

Three microsatellite multiplex PCR assays allowing high resolution genotyping of white spruce, *Picea glauca*

By P. EUSEMANN^{1,*}, P. HERZIG², M. KIEß², S. AHLGRIMM², P. HERRMANN², M. WILMKING² and M. SCHNITTLER²

(Received 11th March 2015)

Abstract

Fifteen previously published microsatellite primer pairs developed for and tested on Canadian white spruce were screened for amplification and polymorphism in Alaskan populations and tested for their suitability in PCR multiplexing. Eleven loci expressing polymorphisms ranging from 7 to 58 alleles were selected for development and optimization of three multiplex assays. Four natural stands containing a total of 1470 trees were used to characterize the selected loci and demonstrate their applicability for genotyping studies and parentage analysis. These assays can be used for studies focusing on population genetics, parentage analysis, provenance research, or individual genetic fingerprinting. The use of multiplex PCR facilitates large-scale studies by simultaneously enabling high resolution and reducing processing time and per sample cost.

Key words: *Picea glauca*; SSR; fingerprinting; parentage analysis; population genetics; population structure

Introduction

For population genetic studies, microsatellites are the marker system of choice (SELKOE and TOONEN, 2006). One of the major drawbacks of the marker system, however, is the need to analyze multiple loci in order to gain a resolution that allows for meaningful statistical analysis. Depending on the degree of individual polymorphism, ecological studies employing microsatellites usually analyze 11–12 loci (GUICHOUX et al., 2011). This number significantly increases laboratory workload and per sample cost and severely limits the applicability of microsatellites, especially when high resolution is required as is the case e.g. for parentage analysis. One way to overcome this drawback is to amplify several loci within a single PCR reaction. The development of such multiplex systems is laborious and time consuming, since multiple loci first have to be tested individually and subsequently in different multiplex PCR reactions, until combinations are found that produce stable amplification. Developing reliable multiplex sets also facilitates the definition of standardized protocols for genotyping studies to aid comparability of results, which has been identified as one of the major shortcomings in the current use of microsatellites (GUICHOUX et al., 2011; LIESEBACH and

EWALD, 2012 and references therein). In this paper, we present the results of a primer screening of previously published microsatellite markers (HODGETTS et al., 2001; RAJORA et al., 2001) and the development of three multiplex assays that are suitable to address questions related to population genetics and parentage analyses of white spruce (*Picea glauca* (Moench) Voss). The aim of this paper is to provide researchers with an accurate, robust, and sensitive tool to perform population genetic studies on large sample numbers of this important forest tree.

Material and Methods

Primer Screening

Primers were chosen from the literature (HODGETTS et al., 2001; RAJORA et al., 2001). Of the 21 primer pairs published, 15 were selected for pre-screening. Selection was based on a high number of alleles, amplification of PCR-products in different size categories to facilitate multiplexing, and identical or similar annealing temperature, also to facilitate simultaneous amplification. Loci selected for pre-screening were:

UAPgCT3, UAPgAC/AT6, UAPgGT8, UAPgCA24, UAPsTG25, UAPgTG64, UAPgTG87, UAPgCA91, UAPgAG105, UAPgCT144 (from HODGETTS et al., 2001) and

PGL7, PGL12, PGL13, PGL14, PGL15 (from RAJORA et al., 2001).

Each primer pair was first tested in a separate PCR following the original protocols. Loci that showed successful and consistent amplification were then arranged into groups of four to develop PCR multiplex assays. Emphasis was put on selecting loci that amplified PCR products in different size categories to minimize overlap within each set of four loci. If possible, only primers with identical annealing temperature were combined into one set.

Plant material

All selected loci were subsequently characterized in 1470 white spruce trees originating from four stands at two sites located in the Alaskan Brooks Range and Alaska Range (Brooks Range: 67°56'N, 149°44'W, Alaska Range: 63°43'N, 149°00'W).

Multiplex PCR assays

DNA extraction was performed using the Invisorb Spin Plant Mini DNA extraction kit (Strattec) following the manufacturer's instructions. Extracted DNA was adjusted to a concentration of 5 ng/µl.

¹) Thünen Institute of Forest Genetics, Eberswalder Chaussee 3a, D-15377 Waldsiedersdorf, Germany.

²) Greifswald University, Department of Botany and Landscape Ecology, Soldmannstraße 15, D-17487 Greifswald, Germany.

*) Corresponding author: PASCAL EUSEMANN. Tel. +48-33433-157176, Fax +48-33433-157199. E-Mail: pascal.eusemann@ti.bund.de

Starting from the 15 pre-screened loci, three multiplex assays, each amplifying four loci, were developed. All primers within an assay were tested for possible primer-primer interactions and hairpin structures using the software AutoDimer 1.0 (VALLONE and BUTLER, 2004). Locus combinations and characteristics are given in *Table 1*.

Amplification was performed using the Qiagen Multiplex PCR Plus Kit (Qiagen). In order to reduce per sample cost, the original protocol was modified to perform the PCR reaction in a total reaction volume of 10 μ l instead of 50 μ l. In order to simplify lab procedures and amplify all sets under identical PCR conditions, primer concentrations were adjusted to produce comparable amplification of all loci within all three sets. A total PCR volume of 10 μ l, containing 1x Multiplex PCR Plus buffer (Qiagen), 0.2 μ M of each primer, and 20.0 ng DNA was used. An equimolar primer concentration of 0.2 μ M

was found to produce balanced signals for all loci within each assay. To reduce processing time, all primers of a specific assay were prepared as a primer mix containing all separate primers (*Table 2*).

PCRs were performed on Eppendorf Mastercycler thermocyclers under the following conditions: A cooling step of 5 minutes at 4°C while the lid of the thermocycler heats up, a denaturing and polymerase activation step of 5 minutes at 95°C, followed by 30 cycles of 95°C for 30 seconds, annealing at 60°C for 90 seconds, elongation at 72°C for 30 seconds and a singular final extension at 68°C for 10 minutes. PCR products were diluted 1:5 in ddH₂O. For analysis, 1 μ l of diluted product was combined with 0.15 μ l of GeneScan 500 LIZ size standard (Life Technologies) and 8.85 μ l of HiDi Formamide (Life Technologies). Fragment Analysis was carried out on a 3130xl Genetic Analyzer (Life Technologies). Genotyping was performed using Peak Scanner 1.0 soft-

Table 1. – Characteristics of the twelve loci used in three multiplex assays. N – number of alleles, N_e – effective number of alleles, H_e – expected heterozygosity, H_o – observed heterozygosity.

Set	Locus	Dye	Size (bp)	Alleles		Heterozygosity		Null allele frequency
				N	N _e	H _e	H _o	
1	UAPgAC/AT6	Hex	110–130	7	4.4	0.77	0.75	0.04
	PGL15	Hex	170–250	26	4.8	0.79	0.37	0.24
	UAPgTG87	Tamra	110–200	58	14.4	0.93	0.56	0.19
	UAPgCA91	Fam	100–160	38	10.0	0.90	0.46	0.23
2	UAPgGT8	Hex	190–245	23	9.5	0.89	0.71	0.10
	UAPgTG64	Hex	90–160	23	7.9	0.87	0.42	0.25
	UAPgAG105	Fam	160–185	12	4.6	0.78	0.76	0.02
	UAPgCT144	Tamra	135–195	17	3.4	0.71	0.68	0.01
3	PGL12	Hex	220–250	14	6.8	0.85	0.46	0.21
	UAPgCA24	Tamra	200–290	54	12.2	0.92	0.39	0.28
	UAPsTG25	Fam	90–140	9	1.3	0.26	0.12	0.15
	PGL14	Hex	120–180					

Table 2. – Primer- and PCR-master mix for the multiplex assays described.

Component	Concentration		Volume (μ l)	Comments
	Stock solution	Final		
Primer mix				
Primer stock solution	100.0 μ M	2.0 μ M	10.0	for each primer in the mix (see <i>Table 1</i>)
TE buffer			420.0	alternatively, ddH ₂ O can be used
PCR Master mix				
PCR Master mix	2x	1x	5.0	
Primer mix	2.0 μ M	0.2 μ M	1.0	
DNA	5.0 ng/ μ l	20.0 ng	4.0	

ware (Life Technologies) and Allelogram (MORIN et al., 2009).

To test for genotyping errors, PCRs for each set were repeated twice in 16 samples and compared. Additionally, 16 samples were isolated a second time, amplified again for all three sets and compared to the original genotypes. Finally, for each 96-well plate one control sample was used that was identical for all PCRs and fragment analysis runs. This allowed standardization of all samples for binning (MORIN et al., 2009) and served as a control to monitor consistency of amplification and fragment analysis.

Population genetic analysis

Estimates of Hardy-Weinberg equilibrium (HWE), observed and expected heterozygosities, probability of identity (P_{ID}), number of effective alleles (GenAEx 6; PEAKALL and SMOUSE, 2006), null alleles and linkage disequilibrium (LDE) (GenePop 4.0; ROUSSET, 2008) were calculated for every locus on the basis of genets (for clones only one tree per genotype was included in the calculation). P_{ID} was furthermore calculated over all loci. Calculations were carried out for each stand independently and over all stands.

Results and Discussion

Primer screening

Of the 15 primer pairs tested, 12 generated polymorphic products that were easily scored and amplified consistently in multiple separate PCR reactions. Using the original protocols, two primer pairs did not amplify at all (PGL7, PGL13) and one primer pair produced unreadable products with multiple signals (UAPgCT3). One of the 12 primer pairs did not amplify at an anneal-

ing temperature of 60°C in the multiplex PCR but produced clearly readable products at an annealing temperature of 57°C (PGL14). Between 7 and 58 alleles were detected in the remaining eleven loci. All of these eleven loci were selected for further characterization.

Locus PGL14 did not amplify at an annealing temperature of 60°C but produced good results at an annealing temperature of 57°C. Since the eleven loci characterized and presented here are adequate for population genetic studies, we chose to remove the locus from the multiplex assay. This enabled us to use only one PCR protocol for all three multiplex assays and therefore removes one potential source of error in the laboratory. However, if the resolution of the eleven loci seems insufficient for a specific problem, locus PGL14 can be added to assay 3 as shown in *Table 1*. The PCR can be performed as described with the exception of the annealing temperature, which has to be set to 57°C. Tests on a reduced set of samples showed that amplification quality of the other three loci in the assay does not suffer from a lower annealing temperature.

Population genetic analysis

Locus characteristics are shown in *Table 1*. The test for Hardy-Weinberg equilibrium and pairwise tests for linkage disequilibrium revealed significant deviation from HWE for all loci except UAPgAC/AT6 and UAPgCT144 ($p < 0.05$). Significant linkage disequilibrium was found for 31 of 55 locus combinations ($p < 0.05$). P_{ID} computed over all loci was 1.25×10^{-15} . P_{ID} for siblings was 1.5×10^{-5} . Null allele frequencies were estimated close to 0 for three loci, but ranged from 0.1 to 0.28 for the remaining eight loci (*Table 1*). All repeat PCRs, repeat DNA isolation and PCR, and control samples consistently produced identical genotypes.

For any genotyping method, the quality of the marker system depends on its ability to reliably discern different individuals. The very low P_{ID} value of the eleven loci presented here indicates that this requirement is fully met. The P_{ID} for siblings is naturally higher, since siblings draw their alleles not from the complete set present in the population but from the same two parents and therefore from a highly reduced gene pool. Nonetheless, the P_{ID} for siblings of the presented marker set can be regarded sufficient to distinguish even full siblings with high confidence (HOFFMAN and AMOS, 2005).

Null alleles have a negligible effect on fingerprinting studies and low impacts both on assignment testing (CARLSSON, 2008) and differentiation estimators as long as null allele frequencies are low to intermediate (<0.20 ; CHAPUIS and ESTOUP, 2007) but may be problematic in parentage analysis due to false exclusion of true parents (DAKIN and AVISE, 2004). Estimated null allele frequencies for our marker set are mostly in the low and intermediate range as defined by CHAPUIS and ESTOUP (2007) and well within the range of frequencies typically reported in microsatellite studies (DAKIN and AVISE, 2004; JARNE and LAGODA, 1996). However, only loci PGL15, UAPgCA24, and UAPgTG64 showed indication of null alleles in a visual inspection of the data. These loci showed missing amplification products when all other loci in the same multiplex assay amplified successfully, therefore ruling out PCR failure. These missing amplification products indicate that the respective tree is homozygous for the null allele at this locus. In the remaining loci, null allele frequencies might be overestimated caused by an observed homozygote excess, indicated by the difference of expected and observed heterozygosity in most loci (Table 1). An excess of homozygotes can cause an increase in estimated null allele frequencies as these tests are based on homozygote excess patterns (DAKIN and AVISE, 2004; GUICHOUX et al., 2011). Apart from null alleles, homozygote excess can be caused by biological factors, e.g. deviation from panmixia, inbreeding or Wahlund effects, or the organism's mating system (JARNE and LAGODA, 1996; OOSTERHOUT et al., 2006; SELKOE and TOONEN, 2006). Still, while null allele frequencies might be overestimated, the occurrence of null alleles cannot be ruled out. If parentage assignment is attempted, care should be taken to choose statistics and software packages that are able to account for null alleles and genotyping errors (DAKIN and AVISE, 2004; JONES et al., 2010; JONES and WANG, 2010; KOCH et al., 2008). Otherwise exclusion of true parents will lead to an underestimation of parent-offspring combinations within the dataset. However, while null alleles can lead to false exclusion of true parents, they will not cause incorrect parent-offspring combinations (DAKIN and AVISE, 2004).

The set of eleven microsatellite loci presented here was demonstrated to reliably identify genotypes in natural populations and did not produce visible errors in 1470 tested plants. Repeat PCRs consistently resulted in identical genotypes, thereby confirming the system's reliability. We show the system to be sufficient for analyses of population genetics as well as parentage assignment, provided that the occurrence of null alleles is

taken into account. Used in large scale studies, multiplex assays can considerably reduce processing time and per sample cost. In contrast to the more traditional single PCR reaction for each locus, the three assays saved a total of almost 12 000 PCR reactions when used on the sample set of 1470 plants described in this paper.

Funding

This study was funded by the German Research Foundation (DFG), grant number EU 132/1-1.

Acknowledgments

Thanks are due to the staff at Denali National Park, CARL ROLAND and LUCY TYRRELL, as well as to MICHAL GAŽOVIČ for logistic help during the field season.

References

- CARLSSON, J. (2008): Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* **99**: 616–623.
- CHAPUIS, M. P. and A. ESTOUP (2007): Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**: 621–631.
- DAKIN, E. E. and J. C. AVISE (2004): Microsatellite null alleles in parentage analysis. *Heredity* **93**: 504–509.
- GUICHOUX, E., L. LAGACHE, S. WAGNER, P. CHAUMEIL, P. LÉGER, O. LEPAIS, C. LEPOITTEVIN et al. (2011): Current trends in microsatellite genotyping. *Molecular Ecology Resources* **11**: 591–611.
- HODGETTS, R. B., M. A. ALEKSIUK, A. BROWN, C. CLARKE, E. MACDONALD, S. NADEEM and D. KHASA (2001): Development of microsatellite markers for white spruce (*Picea glauca*) and related species. *Theoretical and Applied Genetics* **102**: 1252–1258.
- HOFFMAN, J. I. and W. AMOS (2005): Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* **14**: 599–612.
- JARNE, P. and P. J. L. LAGODA (1996): Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution* **11**: 424–429.
- JONES, A. G., C. M. SMALL, K. A. PACZOLT and N. L. RATTERMAN (2010): A practical guide to methods of parentage analysis. *Molecular Ecology Resources* **10**: 6–30.
- JONES, O. R. and J. WANG (2010): COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**: 551–555.
- KOCH, M., J. D. HADFIELD, K. M. SEFC and C. STURMBAUER (2008): Pedigree reconstruction in wild fish populations. *Molecular Ecology* **17**: 4500–4511.
- LIESEBACH, H. and E. EWALD (2012): Optimisation of a multiplex PCR assay of nuclear microsatellite markers for population genetics and clone identification in *Robinia pseudoacacia* L. *Silvae Genetica* **61**, 142–148.
- MORIN, P. A., C. MANASTER, S. L. MESNICK and R. HOLLAND (2009): Normalization and binning of historical and multi-source microsatellite data: overcoming the problems of allele size shift with Allelogram. *Molecular Ecology Resources* **9**: 1451–1455.
- OOSTERHOUT, C. VAN, D. WEETMAN and W. F. HUTCHINSON (2006): Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes* **6**: 255–256.