

Microsatellite-based Genotyping of the Commercial *Eucalyptus* Clones Cultivated in China

By F. LI¹, S. GAN^{1,*}, Z. ZHANG², Q. WENG¹, D. XIANG² and M. LI¹

(Received 12th January 2011)

Abstract

A proper identification of clones is necessary in clonal forestry and will help to protect the legitimate interests of breeders, growers and industry. Twenty-four of the *Eucalyptus* clones most widely cultivated in China were analyzed using a set of 24 microsatellite markers to develop their DNA-based fingerprints and exploit the genetic variations. A total of 286 alleles were detected, averaging at 11.9 alleles per marker locus. All the microsatellites were polymorphic among the clones investigated. The observed heterozygosity (H_o) varied with locus between 0.500 and 1.000 with a mean of 0.885. The 24 clones could be uniquely fingerprinted based on their multilocus genotypes at a minimum of three loci (Embra169, Embra72 and Embra2). The dendrogram constructed from the genotypic similarity coefficients separated the 24 clones into three groups, matching essentially the historically known or speculated clonal origins. Clones T13, Guanglin-5 and Guanglin-9 turned out to be full siblings of cross DH32 while the DH201-2 sampled here appeared to be mislabelled.

Key words: Microsatellite, *Eucalyptus* clone, genotyping, genetic variation.

Introduction

Trees of the genus *Eucalyptus* have been widely planted in the tropical, sub-tropical and temperate climatic regions around the world, with a global total of plantations greater than 17.8 million ha (FAO, 2000). They combine the favorable features of versatile wood properties, fast growth and short rotations, and serve as an important wood source for a large spectrum of industrial products, including furniture, plywood, fibreboard and pulp and paper (ELDRIDGE et al., 1993; QI, 2002). China boasts an area of eucalypt plantations of more than 2.6 million ha (<http://www.chinaeuc.com/show.asp?id=171>), mainly distributed in the southern provinces Guangdong, Guangxi and Hainan, and ranks as the third planter following India (8.0 million ha) and Brazil (3.0 million ha) (FAO, 2000). Moreover, there is an increasing interest in expanding eucalypts to the substantial area in south-central provinces of the country (approx. 25°–30°N), where the established conifer plantations have generally been unsuccessful but huge demands exist for wood and wood products (ARNOLD et al., 2004).

Clonal forestry, in a strict sense, refers to use of a number of tested, selected and identified clones in plantation forestry (ELDRIDGE et al., 1993), in which a clone is typically the basic unit of cultivars representing the original individual tree (ortet) and its vegetative propagules (ramets). Clonal forestry has been largely preferred for those species amenable to mass vegetative propagation as it offers the most effective approach for capturing genetic gains (LIBBY, 1985; GRIFFIN, 2001). Many *Eucalyptus* species and hybrids could be readily macro-(rooted cuttings) and/or micro-propagated (tissue culture), thereby enabling their utility for large-scale clonal cultivation (ELDRIDGE et al., 1993; WATT et al., 1995; QI, 2002). In recent two decades, significant productivity and uniformity gains have been achieved in eucalypts from intensively cultivated clonal plantations. In China, for example, mean annual increment (MAI) could amount to 45 m³/ha for eucalypt clonal plantations, a dramatic increase over ca. 10 m³/ha of the seedling plantations of the 1970s (QI, 2002).

However, incorrect labelling of clones may occur in the process of conservation, propagation, cultivation and/or exchange, which, if unidentified, could limit the efficiency in further breeding and cause lower productivity than expected from plantations (CLARKE, 2001; DE-LUCAS et al., 2008). In *Eucalyptus*, for instance, KEIL and GRIFFIN (1994) found that mislabelling or mis-sampling of clones was considerable during the handling of plant material, suggesting the potentially serious economic consequences. Therefore, a proper identification of clones is necessary in forestry operations with clonally oriented trees, including eucalypts. Also, such an identification will facilitate commercial registration of clones and help to protect the legitimate interests of breeders, growers and industry (RAHMAN and RAJORA, 2002; FOSSATI et al., 2005). Further, this would allow for characterizing the genetic diversity of the clonal germplasm, reducing the risks involved with clonal deployment and promoting the efficiency of a breeding program (PATERSON et al., 1991).

DNA-based molecular markers, compared to phenotypic and phenological characters, are more accurate, reliable, stable and objective for clonal identification. In previous reports with eucalypts, a variety of molecular marker techniques have been applied towards clone/genotype identification and genetic diversity analysis, including random amplified polymorphic DNA (RAPD; KEIL and GRIFFIN, 1994; GAN et al., 1999; MURO-ABAD et al., 2001; ROCHA et al., 2002; TRIPATHI et al., 2006), amplified fragment length polymorphism (AFLP; TRIPATHI et al., 2006), and microsatellites or simple sequence repeats (SSRs; ROCHA et al., 2002; GRATTA-

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) Guangxi Forestry Research Institute, 23 Yongwu Road, Nanning 530001, People's Republic of China.

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

PAGLIA et al., 2004). Of these, microsatellites are stretches of a variable number of tandem repeats with a core unit of two to six nucleotides (TAUTZ, 1989; SCHLÖTTERER, 2000). They are excellent markers for genetic characterization of plant material due to their co-dominant, multi-allelic, reproducible and highly polymorphic nature as well as abundant distribution within the genome (POWELL et al., 1996; VARSHNEY et al., 2005). So far, besides 35 chloroplast microsatellites (STEANE et al., 2005), a total of 374 genomic (see BRONDANI et al., 2006; HE et al., 2011) and 68 genic microsatellites (FARIA et al., 2010, 2011; ZHOU et al., 2010) have been reported for *Eucalyptus* species, enhancing significantly the capacity of molecular characterization of eucalypt germplasm. Additionally, in light of the rapid accumulation of eucalypt expressed sequence tags (ESTs), a large number of EST-derived microsatellites may be available in near future (RABELLO et al., 2005; YASODHA et al., 2008).

There are more than 20 commercial eucalypt clones being cultivated commonly in China. In usual cases, identification of these clones is based on phenotypic and phenological characters, which are mostly subject to environmental influences and thereby make the identification difficult and ambiguous, especially for those genetically related clones displaying similar morphological characteristics. Also, the genetic diversity and relationships of these clones have been little explored. In this study, we used 24 microsatellite markers to geno-

type 24 *Eucalyptus* clones cultivated widely in China and explore their genetic variations. The objectives were to develop the DNA-based fingerprint of each clone and determine the genetic relationship of these clones.

Materials and Methods

Plant material

A total of 24 commercial *Eucalyptus* clones cultivated widely in China were sampled from five state-owned nurseries located in provinces Guangdong, Guangxi and Fujian (Table 1). Each clone was represented by a ramet conserved in stock pools of these nurseries. The stock pools were maintained to supply explants for clonal rejuvenation when depression of a clone occurred after several generations of sub-cultures. Among the clones, 17 represented interspecific hybrids with known parental species, of which eight DH-prefixed clones were bred by National Dongmen Forest Farm with their parentage documented but their parental trees lost (*pers. comm.* GUOFU ZHOU). Specifically, the species *E. urophylla* involved in clonal taxonomic nomenclature denoted the synonym of earlier classification (PRYOR and JOHNSON, 1971), from which two new species, *E. orophila* Pryor and *E. wetarensis* Pryor, were subsequently subdivided (PRYOR et al., 1995). However, it was not possible to assign the *E. urophylla* clones to the new nomenclature due to lack of documentation of their origins.

Table 1. – The 24 commercial eucalypt clones sampled in this study. NYFF, National Yanxi Forest Farm, Fujian Province. YFJ, Yong'an Forestry (Group) Joint-stock Co. Ltd, Fujian Province. RITF, Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangdong Province. NLFB, National Leizhou Forest Bureau, Guangdong Province. NBFF, National Bobai Forest Farm, Guangxi Province.

Number	Clone code	Provider	Species/Hybrid (documented parentage if available)
01	T13	NYFF	<i>E. urophylla</i> S. T. Blake × <i>E. grandis</i> Hill ex Maid.
02	DH32-29	NYFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G46)
03	D29	NYFF	<i>E. urophylla</i>
04	184	YFJ	<i>E. urophylla</i> × <i>E. camaldulensis</i> Dehnh.
05	DH201-2	YFJ	<i>E. grandis</i> × <i>E. camaldulensis</i> (G46 × TM)
06	M1	YFJ	<i>E. urophylla</i> × <i>E. tereticornis</i> Smith
07	BTT10	YFJ	<i>E. urophylla</i> × <i>E. grandis</i>
08	G1	YFJ	<i>E. grandis</i>
09	TI19224	RITF	<i>E. urophylla</i> × <i>E. camaldulensis</i>
10	Guanglin-9	RITF	<i>E. urophylla</i> × <i>E. grandis</i>
11	Guangzhou-1	RITF	<i>E. urophylla</i>
12	U6	RITF	<i>E. urophylla</i> or its natural hybrid
13	Lei-7	NLFB	<i>E. grandis</i> × <i>E. camaldulensis</i>
14	LH88	NLFB	<i>E. urophylla</i>
15	SH1	NLFB	<i>E. exserta</i> F. Muell. or its natural hybrid
16	LH1	NLFB	<i>E. urophylla</i> × <i>E. tereticornis</i>
17	DH32-13	NBFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G46)
18	DII32-28	NBFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G46)
19	DII32-26	NBFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G46)
20	DII33-27	NBFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G33)
21	DII30-1	NBFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G22)
22	DII32-22	NBFF	<i>E. urophylla</i> × <i>E. grandis</i> (U16 × G46)
23	Guanglin-5	NBFF	<i>E. urophylla</i> × <i>E. grandis</i>
24	Guanglin-4	NBFF	<i>E. urophylla</i>

DNA extraction and SSR genotyping

DNA extraction was made with 1 g of young leaves per clone, using a CTAB procedure (DOYLE and DOYLE, 1990) modified by adding 5% polyvinylpyrrolidone (PVP) and 2% β -mercaptoethanol to the extraction buffer (GAN et al., 2003). DNA concentration was determined by electrophoresis in 1.0% agarose gels containing 1:20 GoldView (SBS Genetech Co., Beijing, China) with comparisons to 100 bp standard DNA ladder (Fermentas International Inc., Burlington, Canada).

Twenty-four genomic SSR markers described in BRONDANI et al. (2006) were selected for clonal genotyping (Table 2). All the SSRs contained dinucleotide repeats except Embra70 that harbored a motif of (CATA)₉CAT(AC)₁₇, and were considered to be independent since six of them were unmapped and the rest 18 were sparsely mapped onto 10 linkage groups (BRONDANI et al., 2006). Primer pairs were synthesized by Sangon Co. (Shanghai, China). PCR was carried out following an optimized protocol for fluorescent-dUTP based SSR genotyping (LI and GAN, 2011). PCR reaction of 10 μ L consisted of 1.0 μ L 10 \times buffer (100 mM Tris-HCl pH 9.0, 100 mM KCl, 80 mM (NH₄)₂SO₄ and 0.5% NP-40), 25 μ M each dNTP, MgCl₂ at 1.5, 2.0 or 2.5 mM (Table 2),

0.5 μ M forward primer, 0.5 μ M reverse primer, 1 U *Taq* DNA polymerase, 10 pmol fluorescent-dUTP (Fermentas International Inc.) and about 5 ng DNA template. The reaction was amplified using a DNA Engine thermal cycler (Bio-Rad, Hercules, CA, USA) under touchdown conditions: 94°C for 4 min; 20 cycles of 94°C for 30 s, 70–60°C or 66–56°C depending on microsatellite marker (Table 2) for 30 s with a decrease of 0.5°C per cycle and 72°C for 1 min; 26 cycles of 94°C for 30 s, 60 or 56°C (Table 2) for 30 s and 72°C for 1 min; and a final extension at 72°C for 10 min.

The PCR products (5 μ L) were first checked through electrophoresis in 1.5% agarose gels containing 1:20 GoldView (SBS Genetech Co.) and photographed with Photoprint 215SD (Vilber Lourmat Co., Marne la Vallée, France). SSR genotyping was then performed on an ABI 3130xl genetic analyzer (Applied Biosystems, Foster City, CA, USA). The PCR products (1 μ L) were diluted 1:10.5 with loading buffer [9.34 μ L deionized formamide and 0.16 μ L internal standard GeneScan 500LIZ (Applied Biosystems)] and then denatured at 95°C for 5 min followed by rapid cooling on ice. The genotyping procedure was followed according to the standard module of the manufacturer using software GeneMapper 4.0 (Applied Biosystems). To assure precise and consistent

Table 2. – Microsatellite marker loci and their statistic estimates based on the 24 eucalypt clones. For each microsatellite locus, repeat motif, primer sequences and GenBank accession were as described in BRONDANI et al. (2006). Mg²⁺ concentration in PCR reaction and annealing temperature (T_m) in PCR program were optimized specifically and might differ from those of BRONDANI et al. (2006). *ASR*, allele size range. *Ho*, observed heterozygosity. *NA*, number of alleles. *NG*, number of observed genotypes. *NC*, number of clones with unique genotype. *ENA*, effective number of alleles. *PI*, probability of identity. *PE*, paternity exclusion probability. *PIC*, polymorphism information content.

Marker	Mg ²⁺ (mM)	T_m (°C)	<i>ASR</i> (bp)	<i>Ho</i>	<i>NA</i>	<i>NG</i>	<i>NC</i>	<i>ENA</i>	<i>PI</i>	<i>PE</i>	<i>PIC</i>
Embra2	2.5	56	104–144	0.917	13	13	8	1.926	0.065	0.711	0.835
Embra38	2.0	60	91–135	0.958	12	11	8	1.536	0.094	0.663	0.806
Embra70	1.5	60	130–178	0.958	15	14	10	2.667	0.042	0.776	0.877
Embra72	2.0	56	129–155	0.667	11	10	4	1.852	0.126	0.590	0.747
Embra91	2.5	56	110–154	0.500	16	12	9	2.420	0.123	0.542	0.694
Embra111	1.5	60	95–151	1.000	14	16	13	2.165	0.056	0.739	0.854
Embra115	2.0	56	100–124	0.500	9	10	5	2.014	0.158	0.495	0.659
Embra116	2.0	56	118–156	1.000	9	10	2	2.141	0.075	0.701	0.834
Embra122	2.0	56	212–262	1.000	14	10	7	1.674	0.067	0.712	0.836
Embra139	2.0	56	170–210	1.000	14	12	8	2.250	0.058	0.736	0.854
Embra147	2.0	56	168–216	0.913	12	11	7	2.351	0.053	0.749	0.863
Embra153	2.0	56	181–225	0.917	14	15	11	1.817	0.060	0.723	0.842
Embra154	2.0	56	213–267	1.000	15	14	10	1.485	0.083	0.679	0.814
Embra156	2.0	56	98–130	0.958	9	8	3	1.154	0.143	0.572	0.738
Embra169	2.0	56	103–161	0.792	18	17	14	2.810	0.048	0.741	0.851
Embra179	2.0	56	124–146	0.833	6	8	2	1.260	0.225	0.473	0.666
Embra221	2.0	56	210–254	0.958	11	13	8	1.939	0.079	0.688	0.821
Embra227	2.0	56	294–322	1.000	11	14	10	2.224	0.085	0.679	0.817
Embra233	2.0	56	150–178	0.750	5	7	3	1.241	0.289	0.388	0.586
Embra286	2.0	56	131–161	0.875	11	12	8	1.085	0.132	0.552	0.710
Embra324	2.0	56	114–152	0.917	8	10	6	1.381	0.206	0.486	0.671
Embra333	2.0	56	212–256	0.875	14	15	12	2.233	0.061	0.727	0.846
Embra345	2.0	56	201–255	0.958	16	13	10	2.304	0.050	0.743	0.854
Embra350	2.0	56	201–235	0.958	9	8	4	1.340	0.117	0.619	0.777
Mean / Overall				0.885	11.9	11.8	7.6	1.885	0.104	0.646	0.786
Cumulative					286	283	182		7.063 $\times 10^{-16}$	0.99999	

scoring of the alleles, two replicates of PCR and genotyping were conducted for each marker across the clones.

Data analysis

For each locus, IDENTITY 1.0 (WAGNER and SEFC, 1999) was used to estimate the observed heterozygosity (H_o) (NEI, 1973), the number of alleles (NA), the probability of identity (PI) (PAETKAU et al., 1995) and the probability of paternity exclusion (PE) (WEIR, 1996). Also, overall PI and PE across all loci were calculated with this software. The number of observed genotypes (NG) and number of clones with unique genotype (NC) were computed empirically. The effective number of alleles (ENA) was estimated as $1/\sum p_i^2 = 1/(1-H_e)$, where p_i is frequency of allele i and H_e is the expected heterozygosity (KIMURA and CROW, 1964). The polymorphism information content (PIC) (BOTSTEIN et al., 1980) was calculated with the EXCEL-MICROSATELLITE-TOOLKIT version 3.1 (PARK, 2001). The genetic relationship among the clones was determined by calculating Jaccard's coefficients of similarity (JACCARD, 1908), and a dendrogram was then constructed with UPGMA method (unweighted pair-group method with arithmetic average; SOKAL and MICHENER, 1958) using the software NTSYS-pc 2.02 (ROHLF, 1998). Bootstrap analyses were performed with 1000 replicates.

Genetic relatedness between the DH-prefixed clones was verified based on their multilocus microsatellite data and documented parentage. As the five DH32-prefixed clones were derived from a control-pollinated cross DH32, whose maternal parent was shared by clones DH33-27 and DH30-1 and paternal parent served as the mother of DH201-2 (Table 1), the most likely genotypic combination of DH32 parents was first determined, and potential allelic mutation and clone mislabelling were checked by a sequential progeny exclusion procedure across all the microsatellite loci. Mislabelling was declared only when parental alleles were absent in an offspring clone for at least four independent loci to avoid false exclusions due to mutations at certain alleles (GRATTAPAGLIA et al., 2004). Similarly, other clones were investigated on their possible genetic relatedness with the DH32 series clones.

Results

Microsatellite polymorphisms

All of the 24 microsatellite marker loci were polymorphic among the 24 eucalypt clones analyzed. With two exceptions involving null alleles, namely, clone G1 at locus Embra116 and Lei-7 at Embra147, biallelic genotypes were observed for each of the 24 clones across all the SSR loci (data not presented). A total of 286 alleles were obtained (Table 2), which were all polymorphic among the clones. The number of alleles per locus ranged from five at Embra233 to 18 at Embra169 with an average of 11.9 (Table 2). The frequency at which an allele occurred was highly variable, ranging from 4.2% to 87.5%. All the SSRs amplified unique alleles that appeared only once in the whole set of clones, ranging from one (Embra115, Embra116 and Embra233) to 11

(Embra169) and reaching a total of 116 (40.6%). In addition, H_o , a measure of marker diversity, ranged from 0.500 at Embra115 to 1.000 at Embra111, Embra116, Embra122, Embra139, Embra154 and Embra227 across all the clones, with a mean value of 0.885 over all the 24 microsatellite loci.

Genotyping of clones

For each microsatellite locus, NG , NC , ENA , PI , PE and PIC are given in Table 2. These parameters reflected relatively consistently the discrimination power of a locus. The two loci Embra70 and Embra169 were the most informative as the former had the highest values for PIC (0.877) and PE (0.776) and the latter had the highest NG (17), NC (14) and ENA (2.810). By contrast, Embra233 was the least informative with the lowest PIC (0.586), NG (7) and PE (0.388) besides the highest PI (0.289). In addition, the cumulative PI , which was a measure of the probability in obtaining an identical genotype, was as low as 7.063×10^{-26} . This value corresponded to a cumulative PE of 0.99999, clearly indicating the extremely high power of the microsatellite markers for clonal identification.

The 24 microsatellite markers were successful in genotyping the eucalypt clones. All the primer pairs amplified clear and well-resolved fragment signals with little ambiguity, and all the clones were distinguishable with the combined genotypic profiles at the SSR loci. Consequently, the 24 clones could be uniquely fingerprinted using a minimum of three SSR loci (Embra169, Embra72 and Embra2).

Genetic relationship among clones

For each of the 24 microsatellite loci, the five DH32 series clones had collectively no more than four alleles, confirming their full-sib kinship. Except for clone DH201-2, whose microsatellite profiles deviated from the maternal alleles expected at 17 markers (data not presented), genetic relatedness between the DH-prefixed clones was verified based on their parentage and multilocus assignment analysis. DH33-27 and DH30-1 showed one and two mutated alleles, respectively (data not presented). The clone DH201-2 sampled here appeared to be mislabelled. Further, although clones T13, Guanglin-5 and Guanglin-9 were of uncertain origin, they appeared to share the alleles of DH32 clones and were assigned as full-siblings of the DH32 series clones, albeit the mutation in Guanglin-9 at an allele of Embra333. Consequently, the minimum number of mutated alleles was estimated to be four for all the clones, including one, two and one in clones DH33-27, DH30-1 and Guanglin-9, respectively, and accounting for 0.35% [$4/(24 \times 2 \times 24)$] across all the SSR loci.

Relatively high levels of genetic similarity were observed among all the clones, with Jaccard's coefficients of similarity ranging from 0.70 between clones SH1 and T13 or DH32-29 to 0.99 between T13 and Guanglin-5. An UPGMA dendrogram constructed from the similarity matrix separated at a coefficient of 0.755 the 24 clones into three groups (Figure 1), corresponding approximately to *E. urophylla* × *E. grandis* (Group I),

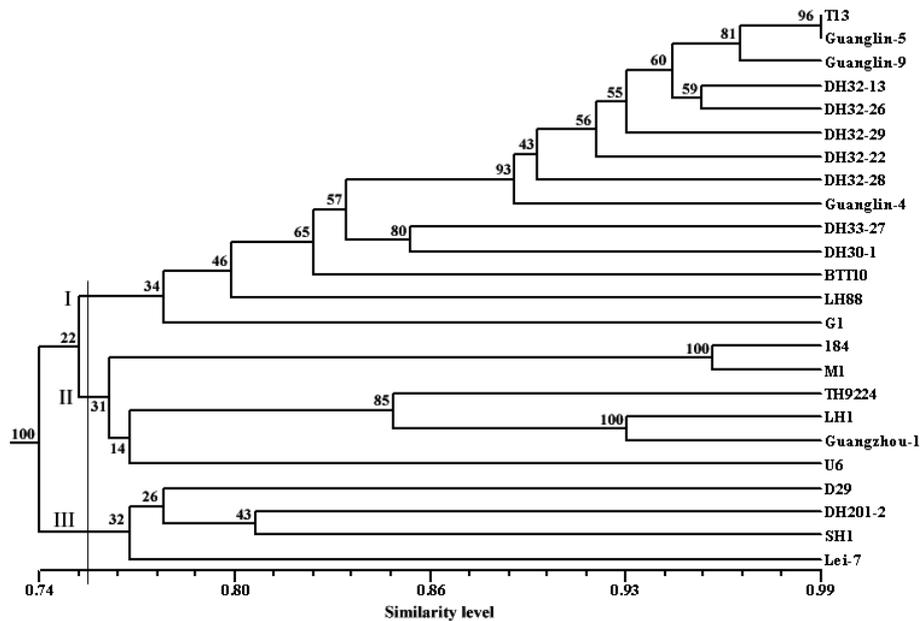


Figure 1. – UPGMA dendrogram of the 24 eucalypt clones analyzed with SSR markers.

E. urophylla × *E. tereticornis* & *E. urophylla* × *E. camaldulensis* (Group II) and *E. grandis* × *E. camaldulensis* (Group III). The clones of known coancestry, e.g. the five DH32 full-siblings and their half-siblings DH33-27 and DH30-1, were confirmed to share high genetic similarity and clustered closely. Together, the placement of clones T13, Guanglin-5 and Guanglin-9 with the DH32 series clones suggested their close relatedness, and Guanglin-4 was most likely a relative of the DH clones. Meanwhile, the three *E. urophylla* pure-species clones LH88, Guangzhou-1 and D29 scattered into different groups and clustered closely with *E. urophylla* × *E. grandis*, *E. urophylla* × *E. tereticornis* and *E. grandis* × *E. camaldulensis*, respectively.

Discussion

This study demonstrates clearly that the 24 eucalypt clones sampled could be accurately and rapidly fingerprinted using a set of microsatellite marker loci. Many of the clones could be uniquely genotyped based on their microsatellite profiles at a single locus without resorting to their multilocus configuration, and all of the clones be distinguished from each other at a minimum of three loci. Meanwhile, a possibly mislabelled case (DH201-2) was detected herein. The success in clone genotyping would allow future verification of clonal identity and the assessment of the correspondence between the reference and the declared or suspected identity of clones, and facilitate significantly the clonal registration and control during any step of vegetative propagation, plant stock handling and commercial exchanges (RAHMAN and RAJORA, 2002; FOSSATI et al., 2005).

This study also corroborates the usefulness of SSR markers in *Eucalyptus* clonal genotyping as demonstrated in ROCHA et al. (2002). All the markers resolved single-locus patterns and showed high allelic diversity in the eucalypt clones studied. However, instead of two loci

seen with EMBRA154 by BRONDANI et al. (2006), only one locus was observed in this study. Given that the SSRs originated from *E. grandis* and *E. urophylla* (BRONDANI et al., 1998, 2006), two species of which gave rise to the majority of clones analyzed here (Table 1), their high transferability is not unexpected. Further, all the five species involved in parentage of the 24 clones fall taxonomically into the same subgenus *Symphomyrtus* (PRYOR and JOHNSON, 1971), and eucalypt SSR markers have been revealed to be transferable among subgenera (STEANE et al., 2001; STOKOE et al., 2001; OTTEWELL et al., 2003) and even among genera (STEANE et al., 2001; OTTEWELL et al., 2003). This suggests that the genomic segments containing the SSR loci investigated, though polymorphic in their basic repeat number, are homologous among the parental species. Nevertheless, two out of the 24 SSR markers detected null alleles (data not presented). Null alleles are possibly the result of point mutations such as single nucleotide polymorphisms or other mutation events occurring at the priming sites residing in the SSR flanking regions, or could originate from large-scale genomic rearrangements (LEIGH et al., 2003).

The mean number of 11.9 alleles per locus obtained here is higher than that of 5.0 reported by ROCHA et al. (2002) for 15 *E. grandis* × *E. urophylla* clones using 16 SSR markers as well as 7.1 by GRATTAPAGLIA et al. (2004) for 256 progeny individuals derived from an *E. grandis* × *E. urophylla* seed orchard using 14 SSRs. However, the number is much lower than those reported for natural populations of most *Eucalyptus* species, e.g. 25 at 12 SSRs for 19 *E. urophylla* populations (PAYN et al., 2008). The discrepancy in these estimates could be attributed to several factors, including sample size, diversity within the samples and techniques for detection and estimation of fragment sizes (ORAGUZZI et al., 2005). Nevertheless, the mean H_o of 0.885 detected here is somewhat higher than those reported elsewhere, such

as 0.786 (GRATTAPAGLIA et al., 2004) and 0.686 (PAYN et al., 2008). The high level of observed heterozygosity, though with only 24 clones, could predominantly reflect the hybrid nature of the majority of the clones investigated here, especially with five species involved in their parentage.

The UPGMA dendrogram matches essentially the historically known or speculated origin of the clones (*Figure 1*). For example, the five full-sibling clones of DH32, together with their half-siblings DH33-27 and DH30-1, were confirmed to share high genetic similarity and clustered very closely in Group I. Therefore, genotypes from different nurseries would be expected to have a high genetic similarity and consequently a close placement in cluster analysis if they had common parents, as demonstrated also in rootstocks of apple (*Malus* spp.; ORAGUZZI et al., 2005). Similarly, in accordance with the results of parent assignment test, clones T13, Guanglin-5 and Guanglin-9 turned out to be offspring of the DH32 cross (no reciprocal mating was made for the parents of DH32), and Guanglin-4 was most likely to be a relative of the DH series clones while the mislabelled DH201-2 appeared to be far from the expectation of an offspring obtained from the male parent of DH32. In addition, the clustering of *E. urophylla* × *E. tereticornis* and *E. urophylla* × *E. camaldulensis* in Group II is not surprising as the two paternal species are classified so closely that *E. tereticornis* has been considered as a subspecies of *E. camaldulensis* (BROOKER and KLEINIG, 1994). On the other hand, a few cases of inconsistency appeared in the dendrogram where the three *E. urophylla* clones LH88, U6 and D29 scattered in different groups. This is probably a reflection of significant divergence of the original seedlots from which these clones were derived. The three clones were selected around the early 1990s when *E. orophila* and *E. wetarensis* were integrated with the taxon *E. urophylla*, and may thus arise from different species per se. Also, there is a possibility that these clones, at least one or two of them, may originate from natural hybrids formulated in previous generations (e.g. the origin of U6 as discussed below).

Though the clone U6 was selected from an *E. urophylla* seedling plantation established with seed from an orchard at National Dongmen Forest Farm, it has been widely suspected as a natural hybrid produced in the seed orchard with pollens from an unknown species. This is also possible as the pollination rate from outside an eucalypt seed orchard could be as high as 29% (GRATTAPAGLIA et al., 2004) and 39.2% (CHAIX et al., 2003). The placement of U6 with *E. urophylla* × *E. tereticornis* and *E. urophylla* × *E. camaldulensis* clones in cluster analysis (*Figure 1*) supports the suspicion of hybrid origin, and *E. tereticornis* or *E. camaldulensis* is probably the paternal species. Nevertheless, we are still not sure which species is exactly the pollen donor mostly due to the taxonomical closeness of the two species (BROOKER and KLEINIG, 1994) as well as the limitation of plant material samples included herein.

In forestry operations with eucalypts, mislabelling or duplicating of clones, though undesirable, may happen easily and subsequent mis-planting can cause loss in productivity to the plantation crop (KEIL and GRIFFIN,

1994; TRIPATHI et al., 2006). In this sense, though no duplicates of clones were found in our investigation, identification of the possibly mislabelled DH201-2 indicates that the true clone has to be further fingerprinted and its identity be scrutinized during exchanging in the market.

There appears to be some allele mutations in clones Guanglin-9, DH33-27 and DH30-1 (data not presented). SSR allelic mutations are most likely due to insertion or deletion of tandem repeat unit(s) generated by slipped-strand mispairing during DNA replication (LEVINSON and GUTMAN, 1987; VIGUERA et al., 2001), and have been widely observed through parent to offspring in plants (e.g. GRATTAPAGLIA et al., 2004; LIA et al., 2007). However, the estimated minimal mutation rate (0.35%) was extremely high as compared to the value of 4.8×10^{-4} per generation detected in *E. grandis* seedlings (GRATTAPAGLIA et al., 2004). This may be attributed to a number of factors on which the mutation frequency depends, such as SSR overall length, repeat type, flanking sequence and recombination rate (SCHLÖTTERER, 2000). Also, di-nucleotide repeat microsatellites, as were used in this study, have been reported to show higher mutation rates following the step-wise mutation model in humans (*Homo sapiens* L.; BRINKMANN et al., 1998) and maize (*Zea mays* L.; VIGOUROUX et al., 2002). Moreover, such mutations may arise from somaclonal variations occurring during the period of vegetative propagation, as is the case for *Eucalyptus* clones detected at RAPD loci (LAIA et al., 2000; TRIPATHI et al., 2006).

However, the level of genetic similarity between the clones was extremely high relative to similar studies in *Eucalyptus*, e.g. an average of 0.39 (mean genetic distance 0.61) for 15 *E. grandis* × *E. urophylla* hybrid clones detected at 20 SSR loci (ROCHA et al., 2002). This is potentially associated with the fairly close relationship of the clones, given that they were bred by several institutions on basis of a limited number of superior trees and, for a specific species, could be thereof related in genealogical origins both within and among the institutions. Moreover, extensive sharing of chloroplast DNA haplotypes within subgenus *Symphyomyrtus* (MCKINNON et al., 2001), to which the five parental species of the clones belong, may contribute to a relatively low level of divergence among the parents and in turn the offspring clones. The high genetic similarity also suggests a narrow genetic base of the eucalypt materials planted widely in China and thus prompts the urgency to recruit diverse parental stocks in deployment of eucalypt clones.

In conclusions, the 24 commercial *Eucalyptus* clones could be accurately and rapidly fingerprinted using microsatellite markers. The clones have generally a high level of genetic similarity, suggesting a narrow genetic base of the eucalypt materials mostly planted in China. Microsatellite markers are highly informative for assignment of possible kinship between clones and identification of mislabelled clone(s). The microsatellite-based clone genotyping would have significant contributions to the clonal registration and identity control during vegetative propagation, plant stock handling and commercial exchanges.

Acknowledgements

This work was financially supported by the National Program of High Technology Development of China (No. 2011AA100202) and Guangxi Academy of Forestry (GAF-Linke200907). The authors would like to thank DEYU GAN, CHANGFU HONG, JIANWEN LI and SHIRAO PENG for their kind help in sample collection.

References

- ARNOLD, R. J., J. LUO and B. CLARKE (2004): Trials of cold-tolerant eucalypt species in cooler regions of south central China. ACIAR Technical Reports No. 57. Australian Centre for International Agricultural Research, Canberra.
- BOTSTEIN, D., R. L. WHITE, M. SKOLNICK and R. W. DAVIS (1980): Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am. J. Hum. Genet.* **32**: 314–331.
- BRINKMANN, B., M. KLINTSCHAR, F. NEUHUBER, J. HÜHNE and B. ROLF (1998): Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.* **62**: 1408–1415.
- BRONDANI, R. P. V., C. BRONDANI, R. TARCHINI and D. GRATTAPAGLIA (1998): Development and mapping of microsatellite based markers in *Eucalyptus*. *Theor. Appl. Genet.* **97**: 816–827.
- 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.
- BROOKER, M. I. H. and D. A. KLEINIG (1994): Field guide to eucalypts, volume 3: northern Australia. Inkata Press, Sydney.
- CHAIX, G., S. GERBER, V. RAZAFIMAHARO, P. VIGNERON, D. VERHAEGEN and S. HAMON (2003): Gene flow estimation with microsatellites in a Malagasy seed orchard of *Eucalyptus grandis*. *Theor. Appl. Genet.* **107**: 705–712.
- CLARKE, C. R. E. (2001): Are *Eucalyptus* clones advantageous for the pulp mill? *South Afr. For. J.* **190**: 61–65.
- DOYLE, J. J. and D. J. DOYLE (1990): Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.
- ELDRIDGE, K. G., J. DAVIDSON, C. HARDWOOD and G. VAN WYK (1993): Eucalypt domestication and breeding. Oxford University Press, Oxford.
- FAO (2000): Global forest resources assessment 2000 – main report. FAO, Rome.
- FARIA, D. A., E. M. C. MAMANI, G. J. PAPPAS and 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. MAMANI, M. R. PAPPAS, G. J. JR PAPPAS and D. GRATTAPAGLIA (2010): A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of *Eucalyptus*. *J. Hered.* **101**: 512–520.
- FOSSATI, T., I. ZAPPELLI, S. BISOFFI, A. MICHELETTI, L. VIETTO, F. SALA and S. CASTIGLIONE (2005): Genetic relationships and clonal identity in a collection of commercially relevant poplar cultivars assessed by AFLP and SSR. *Tree Genet. Genomes* **1**: 11–19.
- GAN, S., J. SHI, J. BAI and J. XU (1999): RAPD fingerprints of clones of *Eucalyptus urophylla* S. T. Blake and *E. tereticornis* Smith. *J. Nanjing For. Univ.* **23**: 11–14.
- 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.
- GRATTAPAGLIA, D., V. J. RIBEIRO and G. D. S. P. REZENDE (2004): Retrospective selection of elite parent trees using paternity testing with microsatellite markers: an alternative short term breeding tactic for *Eucalyptus*. *Theor. Appl. Genet.* **109**: 192–199.
- GRIFFIN, A. R. (2001): Deployment decisions – capturing the benefits of tree improvement with clones and seedlings, pp. 16. *In*: Developing the eucalypt of the future, Proceedings of International Union of Forestry Research Organizations (IUFRO) Symposium, edited by S. BARROS, Valdivia.
- HE, X., F. LI, J. SHI and S. GAN (2011): Seven genomic SSRs revealed in *Eucalyptus* by re-sequencing of DNA sequences from GenBank. *Silvae Genet.* **60**: 92–94.
- JACCARD, P. (1908): Nouvelles recherches sur la distribution florale. *Bul. Soc. Vaudoise Sci. Nat.* **44**: 223–270.
- KEIL, M. and A. R. GRIFFIN (1994): Use of random amplified polymorphic DNA (RAPD) markers in the discrimination and verification of genotypes in *Eucalyptus*. *Theor. Appl. Genet.* **89**: 442–450.
- KIMURA, M. and J. F. CROW (1964): The number of alleles that can be maintained in finite population. *Genetics* **49**: 725–738.
- LAIA DE, M. L., E. A. GOMES, E. J. ESBRISSE and E. F. ARAÚJO (2000): Random amplified polymorphic DNA (RAPD) analysis of genotypic identities in *Eucalyptus* clones. *Silvae Genet.* **49**: 239–243.
- LEIGH, F., V. LEA, J. LAW, P. WOLTERS, W. POWELL and P. DONINI (2003): Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. *Euphytica* **133**: 359–366.
- LEVINSON, G. and G. A. GUTMAN (1987): Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**: 203–221.
- LI, F. and S. GAN (2011): An optimised protocol for fluorescent-dUTP based SSR genotyping and its application to genetic mapping in *Eucalyptus*. *Silvae Genet.* **60**: 18–25.
- LIA, V. V., M. BRACCO, A. M. GOTTLIEB, L. POGGIO and V. A. CONFALONIERI (2007): Complex mutational patterns and size homoplasy at maize microsatellite loci. *Theor. Appl. Genet.* **115**: 981–991.
- LIBBY, W. J. (1985): Potential of clonal forestry, pp. 1–11. *In*: Clonal forestry: its impact on tree improvement and our future forests, Proceedings of the nineteenth meeting of the Canadian Tree Improvement Association, Part 2, edited by L. Zsuffa, R. M. RAUTER and C. W. YEATMAN, Toronto.
- DE-LUCAS, A. I., J. C. CANTANA, P. RECIO and E. HIDALGO (2008): SSR-based tool for identification and certification of commercial *Populus* clones in Spain. *Ann. For. Sci.* **65**: 107p1–107p7.
- MCKINNON, G. E., R. E. VAILLANCOURT, H. D. JACKSON and B. M. POTTS (2001): Chloroplast sharing in the Tasmanian eucalypts. *Evolution* **55**: 703–711.
- MURO-ABAD, J. I., E. A. GOMES, O. N. CANCIO and E. F. ARAÚJO (2001): Genetic analysis of *Eucalyptus urophylla* and *E. grandis* clones selected in commercial crops from the Brazilian Amazon by RAPD markers. *Silvae Genet.* **50**: 177–181.
- NEI, M. (1973): Analysis of genetic diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70**: 3321–3323.

- ORAGUZIE, N. C., T. YAMAMOTO, J. SOEJIMA, T. SUZUKI and H. M. DE SILVA (2005): DNA fingerprinting of apple (*Malus* spp.) rootstocks using simple sequence repeats. *Plant Breed.* **124**: 197–202.
- OTTEWELL, K. M., S. C. DONNELLAN, G. F. MORAN and D. C. PATON (2005): Multiplexed microsatellite markers for the genetic analysis of *Eucalyptus leucocylon* (Myrtaceae) and their utility for ecological and breeding studies in other *Eucalyptus* species. *J. Hered.* **96**: 445–451.
- PAETKAU, D., W. CALVERT, I. STIRLING and C. STROBECK (1995): Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* **4**: 347–354.
- PARK, S. D. E. (2001): Trypano tolerance in West African cattle and the population genetic effects of selection. Ph.D. thesis, University of Dublin, Dublin.
- PATERSON, A. H., S. D. TANKSLEY and M. E. SORRELLS (1991): DNA markers in plant improvement. *Adv. Agron.* **46**: 39–90.
- PAYN, K. G., W. S. DVORAK, B. J. H. JANSE and A. A. MYBURG (2008): Microsatellite diversity and genetic structure of the commercially important tropical tree species *Eucalyptus urophylla*, endemic to seven islands in eastern Indonesia. *Tree Genet. Genomes* **4**: 519–530.
- POWELL, W., G. C. MACHRAY and J. PROVAN (1996): Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1**: 215–222.
- PRYOR, L. D. and L. A. S. JOHNSON (1971): A classification of the eucalypts. Australian National University Press, Canberra.
- PRYOR, L. D., E. R. WILLIAMS and B. V. GUN (1995): A morphometric analysis of *Eucalyptus urophylla* and related taxa with descriptions of two new species. *Aust. Syst. Bot.* **8**: 57–70.
- QI, S. (2002): *Eucalyptus* in China, 2nd Ed. Chinese Forestry Press, Beijing.
- RABELLO, E., A. N. DE SOUZA, D. SAITO and S. M. TSAI (2005): *In silico* characterization of microsatellites in *Eucalyptus* spp.: abundance, length variation and transposon associations. *Gen. Mol. Biol.* **28**: 582–588.
- RAHMAN, M. H. and O. P. RAJORA (2002): Microsatellite DNA fingerprinting, differentiation, and genetic relationships of clones, cultivars, and varieties of six poplar species from three sections of the genus *Populus*. *Genome* **45**: 1083–1094.
- ROCHA, R. B., J. I. M. ABAD, I. E. PIRES and E. F. DE ARAUJO (2002): Fingerprint and genetic diversity analysis of *Eucalyptus* spp. genotypes using RAPD and SSR markers. *Sci. Forestalis* **62**: 24–31.
- ROHLF, F. J. (1998): NTSYS-pc: numerical taxonomy and multivariate analysis system, version 2.02. Exeter Software, Setauket, New York.
- SCHLÖTTERER, C. (2000): Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109**: 365–371.
- SOKAL, R. R. and C. D. MICHENER (1958): Statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* **38**: 1409–1438.
- STEANE, D. A., R. C. JONES and R. E. VAILLANCOURT (2005): A set of chloroplast microsatellite primers for *Eucalyptus* (Myrtaceae). *Mol. Ecol. Notes* **5**: 538–541.
- STEANE, D. A., R. E. VAILLANCOURT, J. RUSSELL, W. POWELL, D. MARSHALL and B. M. POTTS (2001): Development and characterisation of microsatellite loci in *Eucalyptus globulus* (Myrtaceae). *Silvae Genet.* **50**: 89–91.
- STOKOE, R. L., M. SHEPHERD, D. J. LEE, D. G. NIKLES and R. J. HENRY (2001): Natural inter-subgeneric hybridization between *Eucalyptus acmenoides* Schauer and *Eucalyptus cloeziana* F. Muell (Myrtaceae) in southeast Queensland. *Ann. Bot.* **88**: 563–570.
- TAUZ, D. (1989): Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Res.* **16**: 6463–6471.
- TRIPATHI, S. B., N. V. MATHISH and K. GURUMURTHI (2006): Use of genetic markers in the management of micropropagated *Eucalyptus* germplasm. *New For.* **31**: 361–372.
- VARSHNEY, R. K., A. GRANER and M. E. SORRELLS (2005): Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* **23**: 48–55.
- VIGOUROUX, Y., J. S. JAQUETH, Y. MATSUOKA, O. S. SMITH, W. D. BEAVIS, J. S. SMITH and J. DOEBLEY (2002): Rate and pattern of mutation at microsatellite loci in maize. *Mol. Biol. Evol.* **19**: 1251–1260.
- VIGUERA, E., D. CANCELL and S. D. EHRLICH (2001): Replication slippage involves DNA polymerase pausing and dissociation. *EMBO J.* **20**: 2587–2595.
- WAGNER, H. W. and K. M. SEFC (1999): IDENTITY 1.0. Centre for Applied Genetics, University of Agricultural Sciences, Vienna.
- WATT, M. P., E. A. DUNCAN, M. ING, F. C. BLAKEWAY and B. HERMANN (1995): Field performance of micropropagated and macropropagated *Eucalyptus* hybrids. *South Afr. For. J.* **173**: 17–21.
- WEIR, B. S. (1996): Genetic data analysis II. Sinauer Associates Inc. Publishers, Sunderland.
- YASODHA, R., R. SUMATHI, P. CHEZHIAN, S. KAVITHA and M. GHOSH (2008): *Eucalyptus* microsatellites mined *in silico*: survey and evaluation. *J. Genet.* **87**: 21–25.
- 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).