OPEN ACCESS

# International Journal of Applied Sciences and Biotechnology

A Rapid Publishing Journal

ISSN 2091-2609

## Indexing and Abstracting

CrossRef, Google Scholar, Global Impact Factor, Genamics, Index Copernicus, Directory of Open Access Journals, WorldCat, Electronic Journals Library (EZB), Universitätsbibliothek Leipzig, Hamburg University, UTS (University of Technology, Sydney): Library, International Society of Universal Research in Sciences (EyeSource), Journal Seeker, WZB, Socolar, BioRes, Indian Science, Jadoun Science, Jour-Informatics, Journal Directory, JournalTOCs, Academic Journals Database, Journal Quality Evaluation Report, PDOAJ, Science Central, Journal Impact Factor, NewJour, Open Science Directory, Directory of Research Journals Indexing, Open Access Library, International Impact Factor Services, SciSeek, Cabell's Directories, Scientific Indexing Services, CiteFactor, UniSA Library, InfoBase Index, Infomine, Getinfo, Open Academic Journals Index, HINARI, etc.

CODEN (Chemical Abstract Services, USA): IJASKD

Vol-3(2) June, 2015

**Available online at:**
http://www.ijasbt.org
&
http://www.nepjol.info/index.php/IJASBT/index

SEM-Biotech Publishing

Impact factor*: **1.422**
Scientific Journal Impact factor#: **3.419**
Index Copernicus Value: **6.02**

*Impact factor is issued by Universal Impact Factor. Kindly note that this is not the IF of Journal Citation Report (JCR).
#Impact factor is issued by SJIF INNO SPACE.

For any type of query and/or feedback don't hesitate to email us at: editor.ijasbt@gmail.com

ISSN (Online): 2091-2609
DOI Prefix: 10.3126/ijasbt
International Journal of Applied Sciences and Biotechnology
(IJASBT)

Research Article

# DEVELOPMENT OF EST-SSR MARKERS TO ASSESS GENETIC DIVERSITY IN *ELETTARIA CARDAMOMUM* MATON

**N. Anjali[1], Sowmya S. Dharan[2], F. Nadiya[1] and K. K. Sabu[1]***

[1]Biotechnology and Bioinformatics Division, Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Palode, Thiruvananthapuram 695562
[2]Department of Plant Biotechnology, College of Agriculture, Vellayani, Thiruvananthapuram 695522

*Corresponding author's email: sabu@jntbgri.res.in

## Abstract

*Elettaria cardamomum* Maton is one of the most ancient and valuable spice crops. Cardamom is cultivated following intensive pesticide usage where alleles present in the wild cardamom genotypes could positively contribute towards genetic improvement of the cultivars. However, the genetic map or whole-genome sequence of *E. cardamomum* is not available and very limited information on simple sequence repeat (SSR) markers are publicly available. We have tested whether SSRs from *Curcuma longa* can be used to analyze genetic diversity *E. cardamomum*.

**Keywords**: EST-SSR; microsatellite; genetic diversity; *Elettaria cardamomum;* cardamom; marker development

## Introduction

*Elettaria cardamomum* Maton also known as the 'Queen of spices' is one of the most ancient and valuable spice crops. It is the third most expensive spice in the world after saffron and vanilla. Cardamom belongs to Zingiberaceae family which is one of the largest families of monocotyledons, comprising of about 52 genera and 1500 species. Cardamom is believed to be originated in the moist evergreen forest of Western Ghats of Southern India and many wild populations are still confined to this region (Ravindran and Madhusoodanan, 2002). Ecosystem diversity is very limited in cardamom and majority of the diversity in cardamom comes from varietal diversity (Madhusoodanan *et al.,* 1994).

The study of genetic diversity and interspecific or intergeneric relationships among a number of species are now being extensively carried out with the help of molecular markers (Bandopadhyay *et al.,* 2004). SSRs can be classified into genomic SSRs and EST-SSRs based on the original sequences used to identify microsatellite region (Wei *et al.*, 2011). In the past, development of SSR markers following conventional approach of creating genomic library, hybridisation with tandemly repeated oligonucleotides and sequencing of clones have been very expensive and time consuming (Scott *et al.,*2000). Primers used to amplify SSR region in this procedure are species specific so that markers developed in one taxon cannot be transferred to another (Ellis and Burke 2007). EST-derived SSR markers show good transferability across taxonomic boundaries and can be used as anchor markers for comparative mapping and evolutionary studies. In the present study, 20 EST-SSR primer pairs were custom synthesized from 206 set of EST-SSR primers developed from ESTs of *C. longa*, which also belongs to the Zingiberaceae family.

## Materials and Methods

### *EST-SSR primer design*

Using the keyword '*Curcuma longa'* a total of 12,678 EST sequences were downloaded from NCBI Genbank (ncbi.nlm.nih.gov/dbEST/index.html) in Nov. 2013. Due to the vast amount of ESTs deposited in public databases, a large number of EST-SSRs have been developed, and the polymorphism and transferability of EST-SSRs were checked in many plant species (Aggarwal *et al.*, 2007; Luro *et al.*,2008; Poncet *et al.*, 2006) including cereals (Kantety *et al.*,2002), wheat (Yu *et al.*, 2004), citrus (Chen *et al.*,2006), coffee (Aggarwal *et al.*, 2007), rubber (Feng *et al.*, 2009), chickpea (Choudhary *et al.*, 2009), peanut (Liang *et al.*,2009) and grape (Huang *et al.*, 2011).

Retrieved sequences were then uploaded in TRIMEST program of EMBOSS suite (genome.csdb.cn/cgi-bin/emboss/trimest) to remove the poly-A and poly-T stretches of the ESTs corresponding to the poly-A tails of eukaryotic mRNA (Kumpatla and Mukhopadhyay, 2005). The sequences were then assembled using contig assembly program CAP3 in the mobyle portal platform (mobyle.pasteur.fr/cgi-bin/portal.py). Totally, 5050 potential unique ESTs including 3051 contigs and 1999 singletons were generated and 290 SSR regions were identified from this non-redundant dataset using the software WEBSAT (wsmartins.net/websat/). Among the 290 SSRs identified, 84 could not be used to design the primers as the flanking sequences were too short. Hence, 206 primer pairs were designed using WEBSAT. The quality of the designed primers was checked using NETPRIMER (premierbiosoft.com/netprimer/index.html). A selected set of 20 primers were synthesized at Integrated DNA Technologies (IDT), New Delhi.

***Validation of EST-SSR primers in E. cardamomum***

The transferability of 20 EST-SSR primer pairs developed from ESTs of curcuma were tested cardamom. All the cardamom accessions including wild collections (W), landraces (L) and released varieties (RV) were obtained from the cardamom germplasm conservatory of JNTBGRI. They were originally collected from different parts of Kerala and germplasm collection available at Indian Cardamom Research Institute (ICRI) and Cardamom Research Station (CRS) in Idukki. Total genomic DNA was isolated from young leaves using DNeasy plant mini kit (Qiagen, USA). The DNA concentration and quality were analyzed using Biophotometer (Eppendorf India Ltd) and agarose gel electrophoresis (3%) at 70V. The genomic DNA was diluted to a concentration of 50 ng/μl for PCR amplification.

Polymerase chain reaction (PCR) was performed in the selected accessions in a 25 μl reaction mixture that contained 50 ng template DNA, 1× PCR buffer, 200 μM of each of the four dNTPs, 15pm of each of the forward and reverse primers and one unit of Taq DNA polymerase. The following PCR conditions were used: 94°C for 2 mins, followed by 35 cycles of 94°C for 30 sec, specific annealing temperature for 1 min, 72°C for 2 mins, and 7 mins at 72°C for the final extension. Amplification was carried out on Agilent Technologies thermal cycler (Agilent Technologies, Malaysia). The amplified products were resolved in 3% agarose gel at 70 V for 3 hrs and visualised using ethidium bromide in gel documentation system (UVP, UK).

***Analysis of genetic diversity***

The bands obtained with each primer were scored keeping the expected size range of the PCR products as reference. Basic statistics including the number of observations (*No*) for a marker locus calculated based the number of nonmissing genotypes observed in a sample, allele frequency (*Af*), heterozygosity (*H*) as the proportion of heterozygous individuals in the population, gene diversity (*D*) as the probability that two randomly chosen alleles from the population are different and polymorphism information content (PIC) (Bostein *et al.*, 1980) as an estimate of allelic variation which range from 0 to 1 where 0 indicates no allelic variation (Hildebrand *et al.*, 1992) were obtained using PowerMarker Ver3.25 (Liu and Muse, 2005). Nei's genetic distance (Schneiderbauer *et al.*, 1991) was computed for each pair of the accessions. The UPGMA dendrograms based on genetic distance values were constructed using TreeView V1.6.6 (page 1996).

## Results and Discussion

***Type and frequency of Curcuma longa EST-SSRs***

A total of 12,678 ESTs were used to evaluate the presence of SSR motifs. To eliminate redundant sequences, the CAP3 contig assembly program was used to obtain consensus sequences from overlapping clusters of ESTs. Totally, 5050 potential unique ESTs including 3051 contigs and 1999 singletons were generated. A total of 290 SSRs were identified from 270 unique ESTs. Of those, 18 (about 6.67%) ESTs contained more than one SSR. The occurrences of different repeat units were tri- (52.1%), di- (44.1%), tetra- (3.1%), penta- (0.3%) and hexa-nucleotide (0.3%), of which repeat motifs TA and CT were the most abundant.

***Identification and characteristics of EST-SSRs in E. cardamomum***

Among the 290 SSRs identified, 84 could not be used to design the primers as the flanking sequences were too short. Therefore only 206 primer pairs were designed for the SSRs and 20 were custom synthesized to find out cross-amplification in cardamom. The transferability of 20 EST-SSR primer pairs developed from ESTs of *Curcuma longa* were tested in *Eletteria cardamomum*. For each of the 20 EST-SSR primers in *E. cardamomum*, the name, sequence of the forward and reverse primers, the repeat type, annealing temperature and expected size of the PCR products are listed in Table 1. Analysis of the nucleotide sequences of the EST-SSRs showed that, in *E. cardamomum*, the EST-SSRs corresponded to 50% trinucleotide repeats, 37.5% dinucleotide repeats and 12.5% hexanucleotide repeats. Among these primers, 12 EST-SSR primer pairs produced amplicons in the 18 cardamom accessions.The failure of eight primer pairs to produce amplicons might be possibly due to the location of the primers across splice sites, large introns, chimeric primer(s), or poor-quality sequences (Varshney *et al.*, 2005).

***Genetic variation***

The allele frequency (for the major allele) was 0.82 (which ranged from 0.50 to 1.00), number of alleles per loci was 1.62, gene diversity was 0.23 (ranged from 0 to 0.50), heterozygosity was 0.31 (ranged from 0 to 1.00), PIC was 0.18 and inbreeding coefficient was -0.34 (ranged from -1.00 to 0.81). Perusal of the marker data showed that four microsatellite loci were monomorphic in all the accessions (with few missing bands in some of the accessions). These loci were not included in the preceding variability analysis. For the microsatellite loci (at accession level), allele frequency ranged from 0.44 (ICRI-5) to 0.81 (wild), number of alleles ranged from 0.88 (ICRI-1) to 1.75 (wild), and PIC from 0.14 (wild) to 0.48 (ICRI-5 and ICRI-7).

**Table 1:** Details of primers designed for genetic analysis of cardamom

| S. N. | Primer | Primer Sequence (5′→ 3′) | Repeat motif | Annealing temperature | Expected product size (bp) |
|---|---|---|---|---|---|
| 1 | CaSSR 21F | TCACGCTAAATGGATGGTCTAC | AT(AAT)7A | 49.5 | 292 |
| | CaSSR 21R | TATCTACCCACAGCGGAAGTTT | | | |
| 2 | CaSSR 22F | TGAGAGGGGAAGATAAGACCAA | (CGC)7 | 51.5 | 253 |
| | CaSSR 22R | GGAACTGTGGCAGGAGATGTAT | | | |
| 3 | CaSSR 23F | GAGGGAAGAGAGGAGAAGGAGA | (GA)16 | 54.9 | 357 |
| | CaSSR 23R | TCAAGATGTCTGGTATTGGTGG | | | |
| 4 | CaSSR 24F | GTTGAGGAACAGCGACGAG | (GAA)7 | 52.8 | 394 |
| | CaSSR 24R | AACACTTGCTTCCCTTACTCCA | | | |
| 5 | CaSSR 25F | AGGTTTCTTCTTTAGGCGTCGT | (CT)6 | 57.1 | 392 |
| | CaSSR 25R | GAGATGGGGCGAAAATGG | | | |
| 6 | CaSSR 26F | CTGGGATGGGTATCTACAATGA | (ATC)6 | 52.9 | 299 |
| | CaSSR 26R | CAGTGAGTCCACAGAAAGCAAT | | | |
| 7 | CaSSR 27F | AGCAGTTGAGAGGGTTCTTGAG | (CT)7 | 49.5 | 306 |
| | CaSSR 27R | ATTATTCCCTTCCCGCATTTAC | | | |
| 8 | CaSSR 28F | AAGTCCTCCAAGAACAACAACG | (CAG)6 | 49.5 | 372 |
| | CaSSR 28R | CTTTAGCCCAAGTGATAGACGG | | | |
| 9 | CaSSR 29F | TCCTCCTCCACCTCCTCC | C(TGCTCC)7TGCTC | 58.2 | 351 |
| | CaSSR 29R | GGTTTTCACCTTCCCAACTCTA | | | |
| 10 | CaSSR 30F | CAAGAACAAGAAACAACATCGC | (CCG)8 | 51.9 | 358 |
| | CaSSR 30R | GTTCCAGACGATAACAACGACA | | | |
| 11 | CaSSR 31F | ATCTCAACCTGCTGCCTCTG | G(AGG)5AG | 50.5 | 378 |
| | CaSSR 31R | TTCATCGTAACATCCACAATCG | | | |
| 12 | CaSSR 32F | CTCCTCCTCCTCCTCCTCAT | (TCC)7 | 57.1 | 370 |
| | CaSSR 32R | GCATACCTGTTTCAGAGTGGC | | | |
| 13 | CaSSR 33F | TGAAACATAACTTCTTGGAGCG | (TA)13 | 49.5 | 297 |
| | CaSSR 33R | TCTCTCTCTCACACACACACACA | | | |
| 14 | CaSSR 34F | CAGAATCAGCAAAACAAGCAAC | (AT)6 | 50.5 | 260 |
| | CaSSR 34R | TATGGGCAGTCTTAGGCAATCT | | | |
| 15 | CaSSR 35F | CTTGACAGGACAGCAACAGAAC | (AGC)6 | 56.1 | 156 |
| | CaSSR 35R | CGATGACAGAAGAGAGAGAGCA | | | |
| 16 | CaSSR 36F | ACCGCCCTCCTACTTCTTCTC | (CTT)17 | 62.0 | 270 |
| | CaSSR 36R | ACTCCGTGTGATACTTGTGCG | | | |
| 17 | CaSSR 37F | GGAGGAGGAGAAGAAGAAGGAG | (TC)6 | 50.5 | 314 |
| | CaSSR 37R | CGCAACACACAGACATCTATCA | | | |
| 18 | CaSSR 38F | AAACAGCAACATCAGTCAAACG | (CTT)6 | 53.9 | 269 |
| | CaSSR 38R | CAGAGTCACCAGTGCCCTTC | | | |
| 19 | CaSSR 39F | ACCAGTCTTCTCTTTCCGCTC | (CCT)6 | 59.0 | 182 |
| | CaSSR 39R | CTCCACTCCAGGTAGAGCATTC | | | |
| 20 | CaSSR 40F | CAATGGGACTACAGTGGCG | (GCG)6 | 49.2 | 245 |
| | CaSSR 40R | ATAAACAAACTCAACAGCAGCG | | | |

**Table 2:** Genetic variability of the 3 cardamom types estimated using SSR markers

| Group | *N* | *No* | *Af* | *Na* | *D* | *H* | *PIC* | *f* |
|---|---|---|---|---|---|---|---|---|
| L | 6.00 | 5.25 | 0.77 | 1.75 | 0.30 | 0.47 | 0.23 | -0.50 |
| RV | 6.00 | 5.63 | 0.70 | 2.00 | 0.37 | 0.50 | 0.29 | -0.28 |
| W | 6.00 | 4.88 | 0.67 | 1.88 | 0.39 | 0.54 | 0.30 | -0.29 |

There were considerable differences in genetic variability within the 3 genotype groups (Table 2). Even though *N* was 6, *No* varied from 4.88 to 5.63 which indicate extend of missing values. The *Na* ranged from 1.75 to 2, *D* from 0.30 to 0.39. *H* from 0.47 to 0.54, PIC from 0.23 to 0.30 and *f* from -0.50 to -0.28. The present analysis revealed lower diversity for L, moderate for RV and higher for W.

***Phylogenetic analysis***

The pair-wise genetic distance matrix (Schneiderbauer *et al.,* 1991) was computed on the basis of SSR data. The genetic distance values ranged from 0 to 0.45 with a mean value of 0.15. Cluster analysis was performed on the SSR data following the UPGMA (unweighted pair-group method using arithmetic average) method and the dendrograms were constructed through TREEVIEW (Page 1996) showing overall genetic relatedness among the individuals (Fig. 1). The accessions studied were clustered into two main groups, one with 13 and other with the remaining 5 accessions. The first group has 2 subgroups. The first subgroup consists of 4 accessions of which 2 (C24 and C53) are either wild or represents abandoned cultivars where 2 others (C55 and C63) are landraces currently popular among the farmers. The second subgroup is further divided into 2 groups; the first one of which contains 4 accessions where C60 and C56 are landraces, C61 is a released variety and C70 is may be wild or an escape from a plantation which was actually collected from a forest land whereas the second one consists of 5 accessions of which 3 are released varieties and the remaining 2 are landraces. Of the remaining 5 accessions in the second major group, two are released varieties and three are wild accessions. The dendrogram revealed a complex distribution pattern which is in agreement with a previous report (Ashitha *et al.*, 2013) and indicates diverse nature of the accessions.

The study revealed occurrence of different repeat units were tri- (52.1%), di- (44.1%), tetra- (3.1%), penta- (0.3%) and hexa-nucleotide (0.3%), of which repeat motifs TA and CT were the most abundant. The SSR analysis using 12 microsatellite loci generated altogether 211 alleles and based on which, an assessment of genetic diversity was carried out. Eight of these markers revealed low (0.14) to moderate (0.48) polymorphism information content (PIC) across 18 genotypes of *E. cardamomum*, while 4 markers were found to be monomorphic.

This study gives an insight into the frequency, type and distribution of curcuma EST-SSRs and demonstrates successful development as well as utility of EST-SSR markers in cardamom. These EST-SSR markers could contribute to the knowledge and current resource of molecular markers and these markers would be useful for quantitative trait mapping, and marker-assisted selection besides genetic diversity and phylogenetic studies in cardamom.

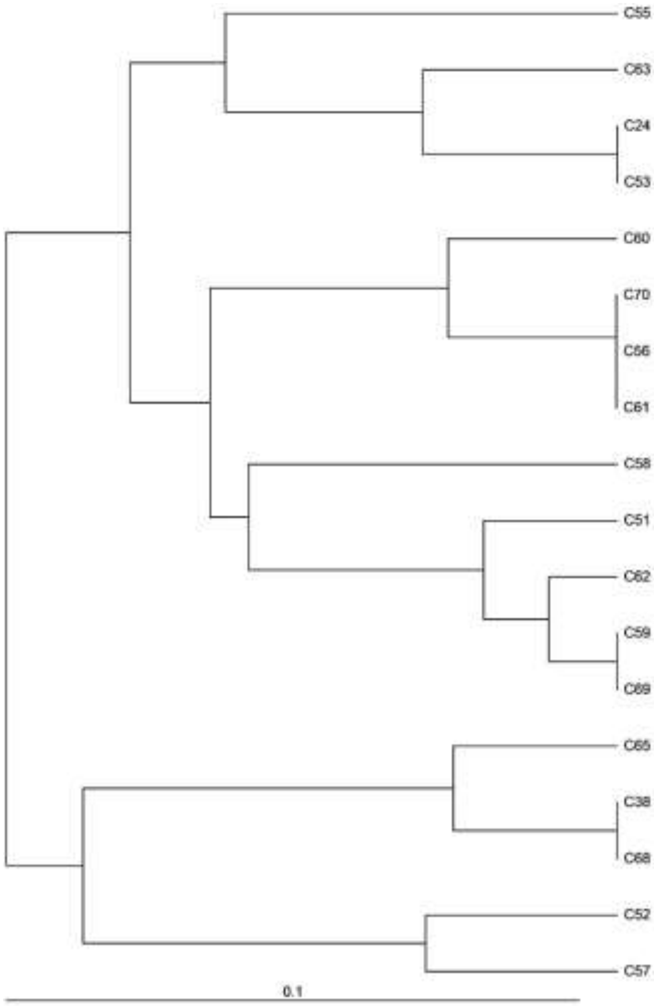


**Fig. 1:** UPGMA based dendrogram of the 18 cardamom accessions

## Acknowledgments

The authors thank the Director, JNTBGRI for providing necessary facilities to carry out the work. Thanks are due to KSCSTE for granting Research Fellowship to AN (010-31/FSHP/10/CSTE) and NF (010-52/FSHP/10/CSTE) to accomplish this study. The authors also thank Dr. Siju Senan for training on SSR development extended to the authors and Mr. Shefeek, S. for collection and management of the cardamom germplasm.

## References

Aggarwal R, Hendre P, Varshney R, Bhat P, Krishnakumar V and Singh L (2007) Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theoretical and Applied Genetics* **114**:359-372. 10.1007/s00122-006-0440-x

Bandopadhyay R, Sharma S, Rustgi S, Singh R, Kumar A, Balyan HS and Gupta PK (2004) DNA polymorphism among 18 species of *Triticum–Aegilops* complex using wheat EST–SSRs. *Plant Science* **166**:349-356. 10.1016/j.plantsci.2003.09.022

Botstein D, White RL, Skalnick MH, and Davies RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics* **32**:314-331.

Chen C, Zhou P, Choi Y, Huang S and Gmitter F (2006) Mining and characterizing microsatellites from citrus ESTs. *Theoretical and Applied Genetics* **112**:1248-1257. 10.1007/s00122-006-0226-1

Choudhary S, Sethy N, Shokeen B and Bhatia S (2009) Development of chickpea EST-SSR markers and analysis of allelic variation across related species. *Theoretical and Applied Genetics* **118**:591-608. 10.1007/s00122-008-0923-z

Feng SP, Li WG, Huang HS, Wang JY and Wu YT (2009) Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). *Molecular Breeding* **23**:85-97. 10.1007/s11032-008-9216-0

Hildebrand CE, Torney DC and Wagner RP (1992) Informativeness of Polymorphic DNA Markers. *Los Alamos Science* **20**:100-102.

Huang H, Lu J, Ren Z, Hunter W, Dowd S and Dang P (2011) Mining and validating grape (*Vitis* L.) ESTs to develop EST-SSR markers for genotyping and mapping. *Molecular Breeding* **28**:241-254. 10.1007/s11032-010-9477-2

Kantety R, La Rota M, Matthews D and Sorrells M (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology* **48**:501-510. 10.1023/A:1014875206165

Liang X, Chen X, Hong Y, Liu H, Zhou G, Li S and Guo B (2009) Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species. *BMC Plant Biology* 9-35. 10.1186/1471-2229-9-35

Luro FL, Costantino G, Terol J, Argout X, Allario T, Wincker P, Talon M, Ollitrault P and Morillon R (2008) Transferability of the EST-SSRs developed on Nules clementine (*Citrus clementina* Hort ex Tan) to other *Citrus* species and their effectiveness for genetic mapping. *BMC Genomics* 9-287. 10.1186/1471-2164-9-287

Madhusoodanan KJ, Kuruvilla KM and Priyadarshan PM (1994) Genetic resources of cardamom. In: Chadha KL and Rethinam P (Eds) *Advances in Horticulture Plantation and Spice Crops* Part I. Vol. 9. New Delhi, India: Malhotra Publishing House; 121–130.

Page RDM (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**:357-358.

Poncet V, Rondeau M, Tranchant C, Cayrel A, Hamon S, de Kochko A and Hamon P (2006) SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. *Mol Genet Genomics* **276**:436-449. 10.1007/s00438-006-0153-5

Ravindran PN and Madhusoodanan KJ (2002) Cardamom**:** the genus *Elettaria*: Taylor and Francis Publishers. 10.4324/9780203216637

Schneiderbauer A, Sandermann Jr H and Ernst D (1991) Isolation of functional RNA from plant tissues rich in phenolic compounds. *Anal. Biochem* **197**:91-95. 10.1016/0003-2697(91)90360-6

Varshney RK, Graner A and Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* **23**:48-55. 10.1016/j.tibtech.2004.11.005

Wei W, Qi X, Wang L, Zhang Y, Hua W, Li D, Lv H and Zhang X (2011) Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC Genomics* 12-451.

Yu JK, Dake TM, Singh S, Benscher D, Li W, Gil B and Sorrells ME (2004) Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. *Genome* **47**:805-818. 10.1139/g04-057