Arch Biol Sci. 2016;68(2):417-425 DOI:10.2298/ABS150608124X

IDENTIFICATION AND CHARACTERIZATION OF GENIC MICROSATELLITES IN *CUNNINGHAMIA LANCEOLATA* (LAMB.) HOOK (TAXODIACEAE)

Yang Xu[1,*], Renhua Zheng[2], Zhanjun Wang[1], Ying Wang[1], Zhou Hong[1], Liwei Yang[1], Ye Lu[1], Yaqi Zhao[1], Jinhui Chen[1] and Jisen Shi[1,*]

[1] *Key Laboratory of Forest Genetics and Biotechnology of Ministry of Education, Nanjing Forestry University*, Nanjing 210037, China

[2] *Fujian Academy of Forestry, Key Laboratory of Timber Forest Breeding and Cultivation for Mountain Areas in Southern China*, Fuzhou, Fujian 350012, China

***Corresponding authors:** xuyang198610@163.com; jshi@njfu.edu.cn

Received: June 8, 2015; **Revised**: September 17, 2015; **Accepted**: October 8, 2015; **Published online**: October 19, 2015

Abstract: Genomic resources for conventional breeding programs are extremely limited for coniferous trees, and existing simple sequence repeat markers are usually identified through the laborious process of hybridization screening. Therefore, this study aimed to identify gene-based microsatellites in the Chinese fir, *Cunninghamia lanceolata* (Lamb.) Hook by screening transcript data. We identified 5200 microsatellites. Trinucleotide motifs were most common (47.94%) and were followed by tetranucleotide motifs (24.92%). The AG/CT motif (43.93%) was the most abundant dinucleotide repeat, whereas AAG/CTT (25.07%) was the most common trinucleotide repeat. A total of 411 microsatellite primer pairs were designed and 97 polymorphic loci were identified by 8 genotypes. The number of alleles per locus (Na) in these polymorphic loci ranged from 2 to 5 (mean, 2.640), the H_o values were 0.000-1.000 (mean, 0.479), and the H_E values were 0.125-0.775 (mean, 0.462). The polymorphic information content (PIC) values were 0.110-0.715 (mean, 0.383). Seventy-two of the 97 polymorphic markers (74.23%) were present within genes with predicted functions. In addition, in genetic diversity and segregation analyses of 16 genotypes, only 5.88% of the polymorphic loci displayed segregation distortion at the $p<0.05$ level. Transferable amplification of a randomly selected set of 30 genic microsatellites showed that transferability decreased with increasing evolutionary distance between *C. lanceolata* and target conifers. Thus, these 97 genic markers will be useful for genetic diversity analysis, germplasm characterization, genome mapping and marker-assisted breeding in *C. lanceolata*, and evolutionary genetic analysis in Taxodiaceae.

Key words: *Cunninghamia lanceolata* (Lamb.) Hook; unigene sequences; microsatellite mining; genic microsatellites; polymorphism

INTRODUCTION

Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook) is mainly distributed in the tropical and subtropical mountainous areas of China and Vietnam. It is an important tree species for timber production and has been extensively planted in southern China for over 3000 years [1,2]. Despite great success in conventional breeding programs [3], highly polymorphic genetic markers remain very limited for coniferous trees [4].

Microsatellites (or simple sequence repeats, SSRs) provide codominant marker types for a wide range of genetic applications [5], but identifying them has traditionally been tedious and labor-intensive, expensive and low-throughput [6-8]. Thus, to date, only 11 genomic SSRs and 28 EST-SSRs in *C. lanceolata* have recently been reported [9,10].

In recent years, next-generation sequencing technologies have provided exciting means of developing genic SSR markers for non-model organisms [11]. Compared with genomic SSRs, genic SSRs are more likely to be linked to gene loci that contribute to morphological phenotypes [5,11]. Moreover, they repre-

sent useful tools for marker-assisted selection. These markers can also facilitate evolutionary analyses as they are more commonly shared across related taxa than "anonymous" SSRs [11,12]. In the present study, we evaluate the frequency and distribution of various types of genic SSRs and develop polymorphic genic SSR markers as genetic tools for *C. lanceolata*. These novel markers will be useful in future genetic studies and breeding applications in the conifer.

MATERIALS AND METHODS

SSR characterization and primer design

UniGene sequences related to *C. lanceolata* [13] from the NCBI database (http://www.ncbi.nlm.nih.gov) were downloaded. The microsatellite identification tool (MISA) [14] was used to find SSRs, and adjustments were made to locate mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeating units with no fewer than 20, 6, 4, 3, 3 and 3 repeats, respectively. After eliminating SSR motifs shorter than18 bp or mononucleotides and fragments that were too short or had inappropriate flanking sequences, primers for the identified SSR loci were designed using BatchPrimer3 software [15]. The parameters for primer design were defined as: 1) a primer length of 18-25 bases (optimum 20); 2) PCR product length of 100-500 bp; 3) primer melting temperature (Tm) of 50°C-60°C (optimum 55°C); and 4) primer GC content of 45%-70%. Primers were synthesized by Invitrogen Biotechnology Co. Ltd (Shanghai, China).

Plant material and data analysis

In this study, a set of eight unique individuals (Table 1), including one sample of *C. konishii* (=*C. lanceolata*. var. *konishii*), from natural populations and a first-generation seed orchard, was used for an analysis of SSR allele diversity and polymorphism screening. The number of alleles (Na), observed heterozygosity (H_o) and expected heterozygosity (H_E) were calculated using POPGEN 1.32 [16]. Polymorphism information content (PIC) was derived according to the following formula [17]:

$$PIC = 1 - \sum_{i=1}^{n} p^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2p_i^2 p_j^2$$

where *n* is the number of alleles at one locus; *Pi* and *Pj* are the frequencies of the *i*th and *j*th alleles at one locus; and *j*=*I* +1.

All unigenes containing polymorphic microsatellites were queried against the GenBank nonredundant protein database using BLASTX with an expected value (E-value) of 10^{-5} for the function of genic SSRs. An additional set of 14 progeny from a "D110" ×"6421" cross and their parents were applied for analysis of SSR allele segregation (Table 1). Chi-square tests were used to analyze the allele segregation of 97 loci by JoinMap 4.1 [18]. Then, the cluster analysis of all polymorphic loci and all 24 genotypes was conducted based on Nei's unbiased measures of genetic distances [19] using POPGEN 1.32 and MEGA 5 [20]. To investigate the transferability of SSR markers in related species, 30 random markers selected from 97 polymorphic SSR markers were used to amplify the genomic DNA from another six species – two Taxodiaceae (*Metasequoia glyptostroboides* and *Glyptostrobus pensilis*), two Pinaceae (*Pinus massoniana* and *Cedrus deodara*), one Cupressaceae (*Platycladus orientalis),* and one Cephalotaxaceae (*Cephalotaxus fortunei*).

All genotypes were conserved at the National Forest Germplasm preservation base of *C. lanceolata* in Yangkou Forest Farm, Fujian Province, China. Genomic DNA from fresh needles from trees of each genotype was extracted using the cetyltrimethylammonium bromide method [21].

Genic SSR marker amplification

Each PCR reaction mixture (10 μl) contained 1μl 10× reaction buffer (100 mM Tris-HCl, pH 9.0, 100 mM KCl, and 80 mM $(NH_4)_2SO_4$), 2 mM of total dNTP, 0.2 μM each of the forward and reverse primers, approximately 30 ng genomic DNA, 0.5 U Taq DNA polymerase (TaKaRa Biotechnology, Dalian, China), and 1.25 mM $MgCl_2$. PCR was performed in a Veriti

Table 1. *C. lanceolata* genotypes used in the SSR diversity and segregation analyses.

Number/ Name	Experimental Forest [a]	Origin		
		Locality	Latitude	Longitude
1	Third-generations of resources collection area	Chun'an, Zhejiang Prov.	29°42'N	118°30'E
2	Third-generations of resources collection area	Tianzhu, Guizhou Prov.	26°74'N	109°36'E
3	Third-generations of resources collection area	Guanxi, Taiwan Prov. (*C. konishii*)	24°48'N	121°10'E
4	Third-generations of resources collection area	Renhua, GuangdongProv.	25°31'N	113°60'E
5	Third-generations of resources collection area	Pengxian, Sichuan Prov.	30°99'N	103°97'E
6	Third-generations of resources collection area	Shunchang, Fujian Prov.	26°74'N	117°50'E
7	Third-generations of resources collection area	JianOu, Fujian Prov.	26°96'N	118°26'E
8	First-generation seed orchard	Shunchang, Fujian Prov.	26°79'N	117°82'E
D110	Female parent	Shunchang, Fujian Prov.	26°79'N	117°82'E
6421	Male parent	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 1	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 2	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 3	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 4	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 5	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 6	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 7	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 8	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 9	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 10	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 11	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 12	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 13	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E
P 14	Hybrid progeny of D110×6421	Shunchang, Fujian Prov.	26°79'N	117°82'E

[a]Experimental Forests were supplied by The National Forest Germplasm preservation base of Chinese fir (Yangkou National Forest Farm in Fujian Province, China). All tree needle samples were deposited in The Key Laboratory of Forest Genetics and Biotechnology of the Ministry of Education at Nanjing Forestry University, Nanjing 210037, China.

96-well thermal cycler (Applied Biosystems, Foster City, California, USA). Samples were incubated at 94°C for 5 min, followed by 20 touchdown cycles, first at 94°C for 45 s, then at Tm+10°C for 45 s (with a 0.5°C reduction in each subsequent cycle), and finally at 72°C for 1 min. Next, the samples were subjected to 20 cycles at 93°C for 45 s, then at Tm for 45 s, and finally at 72°C for 1 min. A final 10-min extension at 72°C was then performed. Tms are listed in Table S1 for different primers.

Fragments resulting from PCR amplifications were detected using 8% polyacrylamide gel electrophoresis (1×TBE buffer at 200 V for 1.5 h). The sizes of bands were compared with those of a 50-bp standard DNA ladder (TaKaRa Biotechnology, Dalian, China). For further verification of the accuracy of novel genic SSR polymorphisms, the forward primers of 22 selected SSR loci were labeled with one fluorescent dye (FAM or HEX) at the 5′end and tested in eight individuals. Then, we performed capillary electrophoresis using an ABI3730xl DNA Automatic Analyzer with a GeneScan-500LIZ size standard (Applied Biosystems). Based on data results, allele sizes were determined using GeneMaker software (Soft Genet-

ics, State College, PA, USA). Polymorphic amplicons detected using polyacrylamide gels with validation by capillary electrophoresis were separated using agarose gel electrophoresis and were then directly sequenced using Big Dye Terminator 3.1 (BDT3.1) in an ABI 3730xl sequencer (Applied Biosystems). Finally, we performed sequence alignment with DNAMAN 5.2.2 (Lynnon Biosoft, Point-Claire, QC, Canada) and MEGA5 software.

RESULTS AND DISCUSSION

A total of 5200 putative SSRs were identified in 4470 sequences from 62895 unigenes using MISA. SSRs occurred at an overall transcript density of 163.79 SSRs/Mbp. The low density of SSRs is similar to previous SSR studies that have generally measured lower SSR frequencies for conifers compared with other plants [22,23]. It is possible that the low SSR density in conifers is associated with the evolutionary rate and/or adaptive evolution level in large, long-lived conifer trees, leading to slow substitution rates and retention of beneficial genomic blocks and/or mutations [24-26].

The SSR repeat types were mostly trinucleotides (n=2,493; 47.94%), which may be attributed to tolerance for frame-shift mutations within coding regions [27], followed by tetra- (n=1,296; 24.92%), hexa- (n=623; 11.98%), penta- (n=357; 6.87%), di- (n=346; 6.65%) and mononucleotides (n=85; 1.63%; Fig. 1). Similar observations were made in previous EST-SSR *in silico* studies of *C. lanceolate* [10].

Among the mononucleotide repeats, A/T was dramatically overrepresented (n=83; 97.6%). The most abundant dinucleotide repeat was AG/CT (n=152; 43.93%), followed by AT/AT (n=132; 38.15%). Conversely, CG/CG was not found in the *C. lanceolata* transcriptome (Fig. 2A). This is consistent with most angiosperms studied, with AG/CT being the dominant EST-SSR dinucleotide repeat [28-30], but differs from most conifer research, including in the loblolly pine [22, 23], spruce [22], and sugi [11], in which AT/AT was the most abundant dinucleotide repeat. When comparing with AT/AT in most angiosperms, a greater percentage was observed in *C. lanceolate*. A trend of a decline in the frequency of AT/AT from gymnosperms to angiosperms was observed. However, the specific functions of SSR motifs within a specific plant genes/genome remain poorly understood.

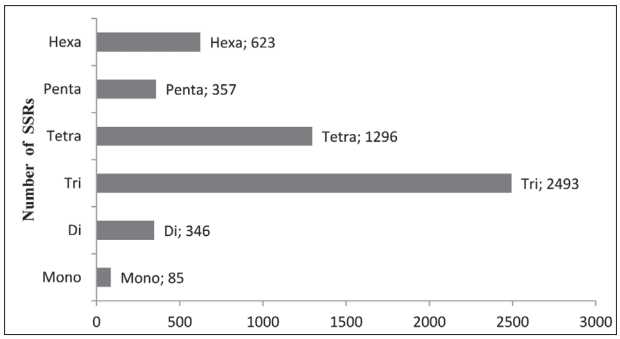

Fig. 1. Frequencies of the various SSR motifs present in *C. lanceolata*.

The most abundant trinucleotide repeat was AAG/CTT (n=625, 25.07%), followed by AGG/CCT (n=412; 16.53%), AGC/CTG (n=399; 16.00%) and ATC/ATG (n=358; 14.36%; Fig. 2B). The motifs AAAG/CTTT (n=255; 19.68%), AAAT/ATTT (n=185; 14.27%) and AATG/ATTC (n=167; 12.89%) were the most abundant tetranucleotide repeats (Fig. 2C), whereas AAGAGG/CCTCTT (n=39; 6.26%) was the most abundant hexanucleotide motif. Finally, although the pentamer repeats contained many different motifs, we did not find any significant patterns in their distribution in this research.

A total of 411 primer pairs were designed. Overall, 97 of the 411 primer pairs were polymorphic among eight genotypes, including *C. konishii* (Table S1 and S2). The other primer pairs were monomorphic or gave no product; therefore, we excluded them from further analysis. The polymorphism was observed in 28 EST-SSRs [11] and 10 genomic-SSRs [10] for *C. lanceolata*. By comparing with previously reported SSRs, all the 97 polymorphic makers were novel primer pairs of new loci, and all were submitted to the NCBI Probe databases (ID: from Pr032066750 to Pr032066846). This demonstrated that our method is more efficient for polymorphic EST-SSR development. The high rate of successful amplification of these 97

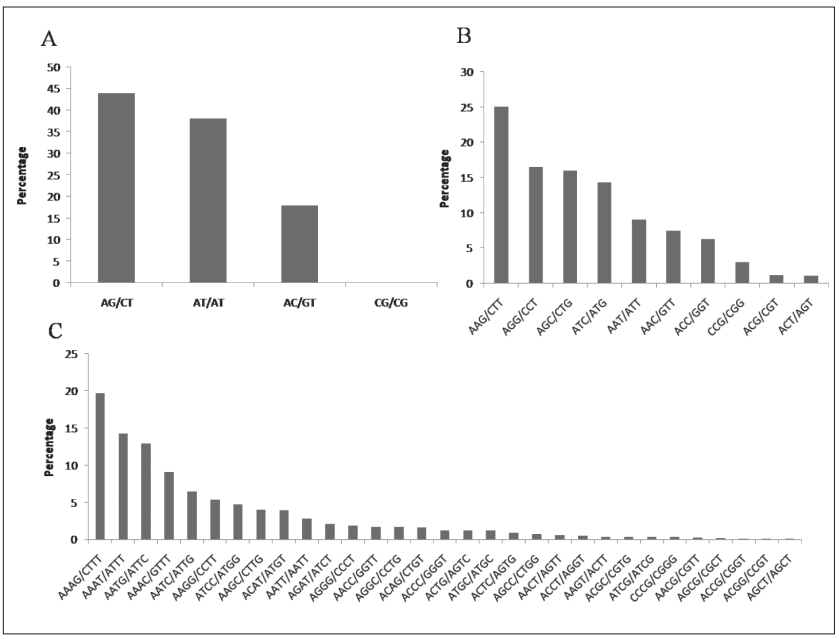

Fig. 2. Characterization and frequency of different motifs among dinucleotide repeat (A), trinucleotide repeat (B) and tetranucleotide repeat (C) EST-SSRs of C. *lanceolata*.

polymorphic primers in *C. konishii* was consistent with previous results describing only a few variations between *C. konishii* and *C. lanceolata* [31].

To verify the accuracy of novel genic SSR polymorphisms after polyacrylamide gel electrophoresis, 22 SSR loci were randomly selected for both capillary electrophoresis and automatic sequencing. In general, the results of three methods were similar (Fig 3). The PCR product size of only one locus (CFeSSR98 in Fig. 3A) from capillary electrophoresis appears smaller than that shown in the photograph of polyacrylamide gel electrophoresis. Fifty-four amplicons of 22 SSR loci were successfully sequenced with genic SSR primer pairs. Multiple sequence alignment reveals the occurrence of a variable number of repeat motifs in different amplicons of allele and expected sequences along with a few point mutations and insertion/deletions (Fig. 3IV).

Of the 97 polymorphic genic SSRs associated unigenes, 72 (74.23%) shared significant homology to known functions, whereas 25 (25.77%) had no significant match (Table S2). These results were in general agreement with previous work that demonstrated that 30%-49% of conifer genes had little or no sequence similarity with plant genes of known function [24, 32]. These genic SSR markers will be more likely to be linked to gene loci that contribute to morphological phenotypes than genomic SSR markers [5,10,12,33].

The number of alleles (Na) per locus of 97 markers (Table S2) ranged from 2 to 5, with an average of 2.64 alleles. In addition, the H_o values were 0.000-1.000 with an average of 0.479, and the H_E values were 0.125-0.775 with an average of 0.462. Lastly, the PIC values were 0.110-0.715 with an average of 0.383. Similar observations were made in previous studies of *Cryptomeria japonica*, for which the average PIC value was 0.33 [11]. However, the PIC was lower, with an average of 0.573, in previous EST-SSR research [10]. Differences in identified polymorphisms may be due to different marker techniques or from different plant populations. However, nearly 32% of the markers in a genetic diversity study contained high levels of genetic information according to the suggested criteria of high (PIC>0.5), moderate (0.25<PIC<0.5) and low (PIC<0.25) [34]. Of the remaining markers, 47% contained moderate and only 20% contained low levels of

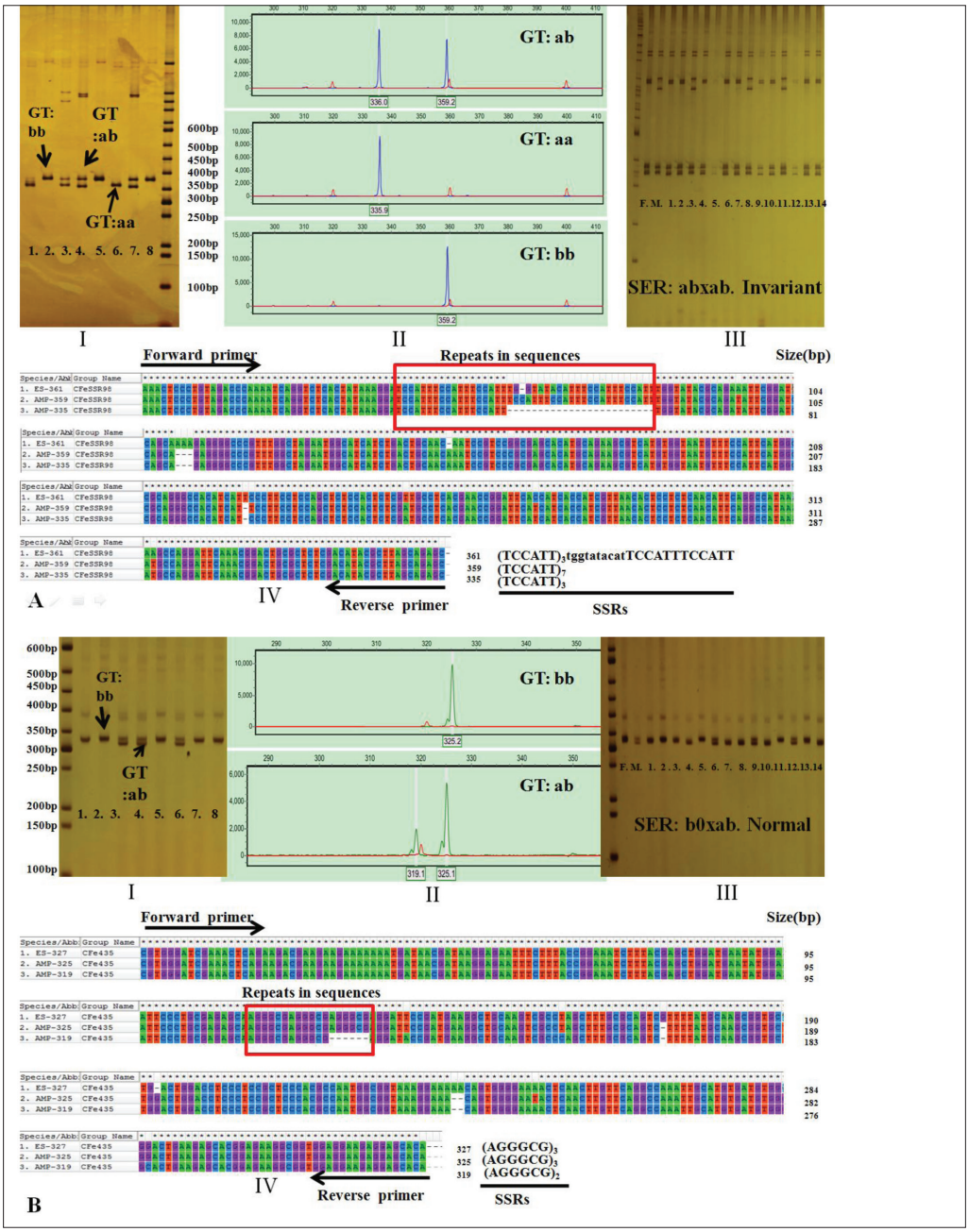

Fig. 3. PCR products amplified by two EST-SSR markers by three comprehensive methods. I: Representative gel showing amplification profiles of microsatellite marker (A: CFeSSR98; B: CFeSSR435) and its fragment length polymorphism among eight unique individuals of *C. lanceolata*. The amplicons are resolved in 8% polyacrylamide gel along with 50 bp DNA size standard. The order of DNA samples from lane 1 to lane 8 within each primer pair image panel is shown in Table 1. GT: Genotype of amplification. II: Representative Genotype of amplification in I (A: CFeSSR98; B: CFeSSR435) performed by capillary electrophoresis. III: Representative gel showing amplification profiles of microsatellite marker (A: CFeSSR98; B: CFeSSR435) among 14 progeny from a "D110" ×"6421" cross and their parents. SER: SSR allele segregation. IV: Multiple sequence alignment of expected sequences and amplicons of microsatellite marker (A: CFeSSR98; B: CFeSSR435) showing the presence of microsatellite repeat motif. Alignment reveals occurrence of variable number of repeat motifs in different amplicons of allele along with a few point mutations and insertion/deletions. ES: expected sequences; AMP: Amplicons of microsatellite marker and size of amplicons.

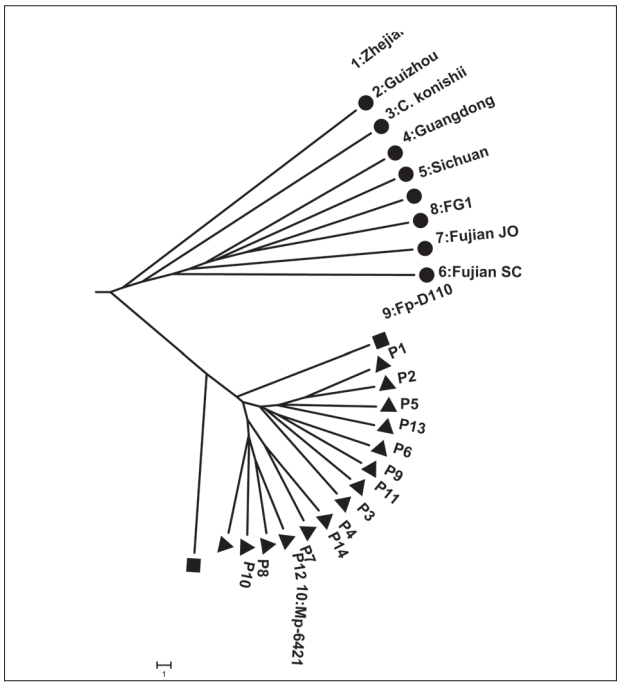

Fig. 4. Dendrogram generated using UPGMA cluster analysis based on the genetic diversity of 24 *C. lanceolata* genotypes.

genetic information. Thus, the SSR loci in this study have higher genetic information content than single biallelic markers, such as single nucleotide polymorphisms (SNPs) [35].

Twenty-nine loci were monomorphic in the segregation analyses of investigated population. In a chi-square test, only 5.88% of the 68 polymorphic markers undergo segregation distortion at the $p<0.05$ level (Tables S2). A dendrogram showed that the 24 *C. lanceolata* individuals fell into two distinct clusters (Fig. 4). In one cluster, 14 offspring were clustered with two parents. The other cluster consisted of 7 unique individuals from natural populations and *C. konishii*. Similar results were also observed for amplified fragment length polymorphism (AFLP) data from *C. lanceolata* and *C. konishii* [31].

Considering the relatively high polymorphic level, low segregation distortion rate, stable codominance and reproducibility, genic SSR markers in the present study will be useful for marker-assisted selection, conducting linkage mapping, quantitative trait locus (QTL) mapping and population genetic studies to improve breeding of *C. lanceolata*.

Transferable amplification of a randomly selected set of 30 genic SSRs showed that 30.0%-33.3% of SSRs could be amplified in *Metasequoia glyptostroboides* and *Glyptostrobus pensilis*, 16.67%-20.0% could be amplified in *Pinus massoniana* and *Cedrus deodara*, whereas 6.67% could be amplified in *Platycladus orientalis* and *Cephalotaxus fortunei* (Tables S3). In general, transferability decreased with increasing evolutionary distance between *C. lanceolata* and target species. Additionally, these markers will be useful tools for comparative genome mapping and evolutionary studies in conifer species in the future because of the greater transferability of gene-based SSRs compared with anonymous SSRs in related taxa.

Previous methods of SSR discovery have been tedious and labor intensive, expensive, and of low throughput [7,8]. Moreover, microsatellite markers from genomic libraries represent only those motifs for which the initial fitting was performed by hybridization or enrichment [6,9,36]. In addition, the low density and unique distribution of SSRs in *C. lanceolata*, as the results of this study show, suggest that it might be more challenging to develop SSR markers for this species by traditional methods. Alternatively, in our study, 97 novel polymorphic microsatellite markers, by mining the EST sequences, have been much more efficiently developed than in a previous study by traditional methods [10] in *C. lanceolata*. These markers will be beneficial for genetic diversity analyses, germplasm characterization, genomic mapping, marker-assisted breeding and evolutionary genetic analysis of *C. lanceolata*. Thus, our results also confirm that the identification of SSRs from the transcriptome is an efficient method for developing gene-based microsatellites for *C. lanceolata*.

Acknowledgments: This project was supported by the National Basic Research Program of China (2012CB114500); the National Science Foundation of China (No. 31170619); the Natural Science Foundation of Jiangsu University (No. 13KJA220001); the Program for New Century Excellent Talents by the Ministry of Education and the leading young scientist project from Ministry of Science and Technology; The Co-Innovation Center for Sustain-

able Forestry in Southern China, at Nanjing Forestry University; The Fujian Province Science and Technology Research funding on the Fourth Breeding Cycle Program of Chinese fir (No. Min Lin Ke 2013-1); the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD); and a Doctorate Fellowship Foundation of Nanjing Forestry University grant.

Authors' contributions: The experiments were conceived and designed by Jisen Shi, Yang Xu, Jinhui Chen and Renhua Zheng. Yang Xu, Zhanjun Wang, Ying Wang, Zhou Hong and Yaqi Zhao performed the experiments. Yang Xu, Ying Wang, Jisen Shi and Jinhui Chen analyzed the data of experiments. Renhua Zheng, Liwei Yang, and Ye Lu contributed the plant materials. Yang Xu and Renhua Zheng contributed equally to this work. Yang Xu and Renhua Zheng wrote the paper. Jisen Shi and Jinhui Chen revised the manuscript.

Conflict of interest disclosure: The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

1. Shi JS, Zhen Y, Zheng RH. Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. J Exp Bot. 2010;61(9):2367-2381.
2. Wang G, Gao Y, Yang L, Shi J. Identification and analysis of differentially expressed genes in differentiating xylem of Chinese fir (*Cunninghamia lanceolata*) by suppression subtractive hybridization. Genome. 2007;50(12):1141-1155.
3. Chen YW, Shi JS. Some fundamental strategies in the genetic improvement of Chinese fir (I). 1983;4:5-19.
4. Shi JS, Ye ZH, Hong JS, Yeh FC, Yang RC. Genetic diversity and multilocus structure in Chinese Fir. Journal of Nanjing Forestry University (Natural Science Edition). 1993;17:9-15.
5. Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 2005;23(1):48-55.
6. Moriguchi Y, Iwata H, Ujino-Ihara T, Yoshimura K, Taira H, Tsumura Y. Development and characterization of microsatellite markers for *Cryptomeria japonica* D.Don. Theor Appl Genet. 2003;106(4):751-758.
7. Zane L, Bargelloni L, Patarnello T. Strategies for microsatellite isolation: a review. Mol Ecol. 2002;11(1):1-16.
8. Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W. How much effort is required to isolate nuclear microsatellites from plants? Mol Ecol. 2003;12(6):1339-1348.
9. Li YX, Wang ZS, Sui JK, Zeng YF, Duan AG, Zhang JG. Isolation and characterization of microsatellite loci for *Cunninghamia lanceolata* (Lamb.) Hook. Genet Mol Res. 2015;14(1):453-456.
10. Wen Y, Ueno S, Han W, Tsumura Y. Development and characterization of 28 polymorphic EST-SSR markers for *Cunninghamia lanceolata* (Taxodiaceae) based on transcriptome sequences. Silva Genet. 2013;62(3):137-142.
11. Ueno S, Moriguchi Y, Uchiyama K, Ujino-Ihara T, Futamura N, Sakurai T, Shinohara K, Tsumura Y. A second generation framework for the analysis of microsatellites in expressed sequence tags and the development of EST-SSR markers for a conifer, *Cryptomeria japonica*. BMC Genomics. 2012;13:136.
12. Ellis JR, Burke JM. EST-SSRs as a resource for population genetic analyses. Heredity. 2007;99:125-132.
13. Wang Z, Chen J, Liu W, Luo Z, Wang P, Zhang Y, Zheng R, Shi J. Transcriptome characteristics and six alternative expressed genes positively correlated with the phase transition of annual cambial activities in Chinese Fir (*Cunninghamia lanceolata* (Lamb.) Hook). Plos One. 2013;(8):e71562.
14. Thiel T, Michalek W, Varshney R, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). Theor Appl Genet. 2003;106(3):411-422.
15. You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Gerard RL, J Dvorak, Anderson OD. Batch Primer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. 2008;9:253.
16. Yeh FC, T Boyle. Population genetic analysis of co-dominant and dominant markers and quantitative traits. Belg J Bot. 1997;129:57.
17. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980, 32: 314-331.
18. Van OJW. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet Res. 2011;93(5):343-9.
19. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978;89(3):583-90.
20. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011;28(10):2731-9.
21. Doyle JJ. DNA protocols for plants-CTAB total DNA isolation. Berlin: Springer –Verlag; 1991.
22. Berube Y, Zhuang J, Rungis D, Ralph S, Bohlmann J, Ritland K. Characterization of EST SSRs in loblolly pine and spruce. Tree Genet Genomes. 2007;3(3):251-259.
23. Echt CS, Saha S, Deemer DL, Nelson CD. Microsatellite DNA in genomic survey sequences and UniGenes of loblolly pine. Tree Genet Genomes. 2011;7 (4):773-780.
24. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013;497(7451):579-84.
25. Buschiazzo E, Ritland C, Bohlmann J, Ritland K. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. BMC Evol Biol. 2012;12.
26. King DG, Kashi Y. Mutability and evolvability: Indirect selection for mutability. Heredity. 2007;99 (2):123-124.

27. Metzgar D, Bytof J, Wills C. Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res. 2000;10:72-80.
28. Zhao Y, Williams R, Prakash CS, He G. Identification and characterization of gene-based SSR markers in date palm (*Phoenix dactylifera* L.). BMC Plant Biol. 2012;12:237.
29. Qiu L, Yang C, Tian B, Yang JB, Liu A. Exploiting EST databases for the development and characterization of EST-SSR markers in castor bean (*Ricinus communis* L.). BMC Plant Biol. 2010;10:278.
30. Li D, Deng Z, Qin B, Liu X, Men Z. De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). BMC Genomics. 2012;13:192.
31. Chung JD, Lin TP, Tan YC, Lin MY, Hwang SY. Genetic diversity and biogeography of *Cunninghamia konishii* (Cupressaceae), an island species in Taiwan: a comparison with *Cunninghamia lanceolata*, a mainland species in China. Mol Phylogenet Evol. 2004;33:791-801.
32. Parchman TL, Geist KS, Grahnen JA, Benkman CW, Buerkle CA. Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. BMC Genomics. 2010;11:180.
33. Yuan QH, Xie RH, Zhang ZC, Ma ZW, Li JQ, Li SL, Chen JB, Lü YH. Identification of expressed resistance gene analogs (RGA) and development of RGA-SSR markers in tobacco. Arch Biol Sci. 2015; 67(2):467-481.
34. Kumar Y, Ranjan A, Asif MH, Mantri S, Sawant SV, Tuli R. EST-derived SSR markers in *Jatropha curcas* L: Development, characterization, polymorphism, and transferability across the species/genera. Tree Genet Genomes. 2010;7:207-219.
35. Hamblin MT, Warburton ML, Buckler ES. Empirical comparison of Simple Sequence Repeats and single nucleotide polymorphisms in assessment of maize diversity and relatedness. PLoS One. 2007;2(12):e1367.
36. Araújo ES., Britto FB., Carvalho AMF, Orasmo GR., Diniz FM. Isolation and identification of a new set of microsatellite loci from Ucides cordatus genome. Arch Biol Sci. 2015; DOI: 10.2298/ABS141205115A

SUPPLEMENTARY MATERIAL (Available online)

Table S1. Characteristics of 97 polymorphic genic SSRs in *C. lanceolata.* The makers were comprehensively validated by polyacrylamide gel electrophoresis, capillary electrophoresis, and automatic sequencing.
Available at: http://www.serbiosoc.org.rs/arch/files/S1-S3/TableS1.docx

Table S2. Diversity statistics, segregation analysis statistics and functional annotations of the 97 genic-SSRs in *C. lanceolata.* Note: [a]: Heterozygosity base on a set of 8 unique individuals. Na = number of alleles detected; Ho = observed heterozygosity; He = expected heterozygosity; PIC = polymorphic information content. [b]: Segregation analysis statistics base on a set of 14 progeny from a "D110" ×"6421" cross and their parents. **: Derived progeny displayed segregation distortion at the $p<0.05$ level.
Available at: http://www.serbiosoc.org.rs/arch/files/S1-S3/TableS2.docx

Table S3. Cross- species transferability of 30 genic-SSRs.
Available at: http://www.serbiosoc.org.rs/arch/files/S1-S3/TableS3.docx