J. Plant Biochem. Biotechnol.
DOI 10.1007/s13562-014-0287-1

SHORT COMMUNICATION

Identification and cross-species amplification of microsatellite markers derived from expressed sequence data of rose species

Ram Kumar Sharma · Anil Chaudhary · Himanshu Sharma · Pankaj Bhardwaj · Vikas Sharma · Rahul Kumar · Paramvir Singh Ahuja

Received: 10 February 2014 / Accepted: 3 August 2014

Abstract Genic SSR markers derived from public expressed sequence tags (ESTs) data are valuable and cost effective marker resources for genome mapping and diversity studies. Owing to their derivation from the transcribed regions which often have putative functions, these markers can be easily associated with desired trait. In the present study, 19 novel SSR markers were identified from 450 non redundant unigenes derived from 3,726 public ESTs of two rose species. Among SSRs, tri-repeats (61.3 %) were most abundant followed by di-repeat (29 %). Newly identified EST-SSR markers recorded significant homology with the known/putative proteins of *Arabidopsis thaliana*. The cross transferability to 12 rose species ranged from 63.2 to 100 %. Novel SSR loci found to be moderately to highly polymorphic with locus wise average number of alleles and polymorphism information content (PIC) were 4.1 and 0.33, respectively. Cloning and sequencing of EST-SSR size variant amplicons of marker locus Rches12 revealed that the variation in the number of SSR repeat-units was the main source of fragment polymorphism. The high polymorphic potential coupled with high cross-transferability rate demonstrates wider applicability of novel SSR markers in genetic diversity, genome mapping and evolutionary studies in various rose species.

Keywords Cross-species amplification · EST-SSR · Microsatellite · *Rosa chinensis* · *R. luciae*

Abbreviations

SSRs	Simple sequence repeats
ESTs	Expressed sequence tags
PIC	Polymorphism information content
PCR	Polymerase chain reaction

Rose is one of the most important crops for floriculture industry. The genus *Rosa* includes 200 species and more than 18,000 cultivars (Gudin, 2000). Roses are recognised as both traditional and modern flowers in India, and two rose viz; Damask and Edourd occupy about 6,000 hectares of the land in the different parts of country (Srivastava and Gupta, 1984). The modern rose varietal registration and protection is based on morphological characteristics as described in the UPOV (Union for Protection of New Varieties of Plants) guidelines. The classical methods become less efficient as the number of varieties to be tested is increasing and the genetic distances between the

---

R. K. Sharma (✉) · A. Chaudhary · H. Sharma · P. Bhardwaj · V. Sharma · R. Kumar · P. S. Ahuja
Biotechnology Division, CSIR-Institute of Himalayan Bioresource Technology, Post Box 6, Palampur, H.P. 176061, India
e-mail: mrk_sharma@yahoo.com

R. K. Sharma
e-mail: ramsharma@ihbt.res.in

H. Sharma
e-mail: himanshusharm@gmail.com

P. Bhardwaj
e-mail: pankajihbt@gmail.com

V. Sharma
e-mail: vikasam@gmail.com

R. Kumar
e-mail: rahulihbt@gmail.com

P. S. Ahuja
e-mail: psahuja@ihbt.res.in

Present Address:
P. Bhardwaj
Central University of Punjab, Bathinda 151001, India

Present Address:
R. Kumar
DAV University, Jalandhar 144012, India

Table 1 Characteristics of 19 polymorphic EST-SSR markers of *Rosa chinensis* and *Rosa luciae* identified in present study

Locus name	Primer sequence	Ta (ºc)	Repeat motif	No. of alleles	Expected size	Size range (bp)	PIC	Arabidopsis proteome hit	Gene bank accession No of contributing ESTs
Rches2	F5′- TGCAAAGTTCTGCGTTTTCA3′ R5′- AATTACACCGTGAGGGGAGA3′	55	(GAA)$_{9}$	4	154	160–185	0.42	Early-responsive to dehydration protein	B1977488, B1977418
Rches3	F5′-ATCTCCACCTCCCAAAATCC3′ R5′- TGCTCCTTGAATCCTCCTTG3′	55	(CACCAT)$_{19}$	3	207	210–250	0.44	60S ribosomal protein L23 (RPL23B)	B1977516, B1977709, B1977291
Rches4	F5′-AACAGGCTGAAGGCAACACT3′ R5′- AATCGAACCATCACATGGAA3′	55	(ATG)$_{18}$	2	172	290–300	0.08	Encodes a protein belonging to the subgroup of HMGB	B1977389, B1977384
Rches5	F5′- CTGCACTCCCTCCTCGTATC3′ R5′- TGCATATGGAATGAATTATTAGGC3′	55	(CTAG)$_{5}$	2	224	260–270	0.23	Encodes 1-aminocyclopropane-1-carboxylate oxidase	B1977243, B1977725, B1977649, B1977626
Rches6	F5′- TCCATTAACACAGCAGTTACCAA3′ R5′- CAGCCAAACCTCAGGAAGAC3′	55	(TTC)$_{12}$	5	150	120–140	0.38	Induced by low temperatures, dehydration and salt stress and ABA.	B1978347, B1977415, B1978273, B1978164, B1977372, B1978975, B1978682, B1978022, B1978985, B1978749, B1977480
Rches7	F5′-AAAAATTAGGTCTTGTGAAATTTGG3′ R5′- AGAGGAGCGAGGAGGAGAGT3′	55	(GAA)$_{10}$	6	221	250–290	0.72	Ubiquinol-cytochrome C reductase complex	B1978495, B1977483
Rches8	F5′- GAAACCACCGAAGAACCAAA3′ R5′- TGGTTCTGCTGCTACCTCCT3′	55	(CAG)$_{26}$	4	177	330–350	0.08	Encodes a protein involved in cell proliferation during leaf and flower development	B1977358, B1977865
Rches9	F5′- AAACTTTGCGATTCCTCTGC3′ R5′- ATCCCATACTTGCGGATCAA3′	55	(TC)$_{15}$	4	191	230–250	0.12	40S ribosomal protein	B1978666, B1978062
Rches11	F5′- TCCGATATACCCGACGACTC3′ R5′- AACAGTCCCAACGAAACCAG 3′	55	(GGA)$_{12}$	7	152	320–355	0.51	*Myb* family transcription factor	B1978838, B1977780
Rches12	F5′- TCCGATATACCCGACGACTC3′ R5′- AACAGTCCCAACGAAACCAG3′	51	(GGA)$_{12}$	5	152	280–300	0.16	*Myb* family transcription factor	B1977927, B1977917
Rches13	F5′- GGCCTAGCAAAGCAACAAAC3′ R5′- AGTGGAGGGCAGTCTCTGAA3′	55	(TC)$_{5}$	5	223	310–335	0.24	VAMP72 Gene family member	B1978291, B1978781
Rches14	F5′- CGTAACATCACCGTCAACGA3′ R5′- CGAGATCCAGACCCCACTAA3′	55	(CGG)$_{72}$	4	389	370–400	0.54	Encodes a small glycine-rich RNA binding protein	B1978636, B1978369
Rches15	F5′- GGATCCTCACACATTTTCAAGC3′ R5′- ATCGAGACCGACCATCAGAA3′	55	(TC)$_{15}$	4	193	175–200	0.68	Gene encoding ADP-ribosylation factor	B1978044,B1978907, B1977312
Rles1	F5′- CATTCTTCGCTCTCCTCCAA3′ R5′- CGGAGAATCCATGAAGCAAT3′	55	(GAA)$_{6}$	2	222	200–210	0.09	Serinc-domain containing serine and sphingolipid biosynthesis protein	EC589706, EC589987
Rles2	F5′- AGGTCACCGTGGAAATGAAG3′ R5′- GGCAGATACGAGCAGTAGGC3′	55	(CTG)$_{7}$	2	195	195–200	0.31	Unknown protein	EC588865, EC589380, EC588647
Rles6	F5′- TTTCAAGCTCTCTCATTCATTCC3′ R5′- GTGACGATCTCACCGAGCTT3′	55	(TC)$_{15}$	3	232	195–210	0.04	Gene encoding ADP-ribosylation factor and similar to other ARFs and ARF-like proteins	EC589130, EC588353, EC588546
Rles7	F5′- GAGAGAAGCCGAGAACCCTAA3′ R5′- GCGATCACAGTAGGCATTCA3′	55	(GA)$_{18}$	3	232	200–220	0.51	Encodes a small ubiquitin-like modifier (SUMO)	EC589083, EC588230, EC589784
Rles8	F5′- CAGAAAAAGCCACAGAACCAA3′ R5′- TGAAGGCTGACAGAGGGTTC3′	55	(TC)$_{13}$	8	245	270–350	0.43	Calcium-dependent protein kinase (CDPK)	EC589816, EC588975
Rles9	F5′- GCTGAGGGAGAAGAATGCTG3′ R5′- AGCTCGGAACTACAGCCTGA3′	55	(GAA)$_{12}$	4	246	460–490	0.28	Nuclear factor Y, subunit B11” (NF-YB11)	EC589709, EC588819

Ta, annealing temperature; *PIC*, Polymorphism information content; Rches: *R. chinensis,* Rles: *R. luciae* varieties is becoming smaller. In past few decades, inferences based on morphological characteristics of rose varieties have been replaced with more reliable RFLP (Hubbard et al, 1992) and RAPD (Martin et al. 2001) markers. However, non-reducibility and underestimation of the recessive allele frequency in case of dominant markers (Nybom 2004), and labour intensive and complexity in RFLP analysis are among the major drawbacks of these markers. In contrast, microsatellites or simple sequence repeats, due to desirable attributes, including genome-wide distribution, hypervariability, co-dominant inheritance and chromosome-specific location gained considerable importance in genetics, breeding, conservation and evolutionary studies (Guichoux et al. 2011). Non-redundant (NR) nucleotide sequences derived from these public databases have become a rapid and cost-effective source for the identification of microsatellite markers in several crop plants and rare species, which were primarily based on tedious, labour-intensive nucleotide sequencing of positive clones from enriched genomic libraries. Considering their derivation from the transcribed genome, EST-SSR markers tend to be widely cross transferred in related species (Bouck and Vision 2007),and hence can be useful in comparative genome mapping (Liu et al. 1999). Eighty one SSR markers available in different rose species are not sufficient for genome mapping studies in rose (Hibrand-Saint Oyant et al, 2008; Zhang et al. 2006). We therefore report, a set of nineteen novel SSR markers derived from publicly available sequence data of two rose species namely *R. chinensis* and *R. luciae.* Furthermore, novel SSR markers successfully validated in several rose species would be valuable resource for future genome mapping and genetic improvement studies.

To develop EST-SSR markers, 3726 ESTs of two rose species namely *R. chinensis* (1794) and *R. luciae* (1932) were retrieved from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). These ESTs were subsequently clustered into 450 unique clusters (unigenes) using SeqMan DNA Star lasergene version 7.1 (DNASTAR, Madison, Wisconsin, USA) as per the search parameters reported earlier by Sharma et al. (2009) . Non-redundant (NR) unigenes were subsequently searched individually for identification of SSRs containing sequences using repeat masker software (http://www.repeatmasker.org/) with search criteria of minimum length of ≥12 bp (di & tri) and ≥15 bp (tetra, penta & hexa). Processing of NR data detected 31 SSR containing sequences with tri-repeat (61. 3 %) being the most abundant, followed by di-repeats (29 %) and only few tetra, penta and hexa repeats. Of these, 26 SSR containing sequences were successfully utilized for designing of EST-SSR primer pairs using PRIMER-3 (www.genome.wi.mit.edu/genome_software/other/primer3.html). The criteria for primer design were; i) nucleotide length of 18–22 base pairs, ii) T_m value of 50 ^{0}C to 60 ^{0}C, iii) the 3' end base with a G or C, preferably and iv) an amplified fragment size of 100–350 bp. Considering their derivation from expressed NR sequence data of *Rosa chinensis* and *R. luciae*, novel EST SSR markers are prefixed as "Rches" and "Rles", respectively. A panel of 12 rose species namely *R. cathayensis, R. multiflora, R. brunonii, R. alba, R. moschata, R. damascena, R. bourboniana, R. microphylla, R. tomentosa, R. canina, R. hybrida, R. wichuraiana* was employed for marker evaluation and cross-species amplification. Total genomic DNA was extracted from fresh young leaves of single plant of each species using CTAB method (Doyle and Doyle 1990). PCR was performed in 10 μL reaction volume containing 20 ng of template DNA, 15 ng of each primer, 200 μM of each dNTP, 10 mMTris-HCL (pH 8.3), 50 mM KCL, 1.5 mM $MgCl_2$, 0.01 % gelatin and 0.2 U of Taq DNA polymerase (Bangalore Genei Pvt. Ltd., Bangalore, India). All PCR reactions were performed on I-cycler PCR system (Bio-Rad, Australia). The PCR Conditions were: 1 step of 5 min at 94 °C, 35 cycles of 1 min at 94 °C, 1 min at respective annealing temperature for each primer (Table 1), 2 min at 72 °C and final extension for 7 min at 72 °C. Amplification products were electrophoresed on 7 % denaturing polyacrylamide gel in 1X TBE buffer. Fragments were then visualized by silver staining (Silver sequence staining reagents, Promega,

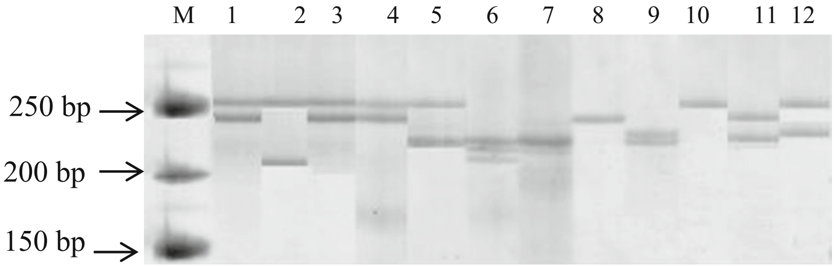
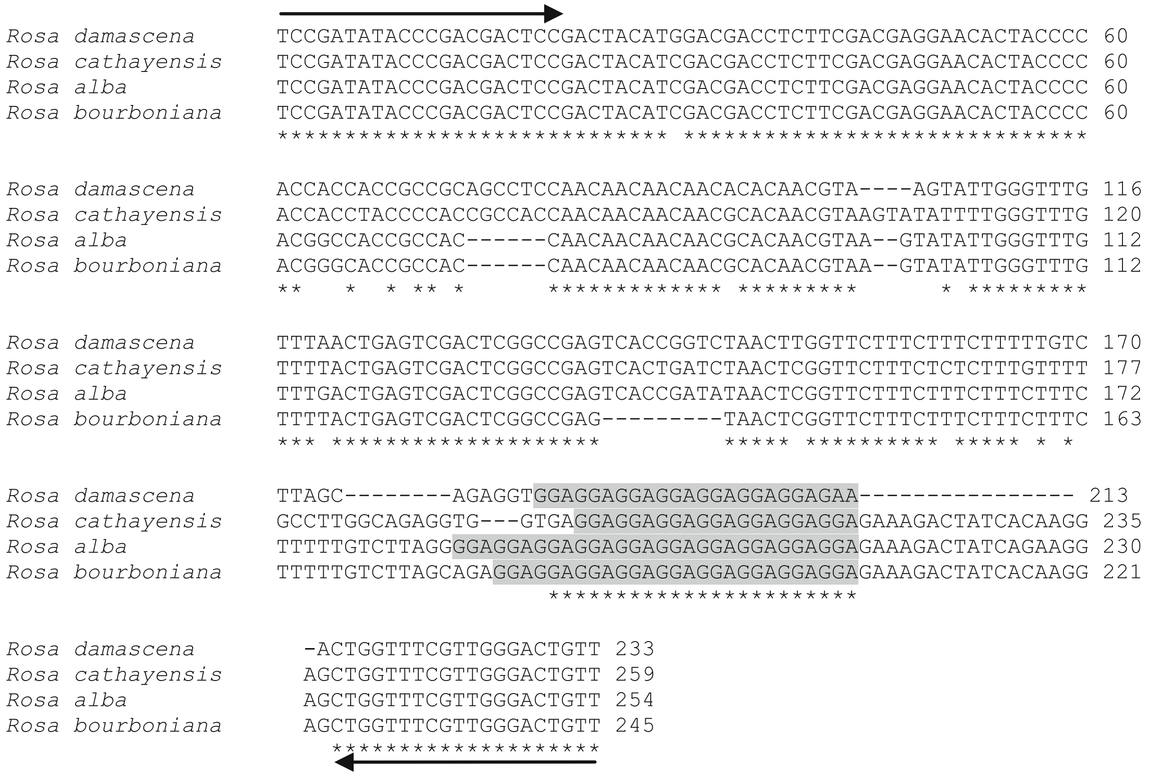


Fig. 1 Representative picture of EST-SSR marker profile generated with primer Rches12. *M*=50 bp ladder standard (MBI fermentas, Lithuania); Lane 1–12 represent different rose species' Lane 1: *Rosa cathayensis;* Lane 2: *R. multiflora*; Lane 3: *R. brunonii;* Lane 4: *R. alba;* Lane 5: *R. moschata;* Lane 6: *R. damascena;* Lane 7: *R. bourboniana;* Lane 8: *R. microphylla;* Lane 9: *R. tomentosa;* Lane 10: *R. canina*; Lane 11: *R. hybrid;* Lane 12: *R. wichuraiana*

Table 2 Amplification pattern and cross transferability details of Novel EST-SSR markers across the 12 rose species utilized in present study. (-) indicates the no amplification or null allele

Marker locus		Rose species											
	No. of alleles	*R.cathayensis*	*R.multiflora*	*R.brunonii*	*R.alba*	*R.moschata*	*R.damascena*	*R.bourboniana*	*R.microphylla*	*R.tomentosa*	*R.canina*	*R.hybrida*	*R.wichuraiana*
Rches2	4	177	160,177	160,177	160,177	160,177	160,177	160,177	160,177	177,185	177,180	–	160,177
Rches3	3	210,230	210,230	210,230	210,230	210,230	210,230	210,230	210,250	210,250	210,250	210,230	210,230
Rches4	2	290,300	300	300	300	290,300	290,300	290,300	290	290	290	290	290,300
Rches5	2	260	260	260	260	260	260	260	260	260,270	260	260,270	260
Rches6	5	120	120	120	120	120	125,140	125,140	125	125,130	125,135	125,140	120
Rches7	6	275,285	275	280	275,280	275,285	250,260	260	260	275	280	250	260
Rches8	3	335	335	335	335	335	335,350	335,350	330	335	335	335	335
Rches9	4	235,240	235	240,250	240	235	250	235,240	240,250	235	230,235	235,240	235,240
Rches11	7	340,355	335	340,355	340,355	340,355	320,330	330	345	330,335	335,350	–	340.355
Rches12	5	245,260	245,260	240,260	240,260	240,260	235	240,210	240,245	245	260	240,245	260,245
Rches13	5	320,335	320	320	325,335	320,335	330,335	330	310,320	320	320,330	310,335	320,335
Rches14	4	390	390	380,390	380,390	–	370,390	390	–	390	380,390	–	400
Rches15	4	175	190	190	190	–	180,190	175,180	180	180	180,190	–	–
Rles1	2	200,210	200	200	200	200,210	200	200	200	200	200	200	–
Rles2	1	195	195	195	195	195	195	195	195	195	195	195	195
Rles6	3	200	200	200	210	200,210	200,210	200	200	200	195,200	200	200,210
Rles7	3	–	215	215	215	215	215	215	220	200	215	220	–
Rles8	8	320	320,330	320,330	310,340	280,320	320	330,340	270,310	345,350	–	280,340	280,320
Rles9	4	460,480	460	460	460,470	460,480	470,490	460,470	470	470,490	–	460,470	–
Cross-transferability (%)		94.7	100	100	100	89.5	100	100	94.7	100	89.5	63.2	73.7



```
Rosa damascena      TCCGATATACCCGACGACTCCGACTACATGGACGACCTCTTCGACGAGGAACACTACCCC 60
Rosa cathayensis    TCCGATATACCCGACGACTCCGACTACATCGACGACCTCTTCGACGAGGAACACTACCCC 60
Rosa alba           TCCGATATACCCGACGACTCCGACTACATCGACGACCTCTTCGACGAGGAACACTACCCC 60
Rosa bourboniana    TCCGATATACCCGACGACTCCGACTACATCGACGACCTCTTCGACGAGGAACACTACCCC 60
                    ***************************** ******************************

Rosa damascena      ACCACCACCGCCGCAGCCTCCAACAACAACAACACACAACGTA----AGTATTGGGTTTG 116
Rosa cathayensis    ACCACCTACCCCACCGCCACCAACAACAACAACGCACAACGTAAGTATATTTTGGGTTTG 120
Rosa alba           ACGGCCACCGCCAC------CAACAACAACAACGCACAACGTAA--GTATATTGGGTTTG 112
Rosa bourboniana    ACGGGCACCGCCAC------CAACAACAACAACGCACAACGTAA--GTATATTGGGTTTG 112
                    **   *  * ** *      ************* *********      * *********

Rosa damascena      TTTAACTGAGTCGACTCGGCCGAGTCACCGGTCTAACTTGGTTCTTTCTTTCTTTTTGTC 170
Rosa cathayensis    TTTTACTGAGTCGACTCGGCCGAGTCACTGATCTAACTCGGTTCTTTCTCTCTTTGTTTT 177
Rosa alba           TTTGACTGAGTCGACTCGGCCGAGTCACCGATATAACTCGGTTCTTTCTTTCTTTCTTTC 172
Rosa bourboniana    TTTTACTGAGTCGACTCGGCCGAG---------TAACTCGGTTCTTTCTTTCTTTCTTTC 163
                    *** ********************         ***** ********** ***** * *

Rosa damascena      TTAGC--------AGAGGTGGAGGAGGAGGAGGAGGAGAA---------------- 213
Rosa cathayensis    GCCTTGGCAGAGGTG---GTGAGGAGGAGGAGGAGGAGGAGAAAGACTATCACAAGG 235
Rosa alba           TTTTTGTCTTAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGAAAGACTATCAGAAGG 230
Rosa bourboniana    TTTTTGTCTTAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGAAAGACTATCACAAGG 221
                                        ***********************

Rosa damascena        -ACTGGTTTCGTTGGGACTGTT 233
Rosa cathayensis      AGCTGGTTTCGTTGGGACTGTT 259
Rosa alba             AGCTGGTTTCGTTGGGACTGTT 254
Rosa bourboniana      AGCTGGTTTCGTTGGGACTGTT 245
                        ********************
```

Fig. 2 Nucleotide sequence alignment of cross species amplicons obtained with marker Rches12 in different rose spp. The shaded nucleotide highlights the conservation of microsatellite motifs in different species, and arrow indicates the respective primer sequences

USA) and sized with 50 bp DNA ladder (MBI fermentas, Lithuania). The PIC of each marker was calculated according to Anderson et al. (1993).

Of the 26 potential SSR markers, 19 (73.1 %) primer pairs produced reliable amplification in most of the tested rose species (Fig. 1) with significantly higher amplification success rate in *Rosa chinensis* (86.6 %) than *R. luciae* (55.4 %). High allelic diversity (2 to 8 with an average of 4.1 alleles per locus) coupled with with polymorphism information content (PIC) values ranging from 0.08 to 0.72 with an average of 0.33 in 12 rose species suggested the highly heterogeneous nature of rose. Five primers namely, Rches-7, Rches-11, Rches-14, Rches-15 and Rles7 with PIC values ≥0.50 were identified to be more informative and thus would be useful for genetic characterization of rose germplasm. A full description of the EST-SSR markers and their characteristics are presented in Table 1. In general, all the newly identified markers have shown moderate to high level of transferability in the tested rose species. Marker wise transferability rate varied from 63.2 to 100 % in 12 rose species (Table 2). Seven markers (36.8 %) of *Rosa chinensis* (Rches-4, Rches-7, Rches-8, Rches-9,Rches-12) and two SSR markers of *R. luciae* (Rles2, Rles6) were found to be conserved in all the tested species. Further, all the newly identified markers have recorded 100 % transferability in 8 of the 12 tested species. Hence, suggested their applicability for large scale rose germplasm characterization, irrespective of the species types. Further, to confirm specificity and conservation of cross-transferability at the sequence level, selected amplicons from *Rosa bourboniana*, *R. damascena*, *R. cathayensis* and *R. alba* were cloned and sequenced for microsatellite primers Rches12 as described earlier (Sharma et al, 2009). The presence of the target microsatellites was observed in all the cases (Fig. 2).

In conclusion, SSR markers reported in the present study derived from the transcribed portion of the rose genome will facilitate the understanding of intra and inter species gene flow, functional diversity analysis and genome mapping studies in wild and cultivated rose species. Further, these markers will provide valuable resource for genetic variation analysis and marker assisted selection in rose breeding programs.

Acknowledgements We gratefully acknowledge Council of Scientific and Industrial Research (CSIR), Government of India, New Delhi for financial support. This is IHBT communication 0960.

References

Anderson JA, Churchill GA, Autroque JE, Tanksley SD, Sorrells ME (1993) Optimising parental selection for genetic linkage maps. Genome 36:181–186

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. Mol Ecol 16:907–924

Doyle JJ, Doyle JE (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Gudin S (2000) Rose: genetics and breeding. Plant Breed Rev 17:159–189

Guichoux E, Lagache L, Wagner S, Chaumeil P, Ger L, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591–611

Hibrand-Saint Oyant L, Crespel L, Rajapakse S, Zhang L, Foucher F (2008) Genetic linkage maps of rose constructed with new microsatellite markers and locating QTL controlling flowering traits. Tree Genet Genomes 4:11–23

Hubbard M, Kelly J, Rajapakse S, Abbott A, Ballard R (1992) Restriction fragment length polymorphisms in rose and their use for cultivar identification. Hortscience 27:172–173

Liu ZJ, Tan G, Li P, Dunham RA (1999) Transcribed dinucleotide microsatellites and their associated genes from channel catfish, *Ictalurus punctatus*. Biochem Biophys Res Commun 259:190–194

Martin M, Piola F, Chessel Jay M, Heizmann P (2001) The domestification process of the modern rose: genetic structure and allelic composition of the rose complex. Theor Appl Genet 102:398–404

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13:1143–1155

Sharma RK, Bhardwaj P, Negi R, Mohapatra T, Ahuja PS (2009) Identification, characterization and utilization of unigene derived microsatellite markers in tea (Camellia sinensis, L). BMC Plant Biol 9:53

Srivastava HP, Gupta MN (1984) Ind Rose Annual 3:44–51

Zhang LH, Byrne DH, Ballard RE, Rajapakse S (2006) Microsatellite marker development in rose and its application in tetraploid mapping. J Am Soc Hortic Sci 131:380–387