Collection: COST Action FP0905
"Biosafety of forest transgenic trees and EU policy directives"
Guest Editors: Cristina Vettori, Matthias Fladung

Differentiation of *Populus* species by chloroplast SNP markers for barcoding and breeding approaches

Hilke Schroeder, Matthias Fladung

About 30 species within the genus *Populus* are classified in six sections. Several species belonging to different sections are cross-compatible, resulting in a high number of naturally occurring interspecific hybrids. Additionally, an even higher number of hybrids has been produced in huge breeding programs during the last 100 years. Hence, determination of poplar species used for the production of "multi-species-hybrids" is often difficult, and a challenge for the development of molecular markers aimed at species identification. Moreover, 14 of the 30 poplar species known are often used for the production of artificial hybrids and clones. In this study, over 20 regions of the chloroplast genome of poplars were analyzed by the use of 23 primer combinations already established for "barcoding" purposes, and 17 new primer combinations previously designed, with the aim of testing their suitability to differentiate 14 poplar species. Only about half of the established barcoding primer combinations yielded PCR amplification products, while new primer combinations had a much higher amplification success. Species-specific SNPs or Indels were detected in 11 out of the 14 poplar species analyzed. In order to set up a low-cost and fast method for species identification, we developed PCR-RFLPs applicable to seven of the species-specific SNPs. Overall, a high variability was found in the chloroplast intergenic spacers, as much as four primer combinations were needed to differentiate 11 species. Based on our results, the use of multi-locus combinations is recommended in barcoding analyses.

Keywords: Chloroplast Genome, SNPs, Indel, Barcoding, Intergenic Spacer

Introduction

The genus *Populus* is one of the world's most important tree genera. High growth rates, particularly of some interspecies hybrids, and a broad applicability ranging from wood and paper to energy production, led to their widespread cultivation in Europe and North America (Dickman & Stuart 1983). Hybrids between various *Populus* species belonging to the same or different sections are commonly used in short rotation coppices (SRC) for biomass production. Further, clones of various poplar species are easily accessible for genetic transformation, and a huge number of transgenic poplar clones have been tested in the field for biosafety reasons and commercial application (Hoenicka & Fladung 2006, Walter et al. 2010).

Especially interspecies-hybrids are well suited for biomass production because of their superior growth and advanced resistance traits. However, due to extreme variability of *Populus* hybrids, the species identification within the genus *Populus* using morphological characters sometimes proved to be difficult. A high number of interspecies-hybrids has been artificially produced in extensive breeding programs during the last 100 years (Stettler et al. 1996), in addition to naturally occurring hybrids also possibly used for breeding purposes. The many years of breeding work, including repetitive crosses and back-crosses, have caused a lack of background information for many of commercially available hybrid clones originally used, though clear species identification is inevitably necessary for breeding activities aimed at licensing new high-efficiency clones.

Due to its small genome size, poplar has been selected as a model species for tree genomics, resulting in the publication of the full genome of western black cottonwood, *P. trichocarpa* (Tuskan et al. 2006). Using this genomic information, we evaluated and tested the usability of both already published plant barcoding regions from the chloroplast (Hollingsworth et al. 2009, Chase et al. 2005, Barcode of Life: http://www.barcodeoflife.org/) and novel developed primers by Schroeder et al. (2012) for their efficacy in differentiating the 14 poplar species listed in Tab. 1.

Material and Methods

Sample information

Species and number of individuals (genotypes) per species used for this study are listed in Tab. 1. Individuals of three *Populus* species of the section *Populus (P. alba, P. tremula, P. tremuloides*) were collected in the arboretum and were provided by colleagues from Bavaria and Saxony, while *P. euphratica* (section *Turanga*) has been taken from the tissue culture of the Thünen-Institute of Forest Genetics in Grosshansdorf (Germany). For the two European species, further individuals have been sampled by colleagues all over the distribution area. Some of the individuals of *Populus* spp. of the sections *Aigeiros (P. nigra* and *P. deltoides*) and *Tacamahaca (P. trichocarpa, P. maximowiczii, P. cathayana, P. szechuanica, P. ussuriensis, P. koreana*) were kindly provided by the NW-FVA (Nordwestdeutsche Forstliche Versuchsanstalt) in Hann. Muenden (Germany). Six individuals of *P. nigra* were from the Ural Mountains and others were taken as samples at riversides in Germany. Nine individuals of *P. maximowiczii* and seven individuals of *P. simonii* have been provided by the Thünen-Institute of Forest Genetics in Waldsieversdorf (Germany). Several individuals of some of the above mentioned species (among others *P. wilsonii*) were received from the Botanical Gardens of Hamburg, Marburg, Tuebingen and Dresden (Germany).

DNA extraction and PCR amplification

One cm^2 of a single leaf was ground to

Thünen Institute of Forest Genetics, Sieker Landstrasse 2, D-22927 Grosshansdorf (Germany)

@ Hilke Schroeder (hilke.schroeder@ti.bund.de)

Received: Apr 28, 2014 - Accepted: Sep 16, 2014

Citation: Schroeder H, Fladung M, 2014. Differentiation of *Populus* species by chloroplast SNP markers for barcoding and breeding approaches. iForest 8: 544-546 [online 2014-11-13] URL: http://www.sisef.it/iforest/contents/?id=ifor1326-007

Communicated by: Elena Paoletti

Tab. 1 - Numbers of species-specific SNPs, number of restriction sites within SNPs, and Indels for 14 *Populus* species in 24 chloroplast fragments. Labels for each species are given in brackets.

Section	Species	No. of individuals	No. SNPs	No. restriction sites	No. Indels
Populus	*P. alba* (alb)	26	4	1	3
	P. tremula (tre)	40	14	3	7
	P. tremuloides (tro)	40	31	4	11
Aigeiros	*P. nigra* (nig)	36	7	2	3
	P. deltoides (del)	19	9	1	4
Tacamahaca	*P. trichocarpa* (tri)	35	6	1	1
	P. maximowiczii (max)	18	2	0	1
	P. simonii (sim)	9	2	0	1
	P. koreana (kor)	4	0	0	0
	P. cathayana (cat)	1	0	0	0
	P. szechuanica (sze)	1	2	1	0
	P. ussuriensis (uss)	1	0	0	0
Leucoides	*P. wilsonii* (wil)	1	4	0	3
Turanga	*P. euphratica* (eup)	1	11	0	2

powder in liquid nitrogen. Total DNA was extracted, following a modified ATMAB protocol by Dumolin et al. (1995). A standard protocol was used for PCR reactions (Schroeder et al. 2012). The PCR reactions contained from 20 to 50 ng template DNA, 50 mM KCl, 20 mM Tris-HCl (pH 8.4), 1.8 mM $MgCl_2$, 200 µM dNTPs, 1 unit Taq polymerase, and 0.4 µM of each primer in a total volume of 25 µl. PCR was carried out in a Biometra Personal Thermocycler (Goettingen, Germany) with a pre-denaturation step at 94°C for 4 min, followed by 30 cycles of 94°C for 1 min, suitable annealing temperature for each primer combination (ranging between 50°C and 65°C) for 45 sec or 1 min, 72°C for 1 min and a final elongation at 72°C for 10 min. PCR amplification products were checked on a 1 % agarose gel stained with Roti-Safe GelStain (Carl Roth GmbH & Co. KG, Karlsruhe, Germany).

Tested chloroplast regions

Twenty three published barcoding primer combinations were used for PCR amplification of coding and non-coding (intergenic spacers) chloroplast regions. Additionally, 17 primer combinations were newly designed taking advantage of the sequence of the *Populus trichocarpa* chloroplast genome (Schroeder & Fladung 2010, Schroeder et al. 2012). The primer combinations are given in detail in Schroeder et al. (2012).

Sequencing and PCR-RFLPs

For sequencing, 15 µl of the PCR product was purified either with 5 µl of 4 M lithium chloride or using the "High Pure PCR Product Purification" kit from Roche (Mannheim, Germany). All the sequences were provided by StarSeq (Mainz, Germany).

Up to ten individuals per species and primer combination have been sequenced. Obtained sequences were aligned and screened for the presence of SNPs by using either the software SeqMan 7.1.0 from DNAStar (Madison, USA) or Sequencher 4.9 (Gene Codes Corporation, Ann Arbor, USA). Differentiating sequences (SNPs, Indels) were checked for a possible application of restriction enzymes using the software NEBcutter V2.0 from New England BioLabs Inc (Ipswich, USA). Individuals already sequenced, plus further individuals given in Tab. 1, have been used to test the usability of PCR-RFLPs for species differentiation.

Results and Discussion

Twelve of the 23 used barcoding primer combinations and 16 out of the 17 newly-designed primer combinations yielded PCR amplification products in all the individuals of the 14 species tested. Twenty-four of these amplification products have been sequenced and checked for species-specific SNPs or Indels for the first seven species in Tab. 1 (*P. alba, P. tremula, P. tremuloides, P. nigra, P. deltoides, P. trichocarpa, P. maximowiczii*), and 19 fragments have been sequenced for the remaining seven species (*P. simonii, P. koreana, P. cathayana, P. szechuanica, P. ussuriensis, P. wilsonii, P. euphratica*). Species-specific SNPs or Indels were not found only in three species (*P. koreana*, *P. cathayana* and *P. ussuriensis* - Tab. 1). For these species, five further primer combinations were tested. In addition, more individuals of the last five species listed in Tab. 1 (*P. cathayana, P. szechuanica, P. ussuriensis, P. wilsonii* and *P. euphratica*) are needed to validate the species-specificity of the genetic variation.

Interestingly, for the two species *P. wilsonii* and *P. euphratica*, each belonging to two different sections, comparatively many SNPs and Indels have been found than for most of the other species of the sections *Tacamahaca* and *Aigeiros* (Tab. 1). Also, the similarities between species within the sections are different. No genetic variation could be found between the two species *P. koreana* and *P. cathayana*. These two species showed a similarity to *P. maximowiczii* (so far differing only by one SNP) much higher than to

Tab. 2 - Number of polymorphisms, SNPs, and Indels detected in PCR amplicons for all successful amplified primer combinations. (Length): length of the PCR product amplified with the given primers in base pairs referred to *P. trichocarpa*; (No. Poly): total number of polymorphisms (SNPs and Indels): (No. SNPs): species-specific SNPs identified; (No. Indels): species-specific Indels identified; (% var): percent of variable sites calculated as the ratio between the total number of polymorphisms and the sequence length.

Chloroplast region	Length (bp)	No. Poly	No. SNPs	No. Indels	% var
trnV-atpE	740	32	18	8	4.3
psaJ-rpl33	617	12	6	3	1.9
rbcL-accD	762	12	4	1	1.6
matK-trnK	851	11	7	0	1.3
trnC-petN	441	11	5	1	2.5
rpoC2-rpoC1	533	10	5	1	1.9
trnG-psbK	545	10	3	1	1.8
psbK-psbI	593	8	6	0	1.3
trnH-psbA	367	8	4	2	2.2
rps8-rpl14	455	7	2	2	1.5
ndhE-ndhG	356	7	2	2	2.0
rpoB2-rpoB4	548	7	4	0	1.3
rbcLa-rbcLajf634R	654	7	2	0	1.1
3F_KIM-1R_KIM (matK)	882	6	4	1	0.7
rps4_trnL	829	5	3	0	0.6
atpI-rps2	570	4	3	0	0.7
rps2-rpoC2	524	4	0	0	0.8
ycf10-petA	359	3	0	1	0.8

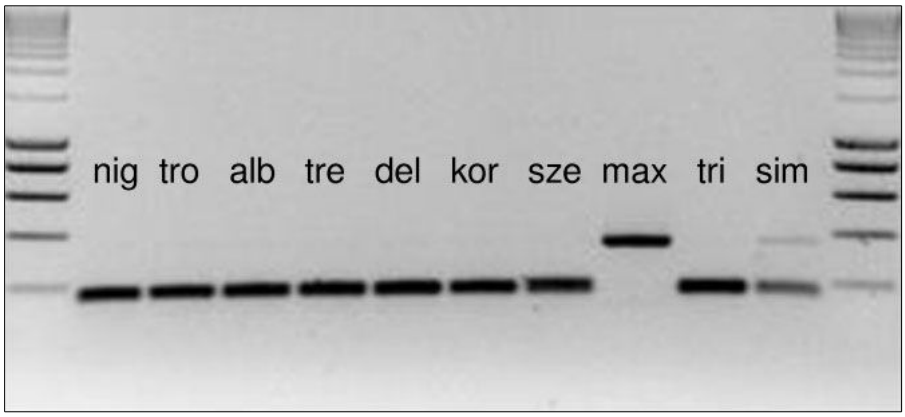

Fig. 1 - PCR-RFLP of the intergenic spacer *ndh*E-*ndh*G. *P. maximowiczii* (max) shows a larger PCR product due to the lack of the restriction site, whereas all other species were cut into two fragments of nearly the same size. Size marker (first and last lanes) is the SMART ladder (Eurogentec, Belgium). For species' labels, see Tab. 1.

all the other species within the section *Tacamahaca*. Analysis of kinship with a wide range of markers for all the species could be of interest.

No species-specific SNPs or Indels were found by the analysis of three chloroplast regions (*atp*F-*atp*H, *rps*2-*rpo*C2 and *psb*J-*psb*F) in the first seven species listed in Tab. 1, though the same chloroplast regions revealed the presence of SNPs or Indels in further seven species. The highest number of species-specific SNPs and Indels were found within the intergenic spacers *trn*V-*atp*E, *mat*K-*trn*K, *psa*J-*rpl*33, *trn*H-*psb*A, *psb*K-*psb*I and *atp*B-*rbc*L, followed by the coding region *rpo*C (Tab. 2). Four primer combinations were needed for the differentiation of 11 species with species-specific SNPs or Indels. Based on our results, we therefore recommend the use of a multi-locus combination of the plastid genome to increase barcoding potential, according to several previous reports (Hollingsworth et al. 2009, Kress & Erickson 2008, Little & Stevenson 2007, Sass et al. 2007).

SNPs and Indels identified in this study could be an important contribution to the project "barcoding of life". Their efficient application in the identification of *Populus* species for breeding purposes requires rapid test methods to be developed, in order to avoid sequencing or the use of capillary electrophoresis. To this purpose, all the sequences with SNPs were checked for restriction sites for use as PCR-RFLPs. Suitable restriction sites were found for 13 SNPs in seven species (Tab. 1). Some further restriction sites offered the possibility of a negative identification, that means all species can be excluded but one. An example of a PCR-RFLP is given in Fig. 1, where *P. maximowiczii* is characterized by the lack of a specific restriction site.

Conclusions

In this study, we demonstrated the applicability of SNP markers in breeding programs. The combination of SNPs with PCR-RFLPs and length polymorphism provides an easy-to-use, fast, and low cost method for identifying and distinguishing closely-related poplar species.

We recommend the use of intergenic spacers for differentiating closely related species within the genus *Populus* because of the higher overall variability (Schroeder et al. 2011, 2012).

Acknowledgments

This project is supported by the Federal Ministry of Food and Agriculture via the "Fachagentur Nachwachsende Rohstoffe e.V." (FNR). We would like to thank the greenhouse staff for harvesting of leaves from high poplar trees, our technical assistant Susanne Bein for the laboratory work, and the colleagues providing us with plant material. Our special thanks go to Dietrich Ewald and Georg von Wühlisch for helping us with their contacts to colleagues from China and the USA, and to Dina Führmann for proof reading of the English grammar and style. We are grateful to the Botanical Gardens from Hamburg, Marburg, Dresden and Tübingen for supplying us with poplar reference material.

References

Chase MW, Salamin N, Wilkinson M, Dunwell JM, Kesanakurthi RP, Haidar N, Savolainen V (2005). Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transaction of the Royal Society of London Series B Biological Science 360: 1889-1895. - doi: 10.1098/rstb.2005.1720

Dickman DI, Stuart K (1983). The culture of poplars in Eastern North America. Department of Forestry, Michigan State University, East Lansing, MC, USA, pp. 168.

Dumolin S, Demesure B, Petit RJ (1995). Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. Theoretical and Applied Genetics 91: 1253-1256. - doi: 10.1007/BF00220937

Hoenicka H, Fladung M (2006). Biosafety in *Populus* spp. and other forest trees: from non-native species to taxa derived from traditional breeding and genetic engineering. Trees 20: 131-144. - doi: 10.1007/s00468-005-0023-5

Hollingsworth ML, Clark AA, Forrest LL, Richardson J, Pennington RT, Long DG, Cowan R, Chase MW, Gaudeul M, Hollingsworth PM (2009). Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. Molecular Ecology Resources 9: 439-457. - doi: 10.1111/j.1755-0998.2008.02439.x

Kress WJ, Erickson DL (2008). DNA barcodes: genes, genomics, and bioinformatics. Proceedings of the National Academy of Science USA 105 (8): 2761-2762. - doi: 10.1073/pnas.0800476105

Little DP, Stevenson DW (2007). A comparison of algorithms for the identification of specimens using DNA barcodes: examples from gymnosperms. Cladistics 23: 1-21. - doi: 10.1111/j.1096-0031.2006.00126.x

Sass C, Little DP, Stevenson DW, Specht CD (2007). DNA barcoding in the Cycadales: testing the potential of proposed barcoding markers for species identification of cycads. PLoS ONE 2 (11): e1154. - doi: 10.1371/journal.pone.0001154

Schroeder H, Fladung M (2010). SSR and SNP markers for the identification of clones, hybrids and species within the genus *Populus*. Silvae Genetica 59 (6): 257-263. [online] URL: http://www.silvaegenetica.de/fileadmin/content/dokument/archiv/silvaegenetica/59_2010/Heft6/_02__Schröder_2430.pdf

Schroeder H, Höltken AM, Fladung M (2011). Chloroplast SNP-marker as powerful tool for differentiation of *Populus* species in reliable poplar breeding and barcoding approaches. BMC Proceedings 5 (Suppl 7): P56. - doi: 10.1186/1753-6561-5-S7-P56

Schroeder H, Höltken AM, Fladung M (2012). Differentiation of *Populus* species using chloroplast SNP-markers - essential for comprehensible and reliable poplar breeding. Plant Biology 14: 374-381. - doi: 10.1111/j.1438-8677.2011.00502.x

Stettler R, Bradshaw T, Heilman P, Hinckley T (1996). Biology of *Populus* and its implications for management and conservation. NRC Research Press. Montreal, Canada, pp. 542.

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313: 1596-1604. - doi: 10.1126/science.1128691

Walter C, Fladung M, Boerjan W (2010). The 20-year environmental safety record of GM trees. Nature Biotechnology 28 (7): 656-658. - doi: 10.1038/nbt0710-656