Isolation and characterization of a first set of nine polymorphic microsatellite loci in *Pongamia pinnata* (Fabaceae)

SHYAM SUNDAR SHARMA[1], MD AMINUL ISLAM[2], MADAN SINGH NEGI[2] and SHASHI BHUSHAN TRIPATHI[2]*

[1]*TERI University, 10 Institutional Area, Vasant Kunj, New Delhi 110 070, India*
[2]*The Energy and Resources Institute, IHC Complex, Lodhi Road, New Delhi 110 003, India*

[Sharma S. S., Islam M. A., Negi M. S. and Tripathi S. B. 2014 Isolation and characterization of a first set of nine polymorphic microsatellite loci in *Pongamia pinnata* (Fabaceae). *J. Genet.* **93**, e70–e74. Online only: http://www.ias.ac.in/OnlineResource/93/e70.pdf]

Introduction

Pongamia pinnata (L.) Pierre, is nonedible oil producing tree legume, which has been recognized as a major biodiesel species in India (Tewari 2003). *P. pinnata* (synonym *Millettia pinnata*) is indigenous to India and Southeast Asia from where it has spread to other parts of the world. Its seed oil content ranges from 24–40% which is at par with other biodiesel species such as *Jatropha*. However, availability of any improved and characterized planting stock has been the major bottleneck in harnessing the biofuel potential of this plant. A large proportion of existing trees do not flower at all and commercially attractive levels of fruiting are observed in only a small fraction of trees. However, there is a large phenotypic diversity in this species, thus providing an opportunity for genetic improvement (Kaushik *et al.* 2007). More recently, initiatives have been taken towards identification of superior genotypes and their characterization.

Characterization of genetic diversity is a prerequisite for efficient conservation and utilization of genetic resources. A number of studies have been conducted on Pongamia using dominant markers such as RAPD, ISSR, AFLP and TE-AFLP (Kesari *et al.* 2010; Sahoo *et al.* 2010; Sharma *et al.* 2011). However, the information content of these dominant markers is less due to presence of only two allelic states. Codominant markers such as microsatellites can display high number of allelic states thus providing higher information per data point. However, these markers need to be developed in species of interest through generating sequence information and primer development. To best of our knowledge, no microsatellite markers have been reported in this species till date. The aim of the present study was to isolate and characterize polymorphic microsatellite markers as to facilitate genetic diversity, clonal identification and conservation in *P. pinnata* germplasm.

*For correspondence. E-mail: sbhushan@teri.res.in.

Materials and methods

Genomic DNA was isolated from lyophilized leaves following modified CTAB-based procedure (Singh *et al.* 1999) from 24 selected accessions (12 accessions with less than 25% and 12 with more than 35% oil content) from National Capital Territory (NCT) of Delhi and other Indian states. Microsatellite loci isolated according to the fast isolation by AFLP of sequences containing repeats (FIASCO) protocol (Zane *et al.* 2002) with minor modifications.

Genomic DNA was simultaneously digested and ligated using *Mse*I restriction enzyme, *Mse*I adapter and T4 DNA ligase (New England Biolab, Massachusetts, USA) at 37°C for 2 h. The library was amplified using *Mse*I primer without any selective nucleotide. The genomic library was enriched for microsatellites using dinucleotide (GA/CT and CA/GT), trinucleotide (GGA/CCT, GAT/CTA, CTT/GAA, TAA/ATT) and tetranucleotide (GACA/CTGT and CATA/GTAT) repeats by hybridization with repeat containing biotinylated oligonucleotide probes. Fragments containing SSR regions were captured with streptavidin-conjugated magnetic beads (Promega, Madison, USA). A total of 108 positive clones containing insert size ranging from 300–800 bp were selected for sequencing (Macrogen, Seoul, Korea). Sequences containing microsatellite repeats were used to design primer pairs from their flanking regions using program BatchPrimer 3 (http://batchprimer3.bioinformatics.ucdavis.edu/cgibin/batchprimer3/batchprimer3.cgi).

Initially, annealing temperature and polymorphism were tested for each primer pair. Nine polymorphic primer pairs selected from this screening were used for genotyping 24 accessions using indirect fluorescent labelling of fragments through a three-primer system (Schuelke 2000). All forward primers were tagged with M13 sequences (5′-TGTAAAACGACGGCCAGT-3′) at 5′ end. The typical PCR mix contained 8 pmol each of the special forward and

Keywords. microsatellites; biodiesel; FIASCO; polymorphic; *Pongamia pinnata.*

normal reverse primer and 2 pmol of FAM-labelled M13 (−21) primer in a 15 μL reaction containing standard 1× PCR reaction buffer, 0.2 mM dNTPs, and 500 ng template DNA, 1 U Taq DNA polymerase (Gene-Bio, Daejon, Korea). PCR amplification conditions were as follows: 94°C for 5 min followed by 30 cycles of 94°C (30 s) / 56°C (45 s) / 72°C (45 s), followed by 8 cycles of 94°C (30 s) / 53°C (45 s) / 72°C (45 s), and a final extension at 72°C for 10 min. The PCR products run on an ABI 3730 Prism Genetic Analyzer (Applied Biosystems, New York, USA). Peak detection and allele scoring were done using STRand software ver. 2.4.59 (Toonen and Hughes 2001). Heterozygosity, Hardy–Weinberg equilibrium (HWE), linkage disequilibrium (LD) and polymorphic information content (PIC) were estimated using power marker ver. 3.25 (Liu and Muse 2005).

Results and discussion

A total of 39 microsatellite containing sequences were isolated and submitted to the GenBank with accession numbers GQ130157–GQ130195. Nineteen of these sequences having repeats with enough flanking nucleotides were used for primer designing, and clone sequences of nine polymorphic microsatellite loci are presented in figure 1. The nine polymorphic microsatellite loci identified in this study showed a clear and single peak for each allele, while remaining 10 markers showed no polymorphism across accessions (table 1). The amplified products were in agreement with the expected size from sequence data for all nine microsatellites analysed (figure 2). The loci exhibited two to five alleles per locus with an average of 3.5 (table 2). A total

>GQ130182
gatgagtcctgagtaatttgctccaaccatgccggaaatgcaacggcgacgcctctgcctctcagatgtcgtcatctctgttgtcaa
gaactcaaccacgccgtttaccgtcgatgttgtgaaacagagcacctccgtgtcttggatctcgagtacctaaagagatgggtgg
ggctggcagtagcatcgaacgacag(For)cagcagtgttgtcattgacgtcgaagttgtatgtaacggtggttgccgaaggaa
aggaaagaggaagcacggggaaaactaataggtttgcggtgaagaaagaagaaacaagagagagagagagagagggag
agagagagagttttctatagccgtggtcgagagaagaggagga(Rev)gtcaaaagtgagaggagaagagggaatgtgaag
gtgagtgttactcaggactcatc

> GQ130187
gatgagtcctgagtaaggacctgcagtcttctgcttcagagggaggactaacttctgagcgcggatccagcacgaaggagattc
ctccccactcttcctct(For)accgagatggaggtaaactcatccgatgatgaggatactgccatgaaagacatagaggaggat
gaggaagaggtagagacagaggaggaggaggaggtagaggctgaggagggctcgagcgtgctgataatgatgtcatc
tacagagaccgaggaggaggaggacctttcagaggatggaggaggggaggagtggccttgatcactagttttctcttgggttga
gattctagatgcctagagtagagtactatcgcacagtgggattatagaccctgtggcacttctcgtgta(Rev)gttagtcactgct
ctaagtcttgcgttagtacttgtggggctcacgtattgttatgtttattatgtatatatgtgtatctactatgctgaccaggtgtgatcagt
ggacagttgttcctagtgtgtgagtatgctgtaatattgtattagttactcaggactcatc

> GQ130160
gatgagtcctgagtaagtgcttatcggtgcaagacaacaaca(For1)gtgaaaaggaggaccttcctgatggcagtggtagtg
atgcttgtgagagagagagagagagagagagagagagagagagagagacaagacaatgttagagagtgtagaaacaatggg
caaagagagaaagagagagcttctcgacagcaatggttgtggcgacagtgcttttag(For2/Rev1)tggcagccattcttttc
gcagcagcaaagcgagagagagagagagataaagagagaaatggagagatagatagagagattagaaggcagagaggag
tttctaggagagagattgcaaggcaaagaggtgagtttggaagagagagagagctcaggagggtttgagggagaga(Rev2)
gagggttagagaaaataacatttcagagaatggagaagaagaagctacacccaagttactcaggactcatc

> GQ130188
gatgagtcctgagtaacaagcaaccaaatggcacctcacaaggtacag(For)agaccgcatacaaccttcaattcaacactatc
tacatgctcacacacacacacacacacacacacacacacacacacacacacaaaaatactacaactcaaaaaccaaaaaggtagg
gaagaagaaggaacccagataggagacaacgattagagacacatataaccaaaaactagaatcagaaagcaacaactgcag
a(Rev)tctcatcacccttgcaccgacccacataaaatcagtaaaatgagcataaagcagataaaggttgaaaattgagttacc
aatggtgtgtagagtcccatagctatatagtgatacatcttcttcccaacaaaaggagcaaagaacacatagaaagcaaatccca
gagccaaaaaaacagcaatcgccaccacctactcaggactcatc

> GQ130173
gatgagtcctgagtaagttagtgttgttttacaccaacctcccagagaaactagagtggcgtggtttgggaaggggtggaggaaa
agagttggaggagaaggtcggtgaaagcacgaacgatgtatttgtgggtatgtttgggattgagagcgaggtagcataagagg
agttcgtggaggtaatcccagtcttttgtgacgtcgaagacttgttggcgagaatggatcatttcctgcattgaatggaggaaatcg
acgtaagggttgagggagtacttggggactctgacgccgccacttgttgttggtgtgaaggagcgagtgtca(For)ggggact

Figure 1 (*continue*)

cgatgatggagttggaggtgccaggggaggagaagaagaagcgttgggaggcgaagatggaggcgaagtcgggttctgcg
gaggaggaggaggaagtggaggtggaggtggaggagaagtggtcgtcggcggtgaaatggaacggggtggaattgtcggt
ggagggtg(Rev)gtgatgggggtggggtgtggttgtggacttct

>GQ130189
gatgagtcctgagtaaaaaaaggaacttgtgttttctttgaaagttgaggtaacggggaatttgcggcgagtattgaacaagcta
acatatccatcctcatcagatacagcaaggatgtggcaaattttgaggtctacaaataaaaaaaatggaaaaataaataaataagg
aaggttgataagagtagggtttagtgaaagtaccttacagaaggagagagaaaggataatggtatctttggcattgtgttgtaggc
caagaacttcggtttcattgaagccaaagacgaatccattcatgcgaggtcat(For)ttttgaactgcgaagaaaaacaaataga
gtgagaaaatgaactcagtaagtaatcggtatagtgaaatttagagagagtgtgtgtgtgtgtgtgtggatggaacgttggaaggaa
ttgaactcccttgaggataagtggttgaagagtgaatggtgttgtggagtttctgttggcattgttcacctaagatgagataattgga
gtagcacaaaaagaggaagaagaagaatgggaattagggatttgatttttagcaactttgaaaattgaaactgagaaagggaag
ggaagg(Rev)aagcaatatattttggtttcgtactaattgaaactgataaaatcaaatgaaagaagcagagaccaaatagagttc
ctcgatgatgcaggttactcaggactcatc

> GQ130164
gatgagtcctgagtaaaccaaacaaaggaggaaaaatcagtgaaaaggaaagaaagagaaactctcactagtggcaatggttg
g(For)ttggtgctagagttggaggcaacaaactaggagagaaagggcttctaagataaaagatagagagtcaaagagatgaa
gagagagaaaaagagagagagagagagagcaaaaattagatgcgagatagagaagagttcttattggggcacagttatggtg
gcaacaagagtgagatgagatgagggtgagagagtgagagaaatgagagagagagagagaaagaaagtgagagagagag
acagagagagagatgagtgtgacagtttcataatgagggagaaaagaggaaggtga(Rev)ccttactcaggactcatc

> GQ130174
gatgagtcctgagtaatattcacttttgtatgcaatgtttcagtcgactttgtaggcctgggacacttgatcccagcaaagcaaatgg
caaaatagtgtcatgcctcagagaaggaaaaataaaatcagttagagaaggtcaagaagctctatctgctggcgccaggggaat
gcttttggcaatcaaaagcaaaatgg(For)gaatacacttcttgctgagcctcatgttttgtccactgtcaaccatctacatccac
atgccggaacacacaaaccatttgacataaccgccaagtaattatatacatacatacatacacacacatatatgtacatatacatac
acacacacacacacacacacatatgccacttgcagaaaccgaatca(Rev)tgacctttagttacatacgcttttttatttatggcttact
caggactcatc

>GQ130180
gatgagtcctgagtaaggttttatttgggcagttgccatccggggttgtggttgtagttgtgtagggcattgaagggtgggtggtgtt
tgaaattctgggtggtttggtttggatgattttctgggtttggttctctaggtcttgctagagatgggggttgg(For)gtgcggagaat
ggtaaggtggtggagtcttgggatgtgtgtaaatcaaaaggggtgaagaagaagaagaagaagaagaagaagaagaagaag
aagaagaaagagggtgaggaagtggtgcaagaggcaggggctgggtgttgggttaggttgaggttcattgggtgctgcattttct
tc(Rev)aagatccaaggttgacagctcagtcagtggcaccagtactcattatggtaatcactatttccttgctagagcagagttatt
attgttcttatgatgatgtgcttggagttggaggttccatctttgtaaatgtgtttgtgttgtgatatgatcttgttgtgtgtaaaagttggt
agtttttcaaatttagtagtttgtttttagcttttgctttgcttgggaaattgccatgcaaaataatgtgaaatgaaacacaaatgatatgc
taataaaaaccaaaggggtctgattactcaggactcatc

Figure 1. Nine polymorphic microsatellite loci containing sequences of *P. pinnata* submitted to GenBank. Where yellow, green and red shaded sequences refer to primer designing, microsatellite repeat and MseI specific adapter sequences, respectively.

number of 32 alleles were detected with nine single locus microsatellite primer pairs. Allele frequencies ranged from 0.2 to 0.89 with a mean value of 0.21. Mean H_e, H_o and PIC was 0.69, 0.53 and 0.68 respectively. PIC value for microsatellite loci ranged from 0.20 to 0.90 (table 2). HWE and LD tests were performed and all nine polymorphic loci were in HWE ($P < 0.005$) and significant LD between pairs of loci were found. The PIC value obtained in this study is significantly higher than earlier study on *Pongamia* using dominant AFLP and TE-AFLP markers (Sharma *et al.* 2011).

In conclusion, we isolated the first set of nine polymorphic microsatellite loci and characterized across 24 *Pongamia* accessions. Despite the modest number of accessions analysed, the results indicate that the new microsatellite loci developed in this study will serve as a very useful tool for genetic diversity analysis, clonal identification and conservation in *P. pinnata* germplasm.

Table 1. List of new 19 microsatellite markers.

Locus	GenBank accession no.	Repeat motif	Forward primer (5′ → 3′)	Reverse primer (5′ → 3′)
PpSSR2238	GQ130170	(AG)...$_7$ (CA)$_{11}$	AAAGGAGGTGAGAGGGCAAG	TTGGAATTTGGGTTCGAGTC
PpSSR2254	GQ130166	(GA)$_{20}$	TGCGAGATAGAGAAGAGTCCTT	CCTTCCTCTTTTCTCCCTCA
PpSSR2255	GQ130172	(GA)$_{13}$	TCACTGTGTGTTGTTTCTGAGG	GATGAGTCCTGAGTAAGGTGTG
PpSSR2265	**GQ130182**	**(AG)$_{16}$**	**GCAGTAGCATCGAACGACAG**	**TCCTCCTCTTCTCTCGACCA**
PpSSR2266	**GQ130187**	**(GAG)$_{17}$**	**TTCCTCCCCACTCTTCCTCT**	**TACACGAGAAGTGCCACAGG**
PpSSR2267	GQ130190	(GAA)$_{10}$	GTGCAGGATGTGGATGTGTC	GCCACCTCCTTCTTCATCAT
PpSSR2270	GQ130159	(GGA)$_8$	AGGCCTTCCATAACATGGTG	CCACCATCACCATCATCTCTT
PpSSR2277	GQ130191	(TG)$_{11...}$(AG)$_{15}$	TTGCTCCATAAATCCCTTCA	CAGGTTCAGTAACAGAGAACGA
PpSSR2279	**GQ130160**	**(GA)$_{19}$**	**ATCGGTGCAAGACAACAACA**	**CTAAAAGCACTGTCGCCACA**
PpSSR2283	GQ130183	(TA)$_{12...}$(TG)$_{25}$	TCCACTTGCAAATACCCAAT	TTTTCTTTGGATCCCCATCA
PpSSR2284	**GQ130188**	**(CA)$_{21}$**	**TGGCACCTCACAAGGTACAG**	**TCTGCAGTTGTTGCTTTCTGA**
PpSSR2288	GQ130167	(AAG)$_{16}$	TCCTTGTGTCGCATTTTCAA	TGCTTGCTGTTCCTCTTCCT
PpSSR2289	**GQ130173**	**(GGA)$_8$**	**GTGTGAAGGAGCGAGTGTCA**	**CACCCTCCACCGACAATTC**
PpSSR2293	GQ130192	(GA)$_{20...}$(GAT)$_6$	AAGCGTACGGAGATGGAGAA	TCAAAACACCTCTATCCCTTTG
PpSSR2295	GQ130161	(AGA)$_{28}$	AGAAGAAGAAGAAGCAGCAGC	TCACGCTCCATTCTTCCTTT
PpSSR2320	**GQ130189**	**(GT)$_{9...}$(GAA)$_4$**	**ATCCATTCATGCGAGGTCAT**	**CCTTCCCTTCCCTTTCTCAG**
PpSSR2323	**GQ130164**	**(AG)$_{28}$**	**TCACTAGTGGCAATGGTTGG**	**TCACCTTCCTCTTTTCTCCCTCA**
PpSSR2325	**GQ130174**	**(AC)$_{14}$**	**GGCAATCAAAAGCAAAATGG**	**TGATTCGGTTTCTGCAAGTG**
PpSSR2326	**GQ130180**	**(GAA)$_{15}$**	**CTTGCTAGAGATGGGGTTGG**	**GAAGAAATGCAGCACCCAAT**

Nine microsatellite markers (bold) showed polymorphism, while remaining 10 showed monomorphic bands across *Pongamia* accessions. All forward primers were tagged with M13 (5′-TGTAAAACGACGGCCAGT-3′)-tailed at the 5′ end.

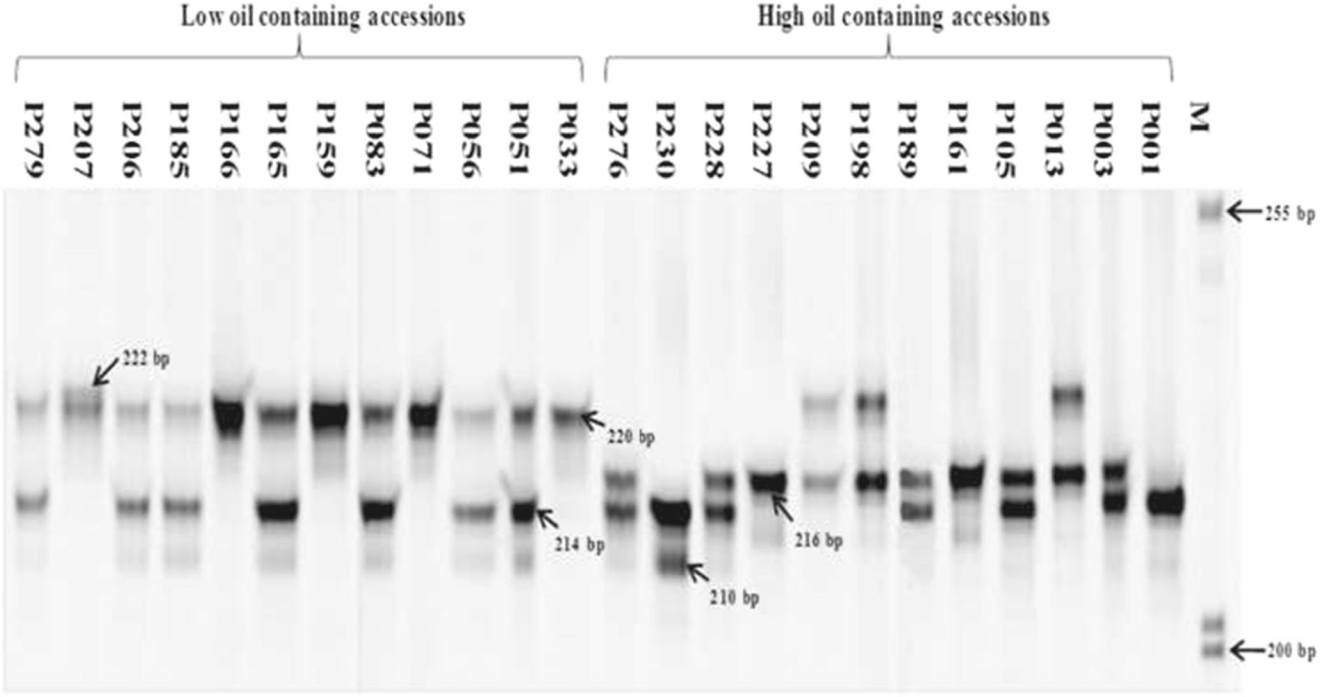

Figure 2. Representative gel profile showing amplification using microsatellite marker PpSSR2279 in 24 *P. pinnata* accessions. Size of alleles also included FAM-labelled M13 primer sequences. Lane M is 50–700 bp IRDye700 labelled used as size markers.

Table 2. Characterization of nine polymorphic microsatellite loci and their marker attributes in *Pongamia pinnata*.

Locus	GenBank accession no.	T_a(°C)	Allele size range (bp)	N_a	H_e	H_o	PIC
PpSSR2265	GQ130182	56	196-202	2	0.68	0.39	0.8
PpSSR2266	GQ130187	63	316-322	3	0.23	0.2	0.2
PpSSR2279	GQ130160	53	193-198	5	0.82	0.73	0.8
PpSSR2284	GQ130188	53	217-224	5	0.85	28	0.8
PpSSR2289	GQ130173	56	195-200	5	0.89	0.82	0.9
PpSSR2320	GQ130189	63	313-322	4	0.72	0.7	0.8
PpSSR2323	GQ130164	63	314-321	3	0.65	0.56	0.6
PpSSR2325	GQ130174	53	207-216	2	0.52	0.43	0.4
PpSSR2326	GQ130180	53	198-202	3	0.87	0.68	0.9

T_a, annealing temperature; N_a, number of alleles; H_e, expected heterozygosity; H_o, observed heterozygosity; PIC, polymorphic information content.

Acknowledgements

Authors sincerely acknowledge the funding received from the Department of Science and Technology, Government of India and Dr Shyam Sundar Sharma thanks CSIR for providing fellowship.

References

Kaushik N., Kumar S., Kumar K., Beniwal R., Kaushik N. and Roy S. 2007 Genetic variability and association studies in pod and seed traits of *Pongamia pinnata* (L.) Pierre in Haryana, India. *Genet. Res. Crop Evol.* **54**, 1827–1832.

Kesari V., Sathyanarayana V. M., Parida A. and Rangan L. 2010 Molecular marker-based characterization in candidate plus trees of *Pongamia pinnata*, a potential biodiesel legume. *AoB Plants* (doi:10.1093/aobpla/plq1017).

Liu K. and Muse S. V. 2005 Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**, 2128–2129.

Sahoo D., Aparajita S. and Rout G. 2010 Inter and intra-population variability of *Pongamia pinnata*: a bioenergy legume tree. *Plant Sys. Evol.* **285**, 121–125.

Schuelke M. 2000 An economic method for the fluorescent labelling of PCR fragments. *Nat. Biotechnol.* **18**, 233–234.

Sharma S. S., Negi M. S., Sinha P., Kumar K. and Tripathi S. B. 2011 Assessment of genetic diversity of biodiesel species *Pongamia pinnata* accessions using AFLP and three endonuclease -AFLP. *Plant Mol. Biol. Rep.* **29**, 12–18.

Singh A., Negi M. S., Rajagopal J., Bhatia S., Tomar U. K., Srivastava P. S. and Lakshmikumaran M. 1999 Assessment of genetic diversity in *Azadirachta indica* using AFLP markers. *Theor. Appl. Genet.* **99**, 272–279.

Tewari D. N. 2003 Report of the committee on development of Biofuel, Planning Commission, Government of India. Available at http://planningcommission.nic.in/reports/genrep/cmtt_bio.pdf.

Toonen R. J. and Hughes S. 2001 Increased throughput for fragment analysis on ABI Prism 377 automated sequencer using a membrane comb and STRand software. *Biotechniques* **31**, 1320–1324.

Zane L., Bargelloni L. and Patarnello T. 2002 Strategies for microsatellite isolation: a review. *Mol. Ecol.* **11**, 1–16.

Received 31 December 2013, in revised form 28 February 2014; accepted 3 March 2014
Published online: 12 August 2014