

# Genotypisierung mit nuklearen Mikrosatellitenmarkern – Möglichkeiten der Datenauswertung am Beispiel von Buchenpopulationen (*Fagus sylvatica* L.) aus einem Herkunftsversuch

Heike Liesebach\*

## Zusammenfassung

Bei sechs ausgewählten Herkünften (zwei aus Deutschland, je eine aus Österreich, Tschechien, Rumänien und Spanien) aus dem 1995 angelegten Internationalen Buchenherkunftsversuch (*Fagus sylvatica* L.) wurden in Ergänzung zu umfangreichen phänotypischen Datenerhebungen Stichproben von je 50 bis 55 Individuen mit acht nuklearen Mikrosatellitenmarkern genotypisiert.

Die festgestellten teilweise starken Abweichungen der Genotypenstrukturen vom Hardy-Weinberg-Gleichgewicht wurden auf zwei wesentliche Ursachen zurückgeführt: 1. auf eine in allen 6 Herkunftsstichproben vorhandene Struktur aus zwei oder drei Subpopulationen, die ihrerseits im Hardy-Weinberg-Gleichgewicht stehen, und 2. auf das Vorkommen von Nullallelen an mehreren Markerloci in teilweise erheblichen Frequenzen. Darüber hinaus konnten aufgrund weiterer Analysen für fünf der sechs untersuchten Herkunftsstichproben relativ ausgewogenen Saatgutanteile der einzelnen Erntebäume angenommen werden. Für die Herkunft aus Spanien dagegen wurden wahrscheinlich unausgewogene Anteile verwendet oder deutlich weniger als zehn Bäume beerntet, was die enge Verwandtschaft von zwölf der untersuchten 50 Individuen nahelegt, die vermutlich eine Halbgeschwisterfamilie bilden.

Beim Vergleich der genetischen Strukturen zwischen den Populationen zeigt sich eine klare Abgrenzung der spanischen Herkunft aus dem iberischen eiszeitlichen Refugium von den anderen fünf Populationen, die aus dem südosteuropäischen Refugium stammen. Eine Clusteranalyse gruppiert drei Herkünfte von relativ kalten Standorten (Deutschland/Harz, Österreich, Tschechien) und zwei Herkünfte von relativ warmen Standorten (Deutschland/Brandenburg, Rumänien).

**Schlüsselwörter:** *Fagus sylvatica* L.; Populationsgenetik; nukleare Mikrosatellitenmarker; Subpopulationsstruktur; Nullallele; Repräsentativität der Saatguternte

## Abstract

### Genetic characterisation of beech populations (*Fagus sylvatica* L.) from a provenance trial with nuclear microsatellite markers

Samples of 50 to 55 beech individuals (*Fagus sylvatica* L.) from each of six selected progenies from a provenance trial (two from Germany, one from Austria, Czech Republic, Romania and Spain each) were genotyped with eight nuclear microsatellite markers to complete comprehensive phenotypical data collections.

The detected partially strong deviations from the Hardy-Weinberg equilibrium within each population could be attributed to two essential sources: 1. A subpopulation structure consisting of two or three subpopulations in Hardy-Weinberg equilibrium per provenance and 2. The existence of null alleles with remarkable frequencies at several marker loci. Furthermore, the representativeness of the seed harvest is indicated for five out of six population samples. In contrast to these, the provenance from Spain was harvested from obviously unbalanced portions or less than ten trees. This was concluded from the close relationship between twelve individuals out of 50, which seem to be a half-sib family.

The comparison of genetic structures among populations demonstrates the clear separation of the Spanish provenance originating from the Iberian glacial refugium at the one side from the remaining five provenances originating from the south-eastern European refugium at the other side. A cluster analysis classified three provenances from more cold sites (Germany, Austria, Czech Republic) together into one group and two provenances from more warm sites (Germany, Romania) into another group.

**Keywords:** *Fagus sylvatica* L.; population genetics; nuclear microsatellite markers; subpopulation structure; null allele; representativeness of seed harvest

\* Johann Heinrich von Thünen-Institut, Institut für Forstgenetik, Sieker Landstr. 2, 22927 Großhansdorf

Kontakt: heike.liesebach@vti.bund.de

## Einleitung

Die Buche (*Fagus sylvatica* L.) als wichtigste Laubbaumart in Deutschland spielt in der forstlichen Forschung eine zentrale Rolle. Ihr großes Verbreitungsgebiet in Europa und ihr entsprechend großes Anpassungspotential an unterschiedliche klimatische Gegebenheiten macht ihre Bedeutung auch unter den Bedingungen des Klimawandels deutlich. Die Buche ist eine weitverbreitete, bestandesbildende Baumart, sie ist monözisch und wird windbestäubt, dabei hauptsächlich fremdbefruchtet (z. B. Merzeau et al., 1994). Zunehmend wird versucht, phänotypische Merkmale wie Wachstum und Phänologie sowie Daten aus der ökologischen und physiologischen Forschung mit Aspekten der Genetik in Verbindung zu bringen (Kramer et al., 2008). Wertvolle Informationen zur Anpassungsfähigkeit und zur phänotypischen Merkmalsausprägung dieser Baumart im Verbreitungsgebiet und unter verschiedenen Umweltbedingungen können aus Herkunftsversuchen gewonnen werden. Das Thünen-Institut für Forstgenetik hat einen internationalen Buchenherkunftsversuch initiiert, der mit insgesamt 128 Herkünften an 23 Standorten in Europa im Jahr 1995 mit zweijährigen Pflanzen angelegt wurde (Liesebach M, 2012).

Ausgewähltes Material von je einer Versuchsfläche in Deutschland und Österreich aus dieser Versuchsserie diente als Basis für umfangreiche Untersuchungen, die in den letzten Jahren gemeinsam von den Thünen-Instituten für Forstgenetik, für Weltforstwirtschaft, für Holzforschung und für Waldökosysteme sowie des Instituts für Genetik der BFW Wien durchgeführt wurden. Diese Kooperation erlaubte eine komplexe Herangehensweise durch die Vielfalt an untersuchten Parametern und eingesetzten Methoden (Wachstum, phänologische Merkmale, ökophysiologische Merkmale, Holzmerkmale, Genotypisierung mit molekularen Markern). Eine ausführliche Beschreibung der sechs für die gemeinsamen Untersuchungen ausgewählten Herkünfte des Internationalen Buchenherkunftsversuchs von 1993/1995 findet sich bei Liesebach, M. (2012). Das Hauptkriterium für die Auswahl dieser sechs Herkünfte war das Vorkommen an relativ warmen (Spanien, Rumänien, Deutschland/Brandenburg: Jahresmitteltemperatur 8,1 bis 11,7 °C) bzw. relativ kalten Standorten (Tschechien, Österreich, Deutschland/Niedersachsen-Harz: Jahresmitteltemperatur 4,9 bis 5,9 °C) des Buchenverbreitungsgebiets. In Ergänzung zu den phänotypischen Untersuchungen sollten populationsgenetische Untersuchungen durchgeführt werden, um Informationen über die genetische Variation innerhalb und zwischen den sechs Populationen zu gewinnen.

Genetische Untersuchungen bei der Baumart Buche wurden bisher mit verschiedenen genetischen Markern durchgeführt. Frühere Analysen mit Isoenzymmarkern ergaben nur geringe Differenzierungen bei dieser Baumart in Europa (Comps et al., 1990) und innerhalb Deutschlands (Konert et al., 2000). Chloroplasten-DNA-Haplotypen zeigten größere Differenzierungen nur in den eiszeitlichen Refugien im Süden Europas, während in West-, Mittel- und Osteuropa im Wesentlichen nur ein cpDNA-Haplotyp gefunden wurde (Demesure et al., 1996).

Nukleare Mikrosatellitenmarker für Buche wurden erstmals für ostasiatische Buchenarten, für *F. crenata* und *F. japonica*, beschrieben (Tanaka et al., 1999). Diese Marker wurden auch für *F. sylvatica* getestet und teilweise angepasst (Vornam et al., 2004). Weitere Mikrosatellitenmarker wurden speziell für *F. sylvatica* entwickelt (Pastorelli et al., 2003). Bisher gibt es erst wenige Untersuchungen von Buchenpopulationen mit diesen Mikrosatellitenmarkern, wobei vier bis sechs Markerloci eingesetzt wurden (Vornam et al., 2004; Buiteveld et al., 2007; Jump und Peñuelas, 2007; Maurer et al., 2008; Nyári, 2010).

Im Gegensatz zu den bisher genannten neutralen Markern spielen adaptive Marker bei der Buche eine noch geringe Rolle. Im Rahmen einer ersten genetischen Kartierung der Buche von Scalfi et al. (2004) wurden einige QTLs für Blattflächenmerkmale gefunden. Einen Hinweis geben auch Jump et al. (2006) mit einem AFLP-Locus, genannt AGTC476, der eine Beziehung zum dendrochronologisch ermittelten Zuwachs in Abhängigkeit von der Temperatur in einer spanischen Population aufweist. Weiterführende Arbeiten dazu sind jedoch bisher nicht erschienen. SNPs in ausgewählten Kandidatengenen untersuchte Seifert (2011), wobei Unterschiede zwischen Populationen nicht nur in potentiell adaptiven Genen, sondern ebenso bei neutralen Markern gefunden wurden.

Aktuelle Forschungsprojekte wie z. B. „Biodiversitäts-Exploratorien – Beechadapt“ (Seifert, Universität Göttingen), „Trockenstress-Sensitivität von Ökotypen der Buche aus süddeutschen Herkünften und Refugialstandorten in Südeuropa“ (Dounavi, FVA Freiburg) sowie „Analyse anpassungsrelevanter Gene bei Rotbuche und Schwarzkiefer mit Schwerpunkt auf der Trockenresistenz als Grundlage für Herkunftsempfehlungen im Klimawandel“ (Fussi und Konert, ASP Teisendorf) beschäftigen sich weiter mit der Untersuchung verschiedener Kandidatengene für die Trockenstress-Resistenz.

In der hier durchgeführten Studie mit begrenztem Umfang wurden nukleare Mikrosatellitenmarker eingesetzt, der aus gegenwärtiger Sicht geeignetste Markertyp zur Charakterisierung der genetischen Variation von Stichproben aus Populationen. Das Auffinden von Assoziationen zu den erfassten phänotypischen Merkmalen kann aufgrund der Neutralität der Marker zunächst nicht erwartet werden. Eine Kopplung zwischen neutralen Mikrosatellitenmarkern und adaptiven Genen ist wegen der geringen Markeranzahl in dieser Untersuchung zwar eher unwahrscheinlich, aber auch nicht ausgeschlossen.

## Material und Methoden

### Material

Von sechs ausgewählten Herkünften aus dem Internationalen Buchenherkunftsversuch von 1993/1995 am Standort Schädtkbek (Schleswig-Holstein) wurden Stichproben von mind. 50 Individuen je Population genetisch charakterisiert (Tabelle 1). Aus allen drei Blöcken der Versuchsanlage mit Überlebensraten von 55 bis 66 % wurden Zufallsstichproben entnommen, wobei je Herkunft ca. 56 bis 64 % der

Tabelle 1

Sechs Buchenherkünfte aus dem internationalen Herkunftsversuch von 1993/1995 und Stichprobenumfang für genetische Untersuchungen

Land	Herkunftsname	Bezeichnung	Geogr. Breite	Geogr. Länge	Höhe über NN	Jahresmitteltemperatur	Stichprobenumfang
Spanien	Anguiano	Anguiano /ES-5	42° 15' N	02° 45' W	950 m	11,7 °C	50
Deutschland	Oderhaus/Harz	Oderhaus /DE-44	51° 40' N	10° 50' O	710 m	5,1 °C	55
Deutschland	Gransee/Brandenburg	Gransee /DE-46	53° 00' N	13° 10' O	70 m	8,1 °C	50
Österreich	Neuberg-Mürzsteg	Neuberg /AT-109	47° 45' N	15° 28' O	1050 m	4,9 °C	54
Tschechien	Kladská	Kladská /CZ-110	50° 02' N	12° 37' O	690 m	5,9 °C	53
Rumänien	Beius-Bihor	Beius /RO-146	46° 41' N	22° 16' O	265 m	9,6 °C	53

noch lebenden Bäume beprobt wurden. Die Beerntung von Ruheknospen erfolgte Anfang März 2011, das Material wurde bei -70°C gelagert.

### Genotypisierung mit DNA-Markern

Die DNA-Extraktion erfolgte mit einer Standardmethode modifiziert nach Dumolin et al. (1995). Insgesamt wurden zehn Mikrosatelliten-Marker getestet, wobei drei Multiplex-PCRs mit dem Qiagen Multiplex PCR-Kit nach Vorschrift des Herstellers durchgeführt wurden. Die Trennung der DNA-Fragmente erfolgte mit dem Kapillar-Sequenzer CEQ 8000 der Firma Beckman-Coulter.

Alle Marker zeigten eine sehr gute Reproduzierbarkeit der Muster. Es wurden jedoch bei zwei Markern (mfc92 und FS111) teilweise bis zu vier Allele je Probe beobachtet, d. h. die Primer binden nicht spezifisch für einen Genort. Diese beiden Genorte scheinen, zumindest unter den verwendeten PCR-Bedingungen, Duplikate zu haben, so dass nur acht kodominante Loci für populationsgenetische Zwecke (Tabelle 2) zur Verfügung stehen.

Teilweise fehlten Peaks bei einzelnen Proben und an einzelnen Loci. Dabei handelt es sich mit hoher Wahrscheinlichkeit um Homozygote von Nullallelen, da die anderen Loci der Probe im selben Multiplex-PCR-Ansatz normale Peaks aufwiesen, ebenso wie andere Proben derselben Serie (gleicher PCR-Mastermix) am betreffenden Genort. Die Untersuchung von 25 betroffenen Proben wurde wiederholt, um das Fehlen der Peaks und damit den Genotyp als Homozygoten eines Nullallels zu bestätigen.

### Datenauswertung

Basisparameter zur Quantifizierung der genetischen Variation von Populationen wie genetische Vielfalt als mittlere Anzahl der Allele pro Locus (A/L) sowie die beobachtete und erwartete Heterozygotie ( $H_{obs}$  und  $H_{exp}$ ) und exakte Tests auf signifikante Abweichungen vom Hardy-Weinberg-Gleichgewicht wurden mit der Software SAS (proc allele) durchgeführt (SAS for Windows, Version 9.2). Auch die Tests auf Kopplungsungleichgewichte (Linkage disequilibrium, 28 verschiedene 2-Locus-Kombinationen für sechs Populationen) wurden mit SAS proc allele durchgeführt. Um bei multiplen

Tabelle 2

Untersuchte nukleare Mikrosatelliten-Marker bei Buche (\* Loci mit bis zu vier Peaks und nicht verwendeten Daten), Zuordnung zu Kopplungsgruppen nach Scalfi et al. (2004)

Mikrosatelliten-Locus	Kopplungsgruppe	Fluoreszenzmarkierung	Fragmentlängenbereich(bp)	Repeat	Referenz
mfc7		BMN-6	111-135	GA	Tanaka et al., 1999
mfs11		Cy5	131-149	AG	Vornam et al., 2004
mfc92*		BMN-6	184-210	GA	Tanaka et al., 1999
mfc5		Cy5	275-327	AG	Tanaka et al., 1999
FS103	III	BMN-6	87-157	GA	Pastorelli et al., 2003
FS115		Cy5	95-141	GA	Pastorelli et al., 2003
FS304	XI	Cy5	195-207	GCT/GTT	Pastorelli et al., 2003
FS125	VII	Cy5	77-135	GA	Pastorelli et al., 2003
FS111*	VI	BMN-6	98-124	GA	Pastorelli et al., 2003
FS446	I	Cy5	217-430	TGA	Pastorelli et al., 2003

Tests (Hardy-Weinberg-Gleichgewicht und Linkage disequilibrium) falsch-positive Ergebnisse zu vermeiden, wurden die erhaltenen p-Werte mit der Software Q-Value (Storey und Tibshirani, 2003) in q-Werte überführt.

Die Software STRUCTURE (Pritchard et al., 2000) wurde verwendet, um Substrukturen in Populationen ohne eine vorgegebene individuelle Zuordnung zu finden. Sie gruppiert Individuen anhand ihres Multilocus-Genotyps in Cluster, so dass das Hardy-Weinberg-Gleichgewicht innerhalb der Cluster maximiert wird. Dabei werden in mehreren Wiederholungen die Zuordnungswahrscheinlichkeiten jedes Individuums zu jedem Cluster ermittelt sowie eine Gesamtwahrscheinlichkeit P für jede vorgegebene Anzahl K von Clustern. Für jede der sechs Populationen mit 50 bis 55 Individuen wurden Werte von K im Bereich von eins bis zwanzig mit je sechs bis zwölf Wiederholungen getestet. Als Parametereinstellungen wurden „Admixture Model“ und „Allele Frequencies correlated“ gewählt, da es sich um Individuen innerhalb einer Population handelt. Für die Gesamtheit von allen 315 Individuen wurde K im Bereich von eins bis zehn mit zwölf Wiederholungen getestet, jedoch mit den Parametereinstellungen „No Admixture Model“, da der genetische Austausch zwischen den sechs untersuchten Populationen als nicht vorhanden oder zumindest als extrem gering angenommen wird, und „Independent Allele Frequencies“, um eine Überschätzung von K zu vermeiden. Die Anzahl der im jeweiligen Datensatz vorhandenen Cluster oder Subpopulationen K ergibt sich dann aus dem Maximalwert oder dem erstem Erreichen eines Plateaus der LnP-Werte (Auswertung mit der Software STRUCTURE Harvester, Earl und von Holdt, 2011).

Es existieren verschiedene Schätzmethoden für die Frequenzen von Nullallelen in den einzelnen populationsgenetischen Softwareprogrammen, die alle auf der Annahme eines Hardy-Weinberg-Gleichgewichts in der Population und Abweichungen der beobachteten von den erwarteten Genotyphenhäufigkeiten basieren. Eine Schätzung der Frequenz von Nullallelen wurde mit der Software Micro-Checker (Van Oosterhout et al., 2004), sowie vergleichsweise CERVUS (Marshall et al., 1998; Kalinowski et al., 2007) und GENEPOP'007 (Rousset, 2008) vorgenommen. Neben der Verwendung verschiedenen Algorithmen unterscheiden sie sich diese vor allem darin, ob fehlende Werte immer als unbekannter Genotyp betrachtet werden (z. B. nach Kalinowski et al. (2007) in der Software CERVUS; nach Van Oosterhout et al. (2004), nach Chakraborty et al. (1992) und bei Methode 1 nach Brookfield (1996) in der Software Micro-Checker), oder ob sie immer als Homozygoter eines Nullallels (Methode 2 nach Brookfield (1996) in der Software Micro-Checker) oder mit unterschiedlicher Kodierung als unbekannter Genotyp bzw. als Homozygoter eines Nullallels in die Kalkulation einbezogen werden (nach Dempster et al. (1977) in der Software GENEPOP'007). Die Software Micro-Checker bietet eine Korrektur der Genotypendaten an, die bei hohen Frequenzen von Nullallelen einen Teil der häufigen homozygoten Genotypen durch Heterozygote mit einem Nullallel ersetzt. Mit solchen korrigierten Daten sind Kalkulationen auf Populationsebene ausführbar, Auswertungen mit individuellen Genotypen sind aber weiterhin nur mit den unkorrigierten Originaldaten möglich.

Mit zwei Verfahren wurde versucht, Informationen über Verwandtschaftsbeziehungen von Individuen innerhalb der Herkunftsstichproben zu gewinnen. Der ermittelte Multilocus-Genotyp jedes Individuums kann während der sexuellen Reproduktion aus verschiedenen Gameten-Haplotypen der Eltern entstanden sein. Für jede denkbare Haplotypenkombination kann mit Hilfe der Allelfrequenzen in der jeweiligen Population im Hardy-Weinberg-Gleichgewicht eine Wahrscheinlichkeit mit einem maximum likelihood-Verfahren berechnet werden (SAS proc haplotype). Haplotypen-Informationen können teilweise zur Rekonstruktion der Abstammung von einem Mutterbaum verwendet werden. Außerdem wurde mit der Software Populations Version 1.2.31 (Langella, 2010) je Herkunft ein sog. „Neighbour-Joining phylogenetic tree“ erstellt, basierend auf genetischen Distanzen zwischen Individuen (Shared allele distance nach Jin und Chakraborty (1994)). Die anschließende graphische Darstellung der Verwandtschaftsbeziehungen erfolgte mit dem Programm Tree-View (Page, 2001).

Zur Schätzung der Differenzierung zwischen Populationen wurden in den letzten Jahren neue Parameter entwickelt, um die Informationen von Loci mit vielen Allelen, hoher Diversität und Heterozygotie, wie sie bei nuklearen Mikrosatelliten üblicherweise auftreten, realistisch abzubilden. Dazu gehören ein auf die Heterozygotie standardisiertes  $G'_{ST}$  (Hedrick, 2005) sowie die aktuelle Differenzierung D zwischen Populationen, die unabhängig von der Anzahl der Allele je Locus und von der Heterozygotie ist (Jost, 2008). Hierbei können  $G'_{ST}$  und D für exakte Allelfrequenzen bzw.  $G_{STest}$  und  $D_{est}$  für Stichproben von Populationen unterschieden werden. Die Verzerrungen des alten  $G_{ST}$  nach Nei (1973), das genaugenommen nur für Loci mit 2 Allelen gilt, sind damit überwunden (Jost, 2008; Gerlach et al., 2010). Die alten sowie neuen Differenzierungsmaße wurden mit der Software SMOGD berechnet (Crawford, 2010). Aus paarweisen Differenzierungen  $D_{est}$  zwischen den Populationen lässt sich eine Abstandsmatrix zur graphischen Darstellung als Dendrogramm erstellen. Ein Dendrogramm wurde über eine UPGMA-Clusteranalyse mit der Software SAS (proc distance, proc tree) erzeugt. Die Korrelation zwischen genetischen und geographischen Distanzmatrizen wurde mit dem Mantel-Test geprüft (Bohonak, 2002).

Für die Erstellung der verschiedenen Inputformate, die die einzelnen Software-Anwendungen erfordern, wurde überwiegend das Programm FormatIn (Buschbom, 2010) verwendet.

## Ergebnisse und Diskussion

### Genetische Variation – Basisdaten

Alle Kalkulationen von populationsgenetischen Parametern gehen von der freien Rekombination der verwendeten kodominanten Marker aus. Bei den Autoren der Markerloci (Tanaka et al., 1999; Pastorelli et al., 2003; Vornam et al., 2004) wurden keine Angaben zur Kopplung gemacht. Teilweise wurden die verwendeten Marker jedoch von Scalfi et al. (2004) auf verschiedenen Chromosomen kartiert (Tabelle 2). Die Tests auf

Kopplungsungleichgewichte (Linkage disequilibrium) ergaben eine schwache, aber signifikante Kopplung zwischen mfc7 und FS103 bei drei von sechs Populationen, und eine weitere Kopplung zwischen FS103 und FS304 bei zwei Populationen, obwohl diese beiden Loci auf den Chromosomen III und XI lokalisiert wurden (Salfi et al., 2004). Das Kopplungsungleichgewicht wurde in diesem Fall von anderen Faktoren als einer realen Kopplung bedingt. Es kann also von einer weitgehend freien Rekombination der Markerloci ausgegangen werden.

Insgesamt wurden 315 Individuen aus sechs Herkunftten an acht Loci vollständig genotypisiert, d. h. es ist ein kompletter Datensatz mit 2520 Einzellocus-Genotypen erfasst worden. Davon wurden 25 Genotypen (ca. 1 %) als Homozygote eines Nullallels identifiziert (reproduzierbar fehlende Peaks: zweimal am Locus FS115, 17-mal am Locus FS125 und 6-mal am Locus FS446).

Alle 315 Multilocus-Genotypen unterschieden sich voneinander. Die Tabelle 3 zeigt eine Übersicht über die Variation an acht nuklearen Mikrosatellitenloci der Buche, die vergleichbar mit der Variation in anderen Untersuchungen ist (Buiteveld et al., 2007; Jump und Peñuelas, 2007; Maurer et al., 2008; Nyári, 2010).

**Tabelle 3**

Genetische Variation an acht nuklearen Mikrosatellitenmarkern der Buche basieren auf 315 Individuen über alle Herkunftte hinweg

Marker	Allele pro Locus (A/L)	Effektive Anzahl der Allele ( $A_e$ )	Beobachtete Heterozygotie ( $H_{obs}$ )	Erwartete Heterozygotie ( $H_{exp}$ )
mfc7	11	2,29	0,5460	0,5638
mfs11	9	2,41	0,4508	0,5853
mfc5	23	11,55	0,6095	0,9134
FS103	29	3,91	0,7238	0,7445
FS115	20	6,86	0,7143	0,8542
FS304	4	1,49	0,3270	0,3272
FS125	25	7,56	0,5492	0,8678
FS446	23	5,39	0,5841	0,8145
Mittelwert	18	5,18	0,5631	0,7088

**Genetische Variation innerhalb der Populationen**

Das Saatgut für Herkunftsversuche soll für den jeweiligen Ausgangsbestand repräsentativ sein, um die verschiedenen phänotypischen Merkmale in ihrer typischen Ausprägung und Variationsbreite zu erfassen. Dies ist am besten gewährleistet, wenn das Saatgut aus einem Erntejahr mit gutem weiblichen Blütenansatz und reichlicher Pollenproduktion für eine gute wechselseitige Bestäubung stammt und von einer ausreichenden Anzahl Erntebäume gesammelt worden ist. Um das nachträglich anhand der genetischen Struktur von Stichproben aus der Nachkommenschaft zu beurteilen, bietet die Überprüfung des Hardy-Weinberg-Gleichgewichtes eine gute Möglichkeit.

Die Tabelle 4 zeigt, dass in vielen Fällen signifikante Abweichungen von einer panmiktischen Population vorliegen (q-Werte < 0,05). Bei diesen Loci kommt meist ein Überschuss an Homozygoten vor. Bei allen sechs Populationen gibt es aber auch jeweils mehrere Loci im Hardy-Weinberg-Gleichgewicht.

**Tabelle 4**

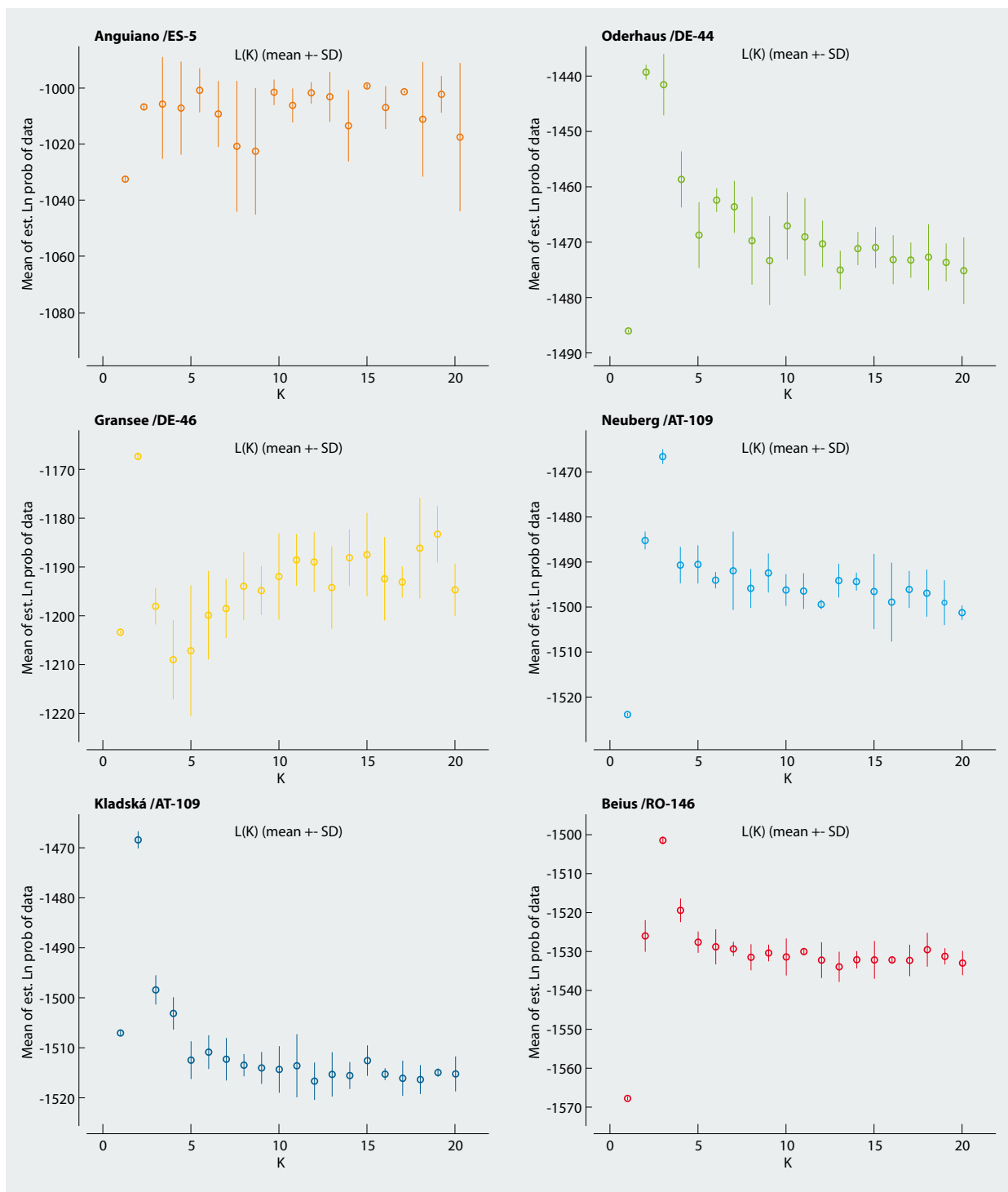
Wahrscheinlichkeiten (q-Werte) für das Hardy-Weinberg-Gleichgewicht an acht Markerloci in Stichproben von sechs Buchen-Populationen (q < 0,05: signifikante Abweichung vom Hardy-Weinberg-Gleichgewicht, gelb unterlegt)

Locus	Anguiano / ES-5	Oderhaus / DE-44	Granssee / DE-46	Neuberg / AT-109	Kladská / CZ-110	Beius / RO-146
mfc7	0,3379	0,1654	0,7498	0,7087	0,7498	0,9671
mfs11	0,0197	0,9671	0,0805	0,0011	0,0709	0,0709
mfc5	0,7087	0,0004	0,0004	0,0004	0,0004	0,0004
FS103	0,0381	0,9671	0,9671	0,9671	0,0404	0,6123
FS115	0,5806	0,0464	0,3317	0,1313	0,0023	0,9671
FS304	0,1724	0,4362	0,7498	0,6015	0,9671	0,0709
FS125	0,0004	0,0004	0,0008	0,0049	0,0004	0,0004
FS446	0,0004	0,0356	0,0616	0,1313	0,0037	0,0004

Verschiedene Ursachen kommen für die Abweichungen vom Hardy-Weinberg-Gleichgewicht in Frage. Das Vorkommen von homozygoten Genotypen von Nullallelen lässt vermuten, dass häufige Nullallele an den Abweichungen vom Hardy-Weinberg-Gleichgewicht beteiligt sind, aber auch die Mischung von unterschiedlich strukturierten Teilpopulationen kann zu solchen Abweichungen führen, was als Wahlund-Effekt bezeichnet wird. Auch Inzucht oder Verwandtenpaarung kann einen Überschuss an Homozygoten verursachen, der jedoch dann an allen untersuchten Loci vorhanden sein muss. Das kann im vorliegenden Fall ausgeschlossen werden, da nicht an allen Loci ein solcher Überschuss vorkommt.

Um zunächst festzustellen, ob es innerhalb der Stichproben aus den sechs Herkunftten Subpopulationsstrukturen gibt, die eher als Hardy-Weinberg-Populationen angesehen werden können als die Gesamtstichprobe einer Herkunft, wurden die Daten mit der Software STRUCTURE analysiert. Die Abbildung 1 zeigt, dass ein Maximalwert oder ein erstes Erreichen eines Plateaus des Ln P-Werts bei K = 2 für Granssee /DE-46, Kladská /CZ-110, Oderhaus /DE-44 und bei K = 3 für Anguiano /ES-5, Neuberg /AT-109 und Beius /RO-146 auftritt. Somit bestehen alle Stichproben der sechs Herkunftten aus jeweils zwei oder drei Subpopulationen.

Die Zuordnung jedes Individuums zu einer der zwei bzw. drei Subpopulationen erfolgte nach ihrer Wahrscheinlichkeit, die auch graphisch dargestellt werden kann. In der Abbildung 2 wird jedes Individuum durch eine Säule veranschaulicht, die die Zuordnungswahrscheinlichkeiten für jede der Subpopulationen in einer anderen Farbe zeigt. Es wird



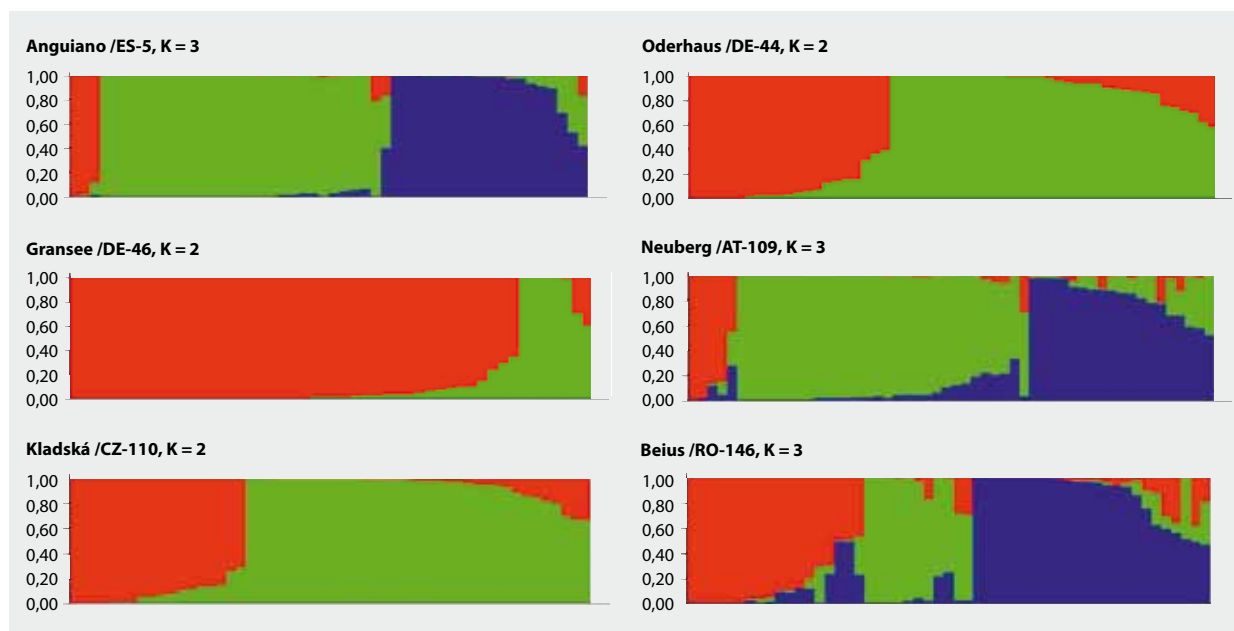
**Abbildung 1**

Ermittlung der Anzahl von Subpopulationen mit STRUCTURE und STRUCTURE Harvester: Mittelwerte und Standardabweichungen der Ln P-Werte für  $K = 1$  bis  $K = 20$  (sechs bis zwölf Wiederholungen) der Analyse von Genotypen aus sechs Buchenherkünften (Erklärung im Text)

deutlich, dass für die meisten Individuen der Stichproben eine recht eindeutige Zuordnung zu einer der Subpopulationen getroffen werden kann, und nur für wenige Individuen mehrere Subpopulationen mit ähnlichen Zuordnungswahrscheinlichkeiten vorkommen. Die Zuordnung der Individuen zu einer von zwei oder drei Subpopulationen erfolgte ent-

sprechend der höchsten Wahrscheinlichkeit. Auf diese Weise konnten insgesamt 15 Subpopulationen gebildet werden, die zu sechs Herkunftsstichproben gehören.

Ein erneuter Test auf signifikante Abweichungen vom Hardy-Weinberg-Gleichgewicht wurde in den so gebildeten 15 Subpopulationen durchgeführt. Die Häufigkeit solcher



**Abbildung 2**

Graphische Darstellung der Zuordnung von Individuen zu Subpopulationen innerhalb von sechs Buchenherkünften mit STRUCTURE

Abweichungen hat sich durch die Berücksichtigung der Subpopulationsstruktur im Vergleich zu den Daten in Tabelle 4 von 44 % schon deutlich auf 16 % reduziert. Es kommen aber weiterhin an einzelnen Loci starke Abweichungen vor, besonders an den Loci, bei denen auch homozygote Genotypen von Nullallelen beobachtet wurden.

Im Folgenden wurden nun in den sechs Herkunftsorten und in 15 Subpopulationen jeweils an acht Loci mit verschiedenen Verfahren die Frequenzen von Nullallelen geschätzt. Es ergeben sich mit den verschiedenen Verfahren überwiegend ähnliche Schätzwerte für die Nullallel-Frequenzen, sofern an dem entsprechenden Locus und in der jeweiligen Population oder Subpopulation kein homozygoter Genotyp für ein Nullallel vorkommt. In den Fällen, bei denen reproduzierbar fehlende Peaks, die als Homozygote für ein Nullallel interpretiert werden, vorhanden sind, gibt es erwartungsgemäß größere Abweichungen zwischen den Methoden nach Brookfield2 und nach Dempster einerseits und den übrigen fünf Methoden andererseits. In der vorliegenden Untersuchung müssen die Schätzungen nach den beiden erstgenannten Methoden wegen des Vorkommens von homozygoten Nullallelen als zuverlässiger gelten. Hinweise auf Nullallele finden sich auch bei Scalfi et al. (2004) für die Loci FS103, FS125 und FS446, bei Maurer et al. (2008) für FS446, bei Oddou-Muratorio et al. (2009) für die Loci mfc5, FS103, FS125 und FS446 und bei Chybicki und Burczyk (2009) für die Loci FS125, FS446 und mfc5.

Die Ergebnisse zeigen, dass signifikante Abweichungen vom Hardy-Weinberg-Gleichgewicht in einer Population oder Subpopulation immer dann vorkommen, wenn ein Nullallel in größerer Frequenz ermittelt wurde. Es wurden jedoch bei häufigeren Nullallelen nicht immer auch homozygote Genotypen gefunden. Bei Existenz von häufigen Nullal-

len muss angenommen werden, dass ein Teil der als scheinbar homozygot identifizierten Genotypen in Wahrheit heterozygot mit einem Nullallel ist. Eine entsprechende Korrektur der Genotypen mit der Software Micro-Checker nach der Methode Brookfield2 für jede der 15 Subpopulationen wies insgesamt 132 Genotypen als Heterozygote mit einem Nullallel aus, das sind ca. 5 % der ermittelten 2520 Einzellocus-Genotypen (Tabelle 5). Solche Korrekturen wurden, speziell für die Baumart Buche mit Nullallelen an vielen Loci, empfohlen, um populationsgenetische Parameter und Differenzierungen zwischen Populationen genauer schätzen zu können (Oddou-Muratorio et al., 2009).

Der Test auf signifikante Abweichungen vom Hardy-Weinberg-Gleichgewicht wurde nun in den gebildeten 15 Subpopulationen nach der Korrektur mit Nullallelen durchgeführt. Jetzt gab es nur noch in 2 % der Fälle signifikante Abweichungen vom Hardy-Weinberg-Gleichgewicht. Damit konnten die Subpopulationsstrukturen innerhalb jeder Herkunft und die Existenz von Nullallelen an mehreren Loci als die beiden wesentlichen Ursachen der zunächst festgestellten großen Abweichungen vom Hardy-Weinberg-Gleichgewicht (Tabelle 4) gefunden werden.

Grundsätzlich führt das relativ häufige Vorkommen von Nullallelen dazu, dass populationsgenetische Auswertungen nur näherungsweise möglich sind. Das trifft auf die Originaldaten zu, aber auch auf die mit Nullallelen korrigierten Daten, obwohl diese der Realität wahrscheinlich näher kommen. Die Tabelle 6 zeigt, dass bei vergleichbarem Stichprobenumfang von allen sechs Populationen die genetische Vielfalt gemessen als mittlere Anzahl der Allele pro Locus bei den Populationen Anguiano /ES-5 und Gransee /DE-46 mit 8,13 bis 9,13 geringer ist als bei den vier anderen Populationen mit 11,0 bis 12,50. Die beobachtete und erwar-

**Tabelle 5**

Geschätzte Frequenzen von Nullallelen je Locus nach Korrektur der Genotypen mit der Software Micro-Checker in Stichproben von sechs Buchenherkünften (Methode Brookfield2 für jede Subpopulation)

Locus	Anguiano /ES-5	Oderhaus /DE-44	Gransee /DE-46	Neuberg /AT-109	Kladská /CZ-110	Beius /RO-146
mfc7	0	0	0	0	0	0
mfs11	0	0	0	0,0556 ± 0,0214	0,0755 ± 0,0246	0
mfc5	0	0,1000 ± 0,0270	0,0900 ± 0,0272	0,1019 ± 0,0274	0,0566 ± 0,0218	0,0472 ± 0,0201
FS103	0	0	0	0	0	0
FS115	0	0,0182 ± 0,0180	0	0	0,0283 ± 0,0207	0
FS304	0	0	0	0	0	0
FS125	0,2400 ± 0,0515	0,0727 ± 0,0299	0,1700 ± 0,0390	0,0278 ± 0,0204	0,1132 ± 0,0344	0,1698 ± 0,0399
FS446	0,1600 ± 0,0386	0,0364 ± 0,0175	0,0300 ± 0,0220	0	0,0755 ± 0,0246	0,0755 ± 0,0337

**Tabelle 6**

Mittelwerte populationsgenetischer Parameter über acht Loci zur Charakterisierung der genetischen Variation innerhalb von sechs Buchen-Populationen (Daten mit Korrektur der Nullallele nach Brookfield2 je Subpopulation)

Population	Anzahl untersuchter Individuen	Allele pro Locus (A/L)	Effektive Anzahl der Allele ( $A_e$ )	Beobachtete Heterozygotie ( $H_{obs}$ )	Erwartete Heterozygotie ( $H_{exp}$ )	Fixierungsindex (F)
Anguiano /ES-5	50	8,13	2,36	0,5225	0,5766	0,0921
Oderhaus /DE-44	55	11,00	3,17	0,6159	0,6842	0,0721
Gransee /DE-46	50	9,13	2,94	0,6050	0,6597	0,0784
Neuberg /AT-109	54	11,00	3,59	0,6574	0,7211	0,0748
Kladská /CZ-110	53	12,50	3,54	0,6462	0,7177	0,0716
Beius /RO-146	53	11,75	3,66	0,6415	0,7265	0,1008
Mittelwert	52,5	10,58	3,21	0,6148	0,6810	0,0816

tete Heterozygotie ist bei der Population Anguiano /ES-5 deutlich kleiner. Die mittleren Fixierungsindizes weisen bei allen sechs Herkunftstypen positive Werte auf, wie es beim Vorliegen von Mischungen aus Subpopulationen durch den Wahlund-Effekt zu erwarten ist.

### Verwandtschaftsbeziehungen innerhalb von Populationen

Mit der Rekonstruktion der Haplotypen, die die Elternbäume zu den vorgefundenen Multilocus-Genotypen jedes Individuums beigetragen haben, wurde nach weiteren Informationen zu Substrukturen innerhalb der sechs Herkunftstypen gesucht. Die Anzahl verschiedener Haplotypen, aus denen sich ein diploider Multilocus-Genotyp gebildet haben kann bzw. die Anzahl verschiedener Gameten, die in der Meiose gebildet werden können, ist von der individuellen Anzahl heterozygoter Loci abhängig. Sie beträgt für ein Individuum maximal  $2^8 = 256$  bei einem Multilocus-Genotyp mit acht heterozygoten Loci. Welche der potentiell möglichen Haplotypen-Kombinationen die wahrscheinlichste ist, berechnet die Software anhand der Allelfrequenzen in der Population. Für jedes Individuum in den sechs Populationen wurden zwei Haplotypen mit Wahrscheinlichkeiten von  $p = 1$  für insgesamt 310 Indi-

viduen und  $p > 0,9817$  für fünf Individuen identifiziert. Dabei mussten nicht erkannte Nullallele im heterozygoten Zustand unberücksichtigt bleiben, was zu einer leichten Unterschätzung der Haplotypenanzahl führt. Die Gesamtanzahl aller möglichen Haplotypen sowie die Anzahl der ermittelten verschiedenen Haplotypen in den Stichproben ist erwartungsgemäß abhängig von der genetischen Variation, der Anzahl der Allele und der Heterozygotie, in der jeweiligen Population. Die Populationen Anguiano /ES-5 und Gransee /DE-46 haben deshalb eine geringe Anzahl Haplotypen als die anderen vier Populationen (Tabelle 7). Bei ungekoppelten Loci und repräsentativen Stichproben muss davon ausgegangen werden, dass keine signifikanten allelischen Assoziationen (Linkage disequilibrium) vorkommen, was mit Ausnahme der Population Anguiano /ES-5 auch zutrifft.

Die meisten mehrfach vorkommenden Haplotypen gibt es zweimal, einige dreimal. Nur in der Population Anguiano /ES-5 kommt ein Haplotyp 5-mal und ein weiterer sogar 7-mal vor, wobei diese beiden Haplotypen an sieben Loci identisch sind. Damit ist die Vermutung naheliegend, dass die zwölf Individuen, die diese beiden Haplotypen tragen, Halbgeschwister sind, d. h. von einem Mutterbaum oder einem Pollenspender mit einem heterozygoten Locus und sieben homozygoten Loci abstammen können. Zwölf Individuen ei-



**Tabelle 7**  
Haplotypenstruktur von Stichproben aus sechs Buchenherkünften

Population	Gesamtanzahl möglicher Haplotypen	Anzahl verschiedener Haplotypen in der Stichprobe	Test auf allelische Assoziationen (p-Wert)	Anzahl einmal und mehrfach vorkommender Haplotypen				
				1mal	2mal	3mal	5mal	7mal
Anguiano /ES-5	987	83	0,0088	76	3	2	1	1
Oderhaus /DE-44	2403	103	1,0000	97	5	1	-	-
Gransee /DE-46	1695	84	1,0000	69	14	1	-	-
Neuberg /AT-109	2314	101	1,0000	94	7	-	-	-
Kladská /CZ-110	1947	102	0,9998	98	4	-	-	-
Beius /RO-146	2054	101	0,9996	96	5	-	-	-

ner Halbgeschwisterfamilie sind deutlich mehr, als in der Stichprobe erwartet werden kann. Wenn 20 Bäume eines Bestandes beerntet wurden und an der Gesamtsaatgutmenge der jeweiligen Herkunft zu etwa gleichen Anteilen enthalten waren, wären bei einer Stichprobe der Nachkommenschaft von 50 bis 55 Individuen jeweils etwa zwei bis drei davon einer Halbgeschwisterfamilie zuzuordnen. Bei der Herkunft Anguiano /ES-5 wurden jedoch nur ca. zehn Bäume beerntet (Liesebach M, 2012), was erwarten lässt, dass jeweils ca. fünf Individuen Halbgeschwister mit identischer Mutter wären. Die vermutete Halbgeschwisterfamilie von zwölf Individuen spricht daher für eine ungleichmäßige Repräsentanz der Erntebäume im Saatgut oder weniger als zehn Erntebäume in der Herkunft Anguiano /ES-5. Es wären auch extreme Selektionen, die zu unterschiedlichen Überlebensraten in Abhängigkeit vom Mutterbaum geführt haben, als Erklärung für eine relativ große Halbgeschwisterfamilie denkbar. Auf der Fläche Schädtkbek betragen die Überlebensraten jedoch bei fünf der sechs Herkünfte einschließlich Anguiano /ES-5 zwischen 55 und 57 % und bei der Herkunft Oderhaus /DE-44 66 %, also relativ geringe Unterschiede. Daten zu Ausfallraten während der Anzucht in der Baumschule liegen leider nicht vor.

Eine andere Möglichkeit der Darstellung von verwandtschaftlichen Strukturen innerhalb der Stichproben von Herkünften bieten sog. Neighbour-Joining trees (Abbildung 3). Bei den fünf Herkünften Oderhaus /DE-44, Gransee /DE-46, Neuberg /AT-109, Kladská /CZ-110 und Beius /RO-146 ist eine fast gleichmäßige sternförmige Darstellung zu sehen, wobei fast alle Individuen genetisch mehr oder weniger gleich weit voneinander entfernt stehen. Ein solches Muster kann als Abbild einer repräsentativen Stichprobe einer Nachkommenschaft aus einer Population aufgefasst werden. Bei der Herkunft Anguiano /ES-5 dagegen gibt es mehrere deutlich sichtbare Gruppierungen mit teilweise sehr eng miteinander verwandten Individuen. Die zwölf Individuen, die aufgrund der Haplotypenanalyse als vermutliche Halbgeschwisterfamilie aufgefallen waren, sind in der Abbildung 3 farbig hervorgehoben. Das Vorkommen enger Verwandtschaftsgruppen steht in deutlichem Kontrast zu den anderen fünf Herkünften.

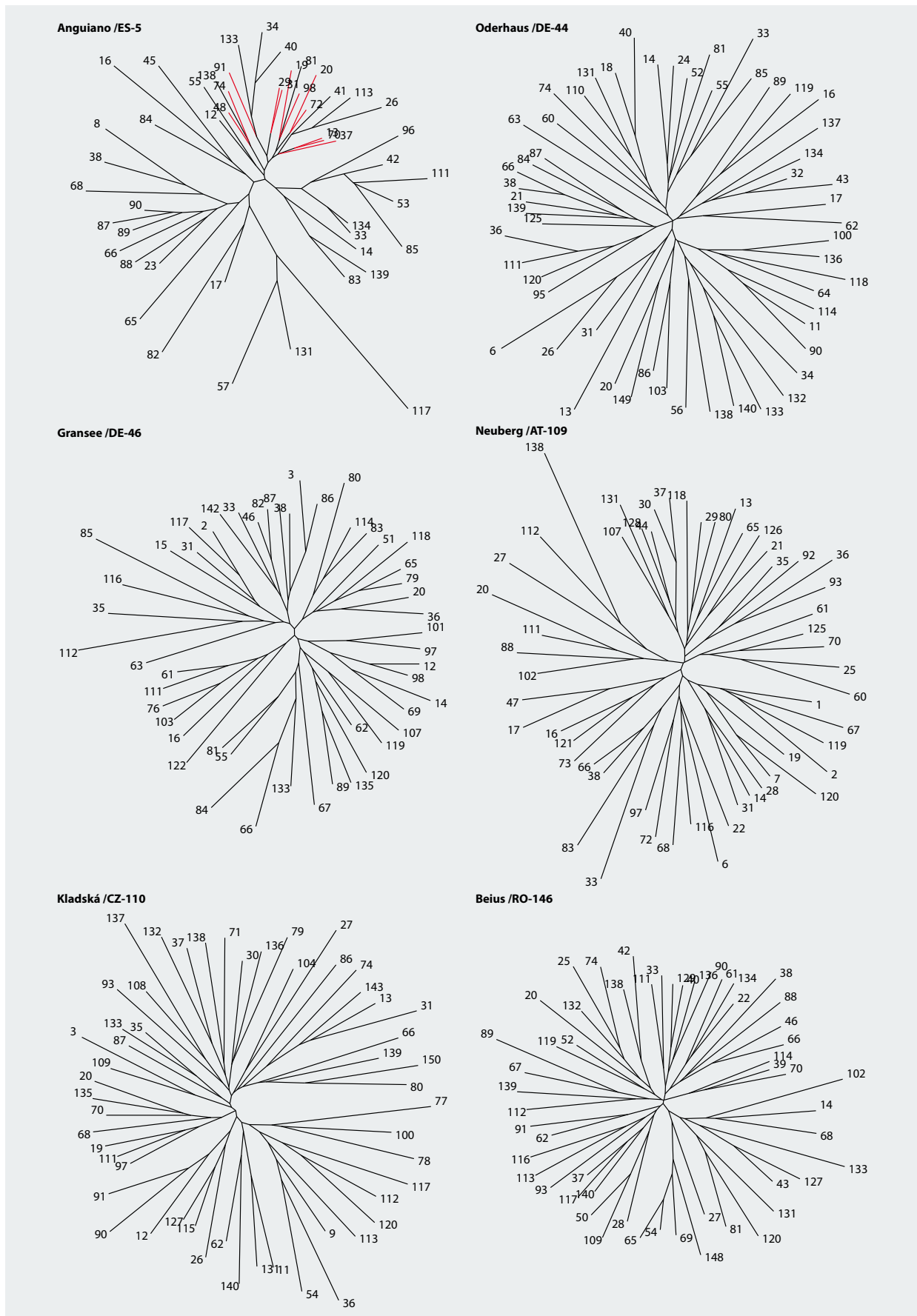
Die zuvor gefundenen zwei oder drei Subpopulationen in jeder Herkunftsstichprobe bilden sich im Wesentlichen, wenn auch nicht vollständig, in den Neighbour-Joining-Dendrogrammen ab (nicht dargestellt).

**Tabelle 8**  
Differenzierung zwischen sechs Buchenherkünften an acht Mikrosatelliten-Loci mit geschätzten Parametern nach verschiedenen Methoden für Stichproben aus Populationen (Daten mit Nullallelkorrektur je Subpopulation)

Locus	$G_{ST\_est}$ (Nei, 1973)	$G'_{ST\_est}$ (Hedrick, 2005)	$D_{est}$ (Jost, 2008)
mfc7	0,0429	0,1026	0,0623
mfs11	0,0522	0,1364	0,0889
mfc5	0,0374	0,4012	0,3779
FS103	0,0198	0,0847	0,0662
FS115	0,0719	0,4070	0,3610
FS304	0,0030	0,0048	0,0018
FS125	0,0450	0,3321	0,3007
FS446	0,0344	0,2054	0,1771
Mittelwert	0,0383	0,2093	0,1795

**Genetische Variation zwischen Populationen**

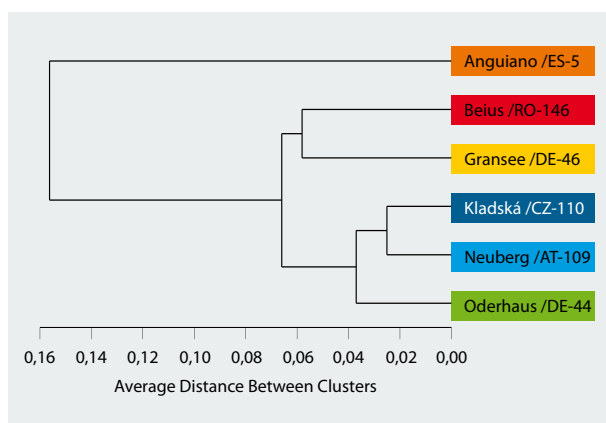
Die Loci mfc5, FS115 und FS125 tragen am meisten zur globalen Differenzierung zwischen den Populationen bei, wenn die Differenzierungsmaße  $D_{est}$  nach Jost (2008) oder  $G'_{ST\_est}$  nach Hedrick (2005) zugrunde gelegt werden (Tabelle 8). Das  $G_{ST\_est}$  nach Nei (1973) kann die vorhandenen Unterschiede zwischen den Loci dagegen nicht widerspiegeln. Das wird besonders deutlich beim Vergleich Allelfrequenzen je Herkunft (Tabelle 9). Beispielsweise zeigt der Locus mfc7 mit ähnlichen Allelfrequenzen in den sechs Herkünften eine entsprechend geringe Differenzierung  $D_{est}$ . Der  $G_{ST\_est}$ -Wert liegt jedoch etwas über dem Mittelwert. Am Locus mfc5 dagegen unterschieden sich die Allelfrequenzen zwischen den sechs Her-



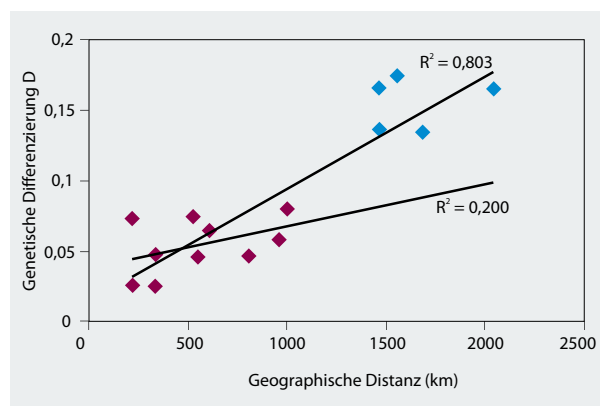
**Abbildung 3**  
Neighbour-Joining-Dendrogramme von Stichproben aus sechs Buchenherkünften

künften erheblich, was ein sehr hohes  $D_{est}$  von 0,378 auch anzeigt. Das  $G_{ST_{est}}$  entspricht jedoch etwa dem Mittelwert.

Die genetischen Ähnlichkeiten bzw. Unterschiede zwischen den sechs Buchenherkünften können graphisch als Dendrogramm dargestellt werden (Abbildung 4), ermittelt aus einer Clusteranalyse mit paarweisen genetischen Differenzierungen. Hier zeigt sich, dass sich die Herkunft aus Spanien deutlich von allen anderen abgrenzen lässt. Die anderen fünf Herkünfte bilden zwei Gruppen, einerseits Beius / RO-146 und Gransee /DE-46 sowie andererseits Kladská /CZ-110, Neuberg /AT-109 und Oderhaus /DE-44.



**Abbildung 4**  
Dendrogramm nach UPGMA-Clusteranalyse mit paarweisen genetischen Differenzierungen zwischen Stichproben aus sechs Buchenherkünften nach Jost, 2008 (Daten mit Nullallelkorrektur je Subpopulation)



**Abbildung 5**  
Lineare Beziehungen zwischen geographischer Distanz und genetischer Differenzierung bei sechs Herkünften der Buche (rote Symbole: paarweise Kombinationen mit Beteiligung der spanischen Herkunft, Erklärung im Text)

Die Abbildung 5 zeigt eine lineare Korrelation zwischen der paarweisen genetischen Differenzierung und dem geographischen Abstand mit dem Bestimmtheitsmaß  $R^2 = 0,803$  ( $p = 0,0170$ ) bei der Einbeziehung aller sechs Herkünfte. Wenn die in dieser Korrelation offensichtlich abseits stehende spanische Herkunft weggelassen wird, ist die Korrelation zwischen den übrigen fünf Herkünften nicht mehr signifikant mit  $R^2 = 0,200$  ( $p = 0,0900$ ).

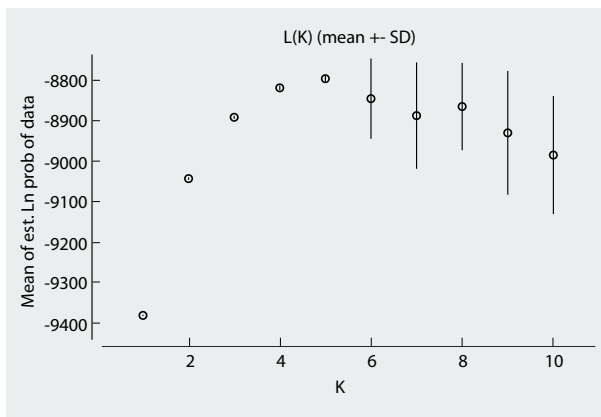
Die Darstellung der genetischen Ähnlichkeiten bzw. Unterschiede als Abstandsmatrix oder Dendrogramm beruht auf der bekannten Zugehörigkeit der einzelnen Individuen zu einer bestimmten Herkunft und Berechnungen auf Populationsebene. Eine andere Möglichkeit der Auswertung

**Tabelle 9**

Allelfrequenzen an Mikrosatelliten-Loci mit kleiner (mfc7) und großer Differenzierung (mfc5), dargestellt sind nur Allele mit einer Mindesthäufigkeit von 0,100 in einer der sechs Populationen

	Anguiano /ES-5	Oderberg /DE-44	Gransee /DE-46	Neuberg /AT-109	Kladská /CZ-110	Beius /RO-146
<b>mfc7</b>						
111	0,820	0,491	0,700	0,463	0,576	0,726
117	0,030	0,173	0,030	0,241	0,170	0,076
123	-	0,046	0,020	0,120	0,028	-
127	0,110	0,246	0,220	0,157	0,170	0,104
<b>mfc5</b>						
285	0,130	0,027	0,020	-	-	0,076
287	0,340	0,082	0,060	0,065	0,028	-
303	0,070	0,064	0,050	0,111	0,151	0,047
305	0,020	0,100	0,330	0,194	0,170	0,189
311	0,130	0,155	0,050	0,046	0,094	0,160
317	-	0,046	0,170	0,046	0,038	0,028
319	-	0,073	0,030	0,056	0,057	0,104
325	0,150	0,091	0,020	0,019	0,019	-
null	-	0,100	0,090	0,102	0,057	0,047

besteht darin, alle 315 Individuen allein mit ihrem Multilocus-Genotyp in hypothetische Populationen einzuordnen. Mit der Software STRUCTURE wurden  $K = 5$  Populationen ermittelt (Abbildung 6). In der Abbildung 7 wird die Zuordnungswahrscheinlichkeit jedes Individuums zu den verschiedenen hypothetischen Populationen jeweils in einer anderen Farbe dargestellt. Fast alle Individuen der spanischen Herkunft werden mit großer Sicherheit einer Population (Farbe „rot“ in Abbildung 7) zugeordnet. Die meisten der Population „blau“ zugeordneten Individuen gehören zur Herkunft Gransee /DE-46, aber teilweise auch zur Herkunft Beius /RO-146, den beiden Nachkommenschaften von warmen Standorten. Weitere solche Gruppierungen lassen sich auch für die hypothetischen Populationen „gelb“, „grün“ und „pink“ finden. Insgesamt jedoch grenzt sich spanische Herkunft mit ihrer fast eindeutigen Zuordnung zur Population „rot“ von den übrigen fünf Herkünften ab, die mehr oder weniger gemischt sind. In diesen Mischungen kommt die größere genetische Ähnlichkeit zum Ausdruck, die auch das Dendrogramm (Abbildung 4) gezeigt hat.



**Abbildung 6**

Ermittlung der Anzahl von Subpopulationen mit STRUCTURE und STRUCTURE Harvester: Mittelwerte und Standardabweichungen der Ln P-Werte für  $K = 1$  bis  $K = 10$  (sechs bis zwölf Wiederholungen) der Analyse der Genotypen von 315 Individuen aus sechs Buchenherkünften

## Schlussfolgerungen

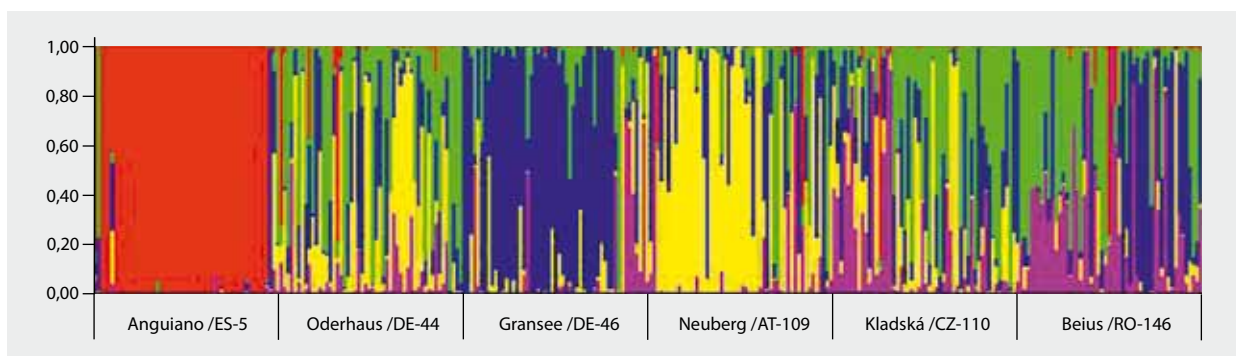
### Methode der Genotypisierung

Die angewandte Methode zur Genotypisierung der Herkunftsstichproben mit nuklearen Mikrosatellitenmarkern ist als sehr gut reproduzierbar und zuverlässig zu beurteilen. Der Anteil von nicht erkannten Heterozygoten mit einem Nullallel musste in dieser Studie jedoch auf ca. 5 % der vollständig erfassten Genotypendaten geschätzt werden. Dabei ist Homoplasie nicht ausgeschlossen, d. h. evtl. vorkommende unterschiedliche DNA-Sequenzen können sich sowohl hinter DNA-Fragmenten gleicher Länge als auch hinter fehlenden PCR-Produkten verbergen. Der evtl. damit verbundene Informationsverlust ist jedoch ohne Sequenzierung aller Fragmente nicht zu vermeiden und muss vorläufig akzeptiert werden (siehe auch Oddou-Muratorio et al., 2009). Neu entwickelte Mikrosatellitenmarker für die Baumart Buche, die weniger Nullallele besitzen sollen, wurden erst nach Abschluss der hier beschriebenen Laboruntersuchungen veröffentlicht (Lefèvre et al., 2012), und konnten nicht mehr berücksichtigt werden.

### Subpopulationsstrukturen in Nachkommenschaften von Buchenpopulationen

In jeder der sechs Herkunftsstichproben wurden mit einer STRUCTURE-Analyse zwei oder drei Subpopulationen gefunden (Abbildung 2), die eher im Hardy-Weinberg-Gleichgewicht stehen als die Gesamtstichproben der jeweiligen Herkunft.

In natürlich verjüngten Buchenbeständen existieren Familienstrukturen, die als größere genetische Ähnlichkeit von Individuen in räumlicher Nachbarschaft mehrfach beschrieben wurden. Eine räumliche Autokorrelation bis ca. 30 m wurden sowohl mit neun Isoenzymmarkern von Dounavi (2000) als auch mit vier bzw. sechs Mikrosatellitenmarkern von Vornam et al. (2004) und Jump und Peñuelas (2007) in je einem Bestand festgestellt. Durch die wesentlich größere Zahl von 250 AFLP-Markern konnten von Jump und Peñuelas (2007) in derselben Untersuchung eine Autokorrelation bis ca. 110 m nachgewiesen werden. Nyári (2010) fand mit fünf Mikrosatellitenmarkern räumliche Autokorrelationen von ca. 30 m im einem deutschen und ca. 70 m in einem ungarischen



**Abbildung 7**

Zuordnung von 315 untersuchten Individuen der Buche zu fünf hypothetischen Populationen mit STRUCTURE (Individuen geordnet nach Herkünften)

schen Buchenbestand. Auch in der Naturverjüngung eines Buchenbestandes wurden genetische Differenzierungen zwischen räumlich getrennten Gruppen nachgewiesen, die sich wegen der begrenzten Pollen- (im Mittel 44 bis 57 m) und Samenverbreitung (im Mittel 11 m) herausgebildet hatten (Oddou-Muratorio et al., 2011).

Insgesamt liegen diese Größenordnungen in einem Bereich, bei dem mit dem gruppierten Auslegen von Netzen in räumlich verschiedenen Bereichen eines Bestandes unterschiedliche Teilpopulationen des Bestandes in einer Ernte erfasst werden. Subpopulationsstrukturen in Nachkommenschaften müssen deshalb als typisch für Saatgut aus natürlich verjüngten Buchenbeständen angesehen werden, die nach in der Praxis üblichen Methoden beerntet wurden.

### **Repräsentativität der Saatguternten bei sechs Buchenherkünften**

Die Ergebnisse der Haplotypenanalyse (Tabelle 7) und die Neighbour-Joining-Dendrogramme (Abbildung 3) der Individuen je Herkunft lassen Rückschlüsse auf die Art der Saatguteinsammlung zu. Bei den fünf Herkunftsorten Oderhaus /DE-44, Gransee /DE-46, Neuberg /AT-109, Kladská /CZ-110 und Beius /RO-146 kann eine Ausgewogenheit der relativen Anteile der Erntebäume angenommen werden. Das Vorliegen einer Halbmast im Bestand Gransee /DE-46 im Vergleich zu Vollmasternten bei den anderen vier Herkunftsorten hat sich nicht nachweisbar auf die Repräsentativität, zumindest des für die Versuchsanlage verwendeten Erntejahrgangs, ausgewirkt. Ebenso scheint trotz des geringen Buchenanteils im Bestand Beius /RO-146 eine ausreichende wechselseitige Bestäubung und repräsentative Saatguternte stattgefunden zu haben. Unter Repräsentativität wird hier eine Beerntung einer ausreichenden Anzahl von Bäumen in ausgewogenen Mengenteilen verstanden. Im Gegensatz dazu steht die Herkunft Anguiano /ES-5, die deutlich ausgeprägte Verwandtschaftsgruppen bis hin zu einer vermuteten Halbgeschwisterfamilie von zwölf Individuen zeigt. Diese Struktur lässt sich nicht allein mit der geringeren Anzahl von nur ca. zehn Erntebäumen im Vergleich zu ca. zwanzig Erntebäumen bei den anderen fünf Herkunftsorten erklären. Hier müssen zusätzliche Faktoren beteiligt gewesen sein. Evtl. wurden deutlich weniger als zehn Bäume beerntet oder der Anteil der Saatgutmengen je Baum war sehr ungleich verteilt. Nähere Informationen zur Bestandesbeschreibung der sechs Herkunftsorte, zur Voll- bzw. Halbmast, zum Buchenanteil in den Erntebeständen und zur Anzahl der Erntebäume befinden sich bei Liesebach M (2012).

### **Genetische Variation innerhalb der Populationen**

Zwei Populationen, Anguiano /ES-5 und Gransee /DE-46, haben eine geringere genetische Vielfalt, ausgedrückt als Anzahl Allele pro Locus, und eine geringere Heterozygotie als die anderen vier Herkunftsorte (Tabelle 6). Diese beiden Herkunftsorte sind auch diejenigen, welche bei einer Halbmast beerntet wurden, während für die anderen vier eine Vollmast dokumentiert wurde (Liesebach M, 2012). Die Stichprobenumfänge (N = 50 vs. N = 53...55) unterscheiden sich nur ge-

ringfügig und können den Unterschied in der Anzahl der Allele nicht erklären.

Ob jedoch das Vorhandensein einer Voll- bzw. Halbmast oder die genetische Zusammensetzung des gesamten Erntebestandes selbst für die festgestellten Ergebnisse verantwortlich sind, kann aus den vorliegenden Daten nicht gefolgert werden. Dazu müssten mehr als nur sechs Bestände unterschiedlicher Zusammensetzung wiederholt beerntet werden. Eine teilweise differenzierende genetische Struktur von Nachkommenschaften aus zwei unterschiedlichen Erntejahren aus einem Altbestand wurde am Beispiel der Weißstanne nachgewiesen, jedoch ohne Informationen zur Anzahl der Erntebäume oder einer Halb- bzw. Vollmast (Hussendörfer und Konnerth, 1998).

Für spanische Herkunftsorte, einschließlich einer Gruppe von Populationen bei Anguiano, wurde auch für Isoenzymmarker eine geringere genetische Vielfalt im Vergleich zu mittel- und südosteuropäischen Herkunftsorten beschrieben (Comps et al., 2001).

Es konnten aus den genetischen Untersuchungen keine Hinweise auf eine mögliche genetische Ursache für das deutlich unter dem Mittel liegende Wachstum der Herkunft Kladská /CZ-110, wie z. B. eine erhöhte Inzuchtrate, festgestellt werden.

### **Genetische Variation zwischen den Populationen**

Zwei unterschiedliche Verfahren, zum einen eine UPGMA-Clusteranalyse (Abbildung 4) mit vorgegebenen Populationen und zum anderen eine STRUCTURE-Analyse, bei der Individuen nur aufgrund ihres Genotyps verschiedenen Populationen zugeordnet werden (Abbildung 7), zeigten die deutliche Abgrenzung der spanischen von den anderen fünf Herkunftsorten. Damit bestätigt sich ein weiteres Mal die unterschiedliche Abstammung aus verschiedenen eiszeitlichen Refugien von mittel- und südosteuropäischen Buchenvorkommen im Vergleich zu iberischen Vorkommen. Diese wurde in einer großen Metaanalyse, bei der Chloroplasten-DNA-Marker und nukleare Isoenzymmarker von ca. 400 Buchenpopulationen aus ganz Europa zusammen mit Pollenprofilen und Makrofossilien ausgewertet wurden, herausgearbeitet, um eiszeitliche Refugien zu identifizieren und die nacheiszeitliche Ausbreitung der Buche zu rekonstruieren (Magri et al., 2006).

Von den fünf Herkunftsorten, die von einem oder mehreren südosteuropäischen Refugien abstammen, bilden die drei Herkunftsorte Kladská /CZ-110, Neuberg /AT-109 und Oderhaus /DE-44 von relativ kalten Standorten eine Gruppe mit größerer Ähnlichkeit und die beiden Herkunftsorte Beius /RO-146 und Gransee /DE-46 von relativ warmen Standorten eine andere Gruppe (Abbildung 4). Eine größere genetische Ähnlichkeit bei geringerer räumlicher Entfernung konnte nicht nachgewiesen werden. Alle Entfernungen zwischen den fünf Herkunftsorten betragen jedoch mehr als 200 km. Unterdurchschnittliche genetische Abstände zwischen Populationen bei Isoenzymmarkern wurden mit geostatistischen Methoden für Entfernungen bis 200 km gefunden (Degen und Scholz, 1998). Räumliche Nachbarschaften bei diesen fünf

Herkünften entsprechen jedoch nicht der Gruppierung nach Jahresmitteltemperaturen, da sie nicht wegen ihrer Lage auf einem Nord-Süd-Gradienten, sondern wegen ihrer Höhenlage als eher kalt bzw. warm eingeordnet wurden.

Ob sich hinter den beiden mit der Temperatur assoziierten Gruppen (Dendrogramm in Abbildung 4) mit drei bzw. zwei Herkünften anpassungsrelevante genetische Strukturen verbergen, kann aus den Ergebnissen noch nicht abgeleitet werden. Es müssten wesentlich mehr Populationen genetisch charakterisiert werden und weitere Marker zum Einsatz kommen, um die Effekte von 1. eiszeitlichen Refugien, 2. ausgleichend wirkendem, wenn auch räumlich begrenztem Genfluss und 3. diversifizierender lokaler Selektion an vorliegende Umweltbedingungen voneinander zu trennen. Bisherige europaweite Untersuchungen mit überwiegend neutralen Markern konnten nur die Abhängigkeit der aktuellen genetischen Struktur von Buchenpopulationen von den jeweiligen eiszeitlichen Refugien (Comps et al., 2001; Magri et al., 2006) oder von der Fragmentierung von Populationen (Jump und Penuelas, 2006) nachweisen.

### Ausblick

Von allen Bäumen liegen individuell zugeordnete Daten zum Durchmesser- und Höhenwachstum, zur Stammform sowie zum Austriebsverhalten und zur Herbstverfärbung für gemeinsame Auswertungen vor. Eine detaillierte Bearbeitung steht noch bevor und wird später dargestellt werden. Wenige Assoziationen deuten sich bereits an und bieten evtl. Ansatzpunkte für weitere Untersuchungen, auch an den eingelagerten DNA-Proben.

Vorläufig bleibt festzustellen, dass genetisch hochwertiges Saatgut für Bestandesbegründungen einerseits eine Anpassung an den vorgesehenen Standort besitzen muss, andererseits aber auch ein hohes Anpassungspotential an sich potentiell ändernde Bedingungen mitbringen soll. Das erste kann mit geeigneten Herkunftsgebieten für die Saatguternte und entsprechenden Anbauempfehlungen gesichert werden. Um ein ausreichendes Anpassungspotential zu Verfügung zu haben, muss auch eine große und repräsentative genetische Variation gewährleistet sein. Ausgehend von hohen Individuenzahlen mit adäquater genetischer Variation ist dann ein Spielraum für Selektion und lokale Anpassungen gegeben. Eine entscheidende Bedeutung der generativen Reproduktion mit der damit verbundenen Rekombination der Erbanlagen im Anpassungsprozess von europäischen Buchenwäldern an Klimaänderungen wurde von Kramer et al. (2010) besonders hervorgehoben. Erkenntnisse, die aus der langjährigen Beobachtung von Herkunftsversuchen gewonnen werden, tragen dazu bei, das Anpassungspotential und die Leistungsfähigkeit von Populationen unter dem Aspekt der Erhaltung genetischer Ressourcen und Ökosysteme sowie dem Aspekt der Leistungsfähigkeit der Forstwirtschaft zu beurteilen.

### Dank

Meiner Kollegin Jutta Buschbom möchte ich für einige kritische Bemerkungen zum Manuskript sowie besonders für den Hinweis zur Existenz des „STRUCTURE Harvesters“ herzlich danken. Ebenso danke ich den Gutachtern für wertvolle Hinweise.

### Literatur

- Bohonak AJ (2002) IBD (Isolation By Distance) : a program for analyses of isolation by distance. *J Hered* 93:153-154
- Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* 5:453-455
- Buiteveld J, Vendramin GG, Leonardi S, Kramer K, Geburek T (2007) Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. *For Ecol Manage* 247:98-106
- Buschbom J (2010) FormatIn - transformation of genotypic data associated with population, family and stage information : version 1.2 [online]. Zu finden in <<http://www.vti.bund.de/de/startseite/institute/fg/software.htm>> [zitiert am 09.10.2012]
- Chakraborty R, Andrade MD, Daiger SP, Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Ann Hum Genet* 56:45-57
- Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100:106-113
- Comps B, Gömöry D, Letouzey J, Thiebaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. *Genetics* 157:389-397
- Comps B, Thiebaut B, Paule L, Merzeau D, Letouzey J (1990) Allozymic variability in beechwoods (*Fagus sylvatica* L.) over central Europe : spatial differentiation among and within populations. *Heredity* 65:407-417
- Crawford NG (2010) SMOGD : software for the measurement of genetic diversity. *Mol Ecol Resources* 10:556-557
- Degen B, Scholz F (1998) Spatial genetic differentiation among populations of European beech (*Fagus sylvatica* L.) in western Germany as identified by geostatistical analysis. *Forest Genetics* 5:191-199
- Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* 50:2515-2520
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J R Stat Soc B* 39:1-38
- Dounavi A (2000) Familienstrukturen in Buchenbeständen (*Fagus sylvatica*). Göttingen : Univ, 142 p
- Dumolin S, Demesure B, Petit RJ (1995) Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theor Appl Genet* 91:1253-1256
- Earl DA, vonHoldt BM (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet* 4(2):359-361
- Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of population differentiation based on GST and D : forget GST but not all of statistics! *Mol Ecol* 19:3845-3852
- Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* 59:1633-1638
- Hussendörfer E, Konnerth M (1998) Untersuchungen zur genetischen Repräsentativität von Prüfgliedern in Provenienzversuchen m Beispiel der Weißtanne. *Allg Forst Jagdzeitg* 169:61-70
- Jin L, Chakraborty R (1994) Estimation of genetic distance and coefficient of gene diversity from single-probe multilocus DNA fingerprinting data. *Mol Biol Evol* 11:120-127
- Jost L (2008) GST and its relatives do not measure differentiation. *Mol Ecol* 17:4015-4026

- Jump AS, Hunt JM, Martinez-Izquierdo JA, Penuelas J (2006) Natural selection and climate change : temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Mol Ecol* 15:3469-3480
- Jump AS, Penuelas J (2006) Genetic effects of chronic habitat fragmentation in a wind-pollinated tree. *Proc Nat Acad Sci USA* 103:8096-8100
- Jump AS, Peñuelas J (2007) Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind-pollinated tree, *Fagus sylvatica*. *Mol Ecol* 16:925-936
- Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 16:1099-1106
- Konnert M, Ziehe M, Tröber U, Maurer WD, Janßen A, Sander T, Hussendörfer E, Hertel H (2000) Genetische Variation der Buche (*Fagus sylvatica* L.) in Deutschland : gemeinsame Auswertung genetischer Inventuren über verschiedene Bundesländer. *Forst Holz* 55:403-408
- Kramer K, Buiteveld J, Forstreuter M, Geburek T, Leonardi S, Menozzi P, Povelion F, Schelhaas MJ, Teissier du Cros E, Vendramin GG, van der Werf DC (2008) Bridging the gap between ecophysiological and genetic knowledge to assess the adaptive potential of European beech. *Ecol Model* 216:333-353
- Kramer K, Degen B, Buschbom J, Hickler T, Thuiller W, Sykes MT, de Winter W (2010) Modelling exploration of the future of European beech (*Fagus sylvatica* L.) under climate change : range, abundance, genetic diversity and adaptive response. *Forest Ecol Manage* 259:2213-2222
- Langella O (2010) Populations : version 1.2.31 [online]. Zu finden in <<http://bioinformatics.org/~tryphon/populations/>> [zitiert am 09.10.2012]
- Lefèvre S, Wagner S, Petit RJ, De Lafontaine G (2012) Multiplexed microsatellite markers for genetic studies of beech. *Mol Ecol Resources* 12:484-491
- Liesebach M (2012) Der Internationale Herkunftsversuch mit Rot-Buche von 1993/95 – Beschreibung der ausgewählten sechs Herkünfte und zwei Versuchsflächen. *Landbauforsch* 62(4):159-168
- Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Lataoava M, Litt T, Paule L, Roure JM, Tantau I, Knaap WOvd, Petit RJ, Beaulieu J-Ld (2006) A new scenario for the quaternary history of European beech populations : palaeobotanical evidence and genetic consequences. *New Phytol* 171:199-221
- Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Mol Ecol* 7:639-655
- Maurer W, Cremer E, Degen B, Konnert M, Tröber U (2008) Erfassung genetischer Strukturen wichtiger Baumarten : Teillos 1: Erfassung der genetischen Struktur der Rotbuche (*Fagus sylvatica*) als Grundlage für ein genetisches Monitoring wichtiger Waldbaumarten in Deutschland [online]. Zu finden in: <[http://www.ble.de/SharedDocs/Downloads/03\\_Forschungsfoerderung/04\\_BiologischeVielfalt/ProjektberichtRotbuche.html](http://www.ble.de/SharedDocs/Downloads/03_Forschungsfoerderung/04_BiologischeVielfalt/ProjektberichtRotbuche.html)> [zitiert am 10.10.2012]
- Merzeau D, Comps B, Thiébaud B, Letouzey J (1994) Estimation of *Fagus sylvatica* L mating system parameters in natural populations. *Ann For Sci* 51:163-173
- Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Nat Acad Sci USA* 70:3321-3323
- Nyári L (2010) Genetic diversity, differentiation and spatial genetic structures in differently managed adult European beech (*Fagus sylvatica* L.) stands and their regeneration. *Forstarch* 81:156-164
- Oddou-Muratorio S, Klein EK, Vendramin GG, Fady B (2011) Spatial vs. temporal effects on demographic and genetic structures : the roles of dispersal, masting and differential mortality on patterns of recruitment in *Fagus sylvatica*. *Mol Ecol* 20:1997-2010
- Oddou-Muratorio S, Vendramin GG, Buiteveld J, Fady B (2009) Population estimators or progeny tests : what is the best method to assess null allele frequencies at SSR loci? *Conserv Genet* 10:1343-1347
- Page RDM (2001) TreeView : tree drawing software for Apple Macintosh and Microsoft Windows. Version 1.6.6 [online]. Zu finden in <<http://taxonomy.zoology.gla.ac.uk/rod/rod.html>> [zitiert am 09.10.2012]
- Pastorelli R, Smulders MJM, Van't Westende WPC, Vosman B, Giannini R, Vetтори C, Vendramin GG (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. *Mol Ecol Notes* 3:76-78
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945-959
- Rousset F (2008) GENEPOP'007 : a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resources* 8:103-106
- SAS for Windows, version 9.2, SAS Inst, Cary, NC, USA
- Scafi M, Troggio M, Piovani P, Leonardi S, Magnaschi G, Vendramin GG, Menozzi P (2004) A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). *Theor Appl Genet* 108:433-441
- Seifert S (2011) Variation of candidate genes related to climate change in European beech (*Fagus sylvatica* L.). Göttingen : Univ, 134 p
- Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Nat Acad Sci* 100:9440-9445
- Tanaka K, Tsumura Y, Nakamura T (1999) Development and polymorphism of microsatellite markers for *Fagus crenata* and the closely related species, *F. japonica*. *Theor Appl Genet* 99:11-15
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535-538
- Vornam B, Decarli N, Gailing O (2004) Spatial Distribution of Genetic Variation in a Natural Beech Stand (*Fagus sylvatica* L.) Based on Microsatellite Markers. *Conservation Genetics* 5: 561-570

