J. Hortl. Sci.
Vol. 16(1) : 125-129, 2021

Short Communication

Isolation and characterization of microsatellite markers from *Garcinia indica* and cross species amplification

Ravishankar K.V.*[1], Vasudeva R.[2], Hemanth B.[1], Nischita P., Sthapit B.R.[3] Parthasarathy V.A.[4] and Rao V.R.[5]

[1]ICAR-Indian Institute of Horticultural Research, Bengaluru - 560 089 India
[2]Department of Forest Biology and Tree Improvement, College of Forestry Sirsi - 581 401 University of Agricultural Sciences (Dharwad), India,
[3] Regional project coordinator (UNEP-GEF), Bioversity International, Pokhara, Nepal,
[4] National project coordinator (UNEP-GEF), ICAR-Indian Institute of Horticultural Research , Bengaluru, India
[5]Bioversity International, Rome
*Corresponding author e-mail : kv_ravishankar@yahoo.co.in, ravishankar.kv@icar.gov.in

ABSTRACT

Garcinia indica **popularly known as '*Kokum*' or *Murugalu*", is a medium sized evergreen tree found in western-ghats of India. This tree species is highly exploited to produce anti-obesity drugs and culinary purposes. Its population is threatened by over exploitation and loss of habitat. Development of microsatellite markers would help in understanding genetic structure and further to develop appropriate conservation strategies. In this study, using next generation sequencing platform Illumina Hiseq 2000, we have sequenced partial genome of *G. indica* and identified 3725 microsatellites. Forty-eight microsatellite markers were analyzed using 30 accessions. Polymorphism information content (PIC) values ranged from 0.718 to 0.968 with a mean value of 0.922. Allele per locus ranged from 3 to 33 per locus. Probability of identity values ranged from 0.00329 to 0.30489. Cross species amplification SSR primers in the related species, showed a moderate transferability from 12.5 % (for *G. morella*) to 18.7%(for *G. gummigutta*)**

Key words : Cross-species amplification *Garcinia indica*; Microsatellite markers and Next-generation sequencing (NGS)

Garcinia indica Choisy (Thouars; Family Clusiaceae), is a perennial tree. *G. indica* is commonly known as a Brindonia Tallow tree or 'Kokum Butter' tree in English. Kokum has many uses in cuisines and an important ingredient in locally prepared medicines. The seeds are a rich source of Kokum butter, which is nutritive, demulcent, agent for smoothening, softening and used for cosmetic, confectionery, culinary purposes. Raw fruits, young leaves and bark are also used as medications against several disorders. The fruit rind is a rich source of Hydroxy Citric Acid (HCA) that prevents accumulation of fat in the human body cells. Therefore, *G. indica* has become the natural source for production of anti-obesity drugs. (Baliga *et al*., 2011). *Garcinia* species are endemic and distributed in tropical rain forests of the Western Ghats. Perceiving the threat of over exploitation, FRLHT (Foundation for Revitalization of Local Health Traditions) and IUCN (International Union for Conservation of Nature) have recognized this species as '*Vulnerable*' and '*Threatened*' category respectively (Hareesh and Vasudeva, 2010). A few studies examined diversity in this species using general DNA markers like RAPD and ISSR markers (Thatte *et al*. 2012; Palkar and Sellappan, 2019). However, so far there are no efforts to develop species specific, highly reproducible microsatellite markers or SSR markers in this species. Keeping this in view, an attempt has been made to develop microsatellite or SSR markers using next generation sequencing

technology. The development of molecular markers would help in studying its diversity, analyzing the genetics of traits, and further help in evolving conservation strategies and improvement.

The plant material was obtained from the germplasm collection of the College of Forestry, Sirsi (University of Agricultural Sciences, Dharwad), Karnataka state, India. Total genomic DNA was isolated from the leaves of *G. indica* genotypes using modified CTAB method (Ravishankar *et al*., 2000). Genomic DNA was sequenced using Illumina HiSeq2000 platform at M/s Genotypic Pvt. Ltd, Bengaluru facility following manufactures instructions. High quality sequence data was used for assembly into contigs. *De novo* assembly of reads into contigs was performed using SOAPdenovo2-src-r240 software (Luo *et al.,* 2012). This has resulted in 92125 contigs. The total assembled size of the contigs is approximately 25.6 Mbp. An SSR survey of genomic sequences using MISA software (http://pgrc.ipk-gatersleban.de/misa), showed that 3590 contigs contained at least one microsatellite (Ravishankar *et al.* 2015). A total of 3725 microsatellite was identified. A total of 1374 microsatellites (ESM1) primers were designed using Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/; Untergrasser *et al*., 2012). From these, randomly 50 loci were selected for initial screening. Finally, 48 SSR primers were selected for genetic analysis based on clear amplification of PCR products. We employed Thirty genotypes of *Garcinia indica* for assessing polymorphism at each locus. The fluorescence based M13 tailed PCR method of Schuelke (2000) was followed to amplify the microsatellites in a quick, accurate and efficient manner. PCR was carried out in the 20µl reaction volume containing 2µl of 10X reaction buffer, 2.0µl of 1 mM dNTPs, 0.9µl (5 pmol) of forward, 0.9µl reverse primers (5 pmol), labeled M13 probe 1.2µl (5 pmol), 5.0 µl (50-75 ng) of template genomic DNA, 0.8 µl (2 U) of Taq DNA polymerase and 7.2 µl of nuclease free water. The PCR cycling profile was: initial denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 30sec., 55°C for 30 Sec., 72°C for 1 min and a final extension at 72°C for 5 min. Amplified products were separated on 96 capillary Automated DNA Sequencer (Applied Biosystems, ABI 3730 DNA Analyzer) at M/S Eurofin facility, Bengaluru.

The raw data generated was analyzed and compiled using Peak Scanner V1.0 software (Applied Biosystems, USA) for estimating the allele size in bp**.** The allele size data was used for genetic analysis using Cervus 3.0 software (Kalinowski *et al*. 2007). We have calculated observed heterozygosity, expected heterozygosity, polymorphic information content(PIC). The probability of identity (PI) was calculated using IDENTITY1.0 software (http://www.uni-graz.at/~sefck/: Wagner and Sefc, 1999). Genetic analysis of 48 SSR loci, showed PIC values ranging from 0.718 to 0.968 with a mean value of 0.922. The mean values of observed and expected heterozygosity are 0.2813 (Table 1) and 0.933 respectively (Table 1 and 2). The allele per locus ranged from 13 to 41 with a mean of 16.395. The probability of identity (PI) values ranged from 0.00329 to 0.304896 with a mean of 0.03506. The total probability of identity is 8.132729×10^{-80}. In cross species amplification, out of 48 SSR primers, 6 amplified in *G. morella ,* accounting 12.5 per cent transferability and 9 amplified in *G. gummigutta* accounting 18.8 percent transferability (ESM2). This relatively low cross-species transferability compared to what has been observed in *G. gummigutta* species (Ravishankar *et al*., 2017).

This is the first report of SSR markers for *Garcinia indica*, where 3725 microsatellites were identified and primers were designed for 1374 microsatellites. The genetic analysis showed that the majority of the SSR primers developed have high PIC values indicating high heterozygosity in the species. The low probability of identity values of many SSR loci is useful for molecular characterization. Finally, the SSR developed will be useful in studying genetic diversity, mapping and fingerprinting of *Garcinia indica* and related species.

Table 1: Genetic analysis of microsatellite markers developed for *Garcinia indica*

Locus	Forward Sequence 5¹→3¹	Reverse Sequence 5¹→3¹	Repeat Type	Number of Allele (k)	Allele size range (bp)	Observed Heterozygosity (Ho)	Expected Heterozygosity (He)	Polymorphic Information Content (PIC)	Probability of Identity (PI)
GI_KVRa577	TTTGGCGAGGGTGTTGGTGAGT	ACACGTGTAGGCTGACACCAACC	$(GT)^{6}$	20	140-230	0.345	0.924	0.902	0.012828
GI_KVRa614	TGTGAGTTGTTTGGCATGGGTGA	GGAGGGTGAGCAAATCACAGCTCA	$(TG)^{22}$	26	197-290	0.185	0.962	0.941	0.005254
GI_KVRa615	TGTGAGGGGTGAGGTTGAGGCT	ACAAACGCATCCCCACTCTCGG	$(AT)^{6}$	27	283-379	0.259	0.953	0.933	0.006829
GI_KVRa651	TGGGTGGCAAATTTGGGAGGAAA	TGCCGCCCAAGGAGAGAGGAAA	$(AC)^{8}$	24	185-277	0.2	0.971	0.95	0.006622
GI_KVRa723	TGCACCAGGAGGGTCACAGACT	ACAACGAGGCCTTCCAACAGGA	$(AC)^{10}$	21	412-488	0.143	0.926	0.904	0.011916
GI_KVRa747	TGACAGATCGACAGGCTAGACTCGAA	TCGCCCCCGTCTATGTATCAGTC	$(AT)^{6}$	25	432-531	0.192	0.962	0.941	0.006535
GI_KVRa748	TGAATGCCGAGAGCAATTGTGCC	TCACATCACAAGGCTTGCTCAAACA	$(TA)^{6}$	33	140-214	0.519	0.979	0.96	0.003290
GI_KVRa834	GTGCACATGTCGCCATAAAGATGGA	ACCTACCCCTCCATAACATGCCTT	$(AT)^{6}$	16	105-180	0.133	0.853	0.828	0.036897
GI_KVRa861	GGCCCATGGCCTCCTCTCATACAA	TGGGGAAGGACAATTAAGTCGGGA	$(TA)^{6}$	15	103-185	0.138	0.721	0.695	0.087401
GI_KVRa862	GGCACATGTGTCTACACCGCAC	TGTGGACAGGTAGGGTCACAGGT	$(AT)^{7}$	9	233-294	0.143	0.855	0.819	0.037316
GI_KVRa961	CCACACACAAAATGCCACAATTCCA	TGTGCGTGTGTGGTTGACAGGT	$(CA)^{6}$	14	99-124	0.286	0.847	0.816	0.036213
GI_KVRb069	AGACATCCGTCACCGGGCTCAT	TGCCATTTGTATGTGTTGTTGGCGG	$(CA)^{7}$	10	99-125	0.214	0.873	0.841	0.029837
GI_KVRb130	ACCCGCATTCACAATGCACATACA	GTGGCGCTATTGGGAAATGAGTACA	$(CA)^{7}$	8	233-341	0.000	0.86	0.823	0.033681
GI_KVRb131	ACCCCTAACGGTGGGTTCGTCA	TCGAGGGTCCTTGAGTTCTCCCCT	$(AT)^{6}$	13	99-190	0.148	0.905	0.879	0.017689
GI_KVRb132	ACCCCTAACGGTGGGTTCGTCA	TGGCCTTCGGTTGAGTTGTCCC	$(AT)^{6}$	10	117-157	0.429	0.774	0.733	0.067668
GI_KVRb174	ACACCGGTAAGGTGGTGAGAAGGA	ACACACAGAGTACCCCATATACGCACA	$(TG)^{7}$	12	101-148	0.25	0.783	0.749	0.054954
GI_KVRb175	ACACCGGTAAGGTGGTGAGAAGGA	ACACAGAGTACCTCACATACGCACA	$(TG)^{7}$	18	100-165	0.517	0.915	0.891	0.016365
GI_KVRb176	ACACCCGATCCCATTCCGACCT	ACACCAACCACGCTCCCTTCCT	$(TA)^{7}$	24	453-524	0.276	0.945	0.925	0.008223
GI_KVRb200	AACTACCATCAAACATCACCAACACGA	TGGAAGGTGTTGAGGTCGGCCA	$(CA)^{6}$	22	430-514	0.32	0.957	0.934	0.009077
GI_KVRb201	AACGGCTAGCTTTTCAACTGACTGT	TGGTAAGTCGATTGTTGGGCTTCG	$(TA)^{6}$	17	116-179	0.16	0.913	0.887	0.017850
GI_KVRa975	CACCCCATACACAACCACATTCCC	GGTGTATGTGCCTGGATAAATGAAGGT	$(CA)^{6}$	23	201-285	0.103	0.938	0.918	0.009212
GI_KVRa976	CACATCCTTACATGTACACGGTCCAC	CTGACCGGCTAAACATACAAGTTCCA	$(TA)^{7}$	20	316-397	0.083	0.926	0.901	0.016775
GI_KVRa977	CACATAAGGAACAACAACAAGGCCTCA	GCCGGAGGCCGTACAATTGTGTT	$(AT)^{7}$	24	99-171	0.433	0.856	0.835	0.031996
GI_KVRa978	CAATCTCATTCCTAGACAACCTGCACA	AGTTGATCCAGGATTTGGCGAGGGT	$(AC)^{6}$	20	99-148	0.414	0.933	0.912	0.011202
GI_KVRa979	CAAGGCTGCTCGGACGTCGAAT	ATCCCACCGGCTCGAGCAAGAA	$(CT)^{6}$	23	428-582	0.286	0.905	0.883	0.015518
GI_KVRa980	CAACATGCTTCAACCAAGCACATACAA	TGCTACTACCTTAGGAGACATGCATCA	$(TG)^{11}$	21	112-198	0.444	0.942	0.92	0.009296
GI_KVRa981	CAACAAAGGGCATTCATGCACACA	TTGGGGGAGGAACCAAGCAAGT	$(AT)^{6}$	24	313-399	0.633	0.955	0.936	0.006817
GI_KVRb047	AGCGAGGACAAGGGAAAGGACG	TGGCGGATATGTGTGCTTGGCG	$(TA)^{7}$	19	323-365	0.36	0.911	0.885	0.018187

Table 1 Contd....

GI_KVRb048	AGCGAATGCATGCGTGTAGCGA	ACGATCACCTTGGGGACGCTCA	$(AT)^6$	19	472-527	0.261	0.871	0.846	0.031785
GI_KVRb204	AACCCAGTGAGTGTAATGCGAATTGT	TGTTGTTGGCTTATAGCCGAATGTGA	$(CA)^7$	21	102-195	0.107	0.948	0.927	0.007728
GI_KVRb205	AACCCAATGAGTGTAATGCCAGTTGT	ACTGTGGTTGGCTTATGGCCTGA	$(CA)^6$	21	103-197	0.5	0.919	0.898	0.015233
GI_KVRb206	AACAGGACCGGTGTGCGGTTGA	TCCGCACATGTGTCCACACCAA	$(TA)^8$	21	201-341	0.423	0.909	0.885	0.016389
GI_KVRb207	AACACGTGGCAGACGCTCAAGG	TGGTGAGGTCGGTCCAAACAGGA	$(AT)^6$	8	117-178	0.233	0.793	0.757	0.070882
GI_KVRb208	AACACGCGCGAGGACATACTGC	CCAAGCCTCCTCTCCCATTTGTGC	$(TA)^6$	7	154-171	0.679	0.774	0.72	0.077586
GI_KVRb209	AACACCTGCACGGGTTTCGTGG	ACTTTCCATCTCGACCACGCCG	$(TA)^7$	10	330-413	0.000	0.89	0.86	0.023726
GI_KVRb213	AAAGGACCGGCGAAGAAAGCGG	CCCAGCTCAAACCGATGCCCAA	$(AG)^6$	10	134-250	00.000	0.881	0.85	0.026089
GI_KVRb214	AAAGAGAGGTCATCTTAGTGAGGGGG	TGTTGGCTTGGTCGTAACGGCT	$(GT)^6$	6	150-251	0.148	0.792	0.742	0.062789
GI_KVRb219	TGTTGGGAAGTAAAAGGAGGGAGCA	TGACCTAGGCATCCATCTCCCCT	$(TGT)^5$	7	113-178	0.5	0.785	0.733	0.063197
GI_KVRb220	TGTGGGGATGGCAAATGAGGTGA	TGCCATTCGGTTGGGGCATACT	$(CAC)^5$	10	143-173	0.115	0.829	0.788	0.044338
GI_KVRb234	TGGCGTGCAGTTCTTCCTCCCA	GGGATCGCATCCAACATTCATTTCCA	$(CAA)^5$	3	173-215	0.154	0.335	0.303	0.304896
GI_KVRb242	TGCAACAACAGGCTCAGGCACA	TGGTGGAGGCACGGGTTGAACA	$(CCA)^5$	15	189-215	0.5	0.907	0.881	0.018089
GI_KVRb243	TGAGCGACCGTGCCTGATGTTG	AGGGCTCCCTCACCCTCTACCTTA	$(CAG)^5$	13	141-171	0.36	0.864	0.83	0.032098
GI_KVRb341	ACAAGCATGCCAAACGTAGCCGA	TGAAGAAGTGCCCAACCCCACT	$(TGG)^5$	12	136-170	0.517	0.78	0.741	0.071213
GI_KVRb352	AAGACGGGTGGCGGTGGAGAAA	AGAAGCGAACCCTCTCCTCCTGA	$(TCT)^8$	13	362-403	0.552	0.866	0.835	0.033609
GI_KVRb357	TGACAATACGTGGGGAGATCCGT	TGTTCAGGCTCAATCCCTTCGTGC	$(AATA)^7$	16	115-191	0.000	0.886	0.861	0.021333
GI_KVRb368	TCCGTGCCAATTCCCTGGCAAC	TGACCTGTCGCCTTAGCTACCCT	$(AAAAT)^5$	17	249-310	0.192	0.925	0.9	0.014054
GI_KVRb373	AGCTAGGGGGCAACCTGTACCA	TGCTATTGAATTCGTGTTGGTGGTGA	$(CAATAC)^5$	8	151-168	0.481	0.818	0.778	0.048049
GI_KVRa011	TCCGTCCATCCGTTCGTCCGTT	ACCGGATGGGATCCAGCGATGT	(CGTC) 6cgtt $(CGTC)^7$	12	100-136	0.172	0.75	0.722	0.074675

Table 2: Summary of Genetic Analysis

	Mean	Range
Polymorphic Information Content (PIC)	0.8416	0.303- 0.96
Observed Heterozygosity (Ho)	0.2813	0.000- 0.679
Expected Heterozygosity(He)	0.8701	0.335- 0.979
Allele per locus	16.395	3- 33
Probability of Identity (PI)	0.03506	0.00329- 0.304896

Total number of Alleles : 787 Total probability of Identity : 8.132729e-080

ACKNOWLEDGEMENT

Authors acknowledge financial support from UNEP/GEF regional project "Conservation and Sustainable Use of Cultivated and Wild Tropical Fruit Diversity: Promoting Sustainable Livelihoods, Food Security and Ecosystem Services"

REFERENCES

Baliga, M. S., Bhat, H. P., Pai, R. J., Boloor, R., and Palatty, P. L. 2011. The chemistry and medicinal uses of the underutilized Indian fruit tree *Garcinia indica* Choisy (kokum): A review. *Food Res. Inter.,* **44**:1790-1799.

Hareesh, T. S. and Vasudeva, R. 2010. Regeneration pattern of *Garcinia indica* Choisy. in the Western Ghats of Uttara Kannada. National Symposium on Garcinia Genetic Resources: linking diversity, livelihood and Management (Eds.) Vasudeva, R., B.S. Janagoudar, B.M.C. Reddy, Bhuwon Sthapit and H.P. Singh., College of Forestry, Sirsi. pp 40-46.

Kalinowski, S.T., Taper, M.L., and Marshall, T.C., 2007, Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molec. Ecol.* **16**: 1099-1006.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y. and Tang, J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler., *Giga Sci.* , **1**:18.

Palkar, R.S. and Sellappan, K. 2019. Genetic diversity between and within the natural populations of *Garcinia indica* (Thouars) Choisy: A high value medicinal plant from Northern Western Ghats of India using ISSR markers. *Journal of Applied Research on Medicinal and Aromatic Plants*, **15:**100219.

Ravishankar K. V., Dinesh M.R., Nischita P. and Sandya B. S. 2015. Development and characterization of microsatellite markers in mango (*Mangifera indica*) using next-generation sequencing technology and their transferability across species., *Mol. Breed.,* **35**:93

Ravishankar, K.V., Anand, L. and Dinesh, M.R. 2000. Assessment of genetic relatedness among a few Indian mango cultivars using RAPD markers. *J of Hortic. Sci. Biotech.* **75**: 198 – 201

Ravishankar, K.V., Vasudeva, R., Hemanth, B., Sandya, B.S., Sthapit, B. R., Parthasarathy, V.A. and Rao, V.R. 2017. Isolation and characterization of microsatellite markers in Garcinia gummi-gutta by next-generation sequencing and cross-species amplification. *J. Genetics*, **96**: 213-218.

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

Thatte, K.S., Khandekar, R.G. and Deodhar, M.A. 2012. Assessment of diversity in *Garcinia indica* (Dupetit-Thouars.) Choisy. using morphological and molecular markers. *J. Tropical Agri.*, **50**(1):30-36.

Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S. G., 2012. Primer3 - new capabilities and interfaces., *Nucl. Acids Res.,* **40:** e115

Wagner, H.W. and Sefc K.M., 1999. IDENTITY 1.0 Centre for Applied Genetics. University of Agricultural Sciences, Vienna. Austria. http://www.boku.ac.at/zag/forsch/identity.htm.

(Received on 12.02.2020, Revised on 15.04.2020, Accepted on 14.05.2021)