Characterization of EST-SSRs for European beech (*Fagus sylvatica* L.) and their transferability to *Fagus orientalis* Lipsky, *Castanea dentata* Bork., and *Quercus rubra* L.

Katrin Burger, Markus Müller*, Oliver Gailing*

University of Goettingen, Faculty for Forest Sciences and Forest Ecology, Forest Genetics and Forest Tree Breeding. Büsgenweg 2, 37077 Göttingen, Germany

*Corresponding authors: Oliver Gailing, E-mail: ogailin@gwdg.de; Markus Müller, E-mail: mmuellef@gwdg.de

Abstract

Due to ongoing climate change, forests are exposed to changing environmental conditions, such as increasing temperatures and lower precipitation, to which trees have to adapt. Successful adaptation to changing and variable environments requires sufficient genetic variation within tree populations. Knowledge of the genetic variation of trees is therefore essential, as it provides information for the long-term conservation, stability and productivity of forests. The genetic variation of a species can be analysed with molecular markers. Despite growing genomic and genetic resources for European beech (*Fagus sylvatica* L.), which is one of the economically and ecologically most important forest tree species in Central Europe, the number of molecular markers for population genetic analyses is still limited. Therefore, the aim of the work is the development of new EST-SSR markers for this species.

A total of 72 DNA samples of European beech from three widely separated regions in Germany were used to test 41 primers for variation and polymorphism, 35 of which were originally developed for American beech (*Fagus grandifolia* Ehrh.) and 6 for red oak (*Quercus rubra* L.). Fifteen of the primers were polymorphic, 13 monomorphic and 13 did not amplify. In addition, the transferability of the markers was successfully tested in the related species *Castanea dentata* Bork., *Fagus orientalis* Lipsky and *Q. rubra*. The EST-SSR markers tested in this study will be useful for future population genetic analyses and extend the set of available markers in European beech.

Keywords: *: Microsatellites, Fagaceae, genetic diversity, transferability*

DOI:10.2478/sg-2018-0019
edited by the Thünen Institute of Forest Genetics

Introduction

European beech (*Fagus sylvatica* L.) would dominate the natural vegetation in Central Europe and currently occupies different habitats, which reach from the mountain regions of Southern and Eastern Europe to the lowlands of Central Europe and Southern Sweden (Paule, 1995; Fang and Lechowicz, 2006). European beech belongs to the beech family (Fagaceae), which contains the economically and ecologically important tree species of the genera oaks (*Quercus*), chestnuts (*Castanea*) and beeches (*Fagus*) (Aldrich et al., 2003). In the course of climate change, trees are exposed to a number of stresses, such as increased temperatures and reduced rainfall (Kölling and Zimmermann, 2007) and hence have to adapt to changing environmental conditions, which may be hampered by their long-life cycles. Genetic variation is the prerequisite for forest trees to be able to adapt to environmental changes and maintain their adaptability for future generations (e.g., Vornam et al., 2004; Gailing et al. 2008).

Microsatellites (short sequence repeats- SSRs) are important genetic markers often used in population genetic analyses (Ellis and Burke, 2007; Guichoux et al., 2011). The advantages of SSR markers are their codominance, their high degree of polymorphism, their uniform distribution across the entire genome and their high reproducibility (e.g., Durand et al., 2010). Among the microsatellite markers, EST-SSR markers are playing an increasingly relevant role (Ellis and Burke, 2007). One of their advantages over nSSR markers is that they are located in regions of the DNA that are strongly conserved within phylogenetically related species. As a result, they often have very high transferability rates within closely related species, especially within the Fagaceae (Durand et al., 2010). Genetic

microsatellite markers can be derived from EST libraries and are located either in coding regions or in the 5' or 3' untranslated regions (UTRs) (Ellis and Burke, 2007). Despite growing genetic and genomic resources for European beech (Lalagüe et al., 2014; Lesur et al., 2015; Müller et al., 2015; Dounavi et al., 2016; Müller et al., 2017; Mishra et al., 2018) the number of available SSR markers for this species is still limited. Especially, only a few gene-based EST-SSRs that were originally developed in other species are available for *F. sylvatica* (e.g., Ueno et al., 2009; Dounavi et al., 2016). Here, we take advantage of EST-libraries and derived primers that were developed in American beech, *Fagus grandifolia* Ehrh. The aim of this study was to establish a new set of EST-SSR markers for European beech, and to analyze genetic variation and differentiation at these markers in selected beech populations in Germany. The specific objectives of the study were: (I) to test different markers developed for *F. grandifolia* and *Quercus rubra* L. for amplification and polymorphism in *F. sylvatica*, and (II) to test the amplified polymorphic markers for their transferability to the species *Fagus orientalis* Lipsky, *Castanea dentata* Borkh. and *Q. rubra*.

Materials and Methods

Plant materials

Leaf samples were collected from beech seedlings that are part of a translocation experiment within the project "Biodiversity Exploratories" (https://www.biodiversity-exploratories.de/1/home/). The seedlings originate from populations from north, middle, and south Germany: Schorfheide Biosphere Reserve (SEW5, latitude: 53°3'25.3", longitude: 13°53'7.3", altitude: 64 m), Hainich National Park (HEW10, latitude: 51°5'24", longitude: 10°27'44.8", altitude: 378 m), and Biosphere Reserve Swabian Alb (AEW5, latitude: 48°25'10.6", longitude: 9°24'52.9", altitude: 788 m). The populations AEW5 and SEW5 are managed stands, whereas HEW10 is an unmanaged stand. For each of the three populations 24 samples were used for the characterization of new EST-SSRs. To test transferability across species 8 *F. orientalis* samples from Turkey (Gailing and von Wühlisch, 2004) and 8 samples of each *Q. rubra* and *C. dentata* were also included.

Marker analysis

A total of 41 EST-SSR loci (Table 1, Supplementary Material 1) were initially tested for amplification and polymorphism in 5 *F. sylvatica* samples. Thereof, 35 primer pairs were originally developed for American beech (*F. grandifolia*), of which 28 came from the 'Hardwood Genomics Project' (https://www.hardwoodgenomics.org/Transcriptome-assembly/1963031?tripal_pane=group_description_download) and seven markers from Kubisiak et al. (2009). The other markers were originally developed for *Q. rubra* (https://www.hardwoodgenomics.org/Transcriptome-assembly/1963023?tripal_pane=group_description_download). The annotation of the genes was obtained by searching the individual EST primer sequences in the respective contigs (*F. grandifolia* and *Q. rubra*) to identify the complete contig sequences (*F. grandifolia* and *Q. rubra*) for similarity searches against the UniProt Viridiplantae database (The UniProt Consortium 2017) using BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1990).

Fifteen markers that amplified polymorphic loci were selected for the population analysis and were tested for amplification and polymorphism in 24 *F. sylvatica* samples from the populations (AEW5, HEW10 and SEW5) described above, and in 8 samples of *F. orientalis* (Gailing and von Wühlisch, 2004), *Q. rubra*, and *C. dentata*, respectively. Thirteen markers that were monomorphic in *F. sylvatica* were also tested for polymorphism in *F. orientalis*.

For PCR amplifications a tailed-primer approach was used (Schuelke, 2000; Kubisiak et al., 2009). A 13.8 µl PCR mix was prepared consisting of 1.5 µl reaction buffer (containing 0.8 M Tris-HCl and 0.2 M $(NH_4)_2SO_4$), 1.5 µl $MgCl_2$ (25 mM), 1 µl dNTPs (2.5 mM each dNTP), 0.2 µl HOTFIREPol *Taq* polymerase (Solis BioDyne, Estonia) (5 units/µl), 5.5 µl H_2O, 0.2 µl tailed forward primer (5 picomole/µl), 0.5 µl PIG-tailed reverse primer (5 picomole/µl) (Schuelke 2000, Kubisiak et al. 2009), 1 µl M13 (6-FAM/HEX) primer (5 picomole/µl), and 2 µl DNA (ca. 0.6 ng/µl). The PCR reaction was performed in a Thermal Cycler (MJ Research PTC 200) with a touchdown program. The PCR profile consisted of an initial denaturation at 95 °C for 15 minutes followed by 10 touchdown cycles at 94°C for 1 min, 1 min at 60°C (decreasing 1 °C each cycle) and 1 min at 72°C, followed by 25 cycles at 94°C for 1 min, 50°C for 1 min and 72°C for 1 min, and a final extension at 72°C for 20 min.

PCR products were separated on an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, USA). Up to 4 primers were simultaneously subjected to fragment length analysis in multiplexes (multiplex I: FgSIC0024, FgSI0006 and FS_C2361, multiplex II: FS_C6563, FS_C5453, FS_C1968 and FS_C5430, multiplex III: FS_C7377, FgSIC0009, FS_C5931 and FgSIC0025, multiplex IV: FS_C6785, FS_C4971 and FgSIC0016, multiplex V: FS_C8040 and FS_C7797). Scoring of alleles was conducted using 'GeneMapper' version 4.1 (Applied Biosystems, Foster City, USA).

Table 1
Primer sequences and descriptions of EST-SSRs

Developed for	Primer Name	Repeat motif	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Observed size (in bp) in *F. sylvatica*	Annotation of sequences
*Fagus grandifolia***	FS_C5756	$(CTC)_5$	GCAGAACTTGGTGTTGAGCG	ATTCGATGGGAAACGGACGG	275	*dof zinc finger protein*
*Fagus grandifolia***	FS_C5675	$(TTC)_5$	TCTTCGGTCCTTCAAAGGCC	TTCTTCTTCGCTGCTGCTCG	263	*Zinc finger protein constant-similar-9-like isoform*
*Fagus grandifolia***	FS_C7313	$(AGA)_5$	GAAGTACAAATGGACGGCGG	TCCAGCTCTCTTGGCAAAGG	160	*molecular chaperone regulator 7-like*
*Fagus grandifolia***	FS_C6090	$(TG)_7$	CAATGATGCAAGGGCTGAGC	CTTCTCCCTCCTGTCAGTCC	254	*thermospermine synthase*
*Fagus grandifolia***	FS_C1924	$(TGG)_5$	TCCAAAGCCAAAGCCAAAGC	GATCGCAGCCATTTCCAACC	275	*Chaperon Isoform*
*Fagus grandifolia***	FS_C8168	$(GAG)_5$	GAAGAAGGTCGTGGGTGATGG	CCTGGCTGTTACATCTCTTTCC	123	*BHLH Protein*
*Fagus grandifolia***	FS_C5931	$(AAT)_5$	CTTTGCCACGTGTCAACTCC	AGTTTCTGAACCTCTCCGGC	284-287	*NAC transcription factor*
*Fagus grandifolia***	FS_C6563	$(GAA)_5$	TAGAAAGGGAGAAGGTCCGGC	GTTCTTTCACCACCACCACC	155-161	*RNA helicase*
Fagus grandifolia	FS_C7797	$(TCT)_5$	TGGAATCAAACACCATTGCC	CTAGCGACTCCGAAGTTCCC	240-248	*serine/threonin protein*
Fagus grandifolia	FS_C5453	$(TGA)_5$	GATGGAGTGCGTAGGAGAGC	TCACTTTGCTGACCAATCCC	225-228	*homeobox protein*
Fagus grandifolia	FS_C6785	$(ACC)_5$	TCGCTCTGAATCACCTGTCG	GCTTGAGCGACTCGGTTAGG	183-192	*ribosomal protein*
Fagus grandifolia	FS_C4971	$(GGT)_5$	CATGTGCAATGTAGCTGGGC	TCAAAGGACCCTGCAATCCC	195-201	*zinc finger protein*
Fagus grandifolia	FS_C7377	$(GAT)_5$	AATCGGACGGTCCATAGTGC	AGATCCGAGCTCAACTCACC	139-166	*plecstrin homology domain*
Fagus grandifolia	FS_C1968	$(TC)_7$	ATCGTTTCCACTTCCTCCGG	TTTCATGCACCCTCTCTAGG	299-301	*auxin-response protein*
Fagus grandifolia	FS_C5430	$(GA)_7$	TGGAGGACTTGTTAAGGTAGGC	CGCCCAACCAAATACAAAGGG	170-172	*uncharacterized protein*
Fagus grandifolia	FS_C8040	$(TTG)_5$	GATGACGTTTGGATCGGTGC	CTACGGCTCCAGTCCTTACG	349-366	*calcium-binding protein*
Fagus grandifolia	FgSI0006	$(TGT)_7$	TTAACACCGCGGTAGAGACC	GCTCCAAGCTCTTGCTCACT	259-264	*Cold regulating 314 thylacoid membrane*
*Fagus grandifolia***	FgSI0009	$(ACC)_5$	TACCCATGCCCATATCCAGT	GGAAAGAAGAAGGGTGGAGG	214-217	*heavy metal-associated isoprenylated protein*
*Fagus grandifolia**	FgSI0016	$(AAC)_5$	CGGAGAAGGACAAGGACAAG	TTCTTCGTAGAGCCTTGATGC	157-172	*Peptidyl-prolyl cis-trans isomerase*
*Fagus grandifolia***	FgSI0024	$(TCG)_5$	GAATCGTCGGAATCGTTGTC	CGGTCGAGGATGATGACTTT	166-174	*RING type E3 ubiquitin transferase*
*Fagus grandifolia***	FgSI0025	$(CTT)_5$	TCCCTGATCAGCTTTTCTCAA	TTCGTAGTATAAGGCCAAAGAAGA	184-193	*Vascular NAC domain protein*
*Fagus grandifolia***	FgSI0026	$(CTT)_5$	TCGATGGTGTCACTGCTCTC	TCGTTCTGTTTCAGGCTTCA	162	*family ubiquitin protein*
*Quercus rubra**	FS_C1702	$(ACA)_5$	CTACCCTGTTGCCTCCTCTG	GAGGCCTCATCACATGGACT	203	*Accelerator of RNA Polymerase II*
*Quercus rubra**	FS_C8183	$(AGC)_5$	TATTCAACCACAGCTGCCTG	ACAGCTGCCTCTGTGGATCT	196	*auxin response factor*
*Quercus rubra**	FS_C2791	$(GA)_7$	CGAAACAGAGAGAACCCAAGA	CTTCAAACATCCAGCGTTGA	292	*ribosomal protein*
*Quercus rubra**	FS_C2660	$(GAG)_5$	AGCAGAATTCGCCAAGTGAT	TGCCTTTGCATTCTCCTCTT	212	*uncharacterized protein*
*Quercus rubra**	FS_C2361	$(GAA)_5$	AGGTCCTTCAGTTTGGGAGC	ATTCCCATGCATCAAAATCC	196-205	*one-helix protein 2*

*: derived from a *Quercus rubra* EST-library (https://www.hardwoodgenomics.org/Transcriptome-assembly/1963023?tripal_pane=group_description_download). All other markers were derived from a *Fagus grandifolia* EST-library (https://www.hardwoodgenomics.org/Transcriptome-assembly/1963031?tripal_pane=group_description_download), **: developed in Kubisiak et al., 2009

Data analysis

Genetic variation in populations was calculated as the number of alleles per locus (N_a), observed heterozygosity (H_o), expected heterozygosity (H_e), and number of private alleles in the 'GenAlEx' software version 6.503 (Peakall and Smouse, 2006; Peakall and Smouse, 2012). GenAlEx was further used to calculate the probability of identity ($P_{(ID)}$) and the probability of identity among siblings ($P_{(ID)sibs}$). Inbreeding coefficients (F_{IS}) and their significance were determined using the ‚Fstat' v. 2.9.3 software (Goudet, 2002). Significant deviations from zero were determined after Bonferroni correction (α= 0.05) to compensate for type I errors. In addition, linkage disequilibrium (LD) was calculated for each pair of loci in the three populations using ‚Genepop' version 4.7 (Rousset, 2008) based on the following settings: dememorization 10000, batches 100 and iterations per batch 5000. Presence and frequency of null-alleles were estimated with the Micro-Checker software 2.2.3 (Van Oosterhout et al. 2004).

To measure the genetic variation among the populations, an analysis of molecular variance (AMOVA) was performed with ‚GenAlEx' using 9999 permutations. The genetic differentiation among populations was also calculated as the fixation index F_{ST} for individual markers and across all markers in 'GenAlEx'.

Results

Out of the 41 primer pairs tested, 28 amplified products in the expected size range and 15 of them were polymorphic (Table 1, see Supplementary Material 1 for primers that did not amplify). The average number of alleles per locus ranged from 3.93 for population SEW5 to 4.40 for population AEW5 (Table 2). For individual loci, N_a ranged from 2 to 8. Inbreeding coefficients across all markers were not significantly different from zero in any population. For polymorphic loci, the expected heterozygosity (H_e) ranged from 0.223 to 0.817 in AEW5 (mean H_e: 0.512), from 0.080 to 0.771 in HEW10 (mean H_e: 0.514), from 0.081 to 0.780 in SEW5 (mean H_e: 0.523) and the observed heterozygosity (H_o) ranged from 0.083 to 0.917 in AEW5 (mean H_o: 0.506), from 0.083 to 0.875 in HEW10 (mean H_o: 0.485), and from 0.083 to 0.833 in SEW5 (mean H_o: 0.514). Genetic differentiation (F_{ST}) among the three populations was relatively low ranging from 0.002 for FS_C7977 to 0.055 for FS_C6785, with a mean differentiation of 0.025 (Table 2). The probability of identity ($P_{(ID)}$) across loci was 2.6E-09 for AEW5, 3.7E-09 for HEW10, and 2.7E-09 for SEW5, whereas the probability of identity among siblings ($P_{(ID)sibs}$) was 1.6E-04 for AEW5, 1.7E-04 for HEW10, and 1.4E-04 for SEW5. The percentage of loci in LD was 8.6 % in AEW5, 3.8 % in HEW10, and 2.9 % in SEW5 (mean over all populations: 5.1 %). The presence of null-alleles was detected only for locus FgSI0009 in the populations AEW5 and HEW10, and for locus FS_C5430 in population HEW10 (Supplementary Material 2). The AMOVA showed that 98 % of the molecular variance was within populations (7 % between individuals and 91 % within individuals) and only 2 % among populations.

Table 2
Summary of genetic diversity parameters for the three beech populations

		AEW5					HEW10					SEW5				
Locus	F_{ST}	N	N_a	H_o	H_e	F_{IS}	N	N_a	H_o	H_e	F_{IS}	N	N_a	H_o	H_e	F_{IS}
FS_C8040	0.004	24	5	0.458	0.595	0.249	24	4	0.708	0.608	-0.145	24	4	0.583	0.581	0.017
FS_C7797	0.002	24	2	0.083	0.080	-0.022	24	2	0.083	0.080	-0.022	24	2	0.125	0.117	-0.045
FgSI0006	0.030	24	3	0.625	0.622	0.016	24	4	0.708	0.654	-0.063	24	3	0.667	0.624	-0.047
FS_C0024	0.042	24	5	0.375	0.416	0.119	24	5	0.417	0.391	-0.045	24	5	0.458	0.611	0.270
FS_C2361	0.018	24	4	0.625	0.631	0.031	24	4	0.708	0.635	-0.094	24	4	0.833	0.663	-0.237
FS_C1968	0.018	23	8	0.739	0.602	-0.206	20	6	0.550	0.623	0.142	24	7	0.458	0.506	0.115
FS_C5430	0.008	24	3	0.500	0.503	0.026	24	3	0.292	0.517	0.453	24	3	0.583	0.596	0.043
FS_C5453	0.040	23	3	0.565	0.519	-0.067	24	3	0.458	0.517	0.135	23	2	0.391	0.485	0.214
FS_C6563	0.029	24	4	0.667	0.650	-0.004	24	3	0.583	0.457	-0.255	24	3	0.542	0.609	0.131
FgSI0009	0.040	24	3	0.125	0.223	0.457	24	3	0.292	0.499	0.433	24	4	0.333	0.506	0.360
FgSI0025	0.029	23	6	0.609	0.646	0.079	24	6	0.542	0.683	0.227	23	6	0.739	0.609	-0.193
FS_C7377	0.024	23	3	0.261	0.373	0.321	24	2	0.292	0.305	0.064	24	3	0.500	0.508	0.037
FgSI0016	0.031	24	7	0.917	0.817	-0.101	24	8	0.875	0.771	-0.114	24	6	0.792	0.780	0.007
FS_C4971	0.008	24	5	0.583	0.543	-0.054	24	5	0.417	0.566	0.283	24	4	0.625	0.565	-0.085
FS_C6785	0.055	24	5	0.458	0.458	0.021	23	4	0.348	0.397	0.146	24	3	0.083	0.081	-0.011
Mean value across loci																
	0.025	23.7	4.400	0.506	0.512	0.033	23.6	4.133	0.485	0.514	0.077	23.8	3.933	0.514	0.523	0.038

Fixation index (F_{ST}), number of alleles per locus (N_a), observed heterozygosity (H_o), expected heterozygosity (H_e) and inbreeding coefficient (F_{IS}). F_{IS} values are not significant different from zero after Bonferroni correction

Private alleles occurred in all three populations. In population AEW5, private alleles with a relative frequency of 0.021 to 0.104 occurred at three loci (FS_C8040, FS_C6563, FS_6785), followed by the population HEW10 with private alleles at two loci (FgSI0006, FgSI0016) with a relative frequency of 0.021 and 0.063, and SEW5 with private alleles at one locus (FgSI0025) with a frequency of 0.022 (Table 3). The private allele 165 at locus FS_C6563 had the highest relative frequency (0.104) in the AEW5 population, followed by the private allele 151 at locus FgSI0016 in population HEW10 (0.063).

Table 3
Relative frequency of private alleles

Population	Loci	Allele	Relative frequency
AEW5	C8040	353	0.021
AEW5	C6563	165	0.104
AEW5	C6785	198	0.021
HEW10	FgSI0006	250	0.021
HEW10	FgSI0016	151	0.063
SEW5	FgSI0025	180	0.022

In addition, the 15 polymorphic markers in *F. sylvativa* were examined for their transferability to other *Fagaceae* species: *C. dentata, F. orientalis* and *Q. rubra* (Table 4). All markers were transferable to *F. orientalis* and 12 of them were polymorphic. Markers FS_C8040 and FgSI0025 were not transferable to *Q. rubra*, all other 13 markers also worked for *Q. rubra* and 11 of them were polymorphic. Only 11 markers were transferable to *C. dentata* (6 of them were polymorphic), while the other markers (FS_C1968, FgSI0009, FS_C8040, FgSI0025) did not amplify.

Table 4
Transferability of EST-SSRs to the three species *Castanea dentata, Fagus orientalis* and *Quercus rubra*

Marker	Species	N_a	Size (bp)	Amplification success
FS_C8040	*Castanea dentata*	-	-	-
	Fagus orientalis	4	350-367	8/8
	Quercus rubra	-	-	-
FS_C7797	*Castanea dentata*	3	244-250	5/8
	Fagus orientalis	4	243-253	8/8
	Quercus rubra	3	232-236	8/8
FgSI0006	*Castanea dentata*	1	250	7/8
	Fagus orientalis	1	261	7/8
	Quercus rubra	4	241-250	8/8
FS_C0024	*Castanea dentata*	2	264-271	7/8
	Fagus orientalis	4	164-178	8/8
	Quercus rubra	1	163	8/8
FS_C2361	*Castanea dentata*	3	194-200	7/8
	Fagus orientalis	4	194-202	7/8
	Quercus rubra	2	194-197	7/8
FS_C1968	*Castanea dentata*	-	-	-
	Fagus orientalis	4	296-321	7/8
	Quercus rubra	4	291-317	7/8
FS_C5430	*Castanea dentata*	1	167	6/8
	Fagus orientalis	1	170	8/8
	Quercus rubra	2	166-176	6/8
FS_C5453	*Castanea dentata*	1	229	1/8
	Fagus orientalis	1	229	8/8
	Quercus rubra	2	225-229	2/8
FS_C6563	*Castanea dentata*	1	170	6/8
	Fagus orientalis	5	155-168	8/8
	Quercus rubra	2	165-170	7/8
FgSI0009	*Castanea dentata*	-	-	-
	Fagus orientalis	4	205-217	8/8
	Quercus rubra	3	203-223	4/8
FgSI0025	*Castanea dentata*	-	-	-
	Fagus orientalis	4	178-200	6/8
	Quercus rubra	-	-	-
FS_CC7377	*Castanea dentata*	4	133-154	5/8
	Fagus orientalis	3	133-146	8/8
	Quercus rubra	5	131-156	7/8
FgSI0016	*Castanea dentata*	2	145-163	7/8
	Fagus orientalis	7	151-175	7/8
	Quercus rubra	1	163	5/8
FS_C4971	*Castanea dentata*	3	192-204	7/8
	Fagus orientalis	5	192-207	8/8
	Quercus rubra	3	211-220	7/8
FS_C6785	*Castanea dentata*	1	192	6/8
	Fagus orientalis	4	190-198	8/8
	Quercus rubra	4	184-200	7/8

Discussion

European beech has been well studied in terms of isoenzyme markers and chloroplast DNA markers (Demesure et al., 1996; Comps et al., 2001). In addition, nuclear microsatellite markers for East Asian beech species (e.g., *Fagus crenata* Blume and *Fagus japonica* Maxim.) were described in Tanaka et al. (1999) and also tested and partially adapted for *F. sylvatica* in Vornam et al. (2004). There are also nuclear microsatellites specifically developed for European beech (Pastorelli et al., 2003). Additionally, EST-SSRs were developed for *F. crenata* (Ueno et al., 2009), eight of them were adapted for use in *F. sylvatica* (Dounavi et al., 2016). However, compared to other important tree species in Europe the number of molecular marker for European beech is still limited. Here, we developed an additional set of 15 polymorphic gene-based EST-SSR-markers for *F. sylvatica*, which were transferred from *F. grandifolia*.

The EST-SSRs tested in this study showed lower variability and polymorphism than nSSR markers in previous studies. For instance, higher N_a, H_e and H_o values on average were revealed by Vornam et al. (2004) (*F. sylvatica*: N_a: 10.75, H_e: 0.765, H_o: 0.572), Tanaka et al. (1999) (*F. crenata*: N_a: 9.88, H_e: 0.615, H_o: 0.525; *F. japonica*: N_a: 8.56, H_e: 0.66, H_o: 0.604), and Liesebach (2012) (*F. sylvatica*: N_a: 10.58, H_e: 0.681, H_o: 0.6148) than the EST-SSR values of the three populations AEW5 (N_a: 4.4, H_e: 0.512 and H_o: 0.506), HEW10 (N_a: 4.13, H_e: 0.514, H_o: 0.485) and SEW5 (N_a: 3.933, H_e: 0.523, H_o: 0.514) in this study. In addition, Rajendra et al. (2014) also found across all beech populations from three geographic regions in Germany lower values in genetic diversity in EST-SSRs that were originally developed for *Quercus robur* L. (Durand et al., 2010) in comparison to the nuclear SSRs. These results show that less variability can be expected in EST-SSRs than at nSSRs likely as result of purifying selection (e.g. Ellis and Burke, 2007; Buonaccorsi et al., 2012; Harmon et al., 2017). The genetic diversity parameters of the new EST-SSR markers are comparable to other studies in *F. sylvatica* using EST-SSRs. Specifically, the values of H_o and H_e (here H_o: 0.502, H_e: 0.516) averaged over all populations and gene loci lie between the values of Rajendra et al. (2014) (EST SSRs: H_o: 0.442, H_e: 0.461) and Dounavi et al. (2016) (EST-SSRs: H_o=0.578, H_e: 0.604).

The unmanaged stand HEW10 showed similar genetic diversity indices as the two managed stands. This is in agreement with other studies that analyzed the same (Rajendra et al., 2014) or different populations (Buiteveld et al., 2007; Shanjani et al., 2010; Paffetti et al., 2012). The fact that none of the F_{IS} values was significantly different from zero (after Bonferroni correction) shows that there are no significant deviations from Hardy-Weinberg-structures. Despite all loci were not originally developed for *F. sylvatica*, they are not much affected by the presence of null-alleles (Supplementary Material 2). Low values of both probability of identity ($P_{(ID)}$) and probability of identity among siblings ($P_{(ID)sibs}$) showed the high discrimination power of the described markers.

One advantage of EST-SSRs as compared to nSSRs is that they are located in areas of the genome that are highly conserved among phylogenetically related species, and thus exhibit high transferability within closely related species (cf. Tanaka et al., 1999; Barreneche et al., 2004; Vornam et al., 2004; Durand et al., 2010). The markers tested in this study could be transferred to the three species *F. orientalis, Q. rubra* and *C. dentata*. All 15 markers were transferred successfully to *F. orientalis*, and 12 of them were polymorphic. Also in Vornam et al. (2004) (transfer of markers developed for *F. crenata* and *F. japonica* in *F. sylvatica*; 3/9 were polymorphic) and Tanaka et al. (1999) (transfer of EST SSR markers from *F. crenata* to *F. japonica*; 100 % transferability), transfers within the genus *Fagus* were successful.

In addition, good transferability between species of the same genus has also been found in other tree species. For example, in Brown et al. (2001) 52% of the microsatellites developed for *Pinus taeda* could be transferred to five other pine species on average. In addition, the microsatellites developed for *Quercus petraea* (Matt.) Liebl. were transferred to six different oak species (*Q. petraea, Q. robur, Quercus pubescens* Willd.: 100 %, *Quercus cerris* L.: 64 %, *Quercus palustris* Münchh., *Q. rubra*: 47 %) in the study by Steinkellner et al. (1997).

Transferability also worked between the genus *Fagus, Quercus* and *Castanea*, but it was observed that the more distantly the species are related, the more difficult it is to transfer the markers (*Quercus*: 13/15, *Castanea*: 11/15). Also, the number of polymorphic markers was higher in *Quercus* (11/13) than in *Castanea* (6/11). Some studies on the transferability of microsatellite markers to related species of different genera can also be found in the literature. For instance, Steinkellner et al. (1997) successfully transferred the microsatellite markers developed for *Q. petraea* to *F. sylvatica* (24 %) and *C. sativa* (47 %). Barreneche et al. (2004) also showed that microsatellites developed for *Q. rubra, Q. petraea*, and *Quercus macrocarpa* Micx. amplified in *Q. robur* (100 %) and *C. sativa* (70 %, 37/53) and markers developed for *C. dentata* amplified in *Q. robur* (83 %, 21/30) and

C. sativa (100 %). In addition, Durand et al. (2010) successfully transferred 63 EST-SSR markers from the 100 markers developed for oak to *C. sativa* (63%).

In summary, the developed markers can be used as efficient tools to characterize the variation in studies with European beech.

Acknowledgements

We thank Alexandra Dolynska for assistance in the lab work. We further thank the managers of the three Exploratories, Swen Renner, Kirsten Reichel-Jung, Kerstin Wiesner, Katrin Lorenzen, Martin Gorke, and all former managers for their work in maintaining the plot and project infrastructure; Simone Pfeiffer and Christiane Fischer for giving support through the central office, Michael Owonibi and Andreas Ostrowski for managing the central data base, and Markus Fischer, Eduard Linsenmair, Dominik Hessenmöller, Daniel Prati, Ingo Schöning, François Buscot, Ernst-Detlef Schulze, Wolfgang W. Weisser and the late Elisabeth Kalko for their role in setting up the Biodiversity Exploratories project.

The work has been partly funded by the DFG Priority Program 1374 „Infrastructure-Biodiversity-Exploratories" (Fi 569/12-2). Field work permits were issued by the responsible state environmental offices of Baden-Württemberg, Thüringen, and Brandenburg.

References

Aldrich PR, Jagtap M, Michler CH, Romeo-Severson J (2003) Amplification of North American red oak microsatellite markers in European white oaks and Chinese chestnut. Silvae Genetica 52(3-4): 176-179

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215(3): 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2

Barreneche T, Casasoli M, ussell K, Akkak A, Meddour H, Plomion C, Villani F, Kremer A (2004) Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). Theoretical and Applied Genetics 108: 558-566. https://doi.org/10.1007/s00122-003-1462-2

Brown GR, Kadel III EE, Bassoni DL, Kiehne KL, Temesgen B, van Buijtenen JP, Sewell MM, Marshall KA, Neale DB (2001) Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. Genetics 159: 799–809

Buiteveld J, Vendramin GG, Leonardi S, Kamer K, Geburek T (2007) Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. Forest Ecology and Management 247: 98–106. https://doi.org/10.1016/j.foreco.2007.04.018

Buonaccorsi VP, Kimbrell CA, Lynn EA, Hyde JR (2012) Comparative population genetic analysis of Bocaccio Rockfish *Sebastes paucispinis* using anonymous and gene-sssociated Simple Sequence Repeat loci. Journal of Heredity 103 (3): 391-399. https://doi.org/10.1093/jhered/ess002

Comps B, Gömöry D, Letouzey J, Thiébaut B, Petit RJ (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. Genetics 157: 389–397

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. Evolution 50: 2515–2520. https://doi.org/10.1111/j.1558-5646.1996.tb03638.x

Dounavi A, Netzer F, Celepirovic N, Ivankovic M, Burger J, Figueroa AG, Schön S, Simon J, Cremer E, Fussi B, Konnert M, Rennenberg H (2016) Genetic and physiological differences of European beech provenances (*F. sylvatica* L.) exposed to drought stress. Forest Ecology and Management 361: 226-236. https://doi.org/10.1016/j.foreco.2015.11.014

Durand J, Bodenes C, Chancerel E, Frigerio J-M, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn H-P, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herran A, Ikaran Z, Cabane C, Ueno S, Alberto F, Dumolin P-Y, Guichoux E, de Daruvar A, Kremer A, Plomion C (2010). A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11(570): 1-13. https://doi.org/10.1186/1471-2164-11-570

Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity 99(2): 125–32. https://doi.org/10.1038/sj.hdy.6801001

Fang J, Lechowicz M (2006) Climatic limits for the present distribution of beech (*Fagus L.*) species in the world. Journal of Biogegraphy 33: 1804– 1819. https://doi.org/10.1111/j.1365-2699.2006.01533.x

Gailing O, von Wühlisch G (2004) Nuclear markers (AFLPs) and chloroplast microsatellites differ between *Fagus sylvatica* and *F. orientalis*. Silvae Genetica 53(3): 105-110. https://doi.org/10.1515/sg-2004-0019

Gailing O, Langenfeld-Heysert R, Polle A, Finkeldey R (2008) Quantitative trait loci affecting stomatal density and growth in a *Quercus robur* progeny: implications for the adaption to changing environments. Global Change Biology 14: 1934- 1946. https://doi.org/10.1111/j.1365-2486.2008.01621.x

Goudet J (2002) Software Fstat 2.9.3. Department of Ecology & Evolution, Biology Building, Lausanne University [online]. Available at <https://www2.unil.ch/popgen/softwares/fstat.htm> [cited 04/15/2018]

Guichoux E, Lagache L, Wagner S, Chaumeil P, Leger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. Molecular Ecology Resources 11: 591-611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Harmon M, Lane T, Staton M, Coggeshall MV, Best T, Chen C-C, Liang H, Zembower N, Drautz-Moses DI, Hwee YZ, Schuster SC, Schlarbaum SE, Carlson JE, Gailing O (2017) Development of novel genic microsatellite markers from transcriptome sequencing in sugar maple (*Acer saccharum* Marsh.). BMC Research Notes 10(369): 1-7. https://doi.org/10.1186/s13104-017-2653-2

Kölling C, Zimmermann L (2007) Die Anfälligkeit der Wälder Deutschlands gegenüber dem Klimawandel. Klimaschutz 67(6): 259-268

Kubisiak T, Carey D, Burdine C, Koch J (2009) Characterization of ten EST-based polymorphic SSR loci isolated from American beech, *Fagus grandifolia* Ehrh. Permanent Genetic Resources Note added to Molecular Ecology Resources Volume: 1-8. https://doi.org/10.1111/j.1755-0998.2009.02759.x

Lalagüe H, Csilléry K, Oddou-Muratorio S, Safrana J, de Quattro C, Fady B, González-Martínez SC, Vendramin GG (2014) Nucleotide diversity and linkage disequilibrium at 58 stress response and phenology candidate genes in a European beech (*Fagus sylvatica* L.) population from southeastern France. Tree Genetics & Genomes 10: 15–26. https://doi.org/10.1007/s11295-013-0658-0

Lesur I, Bechade A, Lalanne C, Klopp C, Noirot C, Leplé J-C, Kremer A, Plomion C, le Provost G (2015) A unigene set for European beech (*Fagus sylvatica* L.) and its use to decipher the molecular mechanisms involved in dormancy regulation. Molecular Ecology Resources 15: 1-13. https://doi.org/10.1111/1755-0998.12373

Liesebach H (2012) Genotypisierung mit nuklearen Mikrosatellitenmarkern – Möglichkeiten der Daten-auswertung am Beispiel von Buchenpopulationen (*Fagus sylvatica* L.) aus einem Herkunftsversuch. Applied Agricultural and Forestry Research 4:221-236.

Mishra B, Gupta DK, Pfenninger M, Hickler T, Langer E, Nam B, Paule J, Sharma R, Ulaszewski B, Warmbier J, Burczyk J, Thines M (2018) A reference genome of the European beech (*Fagus sylvatica* L.). GigaScience 7: 1-8. https://doi.org/10.1093/gigascience/giy063

Müller M, Seifert S, Finkeldey R (2015) Identification of SNPs in candidate genes potentially involved in bud burst in European beech (*Fagus sylvatica* L.) Silvae Genetica 64: 1–20. https://doi.org/10.1515/sg-2015-0001

Müller M, Seifert S, Lübbe T, Leuschner C, Finkeldey R (2017) De novo transcriptome assembly and analysis of differential gene expression in response to drought in European beech. PLoS one 12(9): 1-20. https://doi.org/10.1371/journal.pone.0184167

Paffetti D, Travaglini D, Buonamici A, Nocentini S, Vendramin GG, Giannini R, Vettori C (2012) The influence of forest management on beech (*Fagus sylvatica*

L.) stand structure and genetic diversity. Forest Ecology and Management 284: 34–44. https://doi.org/10.1016/j.foreco.2012.07.026

Pastorelli R, Smulders MJM, van't West-Ende WPC, Vosman B, Giannini R, Vettori C, Vendramin GG (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. Molecular Ecology Notes 3: 76–78. https://doi.org/10.1046/j.1471-8286.2003.00355.x

Paule L (1995) Gene conservation in European beech (*Fagus sylvatica* L.). Forest Genetics 2(3): 161–170

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6(1): 288-295. https://doi.org/10.1111/j.1471-8286.2005.01155.x

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19): 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

Rajendra KC, Seifert S, Prinz K, Gailing O, Finkeldey R (2014) Subtle human impacts on neutral genetic diversity and spatial patterns of genetic variation in European beech (*Fagus sylvatica*). Forest Ecology and Management 319: 138-149. https://doi.org/10.1016/j.foreco.2014.02.003

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecology Resources 8: 103-106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology 18(2): 233-234. https://doi.org/10.1038/72708

Shanjani PS, Vendramin GG, Calagari M (2010) Genetic diversity and differentiation of *Fagus orientalis* Lipsky in Hyrcanian forests revealed by nuclear and chloroplast microsatellite markers. Conservation Genetics 11(6): 2321–2331. https://doi.org/10.1007/s10592-010-0118-4

Steinkellner H, Lexer C, Turetschek E, Glössl J (1997) Conservation of (GA)n microsatellite loci between *Quercus species*. Molecular Ecology 6: 1189–1194. https://doi.org/10.1046/j.1365-294X.1997.00288.x

Tanaka K, Tsumura Y, Nakamura T (1999) Development and polymorphism of microsatellite markers for *Fagus crenata* and the closely related species, *F. japonica*. Theoretical and Applied Genetics 99: 11-15

The Uniprot Consortium (2017) UniProt: the universal protein knowledgebase. Nucleid Acids Research 45: 158- 169. https://doi.org/10.1093/nar/gkw1099

Ueno S, Taguchi Y, Tomaru N, Tsumura Y (2009) Development of EST-SSR markers from an inner bark cDNA library of *Fagus crenata* (Fagaceae). Conservation Genetics 10: 1477-1485. https://doi.org/10.1007/s10592-008-9764-1

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. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vornam B, Decarli N, Gailing O (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. Conversations Genetics 5: 561-570. https://doi.org/10.1023/b:coge.0000041025. 82917.ac