung innerhalb Europas

#### 3.3.1 RAD-Marker

Für die Artdifferenzierung innerhalb von Europa (*Q. robur*, *Q. petraea* und *Q. pubescens*) wurde aus den in 3.2.2 beschriebenen Kern-Markern ein weiteres Set von 60 Kern-SNPs nach Qualitätsparametern (Programm GDA\_NT, Degen 2008) ausgewählt, die in ein Screening mit 190 Individuen verteilt auf die drei europäischen Arten überführt wurden.

Aus diesen konnten fünf SNPs ausgewählt werden, die ausreichen, um über ein statistisches Verfahren *Q. robur* von den beiden anderen europäischen Arten mit 94 % Wahrscheinlichkeit zu unterscheiden. Diese fünf Marker wurden in einem SNaPShot Multiplex-Ansatz zusammengefasst. Die SNaPShot-Methode basiert auf „single base primer extension“. Primer, die direkt vor dem zu analysierenden SNP sitzen, werden um ein einzelnes Nukleotid (das den entsprechenden SNP repräsentiert) verlängert. Da diese Nukleotide fluoreszenzmarkiert sind, können die verschiedenen Varianten, die unterschiedliche Farben zeigen, auf einem Kapillarsequenzierer visualisiert werden. Auf diese Weise können die fünf nukleären SNP-Marker, die aus dem größeren Satz von 60 Markern ausgewählt wurden, in einem Ansatz analysiert werden (Abb. 3).

Das Kern-Genom ist diploid, liegt also in doppelter Ausfertigung vor. Daher ergeben sich bei der Analyse von Kern-Markern für jedes Individuum und jeden Marker zwei Allele. Diese können gleich sein (homozygot) oder unterschiedlich (heterozygot). In Abb. 3 sind daher für jeden der fünf Marker jeweils zwei mögliche Allelvarianten in einem sogenannten „Panel“ vorgegeben (oberhalb der Peaks: 314A, 314T; 425G, 425A; 2063G, 2063A; 3639C, 3639T und 3719C, 3719T). Die dargestellten Muster für die drei Arten in Abb. 3 unterscheiden sich deutlich, jedoch sind diese Muster innerhalb der Arten variabel, so dass, wie oben beschrieben, nur über einen statistischen Ansatz eine Differenzierung verwirklicht wird.



**Abb. 3:** Multiplex der fünf nukleären Marker nach der SNaPShot Methode zur Differenzierung von *Q. robur* (QUROB, oben) gegenüber *Q. petraea* (QUPET, Mitte) und *Q. pubescens* (QUPUB, unten).

### 3.3.2 MiSeq Sequenzierung von *Q. petraea*

Der sehr umfangreiche NGS-Datensatz aus der *Q. petraea* Sequenzierung wurde zunächst - parallel zu den Daten des *Q. robur* Pools - auf die Referenz-Scaffolds von *Q. robur* (Plomion

et al. 2016) gemappt (angeordnet). Aus diesen Mappings wurden nur solche SNPs bzw. InDels gefiltert, welche im Mapping der *Q. petraea*-Daten eine Frequenz von 100% zeigten, wobei im parallelen Mapping der *Q. robur*-Daten keinerlei SNP an derselben Position im Scaffold erkennbar sein durfte (0%-Frequenz des varianten Allels). Daraus ergaben sich noch fast 5000 SNPs und InDels. Von diesen wurden bisher 26 ausgewählt, flankierende Primerpaare entwickelt und die PCR-Amplifikation der entsprechenden Regionen getestet. Nach der ersten Testphase sind zurzeit noch 13 der 26 Primerpaare in der Optimierungsphase. Gut amplifizierbare SNP/InDel-Regionen werden zeitnah in einem erweiterten Set von Individuen hinsichtlich ihrer Eignung zur Differenzierung von *Q. petraea* von den anderen europäischen Weißeichenarten getestet.

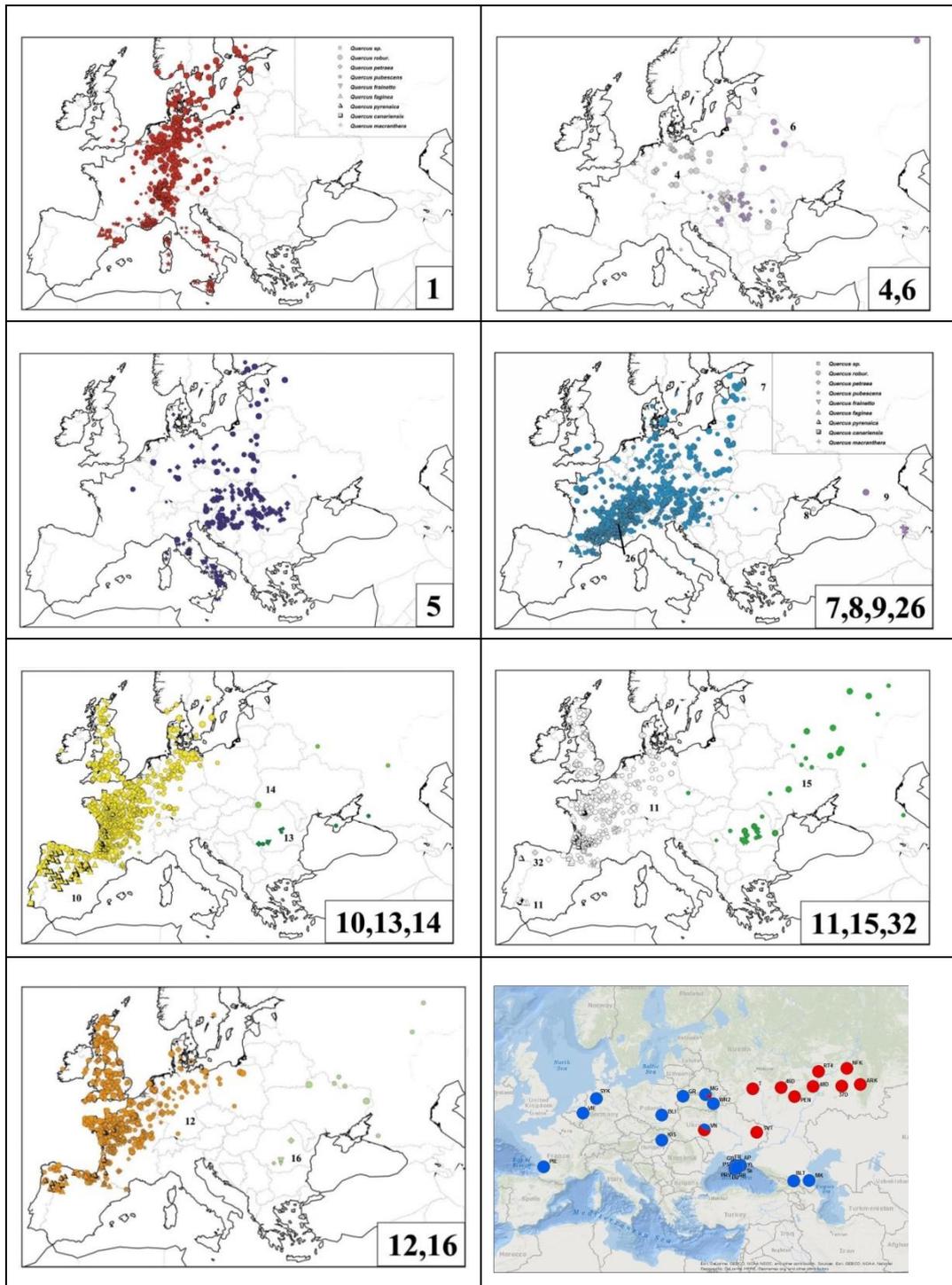
### 3.4 Herkunftsidifizierung

#### 3.4.1 Herkunftsidifizierung innerhalb von Europa

Basierend auf den von Petit et al. (2002 a, b) identifizierten Haplotypen von *Q. robur* innerhalb von Europa und nachfolgenden Anpassungen dieser Marker auf Holz (Deguilloux et al. 2003) wurden in früheren Projekten bereits Sequenzierungen von Chloroplasten-Abschnitten vorgenommen. Die so gesammelten Informationen wurden in dem laufenden Projekt genutzt, um Marker-Kombinationen zu erstellen und zu optimieren, die eine Zuordnung von (Holz-)Proben zu Regionen innerhalb Europas erlauben. Hierbei wurden zunächst, anhand von Sequenzierungen von drei Chloroplasten-Fragmenten, insgesamt sechs SNPs, ein InDel und ein Mikrosatellit zur Analyse herangezogen. Diese Vorgehensweise wurde optimiert, so dass eine Kombination aus Sequenzierung und Restriktionsanalyse mit anschließender Auftrennung der Fragmente auf dem Sequenzierer möglich ist.

Für die unter 2.2 beschriebene MassArray Analyse von 120 SNPs aus der MiSeq Sequenzierung wurden auch 22 *Q. robur* Individuen verwendet. Aus diesen 120 SNPs wurden nach der Qualitätsanalyse 30 plastidäre SNPs ausgewählt und gemeinsam mit 30 Kern-SNPs aus der MassArray Analyse der RAD-Sequenzierung für ein weiteres Screening mit 570 *Q. robur* Individuen verwendet. Aus diesem gemeinsamen screening von Kern- und plastidären Markern ergaben sich neun weitere plastidäre SNPs, von denen einige eine detailliertere Auftrennung zwischen westlichen und im östlichen Verbreitungsgebiet liegenden Populationen ermöglicht. Diese neun Marker wurden anhand vieler Individuen validiert. Fünf der Marker konnten erfolgreich amplifiziert werden und ergaben zu 100% das erwartete Ergebnis. Von diesen wurden zwei ausgewählt, mit deren Hilfe Eichen aus dem Gebiet der Ukraine bis zum Ural („UU“) von den übrigen, bereits bekannten und validierten, europäischen Haplotypen abgegrenzt werden können. Aus diesen bereits etablierten und den neuen SNPs / Indels wurden die aussagekräftigsten in einem Multiplex-Set zusammengefasst. Dieses Set umfasst fünf SNPs und zwei InDels, mit deren Hilfe eine Differenzierung in sieben Haplotypengruppen möglich ist (Tab. 3). Mit weiteren identifizierten Markern ist die Auftrennung auch innerhalb der Haplotypengruppen möglich, da sich jedoch die Verbreitungsgebiete der Haplotypen innerhalb der Gruppen stark überschneiden, ist eine weitere Auftrennung nicht notwendig (Abb. 4). Daher wurde das Markerset nach der Devise hergestellt, eine möglichst hohe Auflösung mit der geringst möglichen Zahl an PCRs und Restriktionen zu erzielen, um die Anwendung des Sets so effektiv wie möglich zu gestalten.

Mit dieser Analyse können Herkünfte von *Q. robur* den Haplotypen zugeordnet werden und damit von Auftragnehmern vorgegebene Deklarationen abgelehnt oder bestätigt werden.



**Abb. 4:** Verbreitungskarten der häufigsten Haplotypen (HT) in Europa (Petit et al. 2002b). HT 1 (rot) ist in Mittel- und Teilen von Westeuropa verbreitet. HT 5 (dunkel lila) ist häufiger in Osteuropa zu finden. Die Haplotypen 4 (grau) und 6 (hell lila) sind seltene Haplotypen, wobei HT 4 in Mittel- und HT 6 in Osteuropa vorkommt. HT 7 (blau) ist der am weitesten verbreitete Haplotyp und kommt von Westeuropa (ohne Iberische Halbinsel) bis nach Osteuropa vor. Die drei Haplotypen 10 (gelb), 11 (weiß) und 12 (orange) haben fast dasselbe Verbreitungsgebiet in Westeuropa (Portugal, Spanien, Frankreich, Großbritannien und Deutschland). Unten rechts sind die neu gefundenen Haplotypen „non\_UU“ in blau und „UU“ in rot abgebildet. Der Haplotyp „UU“ (rot) hat sein Verbreitungsgebiet außerhalb der anderen bisher erfassten Haplotypen.

**Tab. 3:** Zusammenstellung der Marker, die in einem Multiplex-Ansatz zur Differenzierung von *Q. robur* innerhalb Europas verwendet wird. Die Nummerierung der Haplotypen (HT) erfolgt nach Petit et al. (2002). Die Haplotypen „non\_UU“ (= nicht Ukraine/Ural) und „UU“ (= Ukraine/Ural) sind aus den neuen Markern entstanden.

| Haplotyp      | dt72_for | QT5_rev | QT7 | UU_15  | UU_48 |
|---------------|----------|---------|-----|--------|-------|
| HT 1          | 200      | 78      | 100 |        |       |
| HT 4, 5y      | 203      | 88      | 100 |        |       |
| HT 5x, 6      | 203      | 83      | 100 |        |       |
| HT 7          | 203      | 83      | 174 |        |       |
| HT 10, 11, 12 | 142      | 83      | 100 |        |       |
| non_UU        |          |         |     | 70/102 | 160   |
| UU            |          |         |     | 172    | 79/81 |

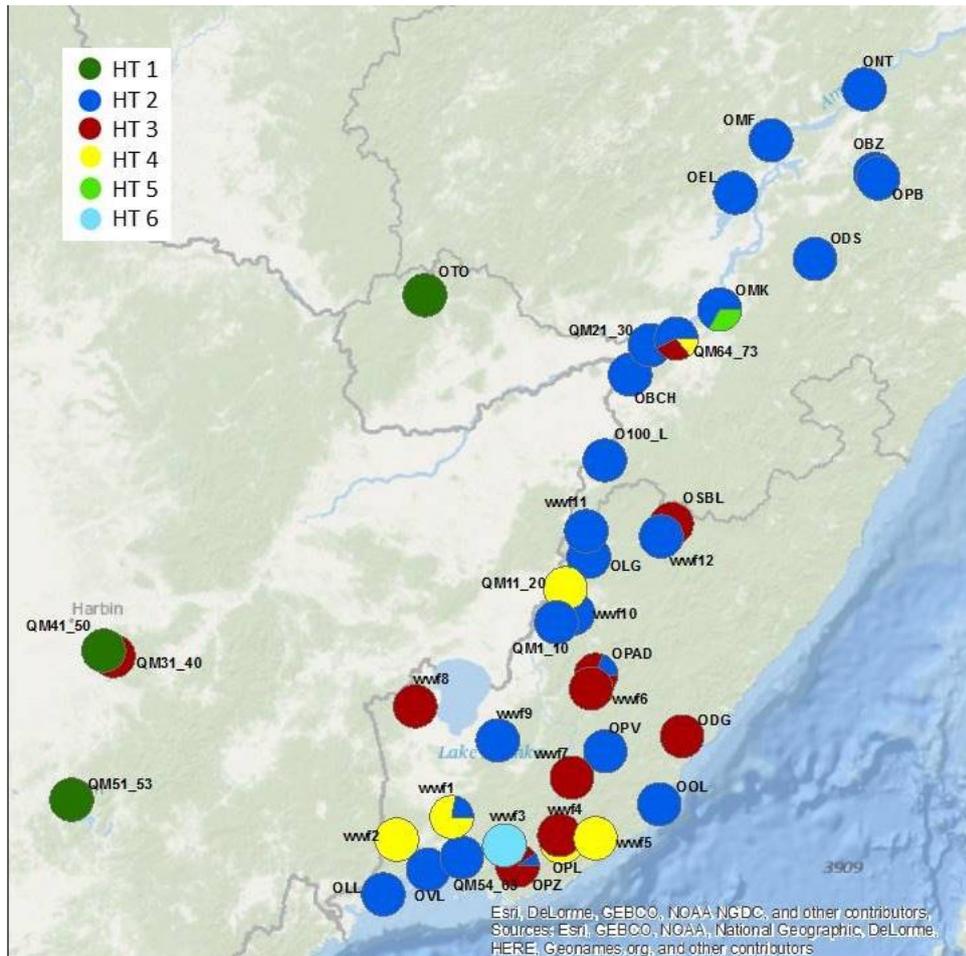
### 3.4.2 Herkunftsidentifizierung innerhalb von Asien

Die MassArray Analyse von 120 cp und mt SNPs aus der MiSeq Sequenzierung enthielt unter anderem auch 20 Individuen von *Q. mongolica* aus 10 Populationen. Aus diesen 120 SNPs wurden nach Qualitätsparametern 27 ausgewählt, die für ein weiteres MassArray screening mit 380 Individuen von *Q. mongolica* eingesetzt wurden, um nach einer höheren Auflösung innerhalb von Asien zu suchen. Hieraus ergaben sich 11 SNPs und ein InDel, die anhand von PCR-RFLP-Analysen überprüft wurden. Zusammen mit den im Antrag bereits aufgeführten SNPs und InDels wurden diese potentiellen Marker weiter getestet und daraus ein Markerset bestehend aus fünf SNPs zusammengestellt (Tab. 4), die anhand von PCR-RFLPs (Restriktionen) als Multiplex-Set auf dem Sequenzierer analysiert werden können. Mit diesen fünf Markern konnten sechs Haplotypen identifiziert werden. Jede Hinzunahme weiterer cp oder mt Marker erhöhte die Auflösung nicht, führte also nicht zu weiteren Haplotypen. Mit diesem Multiplex-Set wurden 40 Populationen von *Q. mongolica* mit 3 bis 10 Individuen pro Population analysiert (Abb. 5, Schröder et al. 2016c).

Die Ergebnisse (Abb. 5, Schroeder et al. 2016c) zeigen, dass Haplotyp 2 (blau in Abb. 5) zwar der häufigste Typ ist, aber in den westlichen Populationen nicht auftritt. Hingegen kommt der seltene Haplotyp 1 (grün in Abb. 5) nur in den westlich gelegenen Populationen vor. Haplotyp 4 (gelb in Abb. 5) ist deutlich häufiger in den südlichen Populationen zu finden und kommt ganz im Norden nicht vor. Der in Abb. 5 rot dargestellte Haplotyp 3 zeigt ein eher verstreutes Auftreten, ist aber im Süden häufiger. Die beiden Haplotypen 5 und 6 sind private Haplotypen, d.h. sie sind jeweils nur in einer Population gefunden worden.

**Tab. 4:** Zusammenstellung der Marker, die in einem Multiplex-Ansatz zur Differenzierung von *Q. mongolica* innerhalb Asiens verwendet werden. Die für die Haplotypen angegebenen Farben entsprechen denen in Abb. 5.

| Haplotyp        | 4067    | 32153  | 60606   | 91100    | 43824 |
|-----------------|---------|--------|---------|----------|-------|
| HT 1 (grün)     | 97 / 96 | 177    | 87 / 77 | 71 / 114 | 125   |
| HT 2 (blau)     | 193     | 71/106 | 164     | 184      | 125   |
| HT 3 (rot)      | 193     | 177    | 87/77   | 71/114   | 125   |
| HT 4 (gelb)     | 193     | 177    | 87/77   | 184      | 125   |
| HT 5 (hellgrün) | 193     | 71/106 | 164     | 184      | 63/62 |
| HT 6 (hellblau) | 193     | 71/106 | 87/77   | 184      | 125   |

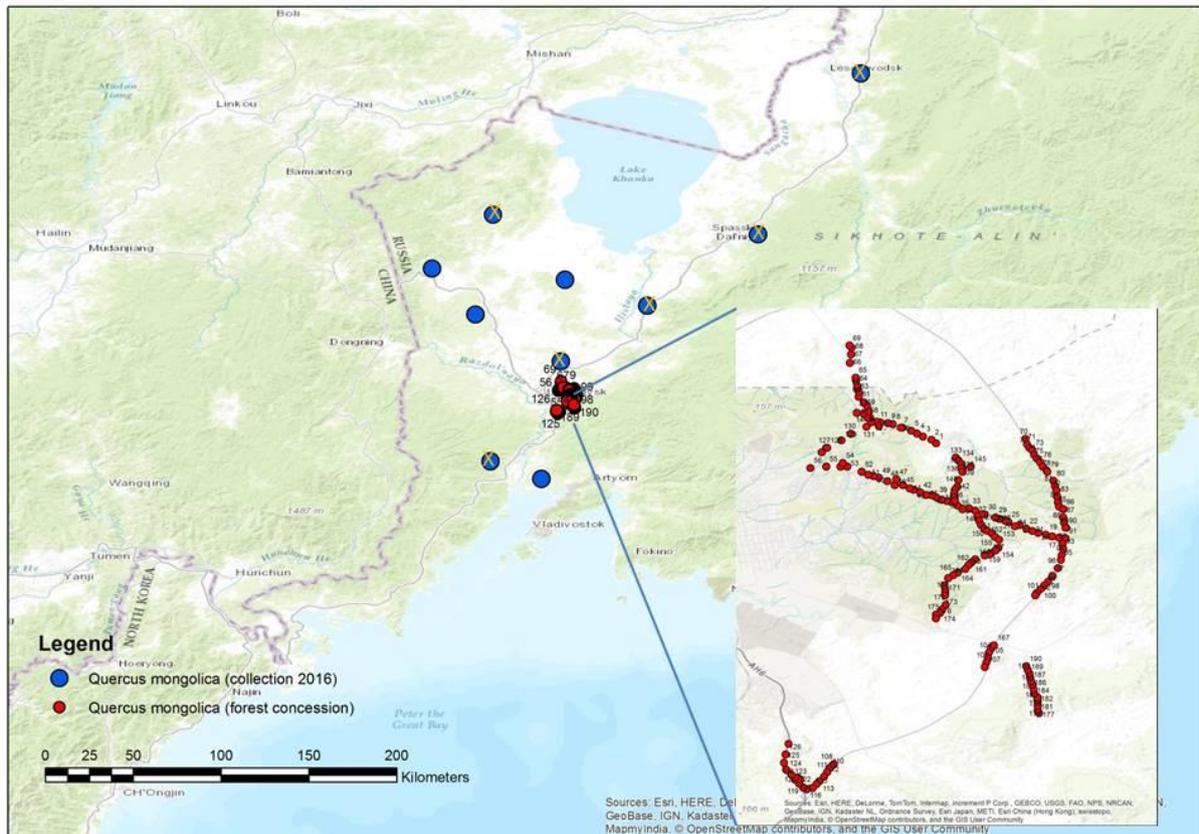


**Abb. 5:** Verteilung der mit Chloroplasten- und Mitochondrien-Markern identifizierten Haplotypen innerhalb von *Q. mongolica* in 40 Populationen (Schroeder et al. 2016c; genauere Erklärungen im Text oben).

Für die unter 2.2 beschriebene MassArray Analyse von 233 SNPs aus der RAD-Sequenzierung wurden neben zahlreichen anderen Arten 90 *Q. mongolica* Individuen aus 30 Populationen verwendet. Auch hier wurde das Ergebnis des MassArray Ansatzes statistisch mit Hilfe des Programms „GDA-NT“ (Degen 2008) analysiert, allerdings wurden explizit nur die *Q. mongolica* Individuen in diese Analyse einbezogen. Nach der Qualitätsanalyse wurden 30 SNPs selektiert und mit 760 *Q. mongolica* Individuen aus allen zur Verfügung stehenden 40 Populationen eine MassArray Analyse durchgeführt. Auch dieses Screening erhöhte die Auflösung zwischen den Populationen nicht weiter.

Trotz der geringen Variation und der daraus resultierenden schwachen Auflösungskapazität von cp-, mt- und Kernmarkern zwischen Populationen von *Q. mongolica* wurde eine weitere Analyse durchgeführt, um zu testen, ob in sehr kleinem Maßstab auf Konzessionsebene eine Differenzierung möglich ist. Aufgrund von Familienstrukturen in natürlichen Populationen können auf einer kleinen geographischen Skala häufig eher Unterschiede gefunden werden als auf einer sehr großen Skala. Zu diesem Zweck sind im Mai 2016 an der Russisch-

Chinesischen Grenze in zehn weiteren Populationen von je 30 Eichen Proben genommen worden. Außerdem wurden in einem Gebiet von 15 x 20 km Blätter von weiteren 190 Eichen entnommen (Abb. 6). Die bereits zuvor selektierten 27 cp\_mt SNPs wurden für ein Screening (MassArray-Analyse) mit den 190 Eichen aus der Konzession sowie weiteren 190 Eichen aus sechs Referenzpopulationen in einem Gradienten von 20 bis 250 km verwendet (Abb. 6).



**Abb. 6:** Karte der 10 in 2016 neu beprobten Populationen und der Konzession bestehend aus 190 Eichenindividuen. Die sechs mit einem gelben Kreuz markierten Populationen wurden für die MassArray-Analyse verwendet.

Mit Hilfe des Programms GeoAssign (Degen et al. 2017) wurden die Daten des Screenings ausgewertet, um zu analysieren wie korrekt eine Zuordnung der Individuen zu der Konzession möglich sind. Die Analyse ergab eine korrekte Selbst-Zuordnung (Self-assignment) zu der Konzession von 86%. Dieses Ergebnis bedeutet, dass mit den entwickelten Markern eine Zuordnung von Individuen zu Konzessionen / Populationen auf sehr kleinem Maßstab mit hoher Wahrscheinlichkeit möglich ist.

### 3.5 Praxistest

Für den Praxistest sind insgesamt 30 Holzproben von je zwei asiatischen und europäischen Arten sowie je eine Holzprobe von sechs verschiedenen amerikanischen Arten verwendet worden. Die Holzproben der amerikanischen Arten wurden von dem Arboretum des Thünen-Instituts in Bergedorf zur Verfügung gestellt, sechs Holzproben von *Q. petraea* kamen von

der LANUV aus Arnberg, je sechs Proben von *Q. mongolica* und *Q. dentata* wurden eigens für den Praxistest von dem russischen Kollegen Yulai Yanbaev eingesammelt und die verbleibenden sechs Holzproben von *Q. robur* wurden aus dem Arboretum in Großhansdorf entnommen. Die 30 Proben wurden von einer technischen Mitarbeiterin präpariert und anschließend einer wissenschaftlichen Mitarbeiterin, die nicht in das Projekt involviert ist, zur Randomisierung übergeben. Die technische Mitarbeiterin extrahierte die DNA und wendete das Markerset zur Überprüfung des kontinentalen Ursprungs an. Die Auswertung erfolgte von der im Projekt beschäftigten Wissenschaftlerin. Die Ergebnisse wurden dann der Kollegin übergeben, die zuvor die Randomisierung vorgenommen hatte, um die Richtigkeit der Ergebnisse zu überprüfen. 29 der 30 Proben wurden dem korrekten Kontinent zugeordnet. Eine als amerikanische Art deklarierte Holzprobe wurde als europäisch identifiziert und bei Nachfrage der Kollegen aus dem Arboretum in Bergedorf war dies tatsächlich keine amerikanische, sondern eine europäische Art, so dass die Zuordnung im Praxistest zu 100% korrekt erfolgt ist.

#### 4. Öffentlichkeitsarbeit / Veröffentlichungen / Vorträge

##### 4.1 Öffentlichkeitsarbeit

In den letzten Jahren ist das Holz-Kompetenzzentrum (<https://www.thuenen.de/de/infrastruktur/thuenen-kompetenzzentrum-holzherkuenfte/>) immer mehr in den Fokus der Öffentlichkeit gekommen. Auch Medienanstalten wurden auf uns aufmerksam und kamen ins Institut, um Beiträge zu drehen (Tab. 5).

**Tab. 5:** Besuche von Medienvertretern zur Erstellung von Fernsehbeiträgen

| Datum         | Mediananstalt                   | Titel Beitrag   |
|---------------|---------------------------------|---|
| April 2015    | ZDF zoom                        | Ikea, Höffner und Co.<br>Woher kommen unsere Billigmöbel? |
| Dezember 2016 | NDR, Schleswig-Holstein Magazin | Forscher arbeiten an weltgrößter<br>Baumdatenbank         |
| März 2017     | Deutsche Welle                  | Germany's wood detectives                                 |

Des Weiteren hat der FSC Deutschland am 29. und 30.06.2016 seine jährliche Sitzung in Hamburg abgehalten. In diesem Zuge haben wir am 30.06.2016 im Thünen-Institut für Forstgenetik eine Vortragsveranstaltung mit etwa 30 Teilnehmern aus der Industrie, der Wissenschaft und den Umweltverbänden gemeinsam mit dem FSC organisiert, auf dem wir auch unsere Arbeiten zur Kontrolle des internationalen illegalen Holzhandels vorgestellt haben.

Vom 19. bis 29.01.2017 fand in Berlin die Grüne Woche statt. Gemeinsam mit dem BMEL und den Kollegen des Thünen-Instituts für Holzforschung haben wir einen Stand betreut, auf dem wir dem Publikum unsere Möglichkeiten der Identifizierung von Holzarten und deren Herkunft demonstriert haben.

## 4.2 Vorträge / Poster

**Tab. 6:** Besuchte Veranstaltungen mit Posterbeiträgen oder Vorträgen zu dem Thema und zu den Arbeiten des Projekts

| Datum                    | Veranstaltung   | Ort                   | Thema   | Person          |
|--------------------------|---|-----------------------|---|-----------------|
| 30.11. bis<br>02.12.2014 | Treffen Global Timber Tracking Network  | Rom,<br>Italien       | Application of DNA fingerprints to control tree species and geographic origin of timber     | Bernd Degen     |
| 04.12.2014               | Meeting of The British Timber Trade Federation                                | London,<br>UK         | Application of DNA fingerprints to control tree species and geographic origin of timber     | Bernd Degen     |
| 10.12. bis<br>11.12.2014 | Conference of The International Consortium on Combating Wildlife Crime (ICWC) | Wien,<br>Austria      | Application of DNA fingerprints to control tree species and geographic origin of timber     | Bernd Degen     |
| 12.06.2015               | Jahrestagung Gesamtverband Deutscher Holzhandel (GD-Holz)                     | Düsseldorf            | Genetische Holzarten- und Herkunftsbestimmung   | Bernd Degen     |
| 09.05. bis<br>11.05.2016 | 4th Plant Genomics Congress   | London,<br>UK         | NGS-based development of molecular markers for determining continental origin of white oaks | Hilke Schroeder |
| 15.06. bis<br>16.06.2016 | Tagung der Sektion Forstgenetik   | Chorin                | Anwenderfreundliche DNA-Marker zur Herkunftsidentifizierung von Eichenholz                  | Hilke Schroeder |
| 21.06.2016               | Deutsch-Chinesischer Workshop zur Holzforschung                               | Hamburg               | Application of DNA fingerprints to control tree species and geographic origin of timber     | Bernd Degen     |
| 15.12.2016               | Plant genetics colloquium and forest genetics seminar                         | Universität Göttingen | GeoAssign: A new genetic approach to assign individuals to species and populations          | Bernd Degen     |
| 08.06.2017               | CITES workshop  | Bergedorf             | Application of DNA fingerprints to control tree species and geographic origin of timber     | Hilke Schroeder |

## 4.3 Publikationen

Schroeder H., Cronn R., Yanbaev Y., Jennings T., Mader M., Degen B., Kersten B. (2016a).

Development of easy-to-use molecular markers for differentiation within the genus *Quercus*. PLoS One 11 (6): e0158221, DOI:10.1371/journal.pone.0158221

Schroeder H., Cronn R., Yanbaev Y., Jennings T., Mader M., Degen B., Kersten B. (2016b).

NGS-based development of molecular markers for determining continental origin of white oaks. 4th Plant Genomics Congress, London, Volume: Poster abstracts

Schroeder H., Degen B., Kersten. (2016c). Anwenderfreundliche DNA-Marker zur

Herkunftsidentifizierung von Eichenholz. In: *Forstgenetik und Naturschutz. 5. Tagung der Sektion Forstgenetik/Forstpflanzenzüchtung am 15./16. Juni 2016 in Chorin*. Liesebach M. (ed.). Thünen-Report 45: 66-73

## 5. Fazit / Ausblick

Die durchgeführten Sequenzierungen (MiSeq und RAD) haben eine große Anzahl an Variationen (SNPs und InDels) ergeben, die über weiterführende Tests als praxistaugliche Marker etabliert werden konnten. Marker, die in ihrer Art- oder Herkunfts-Spezifität bestätigt werden konnten, sind so in Markersets überführt worden. So befinden sich ein Set aus Markern zur Unterscheidung der kontinentalen Herkunft sowie zwei weitere Sets zur Differenzierung von Haplotypen in Europa und in Asien optimiert für Holzproben in der praktischen Anwendung.

Die geringe Auflösung in den Populationen von *Q. mongolica* liegt vermutlich an einer insgesamt geringeren Variation in den ausgewählten Populationen dieser Art verglichen mit den europäischen oder amerikanischen Arten. Für die Chloroplasten- und Mitochondriengenome wurden mit derselben Herangehensweise deutlich mehr SNPs und InDels innerhalb von *Q. robur* identifiziert als innerhalb von *Q. mongolica*. Zum Vergleich, japanische Wissenschaftler haben 33 Populationen von *Q. mongolica* (var. *crispula*) in Japan mit cp-Markern untersucht und eine deutlich höhere Anzahl an Haplotypen identifizieren können (Okaura et al. 2007). Japan war während der letzten Eiszeit ein Refugium für viele Baumarten, u.a. auch für die mongolische Eiche. Refugialgebiete haben häufig eine hohe genetische Variation (Leroy & Arpe 2007). Mit der Rückwanderung sind vermutlich nur wenige der im Refugium vorhandenen Haplotypen bis in den Osten Russlands vorgedrungen. Dies bewirkt die vorgefundene geringe genetische Variation. Auch die Analyse der Variationen im Kerngenom zeigt dieselbe Tendenz. Hier ergaben sich aus insgesamt 179 nach Qualitätsparametern gefilterten Variationen für *Q. robur* 164 polymorphe Stellen (SNPs oder InDels) und für *Q. mongolica* nur 49. Die geringe beobachtete Variation innerhalb der russischen und chinesischen Populationen von *Q. mongolica* erschwerte die Suche nach geeigneten populations-differenzierenden Markern. Die Analysen in diesem Bereich wurden daher weiter intensiviert und auf SNPs des Zellkerns konzentriert. Jedoch erbrachten auch diese keine höhere Auflösung, so dass inzwischen davon ausgegangen werden kann, dass die Variation innerhalb von *Q. mongolica* in den untersuchten Regionen sehr gering ist.

Während des Projekts konnte eine Referenzdatenbank aufgebaut werden, die inzwischen Proben von über 5000 Individuen von insgesamt 16 *Quercus*-Arten enthält. Diese wurden für die Validierung der neu entwickelten Marker eingesetzt. Durch die Zugehörigkeit des Instituts für Forstgenetik zum Holzkompetenzzentrum des Thünen Instituts gehen ständig Prüfaufträge zur Analyse von Holzproben von Weißeichen ein. Somit ist die Überführung in die Praxis erfolgt. Die Validierung der SNPs/InDels von *Q. petraea* (aus der MiSeq Sequenzierung) wird weiter fortgesetzt, um weitere Marker für die Unterscheidung der europäischen Eichenarten zu finden.

## 6. Literaturangaben

- Degen B (2008). GDA\_NT 2.0: Genetic data analysis and numerical tests. Available from [bernd.degen@ti.bund.de](mailto:bernd.degen@ti.bund.de)
- Degen (2017). DNA-Tests an Holzproben im Thünen-Institut. Holz-Zentralblatt 24: 553
- Degen B, Blanc-Jolivet, Stierand K, Gillet E (2017). A nearest neighbour approach by genetic distance to the assignment of individual trees to geographic origin. Forensic Science International: Genetics 27: 132-141
- Degen B, Höltnen AM (2011). DNA-Methoden zur Kontrolle der Holzart und Holzherkunft. Holz-Zentralblatt 19: 461
- Degen B, Ward SE, Lemes MR, Navarro C, Cavers S, Sebbenn AM (2013). Verifying the geographic origin of mahogany (*Swietenia macrophylla* King) with DNA-fingerprints. Forensic Science International: Genetics 7(1): 55-62
- Deguilloux MF, Pemonge MH, Bertel L, Kremer A, Petit RJ (2003). Checking the geographical origin of oak wood: molecular and statistical tools. Mol Ecol 12: 1629-1636
- Gregorius H-R (1996). Differentiation between populations and its measurement. Acta Biotheoretica 44: 23-36.
- Hammer, Ø., Harper, D.A.T., and P. D. Ryan, 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4(1): 9pp.
- Hillis DM, Bull JJ (1993). An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Systematic Biology 42 (2): 182-192
- Hollingsworth PM (2011). Refining the DNA barcode for landplants. Proceedings of the National Academy of Science USA 108 (49): 19451-19452
- Hollingsworth PM, Forrest LL, Spouge JL (2009). A DNA barcode for landplants. Proceedings of the National Academy of Science USA 106: 12794-12797
- Kress WJ, Erickson DL (2007). A two-locus DNA barcode for plants: the coding region rbcL gene complements the non-coding trnH-psbA spacer regions. PLOS one 6: e508
- Krusche D, Geburek T (1991). Conservation of forest gene resources as related to sample size. Forest Ecology and Management 40: 145-150
- Lahaye R, Van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Dutoit S, Barraclough TG, Savolainen V (2008). DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Science USA 105 (8): 2923-2928
- Leroy SAG, Arpe K (2007) Glacial refugia for summer-green trees in Europe and south-west Asia as proposed by ECHAM3 time-slice atmospheric model simulation. Journal of Biogeography 34:2115-2128
- Okaura T, Quang ND, Ubukata M, Harada K (2007). Phylogeographic structure and late Quaternary population history of the Japanese oak *Quercus mongolica* var. *crispula* and related species revealed by chloroplast DNA variation. Genes & Genetic Systems 82: 465-477
- Petit RJ, Csaikl UM, Bordacs S et al. (2002a). Chloroplast DNA variation in European white oaks phylogeography and patterns of diversity based on data from over 2.600 populations. Forest Ecology and Management 156: 5-26
- Petit RJ, Brewer S, Bordacs S et al. (2002b). Identification of refugia and postglacial colonization routes of European white oaks based on chloroplast DNA and fossil pollen evidence. Forest Ecology and Management 156: 49-74
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004). GENECLASS2: A software for genetic assignment and first-generation migrant detection. Journal of Heredity 95: 536-539

- Plomion C, Aury J-M, Amselem J et al. (2016). Decoding the oak genome: public release of sequence data, assembly, annotation and publication strategies. *Molecular Ecology Resources* 16: 254-265
- Simeone MC, Piredda R, Papini A, Vessella F, Schirone B (2013). Application of plastid and nuclear markers to DNA barcoding of Euro-Mediterranean oaks (*Quercus*, Fagaceae): problems, prospects and phylogenetic implications. *Botanical Journal of the Linnean Society* 172 (4): 478-499
- Taberlet P, Coissac E, Pompanon F et al. (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Res* 35: e14