REIS, A., A. BRAGA, M. LEMES, R. GRIBEL and R. COLL-EVATTI (2009): Development and characterization of microsatellite markers for the Brazil nut tree *Bertholletia excelsa* Humb. & Bonpl.(Lecythidaceae). Molecular Ecology Resources **9**: 920–923.

RITLAND, K. (1989): Correlated matings in the partial selfer *Mimulus guttatus*. Evolution 43: 848–859.

RITLAND, K. (2002): Extensions of models for the estimation of mating systems using n independent loci. Heredity **88**: 221–228.

SAMPSON, J. F. (1998): Multiple paternity in *Eucalyptus rameliana* (Myrtaceae). Heredity **81**: 349–355.

SANTOS, C. F. and M. L. AISY (2012): Interactions between carpenter bees and orchid bees (Hymenoptera: Apidae) in flowers of *Bertholletia excelsa* Bonpl. (Lecythidaceae). Acta Amazonica **42**: 84–94.

SEBBENN, A. M. (2006): Sistema de reprodução em espécies arbóreas tropicais e suas implicações para a seleção de árvores matrizes para reflorestamentos ambientais. *In:* HIGA, A. R. and L. D. SILVA. Pomares de sementes de espécies florestais nativas. Curitiba: FUPEF, p. 93–138.

SILVA, C. R. S., P. S. B. ALBUQUERQUE, F. R. ERVE-DOSA, J. W. S. MOTA, A. FIGUEIRA and A. M. SEBBENN (2011): Understanding the genetic diversity, spatial genetic structure and mating system at the hierarchical levels of fruits and individuals of a continuous *Theobroma cacao* population from the Brazilian Amazon. Heredity **106**: 973–985.

SUJII, P. S., P. W. INGLIS, A.Y. CIAMPI, V. N. SOLFERINI and V. C. R. AZEVEDO (2013): Isolation and characterization of microsatellite markers for *Bertholletia excelsa* (Lecythidaceae) population genetic analysis. Genet. Mol. Res. **12**: 5278–5282.

SUJII, P. S., K. MARTINS, L. H. O. WADT, V. C. R. AZEVEDO and V. N. SOLFERINI (2015): Genetic structure of *Bertholletia excelsa* populations from the Amazon at different spatial scales. Conserv. Genet. DOI 10.1007/s10592-015-0714-4.

SURLES, S. E., J. ARNOLD, A. SCHABEL, J. L. HAMRICK and B. C. BONGARTEN (1990): Genetic relatedness in open pollinated families of two leguminous tree species, *Robinia pseudoacacia* L. and *Gleditsia triacanthos* L. Theor. Appl. Genet. **80**: 49–56.

TAMBARUSSI, E. V., D. BOSHIER, R. VENCOVSKY, M. L. M. FREITAS and A. M. SEBBENN (2015): Paternity analysis reveals significant isolation and near neighbour pollen dispersal in small *Cariniana legalis* Mart. Kuntze populations in the Brazilian Atlantic Forest. Ecology and Evolution. doi: 10.1002/ece3.1816

WADT, L. H. O., K. A. KAINER and D. A. P. GOMES-SILVA (2005): Population structure and nut yield of a *Bertholletia excelsa* stand in Southwestern Amazonia. Forest Ecology and Management **211**: 371–384.

WADT, L. H. O., K. A. KAINER, C. STAUDHAMMER and R. SERRENO (2008): Sustainable forest use in Brazilian extractive reserves: Natural regeneration of Brazil nut in exploited populations. Biological Conservation **141**: 332–346.

WARD, M., C. W. DICK, R. GRIBEL and A. J. LOWE (2005): To self, or not to self... A review of outcrossing and pollen-mediated gene flow in Neotropical trees. Heredity **95**: 246–254.

PCR-based detection of single sequence variants from a natural collection of the non-model tree species European Aspen *Populus tremula* (L.)

By S. WINKLER[1], K. LINKE[2], N. GSCHEIDEL[1], M. MEYER[2] and D. KRABEL[2], [*]

Abstract

In the present study we present and discuss the identification of species-specific SNPs to rule out any experimental influence of species-specific primer design (*Populus tremula* vs. the closely related model-species *Populus trichocarpa*) on the detectability of SNPs. Applying a species-optimized method, partial sequences of 14 genes involved in xylem cell development, xylogenesis, pectin formation, and drought stress reaction were analyzed at the genomic level. About 3 Mb of sequence information were generated by Sanger sequencing technology and 258 sequence variants were

[1] MPI of Molecular Cell Biology and Genetics, DNA Sequencing Facility, Pfotenhauerstr. 108, D-01307 Dresden, Germany.

[2] Molecular Physiology of Woody Plants Group, Dresden University of Technology, Pienner Str. 7, D-01737 Tharandt, Germany.

[*] Corresponding author: DORIS KRABEL. E-Mail: krabel@forst.tu-dresden.de

identified. 15 out of these represent insertions/deletions located exclusively in non-coding regions and the remaining 243 are SNPs found in coding and non-coding regions of candidate genes.

The introduction of a species-specific SNP detection pipeline will help to detect nucleotide variants in *P. tremula* and to conduct association mapping in natural *P. tremula* populations.

Key words: drought stress tolerance, natural population, *Populus tremula*, SNP, xylem cell formation

Introduction

Within the last decade the genus *Populus* has become one of the main research objects of molecular genetics on woody plants. Poplar species are easily to propagate compared to other tree or shrub species, which facilitates the cultivation of plants. In addition, the genome is relatively small (2n = 38, ca. 550 Mbp). Both are favorable characteristics for a plant species to become a model organism (STRAUSS and MARTIN, 2004).

European Aspen (*Populus tremula*, L.) is a deciduous tree species widely distributed in the Northern hemisphere (ECKENWALDER, 1996). Its geographic distribution ranges from cool and temperate regions in Northern Europe and North Eastern Asia to warm and dry sites in Southern Europe and North Africa (ECKENWALDER, 1996). *P. tremula* belongs to the family of *Salicacea* and is closely related to poplar species that are mostly used for short rotation forestry. Though, due to its comparably lower biomass production *P. tremula* does not count amongst the typical plantation tree species in Europe. Nevertheless like other widely distributed tree species, adaptive genetic variation, which might reflect the capability of a subpopulation to adapt to different environmental conditions (especially to drought) in combination with an acceptable biomass production has to be expected for European aspen which makes this tree species an interesting object of investigations.

One promising method for identification of drought-adapted genotypes is an association mapping approach based on the characterization of SNPs (single nucleotide polymorphisms) being located in candidate genes. SNP-loci have been identified for a number of plant species including forest trees (e.g. *Populus tremula*:

INGVARSSON, 2005; SEEB *et al.,* 2011; SEIFERT *et al.,* 2012). Although most SNPs do behave selectively neutral and can be used in population genetic analysis in addition to or instead of other neutral markers (SEIFERT *et al.,* 2012), a comparably small number of SNPs are likely to be under selection and can be associated with phenotypic traits (YOSHIURA *et al.,* 2006; SEIFERT *et al.,* 2012).

In the present study, we introduce a reliable and robust method for the detection of single nucleotide polymorphisms within candidate genes related to traits that are mainly involved in drought stress tolerance in *Populus tremula*. In contrast to Ingvarsson (2008), who investigated a total of 77 amplicons distributed randomly over the genome of European aspen, we focused on selected candidate genes. For individuals selected from different populations and environments this approach seemed to be a promising method to detect SNP variation fast and efficiently (CARDON and BELL, 2001; INGVARSSON, 2005) as a basis for a future association mapping study and marker-assisted selection.

Materials and methods

Sample collection

To cover a wide spectrum of environmental conditions such as water availability and soil conditions, 108 *P. tremula* genotypes were collected from eleven natural sites in Germany (9 to 11 trees per subpopulation). Among these eleven sites, ten sites were selected along a North to South gradient beside and 50 km far from the river Oder. Additionally one subpopulation was sampled in the rocky area of the Black Forest region (*Figure 1*). Only trees separated by a distance of 1–2 km were collected to ensure the sampling of genetically distinct trees.

DNA extraction

Total genomic DNA was extracted from frozen or dried leaves according to DOYLE and DOYLE 1987. Several DNA extractions from cambium were done using the DNeasy plant mini-prep kit (QIAGEN, Hilden, Germany). The amount and the quality of the genomic DNA were analyzed in 0.8% agarose gels. For subsequent PCR applications genomic DNA was diluted 1 in 100 in water.
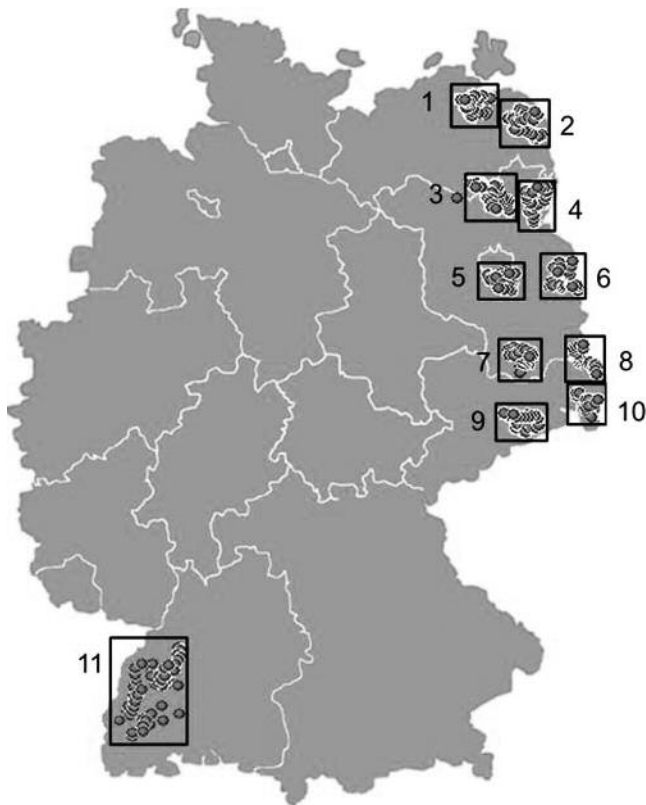
Figure 1. – Collection sites of individual trees of the *P. tremula* natural collection in Germany. GPS data: Greifswald: 54°6'N, 13°23'E; Usedom: 53°52'N, 13°55'E; Templin: 53°7'N, 13°30'E; Schwedt: 53°4'N, 14°17'E; Berlin: 52°31'N, 13°24'E; Frankfurt/Oder: 52°21'N, 14°33'E; Finsterwalde: 51°38'N, 13°43'E; Bad Muskau: 51°33'N, 14°43'E; Dippoldiswalde: 50°54'N, 13°40'E; Zittau: 50°51'N, 14°48'E and Schwarzwald: 48°28'N, 8°10'E.

Candidate gene selection

Candidate genes were selected for their physiological function in the context of drought adaptive behavior. *Table 1* summarizes the genes, which were chosen as candidates. First we focused on genes that are reported to be directly up-regulated upon water deficiency. These were *Drought-sensitive* 1 (*drs1*), *stable protein* (*Sp1*) and *Metallothionein 2A* (*Mt2a*). *SLAC1* and *erecta* homologues are involved in regulation and movement of stomata in leaves as a response to humidity changes, whereas the aquaporin *PIP2.5*-gene is directly involved in water transport. *Alpha-expansin* (α-*exp*) regulates cell-wall extension and stress relaxation. Secondly, other candidate genes were chosen due to their function in xylogenesis, xylem cell development and pectin formation, among these are the K⁺ channel protein *kpt1*, the high-affinity potassium uptake transporter *kup1*, the potassium channel *ptk2*, an outward rectifying potassium channel *ptork*, a MADS box protein *ptm5*, *UDP-glucose dehydrogenase* (*UGDH*), and *UDP-glucose pyrophosphorylase* (*UGP*). In the course of this study, we identified *P. tremula* homologues of *Drs1*, *erecta*, *mt2a*, and *SLAC1* (Accession numbers: *drs1*, KC768063; *erecta*, KC768064; *mt2a*, KC768065; *SLAC1*, KC768066).

P. tremula-specific PCR design

Since *P. tremula*-specific whole genome information is not published yet, protein sequences of candidate genes as summarized in *Table 1* were obtained from Genbank and Refseq, respectively. Protein sequences were blasted against the *P. trichocarpa* genome (TUSKAN *et al.,* 2006), making use of the tools of the Phytozome project (GOODSTEIN *et al.,* 2012: JGI v2.2, *Populus* genome assembly: http://www.phytozome.net/poplar.php). Identified candidate loci (blast hits) were manually examined for a) any duplicated sequence, b) for potential paralogs within the *P. trichocarpa* genome, and c) for the correct annotation of the candidate gene. If more than one candidate locus was identified for one gene, protein and nucleotide sequence comparisons were performed for these hits using CLUSTALW and the closest homologue was chosen for amplicon design.

P. trichocarpa-specific PCR primers were designed using the primer3-tool (ROZEN and SKALETSKY, 2000) applying these criteria: (1) PCR fragments are designed that allow sequencing over internally located primer binding sites that subsequently represent a reference for *P. tremula*-specific primer pairs, (2) primers which are located adjacent to exons to screen as much coding region as possible, (3) if gene annotation allows, primer pairs are designed which are located in coding sequences but not in functional domains to avoid unspecific binding based on high sequence conservation between functional domains. An overview of all primers (trichocarpa-specific, tremula-specific and tremula specific sequencing primers) is summarized in *Table 2*.

PCR and sequencing of PCR fragments has been performed as described in WINKLER *et al.* (2011). *P. trichocarpa*-derived primer pairs were tested on genomic DNA of one *P. trichocarpa* and eleven *P. tremula* trees. PCRs were done in a total volume of 10 µl containing 2 µM primer, 2 mM MgCl₂, 0,2 µM dNTP each, 25 mM

Table 1. – List of candidate genes.

Gene	Accession Number *P. tremula*	Genomic position (*P. trichocarpa* v2.2)	Screened exon	Physiological function	Reference
Drs1	KC768063	LG 18: 12,658,160 - 12,659,729	exon 3	Response to water deprivation, ubiquitin ligase complex	Lee et al., 2008
			exon 4		
Erecta	KC768064	LG 6: 22,390,775 - 22,393,386	exon 23-24	Transpiration, transmembrane receptor protein kinase activity	Masle et al., 2005; Hamanishi et al., 2012
			exon 25-26		
α-exp	AY435099	LG 1: 24,014,878 – 24,017,296 LG 9: 4,301,229 - 4,303,484	not done[1]	Cell-wall extension and stress relaxation	Gray-Mitsumune et al., 2004
Kpt1	AJ344623	LG 4: 14,291.677 – 14.292,374	exon 7-10	Inward potassium channel, potassium dependent xylogenesis	Langer et al., 2002
Kup1	AJ299422	LG 3: 13,771,063 – 13,771,674	exon 2-3	Potassium transporter/uptake, potassium dependent xylogenesis	Langer et al., 2002
		LG 3: 13,767,144 – 13,767,874	exon 9		
Mt2a	KC768065	LG 6: 6,215,377 - 6,215,954	exon 2	Metallothionein, abiotic stress response	Street et al., 2006; Wilkins et al., 2009, Yeh et al., 1995
PIP2.5	AJ849328	LG 16: 6,990,700 – 6,991,987 LG 9: 2,408,396 – 2,409,868	not done[2]	Aquaporin, water conduction	Marjanovic et al., 2005
Ptk2	AJ271447	LG 3: 1,122,870 – 1,123,791	not done (7-8)[3]	Voltage gated potassium channel, potassium dependent xylogenesis	Langer et al., 2002; Arend et al., 2005
		LG 3: 1,122,122 – 1,122,548	exon 9-10		
Ptm5	AF377868	LG 14: 5,289,861 - 5,290,477	exon 1	MADS-box protein, transcription factor activity, wood vascular development	Cseke et al., 2003
		LG 14: 5,283,399 - 5,284,908	exon 3-4		
Ptork	AJ271446	LG 12: 3,217,585 - 3,218,200	exon 3-4	Outward rectifying potassium channel, potassium dependent xylogenesis	Langer et al., 2002
		LG 12: 3,214,335 - 3,214,910	exon 8-9		
SLAC1	KC768066	LG 15: 6,085,516 - 6,085,885	exon 2-3	C4-dicarboxylate transporter/malic acid transport protein, regulation of stomatal movements (humidity response)	Vahisalu et al., 2008
Sp1-1	AJ276517	LG 10: 15,141,391 – 15,142,213	exon 1-2	Heat stable protein, response to stress stimuli (temperature, salt), water transport	Wang et al., 2002
Sp1-2	AJ276517	LG 10: 14,928,636 – 14,929,429	not done[4]		
UDGH	AF053973	LG 4: 10,620,227 – 10,621,029[5] LG 17: 12,927,539 – 12,928,338	exon 1	UDP-glucose 6-dehydrogenase activity, formation of pectin	Johansson et al., 2002
UGP	AY260746	LG4: 6,008,392 – 6,013,141 LG17: 967,805 – 972,392	not done	UDP-glucose pyrophosphorylase, formation of pectin	Johansson et al., 2002

[1] Duplicated region.
[2] Gene family of aquaporins that consists of more than two very similar versions.
[3] PCR failed.
[4] Partially duplicated sp1 locus.
[5] Specific primers for UDGH homolog are located on LG4.

tricine, 85 mM ammoniumacetate, 7% glycerol, 1,6% DMSO, 0,5 U Taq-polymerase, 5–50 ng genomic DNA. Cycling conditions for the PCRs were as follows: 94°C for 2 min, 30 cycles of 94°C for 30 sec, 57°C (or primer optimized annealing temperature) for 20 sec, 72°C for 1 min, and a final extension at 72°C for 10 min. The resulting PCR fragments were visualized and quantified on 1% agarose gels, diluted to a final concentration of 10–20 ng/µl and directly used for DNA Sequencing (Sanger big dye terminator technology). Sanger sequencing was performed according to the manufacturers' protocol making use of low input conditions. The sequencing reaction contains 2–5 ng template DNA, 0.2 µl big dye terminator, 0,4 µl 5x sequencing buffer, and 2 µM sequencing primer in a total volume of 8 µl. Cycling was performed as follows: 96°C for 1 min, 30 cycles of 96°C for 10 sec, 52°C for 5 sec, 60°C for 3 min. Sanger sequencing was performed on an AB3730XL DNA sequencer.

If PCR results were either unspecific or failed in *P. tremula* samples, PCR conditions were optimized for the annealing temperature. If heterozygous indels were found that were not resolved by any of the sequencing primers, internal *P. tremula*-specific sequencing primers were designed to cover as much information as possible.

DNA sequences of all tested individuals were compared and analyzed for differences between both species (cross-species variances) as well as

Table 2. – P. tremula and *P. trichocarpa*-specific primer sequences and location of the primer in annotated locus.

gene, exon	primer pairs *P. trichocarpa*	location	primer pairs *P. tremula*	location	Internal sequencing primer *P. tremula*	location
Drs1, 3	5' AGAGAAGAGGGAATTTGAGGG 5' AGGATTAAACGAGCACTTGTC	exon intron	5' TGTTCGC**G**TTTCTGT**G**GAG**T** 5' AGAACACTGCGGTATGATAAAG	intron intron		
Drs1, 4	5' GGCCACTAGTTCCTCAGATGC 5' ATCATGTCCAGGAAGATTGG	exon intron	5' GCCCACTTTAATTT**G**GAGAG 5' TGATGC**C**CT**T**CAATGGAACC	intron intron		
Erecta, 23+24	5' CAGAGCGAGGTATTCAGATG 5' CTCGATACTGCTCTGCTTTG	exon 3'UTR	5' CGGTAATACCT**T**TTCTGC**A**G 5' CGAGGACTACAATCATGGTG	intron exon	5' GTAGGTGCATGCAAG GAGAG 5' TAATTTTCAACT**C**CAAGACTATC	intron intron
Erecta, 25+26	5' CAAGGATATTCCCTGTCACC 5' GCTTCTCGTGGTCCTACAG	exon 3'UTR	5' TCTCCTTGCATGCACCTAC 5' CTCGATACTGCTCTGCTTTG	intron intron	5' CATTGTCAACTGCTTTCCTT 5' TCTGACACGGATAAGAAACC 5' GGTTTCTTATCCGTGTCAGA	exon intron intron
Kpt1, 7-10	5' ATCAATGTTTGCCCAGTTTC 5' GTTTCGCTATCCTCAACTGC	intron intron	5' GCAGAACGAAGCTCCTACAG 5' CTTTCTTGTAGCCTTTGATCC	exon exon		
Kup1, 2+3	5' TCCAATAAATCATGGTACGC 5' GGTAATAAACACAGCAGCAC	5'UTR exon	5' TGTCTTGATCGTTGCTCTTG 5' CAAACTTCTTGAACCGAGATG	intron exon		
Kup1, 9	5' GAAAGCTGTTCTTACCTTTGC 5' CTGCTCTCCTTATTGCCATC	intron exon	5' AAAGCTGATGTTGCTACTACC 5' TCAGATGACCTCACAACTGC	intron exon		
Mt2a, 1+2[1]	5' CACTTCGCGCAGTTTCC 5' AGAGCTGCAAATCCAAGAAG	5'UTR 3'UTR	5' **T**TGCAGGGATAAGCAT**G**ACG 5' ACAAGAGACAA**GAGACAACC**	intron 3'UTR[2]		
Ptk2, 7+8	5' TGTGATCTTATGTGCCTAACG 5' AATTCCCATGCATGTTCTG	intron intron	not done			
Ptk2, 9+10	5' GATTCCCACCTACTGATTCG 5' CATGAGGTGCATTCTTGTTC	intron exon	5' AGGACATGAAGAGTGTGTGG 5' GC**A**CCGCTCATTACTAGC	exon exon		
Ptm5, 1	5' AAGCTCTCGTGGGTTTATTG 5' ATTATTGGAACGGGCAAAG	5'UTR intron	5' TCCCTCCACTTTCTTACCTG 5' GAGAGAGAGAGGCACACAC	intron intron	5' ATCAAAGGAGCCAGACACG	intron
Ptm5, 3+4	5' TGTTGTAGATGCTTGCGAAC 5' ACATGCAACTGTTGACGAC	intron intron	5' ATAGGATAAGGGTGGAAGG 5' TTCTCTACAGCCTTTGGTTTC	exon intron		
Ptork, 3+4	5' CATTGCCATTAGGTCTGTACC 5' CGTGGATTTGTCTAGCAATG	exon intron	5' GAAATCTCGGTTCTTGGTTG 5' TGCCACTCATCT**G**GACTTC	exon intron		
Ptork, 8+9	5' CCTGTCTCTGAGTGTGTTGG 5' TAATGCAACGGTTCAAGTTC	intron intron	5' GCAAACCTATAGCTCATTTGG 5' AGCCTTGTTAGGATCTGCTC	exon exon		
SLAC, 2+3	5' ATGTCAGGAGGACAGAGAAGG 5' GAGGTCCTGGAGTACGAAT	exon exon	5' TATATTTCTGTCATG**G**ACATTC 5' GTACGAATGCGTGCACT	intron exon		
Sp1, 1+2	5' GGACATTAGGCGTGGAAAG 5' TATCGATGCTGGTTCTTTCC	5'UTR 3'UTR	5' GGCAGAGCATCATGCAG 5' AGCTCATTCGAGGATAATGG	5'UTR 3'UTR		
Sp1, short	5' GGCGGGTCTACCACCAC 5' CACAAACAAGATCAAGACACG	intron intron	not done			
UDGH, 1	5' ATCTACAACGGCAGGTTAGC 5' CCAACACTAGCATTCAAGAAC	intron exon	5' TGAATTTCCCTTCTTTCAGG 5' TCCTTTCCAACAGCATAGGC	intron exon		

[1] A nested PCR approach was used to amplify exon 2 of Mt2a: inner primers were used to amplify from the PCR fragment generated by the outer primer set.

[2] Untranslated regions: considered are 500 bp upstream or downstream of coding sequence since annotation was mostly missing. **Bold:** Nucleotides that differ in the primer binding sites of *P. tremula* and *P. trichocarpa*, shown is *P. tremula* sequence.

between *P. tremula* individuals (within-species variances). *P. trichocarpa*-derived primers served as templates for the design of *P. tremula*-specific primer pairs. These primer pairs were tested again for specificity and performance on the *P. tremula* trees. An overview of all designed primer pairs including sequencing primers is presented in *Table 2*.

Detection of sequence variation and statistical analysis of SNPs in the P. tremula natural populations

PCR fragments were amplified and sequenced from the *P. tremula* collection making use of *P. tremula*-specific PCR and sequencing primers. SNPs in these regions were detected automatically by the Polyphred software package that uses Phred-based base calling and quality assignments with subsequent detection of heterozygous peaks in sequence assemblies

performed by Phred, conditions are as follows: the peak drop ratio is 0.55, a second peak ratio of 0.15 and an average sequence quality of 30 (NICKERSON *et al.,* 1997; BHANGALE *et al.,* 2006). Small and resolvable indels of 1 to 100 bp in size were detected on sequence alignments (CLUSTALW) making use of the MacVector sequence analysis tool (MacVector, Inc., North Carolina, US). Sequence variation was documented for each individual of the *P. tremula* collection. For each SNP in the total population the genetic diversity was estimated by calculating the observed and expected heterozygosities and p-values. All calculations were performed with PowerMarker version 3.25. Tests for Hardy-Weinberg equilibrium (HWE) were performed for each SNP using the exact test implemented in PowerMarker version 3.25 (LIU and MUSE, 2005). Application of Bonferroni correction was done in a conservative manner (corrected α-level: 0.05/243 = 0.0002).

*Determination of an allele bias caused
by cross-species primer and PCR conditions*

The effect of cross-species PCR and sequencing conditions (*P. trichocarpa*-derived primers on *P. tremula* individuals) on SNP detection was compared to within-species conditions (*P. tremula*-derived primers on *P. tremula* individuals). For this experiment, four sets of *P. trichocarpa*- (representing cross-species and tremula-specific primer pairs representing three different genes (*drs1* exon 3, *drs1* exon 4, *erecta* exon 23+24, *mt2a* exon 2) were amplified and sequenced from the *P. tremula* collection. Three independent rounds of PCR were performed with *P. tremula*-specific primers and two independent rounds with *P. trichocarpa*-specific primers. Differences within one data set (PCR- and sequence-derived errors, inner-species dataset) and between *P. trichocarpa*- and *P. tremula*-specific data were determined. The observed and expected heterozygosity values were calculated for each locus (data not shown). The inbreeding coefficient (F_{IS}) and the number of loci that deviate significantly from HWE expectations after conservative Bonferroni correction (corrected α-level: 0.05/81 = 0.0006) are given for each amplicon. All calculations were done with the software PowerMarker version 3.25 (*Table 3*)

Table 3. – Heterozygosity of the wild population, inbreeding coefficient and deviation from HWE for the tested amplicons.

Gene	exon	Total No. of SNPs	Heterozygosity		F_{IS}	Out of HWE[1]
			exp	Obs		
Drs1	3	14	0,204	0,180	0,125	0
	4	19	0,203	0,134	0,443	4
Erecta	23+24	29	0,246	0,183	0,232	1
	25+26	19	0,224	0,221	0,006	0
Kpt1	7-10	17	0,275	0,260	0,024	0
Kup1	2+3	4	0,097	0,092	0,025	0
	9	5	0,332	0,321	0,012	0
Mt2a	2	21	0,239	0,225	0,017	0
Ptk2	9+10	16	0,284	0,237	0,159	3
Ptm5	1	18	0,143	0,111	0,132	0
	3+4	17	0,182	0,156	0,189	0
Ptork	3+4	12	0,318	0,245	0,282	3
	8+9	10	0,117	0,113	0,015	0
SLAC1	2+3	12	0,042	0,043	- 0,022	0
Sp1	1+2	14	0,137	0,133	0,026	0
UDGH	1	16	0,153	0,141	0,053	0

[1] Total number of SNPs that deviate from HWE after multiple test correction.

Results

PCR design

We identified 3 candidate genes as not being suitable for a specific PCR-based screen. The candidate genes *Alpha-expansin*, *UDGP*, and *UDGH* are found in two almost identical copies in the *P. trichocarpa* genome (*Alpha-expansin*: 92,5% identity of coding sequence, LG 1 and 9; *UDGP*: 93,7% identity of coding sequence, LG 4 and 17, *UDGH*: 92,9% identity of coding sequence, LG 4 and 17). *Alpha-expansin* and UDGP do not allow any copy-specific primer design and therefore failed our PCR criteria. However, for UDGH we managed to specifically target the gene copy that is located on LG 4. Another candidate gene *PIP2.5* belongs to the gene family of aquaporins of which several paralogs are found in the *P. trichocarpa* genome. Since it was not possible to resolve the gene orthology for *PIP2.5*, this gene was not used further in this study. The stable protein *Sp1* is found in two copies both located in close proximity on LG 10. Since the *Sp1* locus is duplicated only partially, *P. trichocarpa*- and *P. tremula*-specific primers for both gene products were designed.

P. tremula-specific PCR optimization

We designed 18 *P. trichocarpa*-derived primer pairs covering 11 candidate genes and tested them on genomic DNA derived from a batch of 10 *P. tremula* trees and one *P. trichocarpa* tree. 16 primer pairs resulted in successful amplification of *tremula*-specific PCR fragments that were eventually sequenced and analyzed. Two primer sets (*ptk2*, exon 7 and 8, *sp1*, short fragment) failed in *P. tremula* indicating significant sequence differences between both species. Sequence information derived from each species was compared and in case of differences in the sequence, primers specific for *P. tremula* were designed. These *P. tremula* specific primers flanked the coding region (exons) of our candidate gene and expanded circa 50 to 200 bp into the adjacent sequence of the intron. We established PCR and sequencing conditions for 16 *P. tremula*-specific PCR primer pairs and 6 additional sequencing primers (*Table 2*). Among these 38 primers, 10 showed sequence differences between both species exhibiting 1 to 3 nucleotide mismatches or even a deletion of 6 nucleotides. Most of these primers (9 of 10) were located in non-coding sequence (introns,

UTRs, and non-genic regions) of the gene locus and flanked the informative exons.

Species-specific PCR design is required to reduce any allele bias

Due to our screening focus on informative coding sequence, 63% of our *P. tremula*-specific primer binding sites are located in non-coding regions such as introns and UTRs that are expected to be less conserved than the coding sequence. Indeed 9 of 24 non-coding primer binding sites actually showed differences in the nucleotide sequence when comparing *P. tremula* and *P. trichocarpa*, whereas only 1 of 14 primer binding sites differed in coding sequence (*Table 2*). To test if even minor nucleotide variation in the primer binding sites could result in an allele-specific amplification bias, we detected SNPs for four amplicons making use of either *P. tremula* or *P. trichocarpa* derived primer pairs. These primer binding sites differed either in one, two, or three nucleotides (*drs1*, exon 3; *drs1*, exon 4; *erecta*, exon 23 and 24) or they span a species-specific indel in the 3' region of the reverse primer (*mt2a*). *P. trichocarpa*- or *P. tremula*-specific primer pairs were applied to the *P. tremula* collection. All four primer sets amplified fragments with PCR failure rates between 1.1 and 5.7% irrespective of which primer pairs were used (*Table 4*).

To determine any potential allele-specific bias in the identification of SNPs, we screened all trees individually for SNPs and did an initial statistical analysis to judge the quality of data in our collection (*Table 5*). We identified different numbers of SNPs for a given amplicon when using either *tremula*- or *trichocarpa*-specific primers. Considering that the SNPs amplified in three independent experiments by the *tremula*-specific primers are "real" SNPs in the *P. tremula* collection, the *trichocarpa*-specific primers identified a total of nine additional SNPs and missed twelve SNPs compared to the *tremula*-specific data sets. We assume that sequence differences in the primer binding sites between *trichocarpa* and *tremula* lead to the biased amplification of one of two *tremula*-alleles. This allele-specific PCR bias results in the relatively large number of SNPs missed by the *trichocarpa*-specific primers. Alternatively, additional and unspecific primer binding sites could result in the observed miscalling of SNPs.

Analyzing these SNPs and their distribution among the collection in more detail, we observed that a total of 47 *P. trichocarpa*-amplified SNPs deviate from the Hardy-Weinberg-Equilibrium (HWE) compared to 7 SNPs amplified with the *P. tremula* specific primers. The inbreeding coefficient per amplicon over the collection showed a significantly higher proportion of homozygous alleles for the *P. trichocarpa*- compared to *P. tremula*-amplified SNPs. As the trees were collected from different habitats around Germany (*Figure 1*), we would not expect significant levels of inbreeding and loss of heterozygosity for this collection. These effects are actually observed when applying the *P. trichocarpa*-specific PCR and sequencing conditions. The *P. tremula*-specific SNP detection

Table 4. – SNPs identified by *trichocarpa*- and *tremula*-specific PCR conditions, comparison of PCR failure, false-positive and false-negative SNPs, inbreeding coefficient, and Hardy-Weinberg equilibrium.

Gene, exon	PCR specific for	Failed PCRs	No. of SNPs	Additional SNPs[1]	Missed SNPs[2]	F_{IS}[3]	Out of HWE[4]
Drs1, exon 3	*P. trichocarpa*	5	16	4	2	0,523	11
	P. tremula	1	14			0,095	0
Drs1, exon 4	*P. trichocarpa*	4	21	4	2	0,675	13
	P. tremula	3	19			0,438	3
Erecta, exon 23+24	*P. trichocarpa*	1	24	0	5	0,549	16
	P. tremula	5	29			0,243	4
Mt2a, exon 2	*P. trichocarpa*	2	17	1	3	0,360	7
	P. tremula	2	19			0,030	0

[1] SNPs identified by t*richocarpa*-specific but not by *tremula*-specific primers.
[2] SNPs missed by t*richocarpa*-specific primers but identified by *tremula*-specific primers.
[3] Inbreeding coefficient (F_{IS}) per amplicon, $F_{IS} > 0$ indiciates an excess of homozygous individuals in the analyzed population (N = 88).
[4] Total number of SNPs that deviate from HWE after multiple test correction (Bonferronie correction $\alpha = 0,05/81$).

Table 5. – Sequence variation in the *P. tremula* natural population.

Gene	exon	fragment length (bp)	Variations total	Indels total	SNPs total	Coding sequence				Non-coding sequence			
						Length (bp)[1]	SNPs non-synonymous	SNPs synonymous	Frequency cds	Length (bp)[1]	SNPs (UTR, nongenic)[2]	SNPs (intron)	Frequency non-cds
Drs1	3	631	15	1	14	408	3	4	58	223	0	7	32
	4	555	22	3	19	99	1	1	50	456	0	17	27
Erecta	23+24	997	30	1	29	498	1[3]	3	125	499	0	25	20
	25+26	1175	20	1	19	616[1]	7	11	36	44[1]	0	1	22
Kpt1	7-10	698	18	1	17	397	6	3	40	301	0	8	38
Kup1	2+3	577	4	0	4	394	1	1	197	183	0	2	92
	9	731	5	0	5	676	2	3	135	55	0	0	0
Mt2a	1+2	665	22	1	21	237	4	2	40	428	11	4	29
Ptk2	9+10	427	18	2	16	333	5	7	28	94	0	4	24
Ptm5	1	630	20	2	18	182	0	4	46	448	10	4	32
	3+4	757	18	1	17	204	2	2	51	553	0	13	43
Ptork	3+4	407	13	1	12	343	1	4	80	64	0	3	31
	8+9	576	10	0	10	337	4	3	48	239	0	7	80
SLAC	2+3	484	13	1	12	234	0	2	117	250	0	10	25
Sp1	1+2	869	14	0	14	347	1	2	116	522	9	2	47
UDGH	1	802	16	0	16	761	2	12	54	41	2	0	21
Total	29	11,190	258	15	243	6,122	40	64		4,553	32	107	

[1] Length of screened coding/non-coding sequence: this number represents the actually resolved and finally screened bp, in case of *erecta* (exon 25 + 26) the amplified PCR fragment was sequenced only partially although three different sequencing primers have been used. This is based on the large part of intronic sequence that shows a high content of repetitive regions.

[2] Untranslated regions, 5' UTR: *mta2a*, 160 bp; *ptm5* exon1, 225 bp; *sp1*, 75 bp; and *UDGH*, 41 bp upstream of proposed translation initiation site considered. 3'UTR: *sp1*, 200 bp downstream of stop codon considered.

[3] SNP erecta_1_604 disrupts the 3' donor position of the exon 24 splice site, which might influence the full-length translation of the *erecta* peptide.

Non-synonymous substitutions lead to changes in the amino acid composition or affects predicted splice sites, synonymous substitutions do not affect the amino acid composition, SNPs located in introns are considered as silent mutations.

resulted in a much milder loss of heterozygosity (0.030 to 0.438) and less individuals that deviate from HWE (*Table 4*) for the same batch.

Sequence variation within the P. tremula collection

In an initial set of fourteen candidate genes, we focused on genes with a potential physiological function during stress-mediated responses in the context of drought tolerance and xylogenesis (*Table 1*). We successfully established *P. tremula*-specific amplification and resequencing conditions for 29 exons representing 11 candidate genes (circa 1/3 of all predicted exons).

The established 16 amplicons cover 6.1 kb of coding sequence and additional 4.6 kb of noncoding sequence (*Table 5*). Processing all amplicons on this collection, we generated about 3 Mb of sequence information. We identified 258 genetic variants, 15 of these represent indels that are located exclusively in non-coding regions and 243 are SNPs that are found in coding and non-coding regions of candidate genes.

The overall frequency of SNPs is 1 SNP in 44 bp, which is higher compared to recent studies reporting 1 in 70 bp and 1 in 54 bp, respectively (INGVARSSON, 2005; ISMAIL *et al.,* 2012). On average, we identified twice the number of SNPs in non-coding regions (140 SNPs with an average frequency of 1 SNP in 35 bp) than in coding regions (103 SNPs with an average frequency of 1 SNP in 76 bp) supporting the higher degree of sequence conservation in coding compared to non-coding regions. However, the frequency of SNPs varied between amplicons ranging from 1 SNP in 20 bp in the non-coding region of *erecta* (exon 23 and 24) to 1 in 197 bp in the coding region of *Kup1* (exon 2 and 3). Six of 243 SNPs were found in more than two allelic variants in our collection whereas all other SNPs are bi-allelic.

To get some insights into the potential physiological effect of a SNP located in the coding sequence, we translated the allelic variants based on the predicted annotation of our candidate genes. We identified one SNP in the *erecta* gene that affects the 3' donor site of the pre-

dicted splice site of exon 24, which might cause an early termination of the gene product. In addition, we found 40 SNPs that lead to changes in the translated amino acids, which are classified as non-synonymous substitutions. 64 SNPs in coding areas do not cause changes in the amino acid composition (synonymous substitution), along with 32 SNPs located in UTRs and adjacent non-genic regions and 107 SNPs located in introns.

Diversity of analyzed genes in the Populus tremula collection

Since the *P. tremula* collection is a newly established resource for potential breeding approaches, we wanted to explore its heterogeneity to consider the suitability for ecological and evolutionary studies. As parameters for the genetic heterogeneity of a given population, the expected and observed heterozygosities, and the deviation from HWE were calculated for each detected SNP marker (*Table 3*). Only 4.5% of all detected SNPs deviate from HWE ($\alpha = 0,0002$ after Bonferroni correction). These eleven deviating SNPs are concentrated mainly in three different amplicons with 4 deviating SNPs in *drs1*, exon4 (21%); 3 in *ptk2*, exon 9 and 10 (19%); and 3 in *ptork*, exon 3 and 4 (25%); in addition, 1 deviating SNP is detected in *erecta*, exon 23 and 24 (3.4%).

Discussion

SNPs are considered as the basic detectable unit of genetic variations, which are used as DNA-based markers. Hence, informative SNPs represent an attractive resource for genotyping and breeding strategies due to their wide distribution throughout the genome and the fact that they are easy to score and amenable to large-scale applications (GARVIN *et al.*, 2010; SEEB *et al.*, 2011). Our approach was focused on the detection of SNPs in informative regions of eleven candidate genes that consist of coding (57%) as well as exon-spanning non-coding areas (43%). As stated by SLATE *et al.* (2009) and GARVIN *et al.* (2010), the critical point of this method is the design of species-specific primers and PCR-optimization to avoid any bias in the further analysis.

Applying the method described in the present study on a total of fourteen candidate genes, we managed to design species-specific primers for sixteen amplicons covering mainly coding regions in eleven genes (*Table 1* and *2*). Three genes failed in this approach since they are either found as duplicates or exist as two almost identical paralogs in the *P. trichocarpa* genome, which is partially duplicated (TUSKAN *et al.,* 2006; RODGERS-MELNICK *et al.,* 2012). We started by using *P. trichocarpa* derived primers to amplify from *P. tremula* genomic DNA. This restrictive cross-species primer design strategy resulted in 89% specific amplification products compared to 14% or 61 to 67%, which have been described for other than chloroplast loci in *P. tremula* and other poplar species (INGVARSSON, 2008; JOSEPH and LEXER, 2008; OLSON *et al.,* 2010; SCHROEDER *et al.,* 2012). These *P. tremula* derived amplification products provided the species-specific sequence information that is used for the design and optimization of specific *P. tremula* derived primers. The critical step of our pipeline was the design and experimental validation of *P. tremula* specific primers, a procedure that was proposed for SNP detection in non-model species (SLATE *et al.,* 2009; GARVIN *et al.,* 2010). The systematic comparison of SNP data derived from individual *P. tremula* trees after amplification with either *P. tremula* or *P. trichocarpa* specific primers revealed a significant allele-bias of discovered SNPs depending on the specificity of primers. Even minor sequence variations (one and two nucleotides) in the primer binding sites could lead to miscalling of heterozygous as well as homozygous SNPs. These errors could consequentially result in a misleading population structure for the *P. tremula* collection. This effect was demonstrated by the massive inbreeding tendency and large number of HWE outliers when using *P. trichocarpa* derived primer conditions that represent cross-species conditions. This potentially misleading population structure is massively reduced in our collection when using *P. tremula*-specific primers indicating that a careful design of species-specific PCR conditions is an essential prerequisite for any robust and reliable SNP detection approach no matter which SNP discovery tool is used.

Conclusions

A prerequisite for reliable SNP discovery in non-model organisms such as *P. tremula* is the careful design of PCR amplicons that are highly specific for the gene of interest in this species.

We established a method that combines *in silico* PCR design with experimental confirmation and adjustment of data, which is required to avoid any allele bias in the SNP data. Making use of the present method, we established fourteen new *P. tremula* specific genic regions.

Applying this approach to a newly established collection of *P. tremula* trees led to the identification of 243 SNPs in this population. A total of 95 % of SNPs does not deviate from HWE, which is consistent with the finding that the majority of analyzed regions do not show a significant in- or out-breeding behavior. This initial analysis indicates that a robust species-specific pipeline is required for the reliable detection of SNPs that then can be used for the detailed analysis of population structure and potential adaptive behavior of SNP markers.

References

AREND, M., A. STINZING, C. WIND, K. LANGER, A. LATZ, P. ACHE, J. FROMM and R. HEDRICH (2005): Polar-localised poplar K⁺ channel capable of controlling electrical properties of wood-forming cells. Planta **223**: 140–148.

BHANGALE, T. R., M. STEPHENS and D. A. NICKERSON (2006): Automating resequencing-based detection of insertion-deletion polymorphisms. Nature Genetics **38**: 1457–1462.

CARDON, L. R. and J. I. BELL (2001): Association study designs for complex diseases. Nature reviews. Genetics **2**: 91–99.

CSEKE, L. J., J. ZHENG and G. K. PODILA (2003): Characterization of PTM5 in aspen trees: a MADS-box gene expressed during woody vascular development. Gene **318**: 55–67.

DOYLE, J. J. and J. L. DOYLE (1987): A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19**: 11–15.

ECKENWALDER, J. E. (1996): Systematics and evolution of Populus, pp. 7–32. *In:* Biology of *Populus* and Its Implications for Management and Conservation, edited by R. F. STETTLER, N.R.C. Research Press, Ottawa, Canada.

GARVIN, M. R., K. SAITOH and A. J. GHARRETT (2010): Application of single nucleotide polymorphisms to non-model species: a technical review. Molecular Ecology Resources **10**: 915–934.

GOODSTEIN, D. M., S. SHU, R. HOWSON, R. NEUPANE, R. D. HAYES, J. FAZO, T. MITROS, W. DIRKS, U. HELLSTEN and N. PUTNAM (2012): Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research: 1178–1186.

GRAY-MITSUMUNE, M., E. J. MELLEROWICZ, H. ABE, J. SCHRADER, A. WINZELL, F. STERKY, K. BLOMQVIST, S. MCQUEEN-MASON, T. T. TEERI and B. SUNDBERG (2004): Expansins abundant in secondary xylem belong to subgroup A of the alpha-expansin gene family. Plant Physiology **135**: 1552–1564.

HAMANISHI, E. T., B. R. THOMAS and M. M. CAMPBELL (2012): Drought induces alterations in the stomatal development program in *Populus*. Journal of Experimental Botany **63**: 4959–4971.

HELYAR, S. J., J. HEMMER-HANSEN, D. BEKKEVOLD, M. I. TAYLOR, R. OGDEN, M. T. LIMBORG, A. CARIANI, G. E. MAES, E. DIOPERE and G.. R. CARVALHO (2011): Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources **11**: 123–136.

INGVARSSON, P. K. (2005): Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., Salicaceae), Genetics **169**: 945–953.

INGVARSSON, P. K. (2008): Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula*. Genetics **180**: 329–340.

ISMAIL, M., R. Y. SOOLANAYAKANAHALLY, P. K. INGVARSSON, R. D. GUY, S. JANSSON, S. N. SILIM and Y. A. EL-KASSABY (2012): Comparative nucleotide diversity across North American and European populus species. Journal of Molecular Evolution **74**: 257–272.

JOHANSSON, H., F. STERKY, B. AMINI, J. LUNDEBERG and L. A. KLECZKOWSKI (2002): Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation. Biochimica et Biophysica Acta **1576**: 53–58.

JOSEPH, J. A. and C. LEXER (2008): A set of novel DNA polymorphisms within candidate genes potentially involved in ecological divergence between *Populus alba* and *P. tremula*, two hybridizing European forest trees. Molecular Ecology Resources **8**: 188–192.

KELLEHER, S. R., N. LEVSEN, M. S. OLSON and P. TIFFIN (2012): Local Adaptation in the Flowering-Time Gene Network of Balsam Poplar, *Populus balsamifera* L. Molecular Biology and Evolution **29**: 3143–3152.

LANGER, K., P. ACHE, D. GEIGER, A. STINZING, M. AREND, C. WIND, S. REGAN, J. FROMM and R. HEDRICH (2002): Poplar potassium transporters capable of controlling K$^+$ homeostasis and K$^+$-dependent xylogenesis. The Plant journal: for cell and molecular biology **32**: 997–1009.

LEE, J. H., W. TERZAGHI, G. GUSMAROLI, J. B. CHARRON, H. J. YOON, H. CHEN, Y. J. HE, Y. XIONG and X. W. DENG (2008): Characterization of Arabidopsis and rice DWD proteins and their roles as substrate receptors for CUL4-RING E3 ubiquitin ligases. The Plant Cell **20**: 152–167.

LI, X., C. ZHU, C. T. YEH, W. WU, E. M. TAKACS, K. A. PETSCH, F. TIAN, G. BAI, E. S. BUCKLER, G. J. MUEHLBAUER *et al.* (2012): Genic and nongenic contributions to natural variation of quantitative traits in maize. Genome Research DOI 10.1101/gr.140277.112

LIU, K. and S. V. MUSE (2005): PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics **21**: 2128–2129.

MARJANOVIC, Z., N. UEHLEIN, R. KALDENHOFF, J. J. ZWIAZEK, M. WEISS, R. HAMPP and U. NEHLS (2005): Aquaporins in poplar: what a difference a symbiont makes!. Planta **222**: 258–268.

MASLE, J., S. R. GILMORE and G. D. FARQUHAR (2005): The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. Nature **436**: 866–870.

NICKERSON, D. A., V. O. TOBE and S. L. TAYLOR (1997): PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. Nucleic Acids Research **25**: 2745–2751.

OLSON, M. S., A. L. ROBERTSON, N. TAKEBAYASHI, S. SILIM, W. R. SCHROEDER and P. TIFFIN (2010): Nucleotide diversity and linkage disequilibrium in balsam poplar (*Populus balsamifera*). The New Phytologist **186**: 526–536.

ROZEN, S. and H. SKALETSKY (2000): Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology **132**: 365–386.

SCHROEDER, H., A. M. HOELTKEN and M. FLADUNG (2012): Differentiation of *Populus* species using chloroplast single nucleotide polymorphism (SNP) markers – essential for comprehensible and reliable poplar breeding. Plant Biology **14**: 374–381.

SEEB, L. W., W. D. TEMPLIN, S. SATO, S. ABE, K. WARHEIT, J. Y. PARK and J. E. SEEB (2011): Single nucleotide polymorphisms across a species' range: implications for conservation studies of Pacific salmon. Molecular Ecology Resources 11 Suppl **1**: 195–217.

SEIFERT, S., B. VORNAM and R. FINKELDEY (2012): DNA sequence variation and development of SNP markers in beech (*Fagus sylvatica* L.). European Journal of Forest Research **131**: 1761–1770.

SLATE, J., J. GRATTEN, D. BERALDI, J. STAPLEY, M. HALE and J. M. PEMBERTON (2009): Gene mapping in the wild with SNPs: guidelines and future directions. Genetica **136**: 97–107.

STRAUSS, S. H. and F. M. MARTIN (2004): Poplar genomics comes of age. New Phytologist **164**: 1–4.

STREET, N. R., O. SKOGSTROM, A. SJODIN, J. TUCKER, M. RODRIGUEZ-ACOSTA, P. NILSSON, S. JANSSON and G. TAYLOR (2006): The genetics and genomics of the drought response in *Populus*. Plant Journal **48**: 321–341.

TUSKAN, G. A., S. DIFAZIO, S. JANSSON, J. BOHLMANN, I. GRIGORIEV, U. HELLSTEN, N. PUTNAM, S. RALPH, S. ROMBAUTS, A. SALAMOV *et al.* (2006): The genome of black cottonwood, P*opulus trichocarpa* (Torr. & Gray). Science **313**: 1596–1604.

VAHISALU, T., H. KOLLIST, Y. F. WANG, N. NISHIMURA, W. Y. CHAN, G. VALERIO, A. LAMMINMAKI, M. BROSCHE, H. MOLDAU and R. DESIKAN *et al.* (2008): SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. Nature **452**: 487–491.

WANG, W. X., D. PELAH, T. ALERGAND, O. SHOSEYOV and A. ALTMAN (2002): Characterization of SP1, a stress-responsive, boiling-soluble, homo-oligomeric protein from aspen. Plant Physiology **130**: 865–875.

WILKINS, O., L. WALDRON, H. NAHAL, N. J. PROVART and M. M. CAMPBELL (2009): Genotype and time of day shape the *Populus* drought response. The Plant Journal **60**: 703–715.

WINKLER, S., N. GSCHEIDEL and M. BRAND (2011): Mutant generation in vertebrate model organisms by TILLING. Methods in Molecular Biology **770**: 475–504.

YOSHIURA, K., A. KINOSHITA, T. ISHIDA, A. NINOKATA, T. ISHIKAWA, T. KANAME, M. BANNAI, K. TOKUNAGA, S. SONODA and R. KOMAKI *et al.* (2006): A SNP in the ABCC11 gene is the determinant of human earwax type. Nature Genetics **38**: 324–330.