Indian Journal of Biotechnology
Vol 13, January 2014, pp 75-80

Unique pattern information combinations for clonal discrimination in *Eucalyptus camaldulensis* Dehnh. using microsatellite markers

A Shanmugapriya[1] and R Yasodha[2]*

[1]R & D Center, Bharathiar University, Coimbatore 641 046, India
[2]Division of Plant Biotechnology, Institute of Forest Genetics and Tree Breeding, Coimbatore 641 002, India

Received 22 January 2013; revised 10 August 2013; accepted 21 September 2013

Eucalyptus is the most widely planted hardwood industrial-plantation species in the tropics and is mainly exploited for its pulp wood properties. In India, *E. camaldulensis* Dehnh. is one of the most important cultivated species for its wider adaptability in arid and semi-arid areas. Sixty two microsatellite markers were used for estimation of genetic diversity and unique pattern identification combinations (UPIC) for 28 clones of *E. camaldulensis*. PI (polymorphic information content) estimates, indicating the ability of microsatellite markers to distinguish two closely related genotypes, and UPIC values provide the actual number of SSR loci required to discriminate the clones analyzed. The loci Embra5 was found to be the best informative loci, which could differentiate 26 clonal accessions. Highly informative subset from these 62 SSR loci would be sufficient for the future studies on genetic diversity or fingerprinting of *E. camaldulensis* with limited experimental cost.

Keywords: *Eucalyptus camaldulensis*, fingerprinting, genetic diversity, microsatellites, probability of identity

Introduction

Genus *Eucalyptus* belonging to family Myrtaceae is one of the most widely planted hardwood industrial-plantation species in the tropical countries and is mainly exploited for its pulp wood properties. Plantation estimates indicate that more than 20.0 million ha are planted with eucalypts globally and majority of the eucalypt plantations exists in Brazil and India with 4.3 m ha and 3.9 m ha, respectively[1]. Recently, eucalypts are considered to be the potential feedstock for renewable bioenergy[2]. In India, *E. camaldulensis* Dehnh. is one of the most important cultivated species for its wider adaptability in arid and semi-arid areas. The natural distribution of *E. camaldulensis* (river red gum) spans in all mainland states of Australia[3]. *E. camaldulensis* belongs to the section-*Exsertaria* of subgenus-*Symphyomyrtus* with having fast growth and tolerance to salinity[4]. *E. camaldulensis* is diploid (n=11) and has a genome size of about 590 Mbp/C[5]. Butcher *et al*[6] proposed 7 subspecies in *E. camaldulensis* based on the geographic trends of distribution and genetic diversity analyzed with microsatellite loci.

In India, the domestication programme of eucalypt species are being systematically implemented, and subsequently provenance cum progeny trials, seed production areas, half pedigreed seedling seed orchards and clonal plantations have been established[7,8]. Industrially valuable traits, such as, high wood biomass, disease (gall) resistance, and wood properties for specific end products are considered for selection of elite clones and the individuals are propagated for large-scale plantations by the public and private institutions. Such clones canalso be protected under the Protection of Plant Varieties and Farmers' Rights Act 2001 (PPVFR Act 2001) of India. Institute of Forest Genetics and Tree Breeding, Coimbatore has undertaken the development of DUS (distinctness, uniformity & stability) testing guidelines. It is essential to set high standards for proving the distinctness, uniqueness and stability of eucalypt clones as several organizations including the forest department and paper industries are involved in genetic improvement programmes of eucalypts and tailor made clones are produced for specific pulp wood properties. In cases of alleged infringement, DNA fingerprinting can provide unique pattern information for determining clonal distinctness, essential derivation or hybridity.

SSR markers are considered to be one of the suitable marker tools to assess the genetic diversity, clonal fingerprinting and phylogeny[9,10]. Molecular-genetic characterization of *E. camaldulensis* was

*Author for correspondence:
Tel: +91-422-2484100; Fax: +91-422-2430549
E-mail: yasodha@icfre.org
ΨFor Suppl Table 1 see the paper at www.nopr.niscair.res.in or www.niscair.res.in

carried out at various levels of taxa with different types of DNA markers, such as, ISSRs[11], diversity array technology (DArT) markers[12], RFLPs[3] and SSRs[6,13]. SSRs were found highly informative in eucalypts[14] with extensive transferability across species, providing an appropriate tool for population genetic studies[15]. Twenty five SSR loci were developed *de novo* for *E. camaldulensis* by da Silva *et al*[16]. Recently, Arumugasundaram *et al*[17] used SSR markers to discriminate *E. camaldulensis* from its closely related species *E. tereticornis*. In the present study, microsatellite markers were used for estimation of genetic diversity and unique pattern identification combinations for the clones of *E. camaldulensis*.

Materials and Methods

Plant Material

The germplasms of *E. camaldulensis* consisting of 28 individuals were selected for the analysis from a clonal seed orchard established by the Institute of Forest Genetics and Tree Breeding at Karunya, Coimbatore (Table 1). The selections were carried out from the provenance cum progeny trial established during 1995 from the seed lots obtained from Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia. These 28 clones were assessed for growth performance based on height, diameter and volume production across 3 locations and were ranked as superior performers out of 110 clones selected.

Microsatellite Amplification and Data Analysis

Genomic DNA isolation, microsatellite amplification and electrophoresis were carried out as reported by Arumugasundaram *et al*[17]. Sixty two primers were used for genetic diversity analysis and clonal differentiation (Table 2). The microsatellite data generated in this study were scored (bp) manually using the 50 bp size standard molecular ladder (MBI, Fermentas, USA). The number of alleles, observed and expected heterozygosity estimates, and Roger's genetic distance based dendrogram were obtained using the program Power Marker[18]. Calculations on polymorphic information content (PIC) and unique pattern informative combinations (UPIC) were made to determine the most informative set of SSR markers that could be used for future studies. Similarly, knowledge on the degree of heterozygosity of the clonal individuals would allow selection of potential individuals for future breeding. PIC, UPIC values of the SSR loci and heterozygosity estimates of the clones were calculated using UPIC Perl scripts[19]. Another important measure in clonal differentiation is the probability of identity (PI), useful to test identity among the individuals by chance of genotypes. The Gimlet software[20] was used to measure the PI. The discriminatory power of each microsatellite locus was calculated by estimating sib-based and unbiased corrected PI estimates and cumulative power of discrimination was calculated as products of PIs of successive informative markers arranged in decreasing order as described by Waits *et al*[21].

Results

In the present study, 62 SSR loci were analyzed with 28 clones of *E. camaldulensis* and the details of the diversity statistics like PIC, number of alleles, heterozygosity (observed and expected), PI and UPIC

Table 1—Details on *E. camaldulensis* clonal accessions used in the study and their heterozygosity status

No.	Sub-species of *E. camaldulensis*	Acc. ID	Provenance name	Heterozygosity (%)
1	*obtusa*	1	Katherine	80.33
2	*obtusa*	9	Katherine	81.67
3	*obtusa*	7	Victoria River	77.97
4	*obtusa*	14	Victoria River	78.33
5	*obtusa*	10	Victoria River	90.00
6	*acuminata*	16	Gilbert River	73.33
7	*obtusa*	17	Katherine	71.67
8	*obtusa*	19	Katherine	64.52
9	*acuminata*	31	Morehead River	65.57
10	*simulata*	53	Kennedy River	64.41
11	-	63	Bulked	70.49
12	-	75	Bulked	60.00
13	-	66	Bulked	63.33
14	-	76	Bulked	63.33
15	*simulata*	100	Kennedy River	68.09
16	*simulata*	115	Kennedy River	81.67
17	*simulata*	101	Kennedy River	73.77
18	*simulata*	111	Kennedy River	67.24
19	*simulata*	116	Kennedy River	71.19
20	*simulata*	118	Kennedy River	60.34
21	*simulata*	124	Kennedy River	65.52
22	-	154	Wrotham	62.30
23	*simulata*	186	Kennedy River	60.66
24	*obtusa*	188	Victoria River	65.57
25	*obtusa*	191	Katherine	71.19
26	-	198	Bulked	75.44
27	-	196	Bulked	65.52
28	*simulata*	123	Kennedy River	56.14

are provided in Table 2. A total of 795 alleles were detected and the number of alleles ranged from 4 to 22 with an average of 12.8. The highest numbers of alleles, 22 and 20 were found at locus Embra5 and Embra35, respectively. About 72.5% of the loci (45 SSRs) showed between 10 and 20 alleles, and the lowest number was found for the locus Eg24. The PIC calculated for the 62 markers was plotted in Fig. 1. Except two loci, the number of alleles was above 5, and PIC values ranged from 0.273 to 0.979. PIC values were less when the numbers of alleles were below five.

The degree of heterozygosity of clones analyzed allows choosing the potential individuals for further utilization like seed orchard establishment or parents for hybrid production. The heterozygosity values (Table 1) of the clones analyzed ranged from 90 (clone ID 14) to 56% (clone ID 123), showing the high diversity existing among the clones. Roger's genetic distance based dendrogram also revealed the genetic diversity among the clones (Fig. 2).

Two types of PI estimates, sib-based and unbiased considering the clonal accessions as related or unrelated, respectively were calculated (Ψ Suppl Table 1).

Table 2—SSR loci used in the study and information on their number of alleles (Na), polymorphic information content (PIC), observed (Ho) and expected (He) heterozygosity, and unique pattern identification (UPIC) value

SSR loci	Na	PIC	Ho	He	UPIC	SSR loci	Na	PIC	Ho	He	UPIC
Embra11	16	0.962	0.919	0.577	20	Embra28	13	0.954	0.865	0.964	15
Embra56	12	0.962	0.888	0.818	9	Embra8	17	0.979	0.929	0.964	22
Embra6	16	0.967	0.914	0.750	19	Embra50	12	0.928	0.882	0.444	10
Embra70	12	0.956	0.885	0.929	8	Embra25	8	0.887	0.816	0.667	11
Embra12	11	0.927	0.823	0.857	9	Embra20	15	0.967	0.904	0.889	17
Embra35	20	0.969	0.934	0.520	19	Embra167	18	0.950	0.885	0.607	16
Embra100	19	0.973	0.922	0.815	22	Embra7	10	0.951	0.881	0.929	13
Eg117	13	0.945	0.877	0.714	17	Embra226	15	0.961	0.901	0.741	20
En10	8	0.842	0.807	0.407	6	Embra119	10	0.915	0.843	0.643	8
Embra55	10	0.832	0.736	0.458	7	Embra3	8	0.810	0.701	0.630	8
Embra63	13	0.934	0.889	0.444	9	Embra17	13	0.922	0.827	0.679	11
Embra172	9	0.869	0.813	0.464	6	Embra18	9	0.900	0.821	0.654	9
Embra43	17	0.968	0.905	0.852	23	Embra204	12	0.909	0.853	0.417	13
Embra207	17	0.974	0.922	0.880	21	Embra58	18	0.976	0.925	0.893	22
Embra195	18	0.970	0.916	0.769	24	Embra10	12	0.959	0.889	0.889	16
Embra227	11	0.912	0.861	0.519	9	Embra23	14	0.940	0.864	0.667	15
Embra122	18	0.975	0.920	0.929	18	Embra61	13	0.925	0.882	0.385	11
Embra77	15	0.945	0.876	0.625	12	Embra101	17	0.965	0.891	0.926	16
Eg24	4	0.273	0.535	0.536	1	Embra40	14	0.954	0.882	0.786	14
Embra19	12	0.929	0.888	0.440	12	Embra155	13	0.916	0.815	0.679	16
Embra78	15	0.945	0.872	0.667	15	Embra29	17	0.967	0.904	0.821	17
Embra66	17	0.967	0.905	0.852	19	Embra39	15	0.964	0.895	0.893	17
Embra36	15	0.947	0.882	0.643	15	Embra2	9	0.861	0.752	0.667	7
Embra179	9	0.889	0.811	0.607	10	Eg128	13	0.958	0.909	0.619	13
Embra41	18	0.975	0.922	0.893	20	Embra243	5	0.446	0.565	0.429	22
Embra143	15	0.950	0.874	0.750	17	Embra207	8	0.862	0.781	0.607	21
Embra54	15	0.946	0.895	0.556	17	Eg16	14	0.955	0.897	0.679	3
Embra9	10	0.922	0.865	0.542	10	Embra72	8	0.737	0.615	0.571	11
Embra24	8	0.854	0.834	0.333	8	CD669383	7	0.833	0.767	0.643	8
Embra168	6	0.791	0.685	0.885	2	CD668471	11	0.925	0.813	0.893	7
Embra5	22	0.979	0.938	0.786	26	CD668519	6	0.793	0.669	1.000	11
						CD668704	16	0.962	0.847	0.695	1

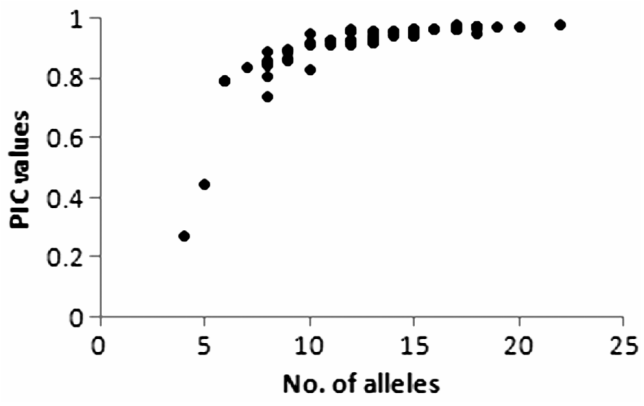

Fig. 1—Polymorphic information content (PIC) *vs* number of alleles.

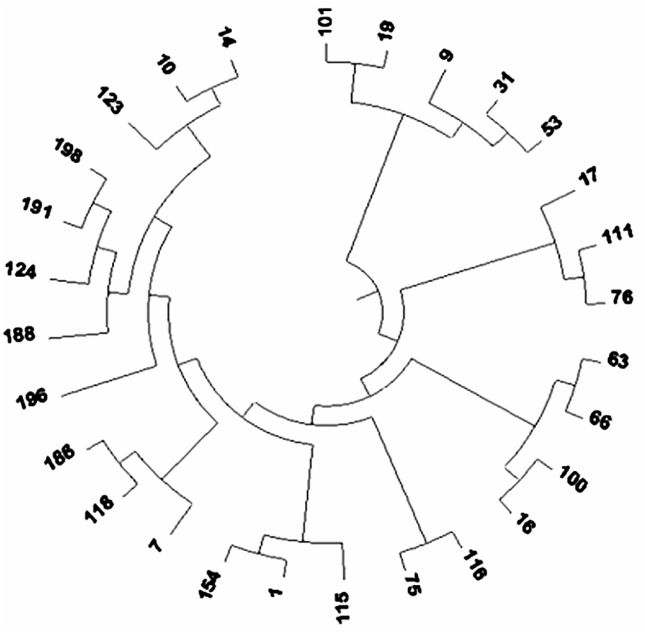

Fig. 2—Dendrogrm showing the genetic diversity among the 28 *E. camaldulensis* clones. [Numbers on the node indicate accession ID as in Table 1]

The cumulative PIs indicating discriminatory power of the markers was found to be high as compared to individual SSR loci, wherein the sib-based cumulative PIs calculated over 62 SSR markers were 2.938×10^{-01} to 3.030×10^{-30} and unbiased cumulative PI estimates were ranged from 4.329×10^{-03} to 7.965×10^{-112}. Since a low PI value indicates high marker efficiency, the best markers for clonal discrimination should be selected accordingly. UPIC values were ranged from 1 (for the EST-SSR loci CD668704) to 26 (for the loci Embra5), indicating that the loci CD668704 could differentiate only one clone, whereas Embra5 could differentiate 26 clonal accessions.

Discussion

Morphological descriptors, such as, leaf shape, bark texture, fruit and flower characters are generally considered in the traditional procedure of clonal identification and accepted by the International Union for the Protection of New Varieties of Plants (UPOV). However, the evaluation of morphological descriptors is prone to errors, and to ascertain flower and fruit characters is also time consuming[22]. Because of these constraints, countries, such as, Brazil and Uruguay have included the use of microsatellite markers as additional descriptors in their legislation for clonal discrimination of eucalypts. Individual identification of clones finds several applications in tree breeding and clonal forestry[23]. Genotype x environment interaction studies in out-crossing tree species require clonally replicated propagules, wherein the proper identity of the test individuals gives results with greater precision. The successful application of marker-aided selection to tree improvement, particularly for the traits with moderate heritability, depends upon precise estimation of phenotypic variation attributable to marker genotype and clonal testing is the best option. Clonal forestry uses specific clones with specific wood properties, which results in uniform raw material for the industry and hence material traceability becomes important. Seed orchards established with the clones are tested for their breeding value and specific combing ability. Thus genetic structure and clone combination can be designed in an improved seed orchard, in which clones have the most effective genetic contribution to the seed lots[24]. As a result, both the nursery and field plantations require suitable tools, such as, DNA markers for easy identification of individuals. DNA markers particularly SSR markers are one of the best tools for the assessment of genetic diversity among the clones in seed orchards and clonal plantations, wherein the maintenance of suitable genetic diversity is highly advocatedto minimize the losses of clonal plantations[25]. However, given the cost of detection of SSR polymorphism, cost-effective planning of genotyping and fingerprinting has to be carried out by selecting highly informative markers.

Similarity threshold for identification of clones to avoid misidentification of genetically similar seedlings as different clone and of dissimilar fingerprints of clones as genetically similar individuals has been worked out for a few plant species[26]. PI estimate provides the power of discrimination of a SSR locus with the given genotypes. This estimator can be used as a conservative upper bound for the probability of

observing identical multilocus genotypes sampled from a population[21]. Two types of PI estimates were calculated (ᵂSuppl Table 1), which showed that the sib-based PI values for individual markers were around 10^{-1} however the unbiased PI estimates ranged from 10^{-1} to 10^{-3}, thus indicating a low probability that any two individuals drawn from these accessions will share the same multi-locus genotype. The cumulative PI indicating discriminatory power of the markers were found to be higher wherein the sib-based cumulative PI calculated over 62 SSR markers was 2.938×10^{-01} to 3.030×10^{-30} and unbiased cumulative PI estimates were ranged from 4.329×10^{-03} to 7.965×10^{-112}. These PI values indicate that the risk of falsely inferring clonal identity of tested individuals is very low. Further, the cumulative use of markers is useful in absolute identification of clonal accessions compared to the use of single marker. Similar studies were made in coffee germplasms and it was suggested that the use of sib-based PI for the clonal discrimination would be more authentic because the sib-based PI accounts the possible relatedness in the target germplasm arising due to common parentage[27].

UPIC values indicate the maximum number of accessions could be differentiated with a particular SSR loci and they are more informative for selecting subsets of SSRs than the use of its PIC alone[19]. UPIC values were ranged from 1 (for the EST-SSR loci CD668704) to 26 (for the loci Embra5), indicating that the loci CD668704 could differentiate only one clone, whereas Embra5 could differentiate 26 clonal accessions. Hence, in future studies on genetic diversity or fingerprinting of *E. camaldulensis*, selection of a small and highly informative subset from these 62 SSR loci would be sufficient and also lead to limited experimental cost.

In conclusion, PI estimates indicated the probability of identity of the accessions or the ability of microsatellite markers to distinguish two closely related genotypes, while UPIC values provided the actual number of SSR loci required to discriminate the clones analyzed. In the present study, all the clonal accessions were genetically distinct having a unique pattern information combination. These clones were already proven to be the best for growth performance and they could be selected as base population for clonal seed orchard establishment.

Acknowledgement

The first author acknowledges with thanks the award of Junior Research Fellowship from the Indian Council of Forestry Research and Education (ICFRE), Dehradun, India.

References

1 Iglesias-Trabado G & Wilstermann D, *Eucalyptus universalis*. Global cultivated eucalypt forests map 2008, version 1.0.1, in *GIT Forestry Consulting's Eucalyptologics: Information resources on Eucalyptus cultivation worldwide, 2008.* [Retrieved from http://www.git-forestry.com/ (September 2009)]

2 Rocha J D, Coutinho A R & Luengo C A, Biopitch produced from *Eucalyptus* wood pyrolysis liquids as a renewable binder for carbon electrode manufacture, *Braz J Chem Eng*, 19 (2002) 127-132.

3 Butcher P A, Otero A, McDonald M W & Moran G F, Nuclear RFLP variation in *Eucalyptus camaldulensis* Dehnh. from northern Australia, *Heredity*, 88 (2002) 402-412.

4 Benyon R G, Marcar N E, Crawford D F & Nicholson AT, Growth and water use of *Eucalyptus camaldulensis* and *E. occidentalis* on a saline discharge site near Wellington, NSW, Australia, *Agric Water Manage*, 39 (1999) 229-244.

5 Grattapaglia D & Bradshaw H D, Nuclear DNA content of commercially important *Eucalyptus* species and hybrids, *Can J For Res*, 24 (1994) 1074-1078.

6 Butcher P A, McDonald M W & Bell J C, Congruence between environmental parameters, morphological and genetic structure in Australia's most widely distributed eucalypt, *Eucalyptus camaldulensis*, *Tree Genet Genom*, 5 (2009) 189-195.

7 Varghese M, Harwood C E, Hegde R & Ravi N, Evaluation of provenances of *Eucalyptus camaldulensis* and clones of *E. camaldulensis* and *E. tereticornis* at contrasting sites in southern India, *Silvae Genet*, 57 (2008) 170-179.

8 Varghese M, Kamalakannan R, Harwood C E, Lindgren D & McDonald M W, Changes in growth performance and fecundity of *Eucalyptus camaldulensis* and *E. tereticornis* during domestication in southern India, *Tree Genet Genom*, 5 (2009) 629-640.

9 Fossati T, Zapelli I, Bisoffi S, Micheletti A, Vietto L *et al*, Genetic relationships and clonal identity in a collection of commercially relevant poplar cultivars assessed by AFLP and SSR, *Tree Genet Genom*, 1 (2005) 11-20.

10 Riaz S, Tenscher A C, Smith B P, Ng D A & Walker M A, Use of SSR markers to assess identity, pedigree, and diversity of cultivated muscadine grapes, *J Am Soc Hortic Sci*, 133 (2008) 559-568.

11 Balasaravanan T, Chezhian P, Kamalakannan R, Ghosh M, Yasodha R *et al*, Determination of inter- and intra-species genetic relationships among six *Eucalyptus* species based on inter-simple sequence repeats (ISSR), *Tree Physiol*, 25 (2005) 1295-1302.

12 Steane D A, Nicolle D, Sansaloni C P, Petroli C, Carling J *et al*, Population genetic analysis and phylogeny reconstruction in *Eucalyptus* (Myrtaceae) using high-throughput, genome-wide genotyping, *Mol Phylogenet Evol*, 59 (2011) 206-224.

13 Faria D A, Mamani E M, Pappas M R, Pappas G J Jr & Grattapaglia D, A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of eucalyptus, *J Hered*, 101 (2010) 512-520.

14 Steane D A, Vaillancourt R E, Russell J, Powell W, Marshall D *et al*, Development and characterization of microsatellite loci

in *Eucalytpus globulus* (*Myrtaceae*), *Silvae Genet*, 50 (2001) 89-91.

15 Ochieng J W, Steane D A, Ladiges P Y, Baverstock P R, Henry R J *et al*, Microsatellites retain phylogenetic signals across genera in eucalypts (Myrtaceae), *Gen Mol Biol*, 30 (2007) 1125-1134.

16 da Silva J M, de Sousa A C B, de Souza A P, Mori E S, Freitas M L M *et al*, Development and characterization of 14 microsatellite loci from an enriched genomic library of *Eucalyptus camaldulensis* Dehnh, *Conserv Genet Resour*, 1 (2009) 465-469.

17 Arumugasundaram S, Ghosh M, Veerasamy S & Ramasamy Y, Species discrimination, population structure and linkage disequilibrium in *Eucalyptus camaldulensis* and *Eucalyptus tereticornis* using SSR markers, *PLoS ONE*, 6 (2011) e28252. [doi:10.1371/journal.pone.0028252]

18 Liu K & Muse S V, PowerMarker: An integrated analysis environment for genetic marker analysis, *Bioinformatics*, 21 (2005) 2128-2129.

19 Arias R S, Ballard L L & Scheffler B E, UPIC: Perl scripts to determine the number of SSR markers to run, *Bioinformation*, 3 (2009) 352-360.

20 Valière N, GIMLET: A computer program for analyzing genetic individual identification data, *Mol Ecol Notes*, 2 (2002) 377-379.

21 Waits L P, Luikart G & Taberlet P, Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines, *Mol Ecol*, 10 (2001) 249-256.

22 De-Lucas A I, Santana J C, Recio P & Hidalgo E, SSR-based tool for identification and certification of commercial *Populus* clones in Spain, *Ann For Sci*, 65 (2008) 107.

23 Torres-Dini D, Bennadji Z, Cabrera M, Centurian C, Resquin F *et al*, Use of SSR-tools for clone certification in Uruguayan *Eucalyptus grandis* and *Eucalyptus dunnii* breeding programs, *BMC Proc*, 5 Supl 7 (2011) P58.

24 Li x, Zhou J, Jing Y, Liu x & Wang x F, Expression of a *GALACTINOL SYNTHASE* gene is positively associated with desiccation tolerance of *Brassica napus* seeds during development, *J Plant Physiol*, 168 (2011) 1761-1770.

25 Bishir J & Roberds J H, On number of clones needed for managing risks in clonal forestry, *For Genet*, 6 (1999) 149-155.

26 Douhovnikoff V & Dodd R S, Intra-clonal variation and a similarity threshold for identification of clones: Application to *Salix exigua*, *Theor Appl Genet*, 106 (2003) 1307-1315.

27 Hendre P S, Phanindranath R, Annapurna V, Lalremruata A & Aggarwal R K, Development of new genomic microsatellite markers from robusta coffee (*Coffea canephora* Pierre ex A. Froehner) showing broad cross-species transferability and utility in genetic studies, *BMC Plant Biol*, 8 (2008) e51.