SEDGLEY, M., J. HARBARD, R. M. M. SMITH, R. WICKNESWARI and A. R. GRIFFIN (1992): Reproductive biology and interspecific hybridization of *Acacia mangium* and *Acacia auriculiformis* A. Cunn. ex Benth (Leguminosae: Mimosoideae). Australian Journal of Botany **40**: 37–48.

SHIRAISHI, S and A. WATANABE (1995): Identification of chloroplast genome between *Pinus densiflora* Sieb. *et* Zucc and *P. thunbergii* Parl. based on polymorphisms in *rbc*L gene. Journal of Japanese Forest Society **77**: 429–436. (in Japanese with English abstracts)

TURNBULL, J. W. (1986): Multipurpose Australia trees and shrubs: lesser known species for fuel wood and agroforestry. ACIAR Monograph No. 1, pp316.

VALLONE, P. M., R. S. JUST, M. D. COBLE, J. M. BUTLER and T. J. PARSONS (2004): A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome. Int. J. Legal Med **118**: 147–157.

VALLONE, P. M and J. M. BUTLER (2004): Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension. Journal of Forensic Sciences **49**: 723–732.

VALLONE, P. M, A. E. DECKER, M. D. COBLE and J. M. BUTLER (2006): The evaluation of an autosomal SNP 12-plex assay. International Congress Series **1288**: 61–63.

VARSHNEY, R. K., U. BEIER, E. K. KHLESTKINA, R. KOTA, V. KORZUN, A. GRANER and A. BORNER (2007): Single nucleotide polymorphisms in rye (*Secale cereale* L.): discovery, frequency, and applications for genome mapping and diversity studies. Theoretical Applied Genetics **114**: 1105–1116.

VIGNAL, A., D. MILAN, M. SANCRISTOBAL and A. EGGEN (2002): A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution **34**: 275–305.

WENG, C., T. L. KUBISIAK and M. STINE (1998): SCAR markers in a longleaf pine x slash pine F1 family. Forest Genetics **5**: 239–247.

WERNER, F. A. O., G. DURSTEWITZ, F. A. HABERMANN, G. THALLER, W. KRAMER, S. KOLLERS, J. BUITKAMP, M. GEORGES, G. BREM, J. MOSNER and R. FRIES (2004): Detection and characterization of SNPs useful for identity control and parentage testing in major European dairy breeds. Animal Genetics **35**: 44–49.

WILLIAM, J. G., A. R. KUBELIK, K. J. LIVAK, J. A. RAFALSKI and S. V. TINGEY (1990): DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research **18**: 6531–6535.

YOON, M. S., Q. J. SONG, I. Y. CHOI, J. E. SPECHT, D. L. HYTEN and P. B. CREGAN (2007): BARCSoySNP23: a panel of 23 selected SNPs for soybean cultivar identification. Theoretical Applied Genetics **114**: 885–899.

Short Note: Seven Genomic SSRs Revealed in *Eucalyptus* by Re-sequencing of DNA Sequences from GenBank

By X. HE[1),2)], F. LI[1)], J. SHI[2)] and S. GAN[1),*)]

(Received 28th January 2010)

Abstract

Seven genomic SSR markers of *Eucalyptus* were developed from DNA sequences of *E. grandis* deposited in GenBank. Their repeat motifs were revealed by re-sequencing with an individual tree of *E. urophylla* or *E. tereticornis*, and five out of the seven markers turned out to be heterozygous within the specific tree sequenced. The sequence identity ranged from 75.06% to 96.66%, with an average of 87.31%. These markers could be valuable in genetics studies in *Eucalyptus*. This report demonstrates the advantages of re-sequencing in developing SSR markers from publicly accessible databases.

Key words: Simple sequence repeats (SSRs), microsatellites, molecular markers, *Eucalyptus*, re-sequencing.

Simple sequence repeats (SSRs), as known as microsatellites, are useful markers for a wide spectrum of biological applications (POWELL et al., 1996; GUPTA and VARSHNEY, 2000). To date, a large number of SSR markers have been produced for a number of plant species, e.g. more than 5000 SSRs available for sorghum (*Sorghum bicolor*) (YONEMARU et al., 2009). In the woody genus *Eucalyptus* (family Myrtaceae), however, only 367 genomic SSRs (as reviewed in BRONDANI et al., 2006) and 68 EST-SSRs (FARIA et al., 2010, 2011; ZHOU et al.,

[1)] Research Institute of Tropical Forestry and National Key Laboratory of Forest Tree Genetics and Breeding, Chinese Academy of Forestry, Longdong, Guangzhou 510520, People's Republic of China.

[2)] The Key Laboratory of Forest Genetics and Biotechnology of Ministry of Education, Nanjing Forestry University, 159 Longpan Road, Nanjing 210037, People's Republic of China.

[*)] Corresponding author: SIMING GAN. Tel: +86 20 87032402, Fax: +86 20 87031622. E-mail: Siming.Gan@ritf.ac.cn.

DOI:10.1515/sg-2011-0013
edited by Thünen Institute of Forest Genetics

2010) have been reported, and the SSR marker resources are still limited. Thus, it may be necessary to develop more eucalypt SSR markers, especially from publicly accessible genomic databases, such as GenBank.

In this study we explore the possibility of developing SSR markers in *Eucalyptus* from publicly available DNA sequences. A total of 167 *E. grandis* DNA fragments with three definitions (*Table 1*) were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/nucleotide/), 47 of which were identified to contain tandem repeats using a Perl script Microsatellite Identification Tool (MISA, http://pgrc.ipk-gatersleben.de/misa), in which a minimum of 12, 12, 12, 15 and 18 bases were set for di-, tri-, tetra-, penta- and hexa-nucleotide repeats, respectively. After excluding the fragments with too short or inappropriate flanking sequences, thirty-three primer pairs were designed using Primer 3 (ROZEN and SKALETSKY, 2000) and synthesized by Invitrogen Co. (Shanghai, China).

PCR was conducted using the maternal *E. urophylla* (P_1, UX-30) and paternal *E. tereticornis* (P_2, T4305) parents of a mapping population (GAN et al., 2003). PCRs of 10 µL consisted of 1.0 µL 10x buffer (100 mM Tris-HCl pH9.0, 100 mM KCl, 80 mM $(NH_4)_2SO_4$ and 0.5% NP-40), 200 µM each dNTP, 2.0 mM $MgCl_2$, 0.5 µM forward primer, 0.5 µM reverse primer, 1 U *Taq* DNA polymerase (Biocolors Technology Co., Shanghai, China) and about 5 ng DNA template. The PCRs were done in 96-well plates on a DNA Engine thermal cycler (Bio-Rad, Hercules, CA, USA) using the following program: 94°C for 4 min; 35 cycles of 94°C for 30 s, 56°C for 30 s, and

Table 1. – Seven genomic SSR markers developed in *Eucalyptus*. P_1, *E. urophylla*; P_2, *E. tereticornis*.

Marker	GenBank accession (ProbeDB ID)	Repeat motif revealed in re-sequencing [d]	Repeat motif in original DNA sequences	Forward primer (5'-3') Reverse primer (5'-3')	Expected length (bp)	Re-sequenced length (bp)	Identity
EUCgSSR01	BD272940[a] (10703611)	P_2: $(GA)_{21/25}$	$(GA)_{22}$	F: ACGCCACATTGGACCTTT R: CCGAGCTGCTGGATAACG	333	294/302	76.82% (232/302)
EUCgSSR05	BD262196[b] (10703612)	P_1: $(CT)_{12/13}$	$(CT)_{14}$	F: ACAAAGTGAAACCCAAGC R: GGAAGAGCAGGACCAGTA	462	431/433	75.06% (325/433)
EUCgSSR20	DD458136[c] (10703613)	P_1: $(AG)_n$+$(GAGCGA)_3$	$(AG)_6$+$(GAGCGA)_3$	F: ACTGAAGTGACTGACTGATG R: GTCGGAGTCGTGGCGTAT	334	299	96.66% (289/299)
EUCgSSR21	DD458137[c] (10703614)	P_1: $(TCT)_n(TCG)_9$	$(TCT)_6(TCG)_9$	F: TCGCCTCGTTCGTCATCTT R: CGCTGCTTGTTCCACCATT	223	192	96.35% (185/192)
EUCgSSR23	DD458141[c] (10703615)	P_1: $(CT)_{8/9}$	$(CT)_9$	F: CCCCGTATCACTCATCTCC R: TTTAGCCGAGTCCCAGAA	264	228/230	80.43% (185/230)
EUCgSSR25	BD262179[b] (10703616)	P_1: $(GA)_{13/14}$	$(GA)_{18}$	F: ATCACATCCATCCAGCCTCCAC R: CCGCCATCACCTTCCAAGAC	232	183/185	92.97% (172/185)
EUCgSSR26	BD262120[b] (10703617)	P_2: $(GA)_{>4/5}$	$(GA)_{17}$	F: AAAGGTCAGATGTCATCCCACG R: ACCAACCCTCCCGCAAAA	153	110/112	92.86% (104/112)

[a] Definition: Materials and methods for the modification of isoprenoid content, compostition and metabolism.

[b] Definition: Composition and methods for the modification of gene expression.

[c] Definition: Compositions and methods for regulating polysaccharides of a plant cell.

[d] The symbols "n" and ">" denote no and partial repeats, respectively, visible in the re-sequencing profile as the SSR motifs were included in (e.g. four bases for EUCgSSR21) or very close to (e.g. three bases for EUCgSSR20 and zero base for EUCgSSR26) the end of the forward sequencing primer site.

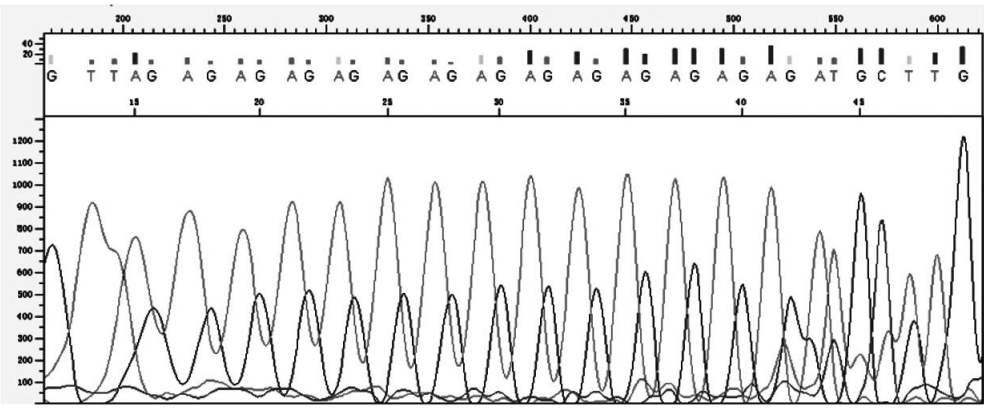

Figure 1. – Partial re-sequencing profile of marker EUCgSSR25 with P_1. The higher and lower bases correspond to the larger (185 bp) and shorter (183 bp) alleles sequenced, respectively. The underlined letters show the identical bases downstream of the SSR motifs.

DOI:10.1515/sg-2011-0013
edited by Thünen Institute of Forest Genetics

72 °C for 1 min; and a final extension at 72 °C for 10 min. The PCR products were checked through electrophoresis in 1.5% agarose gels containing 1 : 20 GoldView (SBS Genetech Co., Beijing, China) and photographed with Photoprint 215SD (Vilber Lourmat Co., Marne la Vallée, France).

PCR products of either P_1 or P_2 were sequenced using BigDye Terminator Version 3.1 (BDT3.1) and an ABI 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions except that 0.5 µL of BDT3.1 was used in the sequencing reaction instead of the standard 8.0 µL (ZHANG et al., 2009). Seven SSRs were verified in re-sequencing, including five that were heterozygous within the specific tree that was sequenced (*Table 1*). *Figure 1* shows a partial re-sequencing profile of marker EUCgSSR25 with P_1.

The nucleotide sequences generated in re-sequencing were aligned with the target sequences using DNAMAN version 5.2.2 (Lynnon Biosoft, Quebec, Canada). The identity ranged from 75.06% in EUCgSSR05 to 96.66% in EUCgSSR20 (*Table 1*), with an average of 87.31%. BLASTN searches within the GenBank database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) revealed that EUCgSSR01 was highly similar to an EST-SSR Embra1868 (FARIA et al., 2010) while there were no similar markers for the rest six SSRs. The seven SSRs were deposited in ProbeDB of GenBank with IDs 10703611-10703617 (*Table 1*; http://www.ncbi.nlm.nih.gov/sites/entrez?db=probe).

The set of seven SSR markers reported here will make a valuable addition to the growing collection of SSR markers for genetic studies in *Eucalyptus*. This report demonstrates several advantages of re-sequencing in developing SSR markers from publicly accessible databases: (i) relatively low cost and high efficiency; (ii) direct identification of heterozygous markers, which may be highly variable; and (iii) potential to identify other variations within the flanking sequences, e.g. single nucleotide polymorphism (SNP) and insertion/deletion (indel).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 31070592), the Guangdong Natural Science Foundation (No. 10151052001000000), National Program of High Technology Development of China (No. 2011AA100202) and the Ministry of Finance of China through a specific program for national non-profit scientific institutions (Nos. RITF2007-6 and RITF2008-4). We thank YU WANG for her technical assistance.

References

BRONDANI, R. P. V., E. R. WILLIAMS, C. BRONDANI and D. GRATTAPAGLIA (2006): A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. BMC Plant Biol. **6**: 20.

FARIA, D. A., E. M. C. MAMANI, G. J. PAPPAS, D. GRATTAPAGLIA (2011): Genotyping systems for *Eucalyptus* based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and their use for individual fingerprinting and assignment tests. Tree Genet. Genomes **7**: 63–77.

FARIA, D. A., E. M. C. MAMANI, M. R. PAPPAS, G. J. PAPPAS JR. and D. GRATTAPAGLIA (2010): A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of *Eucalyptus*. J. Hered. **101**: 512–520.

GAN, S., J. SHI, M. LI, K. WU, J. WU and J. BAI (2003): Moderate-density molecular maps of *Eucalyptus urophylla* S. T. Blake and *E. tereticornis* Smith genomes based on RAPD markers. Genetica **118**: 59–67.

GUPTA, P. K. and R. K. VARSHNEY (2000): The development and use of microsatellite markers for genetic analysis and plant breeding with special emphasis on bread wheat. Euphytica **113**: 163–185.

POWELL, W., G. C. MACHRAY and J. PROVAN (1996): Polymorphisms revealed by simple sequence repeats. Trend. Plant Sci. **1**: 215–222.

ROZEN, S. and H. J. SKALETSKY (2000): Primer3 on the WWW for general users and for biologist programmers. *In:* Bioinformatics methods and protocols: methods in molecular biology, eds KRAWETZ, S. and S. MISENER. Humana Press, Totowa, NJ. 365–386.

YONEMARU, J.-I., A. TSUYU, M. TATSUMI, K. SHIGEMITSU, M. TAKASHI and Y. MASAHIRO (2009): Development of genome-wide simple sequence repeat markers using whole-genome shotgun sequences of sorghum (*Sorghum bicolor* (L.) Moench). DNA Res. **16**: 187–193.

ZHANG, X., F. LI, Y. WANG, L. XU, M. LI and S. GAN (2009): An optimized protocol for sequencing EST-PCR products in *Eucalyptus*. Genomics Appl. Biol. **28**: 535–543.

ZHOU, C., F. LI, Q. WENG, X. YU, M. LI and S. GAN (2010): Comparison between direct sequencing and pool-cloning-based sequencing of PCR products in EST-SSR marker development in Eucalyptus. Mol. Plant Breed. **8**: 1 (doi:10.5376/mpb.cn.2010.08.0001).

DOI:10.1515/sg-2011-0013
edited by Thünen Institute of Forest Genetics