ONLINE RESOURCES

Development and characterization of genomic microsatellite markers in *Tinospora cordifolia*

MADHURANJANA GARGI, SAPNA THAKUR, SHASHI SHEKHAR ANAND, SHRUTI CHOUDHARY and PANKAJ BHARDWAJ*

Molecular Genetics Laboratory, Centre for Plant Sciences, Central University of Punjab, Bathinda 151 001, India

[Gargi M., Thakur S., Anand S. S., Choudhary S. and Bhardwaj P. 2017 Development and characterization of genomic microsatellite markers in *Tinospora cordifolia*. *J. Genet.* **96**, e25–e30. Online only: http://www.ias.ac.in/jgenet/OnlineResources/96/e25.pdf]

Introduction

Tinospora cordifolia (Giloy, Guduchi, Amrita and heart-leaved moonseed, $2n = 22$) is a herbaceous vine, native to the tropical areas of India, Sri Lanka and Myanmar (Rana *et al.* 2012). It is a large, perennial, deciduous massively spreading climbing shrub with succulent stem. It is a well-known medicinal plant, widely used as a key component in various plant-based drugs and herbal remedies. In ancient ayurveda system, this herb has been prescribed in the treatment of various diseases, i.e. diabetes, jaundice, fever, dyspepsia, skin diseases, urinary problems, chronic diarrhoea and dysentery (Rege *et al.* 1999; Ahmad *et al.* 2009). The genetic data available on this species is insufficient to meet the future productivity and availability of plants for drugs, towards conservation of genetic resources and evaluation of the genetic impacts on cultivar productions. Understanding the genetic diversity and population structure of a species is essential for its conservation and effective management (Ayad *et al.* 1997). There has been a convincing increase in the appliance of molecular genetic methods for appraising the conservation and usage of plant genetic resources. Codominant nature of SSR markers allow the revelation of a high number of alleles per locus and also contributes to higher levels of expected heterozygosity (Morgante and Olivieri 1993) and therefore, often employed in the genetic studies.

In the present study, we focussed on the development of genomic microsatellite marker for medicinally important plant species, *T. cordifolia* and their characterization across its genotypes, collected from Indo-Gangetic plains. Based on biotin–streptavidin hybridization technique, $(AG)_{10}$ enriched library was constructed. Thirty-one microsatellite markers were developed, of which 23 were successfully characterized. A total of 76 alleles were identified with a mean of 3.304 alleles per locus and a mean of polymorphic content of 0.4633 per locus. Twenty-two loci showed significant deviations from Hardy–Weinberg equilibrium (HWE).

Materials and methods

A total of 30 genotypes of *T. cordifolia* were collected from different regions of Indo-Gangetic plains (figure 1). Details of sample collection along with their GPS location are provided in table 1. Genomic DNA was isolated from young leaves using the modified CTAB procedure (Doyle 1990), quantity and quality were estimated on Nanodrop 2000 Spectrophotometer and 0.8% agarose gel.

A library enriched for the dinucleotide repeat $(AG)_n$ was constructed from the whole genome using restriction endonuclease: *Eco*R1 and *Mse*1 (Bhardwaj *et al.* 2013). Following enzyme digestion and adapter ligation, and amplification using respective primers, the fragments within the size range 200–1000 bp were eluted. The selected fragments were then enriched by biotinylated $(AG)_{10}$ probe, amplified and cloned through pGEM-T EASY vector into chemically competent DH-5α (*Escherichia coli* strain). Positively transformed cells were screened and selected by blue–white screening (following β-galactosidase interruption), and confirmed by secondary enrichment through colony PCR amplification. Clones amplifying multiple bands were selectively grown overnight in LB broth containing 100 μg/mL ampicillin

*For correspondence. E-mail: pankajihbt@gmail.com.

Keywords. microsatellite markers; genetic diversity; *Tinospora cordifolia*.

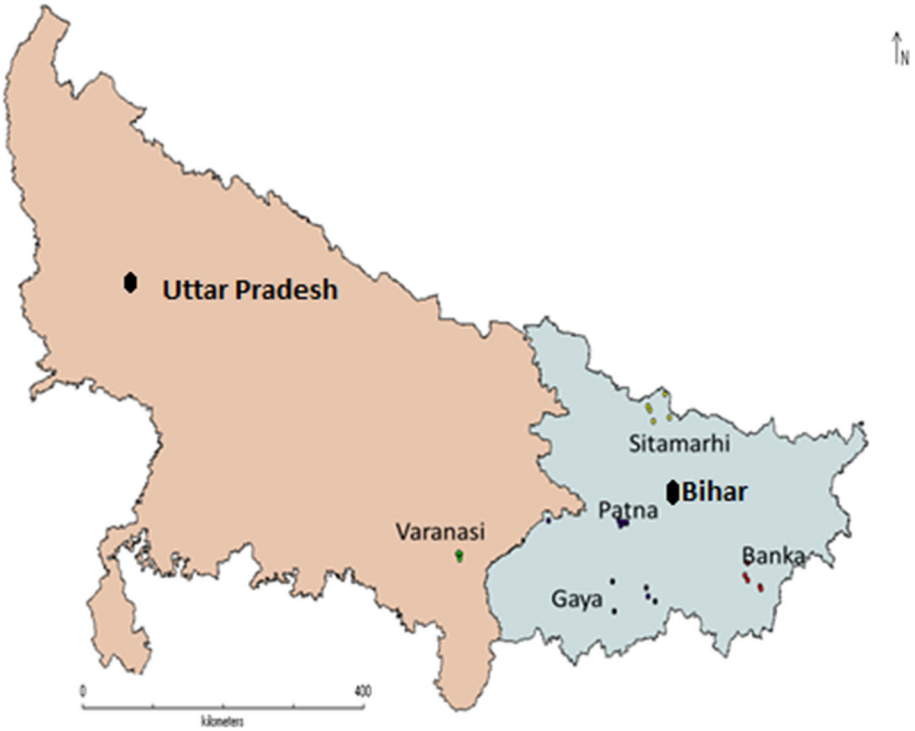

Figure 1. Sampling sites on the map of Bihar and Uttar Pradesh, India.

Table 1. Geographical locations of *T. cordifolia* populations from north India.

Individual ID	Region	Site	Latitude	Longitude
1	Patna	GM Road	25.6160634	85.1597281
2	Patna	Vivekanand Marg	25.6206403	85.1120348
3	Patna	Gayaghat	25.6322245	84.142854
4	Patna	Sheikhpura	25.6058695	85.0921648
5	Patna	ShriKrishna Puri	25.6096881	85.118112
6	Patna	Darbhanga House	25.621648	85.164382
7	Patna	BIT Campus	25.5961217	85.0876228
8	Patna	Gola Road	25.6388389	85.0547072
9	Patna	Mainpura	25.6274892	85.1207061
10	Patna	Phulwarisharif Jn.	25.586539	85.079754
11	Banka	Kaushalpur	25.0290341	86.8897655
12	Banka	Kastikri	25.1241855	86.7001311
13	Banka	Sambhuganj	25.0804267	86.730321
14	Banka	Sultanganj Jn.	25.241744	86.735822
15	Banka	English More	24.9960278	86.9054306
16	Sitamarhi	Hanuman Chawk	26.811389	85.661389
17	Sitamarhi	Baghari	26.5927	85.7192
18	Sitamarhi	Dumra	26.559722	85.504722
19	Sitamarhi	Riga	26.658	85.4681
20	Sitamarhi	Nazarpur	26.6943	85.4407
21	Gaya	Rajgir, Nalanda	25.0172473	85.4161642
22	Gaya	Nawada, Hilsa	24.8910775	85.5354724
23	Gaya	Makdumpur, Gaya	25.0655674	84.9739521
24	Gaya	Nagmita colony	24.7954523	84.999431
25	Gaya	NH 82, Nardiganj	24.9325215	85.436039
26	Varanasi	Lahurabir	25.3217454	83.0015568
27	Varanasi	Lanka	25.2814092	82.9978626
28	Varanasi	Railwayganj Colony	25.32742	82.9860535
29	Varanasi	Rasulghad, Var. Cantt.	25.3272698	82.9713249
30	Varanasi	Chaukaghat	25.3328555	82.9878044

Table 2. Characteristic of 23 SSR loci developed from *T. cordifolia*.

S.no	Locus name	Primer sequences	Repeat motif	T_a (°C)	N_a	Heterozygosity		PIC	I	Approx. size range (bp)	No. of genotypes amplified	Accession no.
						H_o	H_e					
1	TC1	F-5′ATTACGTCTTGTTTTGCATTC R-5′AGAGACTGAGAGGGAGAGTTG	$(CT)_6$	54	4	0.3333***	0.7028	0.635	1.2616	153	18	KX602220
2	TC2	F-5′TTTGCCACATATTTACCACTT R-5′CTAGTGATTGGAGGGAGGA	$(TC)_{10}$	54	3	0.1333***	0.5605	0.456	0.8823	150	14	KX602221
3	TC3	F-5′CGAGAGAGAGATTGTGAGAGA R-5′CGACTTCATCTTCTTCCAGA	$(GA)_7$	54	3	0.2000***	0.5672	0.478	0.9183	167	20	KX602222
4	TC5	F-5′CTACATCTTCTTCCAGAACTCC R-5′GCGAGAGAGAGAGAGAGAGAG	(TC)7	54	5	0.1333***	0.5718	0.532	1.1493	154	11	KX602224
5	TC6	F-5′ACAGAGAGAGAGAGAGAGAGAGA R-5′TAGTTGTTGCAAAAAGCCATT	$(AG)_{11}$	54	3	0.1333***	0.4249	0.370	0.7299	103	8	KX602225
6	TC7	F-5′GATCCTCAAAAGTGAAAACTC R-5′GCCAGAGAGAGAGAGAGAGAG	$(TC)_{11}$	53	4	0.4000***	0.5175	0.466	0.9510	143	26	KX602226
7	TC8	F-5′TGAGTGATGAGTCCTGAGTAAA R-5′TGAGAGAGAGAGAGAGTCGAA	$(TC)_{10}$	54	4	0.5333***	0.6706	0.591	1.1384	159	21	KX602227
8	TC9	F-5′AGAGAGAGGTGACGTAGGAAT R-5′TAATGGGAAAGGGTGTAGAAT	$(GA)_8$	54	3	0.7667***	0.6503	0.563	1.0542	162	24	KX602228
9	TC10	F-5′ACTAGTGAGTGACTGCGTACC R-5′AGAGAGAGAGAGAGGAATTGG	$(CT)_{10}$	53	3	0.3667***	0.4062	0.356	0.7073	152	28	KX602229
10	TC12	F-5′ATTCGATTCAGGACTCATCAT R-5′TCTCTGATCGAATTCCCTAGT	$(CT)_6$	55	3	0.5333***	0.5898	0.513	0.9730	137	25	KX602231
11	TC14	F-5′ACTAGTGATTGACTGCGTACC R-5′AGTGAGATTGATGAGTCCTGA	$(CT)_{10}$	54	3	0.4667***	0.6350	0.554	1.0386	147	22	KX602233
12	TC15	F-5′TGAGATTGATGAGTCCTGAGT R-5′CGACTGTGTACCAATTCTACC	$(AG)_{10}$	54	3	0.1333***	0.4994	0.433	0.8401	176	10	KX602234
13	TC16	F-5′AGTGAGATTGATGAGTCCTGA R-5′ACTAGTGATTGACTGCGTACC	$(AG)_{10}$	54	2	0.4333^{NS}	0.3452	0.282	0.5227	134	30	KX602235
14	TC17	F-5′CCTGAGTAATATGGACAGTGC R-5′AAGCTGATTCAAGGAGAAAGT	(AG)12	54	4	0.3000***	0.5000	0.452	0.9268	155	26	KX602236
15	TC19	F-5′CTCATTGCTCACAACAAACTA R-5′AGAGAGAGAGAGAGAGGGTGA	(CT)8	53	3	0.4667***	0.4994	0.433	0.8401	147	27	KX602238
16	TC22	F-5′TGGTACGCAGTCAATCACTA R-5′ACTCAGGACTCATCAATCTCA	(AG)10	54	3	0.4000***	0.5695	0.499	0.9503	122	24	KX602241
17	TC23	F-5′TGGACAGTGTAGAGAGAGAGAG R-5′CTAGTGATTTGACTGCGTACC	(AG)10	53	4	0.6667***	0.7299	0.664	1.3101	125	20	KX602242
18	TC24	F-5′AGAGAGAGAGGAATTGGTACG R-5′TCAGGACTCATCAATCTCACT	(AG)10	54	3	0.4667***	0.5605	0.485	0.9276	161	26	KX602243

Table 2 (*contd*)

S.no	Locus name	Primer sequences	Repeat motif	T_a (°C)	N_a	Heterozygosity		PIC	I	Approx. size range (bp)	No. of genotypes amplified	Accession no.
						H_o	H_e					
19	TC27	F-5′CGCGGGAATTCGATTAGAG R-5′CAAAACAAACCAAGTCTCTCTCTC	(AG)8	58	3	0.1000***	0.1881	0.177	0.3944	153	3	KX602246
20	TC28	F-5′GCGGGAATTCGATTAGAGAG R-5′TGGCATAATTGCCGTAGGAT	(AG)7	58	3	0.1333***	0.5808	0.486	0.9334	130	19	KX602247
21	TC29	F-5′CACTAGTGAATTCGATTAGAGAGAGA R-5′CCGTCCCTTGAATCTCAGTG	(GA)5	58	4	0.2000***	0.4198	0.385	0.8147	179	26	KX602248
22	TC30	F-5′GCGGGAATTCGATTAGAGAG R-5′TGCGTACCAATTCAATGGTCT	(AG)7	58	3	0.4000***	0.3774	0.323	0.6390	162	29	KX602249
23	TC31	F-5′TGGACAGTGTAGATGAGAGAGAGAG R-5′GCGTTGGTGGTCAATTAGGT	(GA)5	57	3	0.8667***	0.6169	0.523	0.9934	112	26	KX602250
	Mean				3.3043	0.3725	0.5297	0.4633	0.9085			
	SD				0.6350	0.2129	0.1263		0.2162			

T_a, annealing temperature; N_a, total number of alleles; H_o, observed heterozygosity; H_e, expected heterozygosity; PIC, polymorphic information content.
Significant deviations from HWE at *** $P < 0.001$; ^{NS} nonsignificant.

at 37°C. Plasmid DNA from the overnight culture was isolated using Thermo Plasmid miniprep kit, and sequenced with ABI 3730 xl DNA Analyzer using BigDye terminator cycle sequencing kit v3.1 (Applied Biosystems, Foster City, USA) as per the manufacturer's protocol. Sequences containing SSRs were analysed using SSR identification tool (SSRIT). Plasmid chimeras were trimmed from sequenced fragments using Gene Runner and primers were designed from region flanking SSRs using Batch Primer3 v1.0 software and amplified in a reaction volume of 20 μL, consisting of 1× Taq buffer (1 mM Tris pH 9.0, 50 mM KCl, 0.01% gelatin, 1.5 mM $MgCl_2$), 25 ng template DNA, primer 2.5 mM of dNTPs mix, 5 ng each of forward and reverse primers. Thermal profile constituted of a 94°C predenaturation step for 3 min proceed by 35 cycles of denaturation at 94°C, annealing at T_a (specific for each loci as mentioned in table 2 and elongation at 72°C for 1 min, respectively, and a 8 min final elongation step at 72°C. The developed primer pairs were then characterized among 30 genotypes on 6% denaturing gel and visualized using silver stain (Thakur *et al.* 2016).

On the basis of genetic distance, genetic diversity was estimated at different parameters, namely, effective number of alleles (N_e), observed and expected heterozygosity (H_o, H_e), Shannon's information index *(I)* using PopGene software ver. 1.31 (Yeh *et al.* 1999). The Structure ver. 2.3.4 (Pritchard *et al.* 2000) and DARwin ver. 5.0.158 (Perrier and Jacquemoud-Collet 2006) was used for deciphering phylogenetic relationship among characterized individuals. The clusters of hidden population were distinguished on the basis of multilocus Bayesian analysis. For evaluation of optimum number of populations (K), a simulation was coordinated using parameters; K (1–20) with a random start for each K value and 10 independent runs (Evanno *et al.* 2005). Number of distinct population clusters *(K)* was determined using a web-based program, Structure Harvester (Earl 2012), for envisioning structure output and implementing the Evanno method.

Results and discussion

A total of 1800 recombinant clones isolated from $(AG)_{10}$-enriched genomic library were screened for presence of insert using secondary enrichment through PCR amplification. On sequencing of 380 (21%) recombinant clones exhibiting presence of insert in secondary enrichment, 116 (31%) were found to be positive for SSRs, which show the success of enrichment procedure. These 116 SSRs containing sequences have dinucleotide repeats of length varying from 6–14 bp. Hundred per cent of SSRs constitute dinucleotide motif of AG/GA or CT/TC as the genomic library was enriched using $(AG)_{10}$. Seventy six (65%) sequences could not be used for designing primer pairs either due to short sequences or presence of SSR loci

Table 3. Statistic for library enriched for G-SSR development for *T. cordifolia.*

Fact	Total
Total number of positive colonies obtained after blue–white screening	1800
Number of clones sequenced (percentage of the whole library)	380 (21%)
Clones containing SSR loci (percentage of clones sequenced)	116 (31%)
Nonsuitable SSR clones (percentage of clones sequenced)	76 (65%)
Unique SSR clones (percentage of clones sequenced)	40 (34.5%)

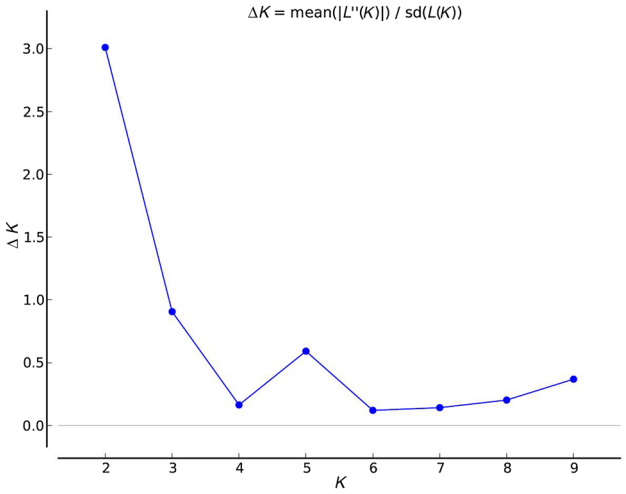

Figure 2. Most appropriate value of K using the second order statistics (ΔK) given by Evanno *et al.* (2005).

closed to trimmed ends. With the remaining 40 sequences, 54 primer pairs were designed, of which 23 were found to be homologous and in remaining 31 primers, unique primer pairs were developed and synthesized. A statistic of genomic library constructed for *T. cordifolia* is represented in table 3. Of the 31 unique primer pairs, 30 showed successful amplifications in selected germplasm. Seven (23.3%) of them showed monomorphic amplifications, and the other 23 (76.6%) were polymorphic. The polymorphic potential of the developed primer pairs is represented in table 2. Number of alleles ranged from 2 to 5 with an average of 3.3043 having an approximate fragment size ranges from 103 to 179 bp. The observed and expected heterozygosity values varied from 0.1333 to 0.8667 (average of 0.3725) and from 0.1881 to 0.7299 (mean of 0.5297). Polymorphic information content and Shannon's information index varied within the range of 0.177–0.664 (average of 0.4633) and 0.3944–1.3101 (with a mean of 0.9085), respectively (table 3). Eight loci were found highly informative (PIC > 0.5) and other loci were with moderate PIC (PIC > 0.25). Twenty two of the 23 loci showed significant deviations from HWE (table 2). Structure based on Bayesian model and unweighed neighbour-joining (UNJ) tree constructed using DARwin exhibited the presence of two clusters (figures 2 and 3). As established from ΔK distribution, the graphical method detected the highest value of ΔK at $K = 2$ indicating true number of subpopulations (K) present. Clustering or grouping of the

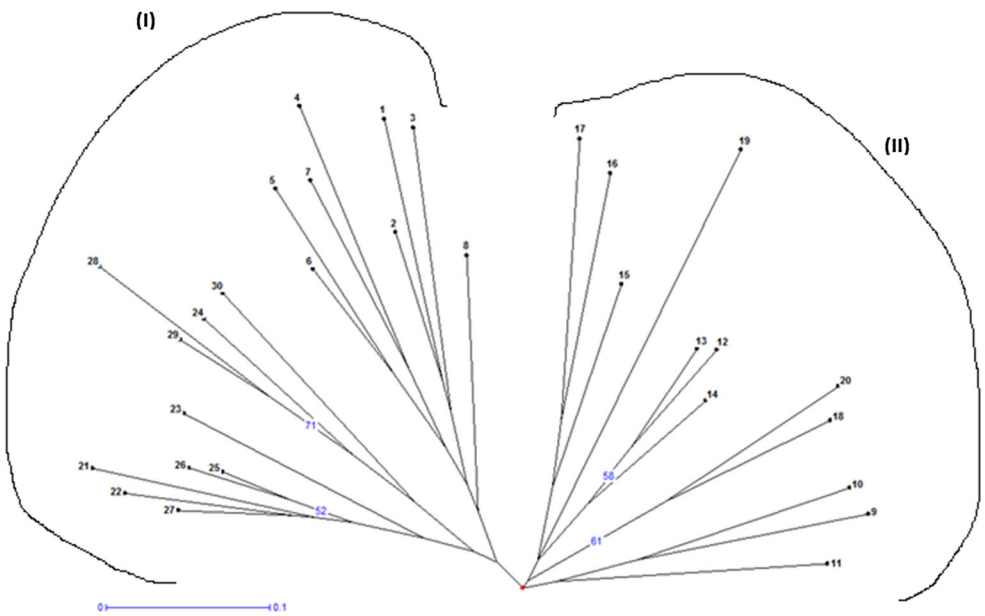

Figure 3. UWN tree of 30 individuals as constructed by DARwin 6.0.13; each branch represents single individual.

populations was done on the basis of genetic similarity among the groups irrespective of geographical locations of sampling. Most of the individuals from regions Patna, Gaya and Varanasi formed cluster (I) and individuals from Banka and Sitamarhi were grouped under cluster (II) (figure 3). Results of the present study indicate presence of two genetic subpopulation for *T. cordifolia* in northern India. The polymorphic loci developed in the present study adds on the available genetic resource for this species and will help in understanding its genetic diversity and population structure. These markers will be helpful in planning strategies for effective management and conservation of this medicinally important species.

Acknowledgment

Sapna Thakur and Shruti Choudhary acknowledge the fellowship received from ICMR towards Ph.D.

References

Ahmad S., Hoot S., Qazi P. and Verma V. 2009 Phylogenetic patterns and genetic diversity of Indian Tinospora species based on chloroplast sequence data and cytochrome P450 polymorphisms. *Plant Syst. Evol.* **281**, 87–96.

Ayad W. G., Jaradat T. and Rao V. R. 1997 *Molecular genetic techniques for plant genetic resources: report of an IPGRI Workshop, 9–11 October 1995*. IPGRI, Rome, Italy.

Bhardwaj P., Kumar R., Sharma H., Tewari R., Ahuja P. S. and Sharma R. K. 2013 Development and utilization of genomic and genic microsatellite markers in Assam tea (*Camellia assamica* ssp. *assamica*) and related Camellia species. *Plant Breeding* **132**, 748–763.

Doyle J. J. 1990 Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15.

Earl D. A. 2012 STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361.

Evanno G., Regnaut S. and Goudet J. 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620.

Morgante M. and Olivieri A. 1993 PCR-amplified microsatellites as markers in plant genetics. *Plant J.* **3**, 175–182.

Perrier X. and Jacquemoud-Collet J. P. 2006 DARwin software: dis-similarity analysis and representation for windows (accessed 1 March 2013) (http://darwin.cirad.fr/).

Pritchard J. K., Stephens M. and Donnelly P. 2000 Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Rana V., Thakur K., Sood R., Sharma V. and Sharma T. 2012 Genetic diversity analysis of *Tinospora cordifolia* germplasm collected from northwestern Himalayan region of India. *J. Genet.* **91**, 99.

Rege N. N., Thatte U. M. and Dahanukar S. A. 1999 Adaptogenic properties of six rasayana herbs used in Ayurvedic medicine. *Phytother. Res. PTR* **13**, 275–291.

Thakur S., Choudhary S., Singh A., Ahmad K., Sharma G., Majeed A. and Bhardwaj P. 2016 Genetic diversity and population structure of *Melia azedarach* in north-western plains of India. *Trees* 1–12.

Yeh F., Yang R. and Boyle T. 1999 POPGENE Microsoft windowsbased freeware for population genetic analysis: release 1.31 (Accessed 1 May 2014). University of Alberta, Edmonton (https://sites.ualberta.ca/~fyeh/popgene_download.html).

Received 1 September 2016; accepted 13 October 2016
Unedited version published online: 18 October 2016
Final version published online: 26 June 2017

Corresponding editor: Rajiva Raman