*Full Length Research Paper*

# Genetic diversity in cocoa (*Theobroma cacao* L.) plus trees in Tamil Nadu by simple sequence repeat (SSR) markers

**Velayutham Thondaiman[1], K. Rajamani[1], N. Senthil[2], N. Shoba[1] and A. John Joel[3]**

[1]Department of Spices and Plantation Crops, TNAU, Coimbatore, India.
[2]Centre for Plant Molecular Biology, TNAU, Coimbatore, India.
[3]Department of Plant Genetic Resources, TNAU, Coimbatore, India.

The range of polymorphism of about 27 cocoa plus trees screened in the major cocoa growing regions of Tamil Nadu were assessed using 10 simple sequence repeat (SSR) markers. The gene diversity, genetic differentiation and genetic similarities were analyzed for the cocoa trees. The number of alleles detected by different primers ranged from 0 to 3 and the level of polymorphism was 0 to 100%. The polymorphism information content (PIC) value ranged from 0.000 to 0.677. The higher the PIC value, the more informative was the SSR marker. Hence, primer mTcCIR33 was found to be highly informative. The Jaccard's similarity coefficient for the SSR data set varied from 0.39 to 1.00. The SSR marker profiles resulted in nine clusters at nearly 54% similarity. From this study, it could be inferred that the diversity exists in cocoa plantations in Tamil Nadu and can be exploited in crop improvement research.

**Key words:** Cocoa, simple sequence repeat (SSR) markers, polymorphism, genetic diversity, breeding.

## INTRODUCTION

The cocoa, botanically known as *Theobroma cacao* L. belongs to the family Malvaceae (Alverson et al., 1999). In India, cocoa is cultivated predominantly in four states viz., Kerala, Andhra Pradesh, Tamil Nadu and Karnataka. It occupies an area of 46,318 ha with the production of 12,954 MT and the national productivity is 550 kg dry beans per ha. Kerala leads in production of cocoa from an area of 11,044 hectares with production of 6344 MT.

The productivity of cocoa beans in the state is 592 kg per hectare. Tamil Nadu occupies third position in cocoa cultivation with an area of 9347 ha. It produces 900 MT cocoa beans with the productivity of 443 kg dry beans per hectare (DCCD, 2012). Three distinct varieties of cocoa viz., Forastero, Criollo and Trinitario are cultivated of which, Forastero accounts for 90% of the cocoa beans produced in the world. It is found widely in West Africa

and Brazil. The second type is Criollo, which produces fine and flavor beans, mostly grown in parts of the Caribbean, Venezuela, Papua New Guinea the West Indies, Sri Lanka, East Timor and Java. The third type is the Trinitario variety, which is a cross between Criollo and Forastero (Anonymous, 2010). Cocoa is highly cross pollinated due to the prevalence of self-incompatibility, thereby paving the way for high variability in cocoa population. In India, the early introductions made in the early 20th century included presumably Criollos and Forasteros.

The exact sources from which cocoa was introduced into India are not known. Around 1930, all the Forastero plants available in the country were removed and only Criollos were maintained. When large-scale cultivation of cocoa was taken up in the 1960's, seeds of presumably fresh cocoa beans were introduced from selected plants from Malaysian

---

estates and other sources and subsequent introductions were made in 1970's from Cocoa Research Institute of Ghana and Kew Botanical Garden of United Kingdom. Further systematic introductions were made since 1990 from University of Reading, United Kingdom. Crop improvement research in cocoa is in progress in India especially at Kerala Agricultural University and Central Plantation Crops Research Institute, Vittal for more than 40 years and many improved selections and hybrids were evolved (Nair et al., 1990).

The F$_1$ progenies of these cultivars have been introduced in most of the plantations in Tamil Nadu of India. Tamil Nadu happens to be the potential state for cocoa production on account of more area under irrigated coconut and cocoa Coimbatore and Erode districts. The trees have adapted to the local climatic conditions of the region for many years. High degree of genetic variability has been observed in these plantations. However, the genetic diversity of cocoa existing in Tamil Nadu state has not been scientifically documented. Systematic identification, documentation and conservation of genetic diversity of cocoa either *ex situ* or *in situ* will be the viable approach for improvement of cocoa in Tamil Nadu. Therefore, this study was conducted with the following objectives: to assess the genetic diversity of cocoa in Tamil Nadu using molecular markers and to construct phylogenetic tree to display the relationship among the trees for further crop improvement work.

## MATERIALS AND METHODS

### Cocoa plus trees

About 27 cocoa plus trees screened from the farmer's fields in Tamil Nadu were selected to constitute the subset for the study. The trees were selected in different farms distributed widely in five different plantations of major cocoa growing regions of Tamil Nadu viz, KUL-2, 18, 25, SMJ- 3, 4, 10, 15, 18, 21, 25, 33, 34, 37, 50, SME - 2, 5, 6, 9, 16, 21, 24, 26, 28, 29, VPS- 12, 13 and 15; were identified by metroglyph analysis.

### DNA extraction

The DNA was extracted from young leaves of cocoa following the protocol of Echevarria-Machado et al. (2005) for Malvaceae plant species with slight modification which was based on the extraction procedure of Dellaporta et al. (1983). 0.1 g of tender leaf material was homogenized with 1 ml of extraction buffer, transferred to a 2.0 ml eppendorf tube, and 100 µL of 20% Sodium dodecyl sulfate (SDS) was added. After mixing, the mixture was incubated at 65°C for 10 min. 500 µl of 5 M potassium acetate was added. The tube was shaken vigorously and incubated for 20 min on ice. The tubes were centrifuged at 12,000 rpm for 20 min. The supernatant was transferred to a new 1.5-ml tube to which 300 µl of silica was added and mixed manually for 3 to 5 min. The tubes were centrifuged at 12,000 rpm for 30 s. The pellet was washed twice with 70% ethanol and dried. The pellet was re-suspended in 50 µl of distilled water and incubated at 55°C for 5 min. The tubes were centrifuged at 12,000 rpm for 2 min and supernatant transferred to a new 500-µl eppendorf tube. Aliquot was used for quantification of total DNA in gel electrophoresis.

DNA is visualized and quantified on 0.8% (w/v) agarose gels containing 3 µl / 40 ml of agarose solution in 1 X TBE buffer. The gel was run at 80 V current for 45 min. Then, the gel was visualized under U-V lamp and documented using Alpha Imager. DNA concentration for polymerase chain reaction (PCR) amplification was estimated by comparing the band intensity of a sample with the standard lamda DNA concentration. The dilutions were carried out by dissolving the genomic DNA in appropriate quantity of TE buffer (pH 8.0).

### DNA amplification

DNA from 27 cocoa candidate plus trees was amplified using a set of 10 simple sequence repeat (SSR) primers. The sequence details of the SSR primers were as follows: 2 µl of template DNA, 1.2 µl of dNTP were mixed, and 2 µl of 10 µM primer, 1.5 µl of 10X assay buffer, 0.20 µl of 1U Taq DNA polymerase and sterile water (9.5 µl) were used as reaction mixture. Annealing temperature was calculated based on melting temperature of primer. Denaturation (95°C) and extension (72°C) temperatures were set. After running 35 cycles in PCR (Make:MJ Research), the products were run in 3% agarose gel containing 5 µl of ethidium bromide dye. Gels were documented using Alpha imager TM 1200-Documentation and analysis System of Alpha Innotech Corporation, USA. The primers that produced amplification were used for analyzing diversity in all the selected trees. Allele sizes of PCR products were determined by referring to the 100 bp ladders (Bangalore Genei Cat. No. 105656).

### Data analysis

#### *Polymorphism survey of SSR markers*

Polymorphism survey of SSR markers was carried out by considering only the clear and unambiguous bands. Markers were scored for the presence and absence of the corresponding band among the different trees. The scores '1' and '0' were given for the presence and absence of bands, respectively.

#### *Cluster analysis*

The data obtained by scoring the SSR profiles of different primers were subjected to cluster analysis. Similarity matrix was constructed using Jaccard's coefficient and the similarity values were used for cluster analysis and dendrogram was constructed by unweighted pair-group method using arithmetic averages (UPGMA) with the Sequential Agglomerative Hierarchical and Nested (SAHN) function (Sneath and Sokal, 1973). Data analysis was done using NTSYS-pc version 2.02i (Rohlf, 1998).

#### *Genetic diversity estimation*

After visualizing the gel, amplified fragments of each SSR marker were scored as "1" and "0", where "1" indicated the presence of a specific allele (band) and "0" indicated its absence. Polymorphism information content (PIC) of SSR markers was calculated using the formula developed by Anderson et al. (1993). A PIC value of each locus was calculated as

$$PIC_j = 1 - \sum\nolimits_{l=1\,to\,L} P^2_{ij},$$

Where, $P_{ij}$ is the relative frequency of the i$^{th}$ allele for the locus j and was summed across all the alleles (L) over all lines. PIC provided an estimate of the discriminatory power of a locus by taking into

**Table 1.** List of primers used for SSR analysis.

| Primer name | Sequence |
| --- | --- |
| mTcCIR6 F | TTCCCTCTAAACTACCCTAAAT |
| mTcCIR6 R | TAAAGCAAAGCAATCTAACATA |
| mTcCIR11 F | TTTGGTGATTATTAGCAG |
| mTcCIR11 R | GATTCGATTTGATGTGAG |
| mTcCIR12 F | TCTGACCCCAAACCTGTA |
| mTcCIR12 R | ATTCCAGTTAAAGCACAT |
| mTcCIR15 F | CAGCCGCCTCTTGTTAG |
| mTcCIR15 R | TATTTGGGATTCTTGATG |
| mTcCIR18 F | GATAGCTAAGGGGATTGAGGA |
| mTcCIR18 R | GGTAATTCAATCATTTGAGGATA |
| mTcCIR22 F | ATTCTCGCAAAAACTTAG |
| mTcCIR22 R | GATGGAAGGAGTGTAAATAG |
| mTcCIR24 F | TTTGGGGTGATTTCTTCTGA |
| mTcCIR24 R | TCTGTCTCGTCTTTTGGTGA |
| mTcCIR28 F | GATCAATCAGAAGCAAACACAT |
| mTcCIR28 R | TAAAGCAGCCTACCAAGAAAAG |
| mTcCIR33 F | TGGGTTGAAGATTTGGT |
| mTcCIR33 R | CAACAATGAAAATAGGCA |
| mTcCIR40 F | AATCCGACAGTCTTTAATC |
| mTcCIR40 R | CCTAGGCCAGAGAATTGA |

account, not only the number of alleles that are expressed, but also the relative frequencies of those alleles. PIC values ranged from 0 (monomorphic) to 1 (very highly discriminative), with many alleles in equal frequencies. Genetic diversity estimate related analyses were done using NTSYSpc ver.2.02i (Rohlf, 2000). Genetic similarities (GS) between pairs of trees were measured by the DICE similarity coefficient based on the proportion of shared alleles with SIMQUAL module. Genetic distances between pairs of lines were estimated as GD or D = 1 - GS. The clustering of trees was done based on a similarity matrix using UPGMA algorithm following SAHN module. The clustering result was used to construct a dendrogram following TREE module (Ali et al., 2008).

## RESULTS AND DISCUSSION

All the SSR markers used in this study was developed at CIRAD by Lanaud et al. (1999). Among the 10 SSR primers used in the present study, 7 primers produced discrete, scorable and unambiguous bands. The details of SSR primers used for assessing the mole-cular diversity among 27 candidate plus trees are pre-sented in Table 1. The PCR product size obtained by the ampli-fication of SSR primers ranged from 190 to 350 bp. The amplification pattern of SSR markers namely mTcCIR11 and mTcCIR33 in 27 candidate plus trees are presented in Figures 2 (mTcCIR11) and 3 (mTcCIR33), respectively.

### Allele diversity of SSR marker analysis

The allele diversity and polymorphism information content (PIC) was calculated for all the SSR markers used in this study (Table 2). Highest PIC was recorded by the primer

TcCIR33 (0.667) and it was found to be the lowest for the primers mTcCIR22, mTcCIR28 and mTcCIR40 (0.000). The higher the PIC value, the more informative is the SSR marker. Hence, primer mTcCIR33 was found to be highly informative.

### Cluster analysis

The banding pattern of the SSR markers scored in the form of binary data was used for computing Jaccard's similarity index. The similarity index values obtained for each pair wise comparison among the 27 candidate plus trees are presented in Table 4. The similarity coefficients based on 10 SSR markers ranged from 0.39 to 1.00. Among the 27 candidate plus trees studied, the highest similarity index (1.00) was recorded between the candi-date plus trees SMJ 10 and VPS 13 and also between SME 21 and SME24. The lowest similarity index (0.39) was recorded by the candidate plus trees SMJ 25 and KUL 25. The similarity values obtained for each pair wise comparison of SSR markers among the 27 candidate plus trees were used to construct dendrogram based on hierarchical clustering and the results are presented in Figure 1. 27 candidate plus trees were grouped into eight clusters at nearly 54% similarity level. The cluster size varied from eight (cluster IV) to 1 (Clusters II, V and VI). The list of all the eight clusters along with the trees included is presented in Table 3. The cluster I consisted of the candidate plus trees of SMJ 33, SMJ 34 and SMJ 50. The cluster II consisted of SMJ 3. The cluster III

**Table 2.** SSR marker profile across candidate plus trees of cocoa.

| Primer code | Number of alleles | Number of polymorphic alleles | Polymorphism (%) | PIC |
|---|---|---|---|---|
| mTcCIR6 | 3 | 3 | 100 | 0.389 |
| mTcCIR11 | 3 | 3 | 100 | 0.577 |
| mTcCIR12 | 3 | 3 | 100 | 0.599 |
| mTcCIR15 | 3 | 3 | 100 | 0.607 |
| mTcCIR18 | 3 | 3 | 100 | 0.541 |
| mTcCIR22 | 1 | 0 | 0 | 0.000 |
| mTcCIR24 | 2 | 1 | 50 | 0.133 |
| mTcCIR28 | 1 | 0 | 0 | 0.000 |
| mTcCIR33 | 3 | 3 | 100 | 0.667 |
| mTcCIR40 | 1 | 0 | 0 | 0.000 |
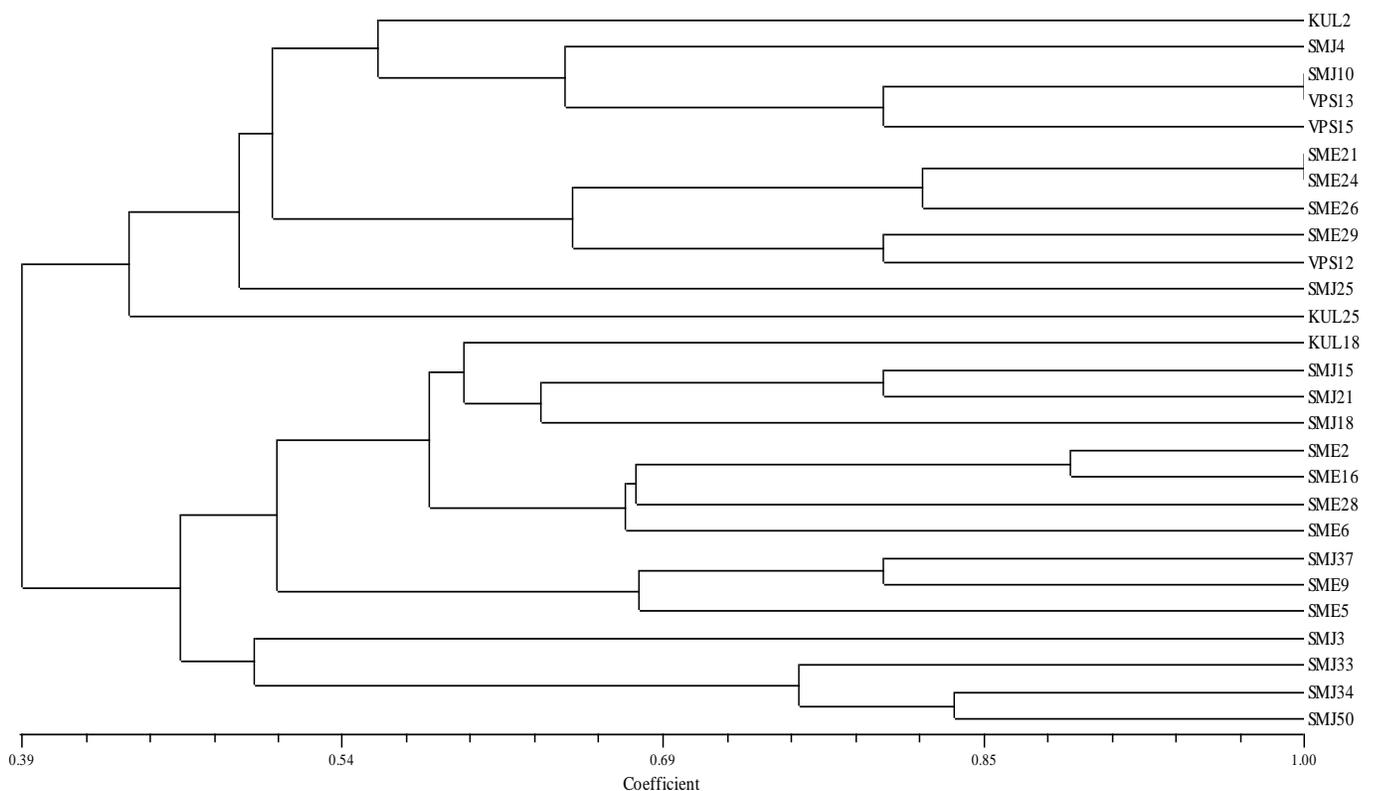


**Figure 1.** Dendrogram of cocoa plus trees based on SSR analysis.

consisted of SMJ 37, SME 9 and SME 5. The cluster IV consisted of KUL 18, SMJ 15, SMJ 21, SMJ 18, SME 2, SME6, SME 16 and SME 28. The cluster V consisted of KUL 25. The cluster VI consisted of SMJ 25. The cluster VII consisted of SME 21, SME 24, SME 26, SME 29 and VPS 12. The cluster VIII consisted of KUL 2, SMJ 4, SMJ 10, VPS 13 and VPS 15.

Among the primers used in the study, the primer mTcCIR11 was able to show 330 bp allele specific to SME 5 candidate plus tree and the primer mTcCIR15

showed 250 and 230 bp allele specific to SMJ 3. Similar studies were carried out by Motamayor (2001), Motamayor et al. (2002) in Central America, Efombagn et al. (2006) in South Cameroon, Johnson et al. (2009) in Trinidad and Tobagao, and Zhang et al. (2009) and Aikpokpodion et al. (2009) in West Africa. These findings suggest that ample diversity exists in cocoa in Tamil Nadu. These trees have to be observed for both yield and quality parameters for few more years and promising trees have to be clonally raised and tested before using

**Table 4.** Genetic similarity co-efficient among the candidate plus trees of cocoa based on SSR analysis.

| | KUL 2 | KUL 18 | KUL 25 | SMJ3 | SMJ4 | SMJ10 | SMJ15 | SMJ18 | SMJ21 | SMJ25 | SMJ33 | SMJ34 | SMJ37 | SMJ50 | SME2 | SME5 | SME6 | SME9 | SME16 | SME21 | SME24 | SME26 | SME28 | SME29 | VPS12 | VPS13 | VPS15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KUL 2 | 1.000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| KUL 18 | 0.313 | 1.000 | | | | | | | | | | | | | | | | | | | | | | | | | |
| KUL 25 | 0.444 | 0.316 | 1.000 | | | | | | | | | | | | | | | | | | | | | | | | |
| SMJ3 | 0.500 | 0.438 | 0.400 | 1.000 | | | | | | | | | | | | | | | | | | | | | | | |
| SMJ4 | 0.692 | 0.400 | 0.529 | 0.500 | 1.000 | | | | | | | | | | | | | | | | | | | | | | |
| SMJ10 | 0.643 | 0.467 | 0.688 | 0.471 | 0.769 | 1.000 | | | | | | | | | | | | | | | | | | | | | |
| SMJ15 | 0.438 | 0.692 | 0.421 | 0.563 | 0.438 | 0.500 | 1.000 | | | | | | | | | | | | | | | | | | | | |
| SMJ18 | 0.438 | 0.692 | 0.421 | 0.563 | 0.438 | 0.600 | 0.714 | 1.000 | | | | | | | | | | | | | | | | | | | |
| SMJ21 | 0.533 | 0.692 | 0.350 | 0.667 | 0.533 | 0.500 | 0.846 | 0.714 | 1.000 | | | | | | | | | | | | | | | | | | |
| SMJ25 | 0.467 | 0.615 | 0.529 | 0.412 | 0.571 | 0.643 | 0.643 | 0.643 | 0.643 | 1.000 | | | | | | | | | | | | | | | | | |
| SMJ33 | 0.263 | 0.533 | 0.474 | 0.529 | 0.412 | 0.389 | 0.471 | 0.471 | 0.563 | 0.412 | 1.000 | | | | | | | | | | | | | | | | |
| SMJ34 | 0.368 | 0.563 | 0.500 | 0.647 | 0.444 | 0.500 | 0.500 | 0.688 | 0.588 | 0.444 | 0.750 | 1.000 | | | | | | | | | | | | | | | |
| SMJ37 | 0.375 | 0.750 | 0.300 | 0.412 | 0.375 | 0.438 | 0.643 | 0.533 | 0.643 | 0.467 | 0.412 | 0.444 | 1.000 | | | | | | | | | | | | | | |
| SMJ50 | 0.333 | 0.643 | 0.556 | 0.529 | 0.412 | 0.563 | 0.563 | 0.667 | 0.563 | 0.500 | 0.733 | 0.867 | 0.500 | 1.000 | | | | | | | | | | | | | |
| SME2 | 0.533 | 0.692 | 0.350 | 0.563 | 0.353 | 0.500 | 0.714 | 0.714 | 0.714 | 0.533 | 0.471 | 0.588 | 0.643 | 0.667 | 1.000 | | | | | | | | | | | | |
| SME5 | 0.313 | 0.538 | 0.316 | 0.533 | 0.313 | 0.375 | 0.467 | 0.375 | 0.467 | 0.313 | 0.438 | 0.471 | 0.750 | 0.533 | 0.571 | 1.000 | | | | | | | | | | | |
| SME6 | 0.412 | 0.643 | 0.333 | 0.625 | 0.333 | 0.471 | 0.563 | 0.667 | 0.563 | 0.412 | 0.444 | 0.647 | 0.600 | 0.625 | 0.786 | 0.643 | 1.000 | | | | | | | | | | |
| SME9 | 0.412 | 0.643 | 0.400 | 0.529 | 0.500 | 0.563 | 0.667 | 0.563 | 0.667 | 0.500 | 0.529 | 0.556 | 0.846 | 0.625 | 0.667 | 0.769 | 0.625 | 1.000 | | | | | | | | | |
| SME16 | 0.467 | 0.750 | 0.300 | 0.600 | 0.375 | 0.438 | 0.769 | 0.643 | 0.769 | 0.571 | 0.500 | 0.529 | 0.692 | 0.600 | 0.917 | 0.615 | 0.714 | 0.714 | 1.000 | | | | | | | | |
| SME21 | 0.600 | 0.438 | 0.474 | 0.444 | 0.500 | 0.667 | 0.471 | 0.471 | 0.471 | 0.500 | 0.368 | 0.474 | 0.600 | 0.529 | 0.667 | 0.643 | 0.625 | 0.733 | 0.600 | 1.000 | | | | | | | |
| SME24 | 0.600 | 0.438 | 0.474 | 0.444 | 0.500 | 0.667 | 0.471 | 0.471 | 0.471 | 0.500 | 0.368 | 0.474 | 0.600 | 0.529 | 0.667 | 0.643 | 0.625 | 0.733 | 0.600 | 0.500 | 1.000 | | | | | | |
| SME26 | 0.500 | 0.533 | 0.474 | 0.444 | 0.500 | 0.667 | 0.471 | 0.563 | 0.471 | 0.600 | 0.444 | 0.556 | 0.500 | 0.625 | 0.667 | 0.533 | 0.625 | 0.625 | 0.600 | 0.857 | 0.857 | 1.000 | | | | | |
| SME28 | 0.500 | 0.533 | 0.400 | 0.625 | 0.412 | 0.563 | 0.667 | 0.667 | 0.667 | 0.500 | 0.444 | 0.556 | 0.600 | 0.625 | 0.786 | 0.643 | 0.733 | 0.733 | 0.714 | 0.733 | 0.733 | 0.733 | 1.000 | | | | |
| SME29 | 0.600 | 0.438 | 0.556 | 0.529 | 0.500 | 0.667 | 0.563 | 0.563 | 0.563 | 0.714 | 0.368 | 0.474 | 0.412 | 0.529 | 0.667 | 0.438 | 0.625 | 0.529 | 0.600 | 0.733 | 0.733 | 0.733 | 0.733 | 1.000 | | | |
| VPS12 | 0.571 | 0.500 | 0.444 | 0.600 | 0.571 | 0.643 | 0.643 | 0.533 | 0.643 | 0.692 | 0.412 | 0.444 | 0.467 | 0.500 | 0.643 | 0.500 | 0.600 | 0.600 | 0.692 | 0.714 | 0.714 | 0.714 | 0.714 | 0.846 | 1.000 | | |
| VPS13 | 0.643 | 0.467 | 0.688 | 0.471 | 0.769 | 1.000 | 0.500 | 0.600 | 0.500 | 0.643 | 0.389 | 0.500 | 0.438 | 0.563 | 0.500 | 0.375 | 0.471 | 0.563 | 0.438 | 0.667 | 0.667 | 0.667 | 0.563 | 0.667 | 0.643 | 1.000 | |
| VPS15 | 0.643 | 0.571 | 0.588 | 0.389 | 0.643 | 0.846 | 0.500 | 0.600 | 0.500 | 0.643 | 0.389 | 0.500 | 0.533 | 0.563 | 0.600 | 0.375 | 0.471 | 0.563 | 0.533 | 0.667 | 0.667 | 0.667 | 0.471 | 0.563 | 0.533 | 0.846 | 1.000 |

**Table 3.** Cluster composition of candidate plus trees of cocoa for SSR markers.

| Cluster number | Number of genotype | List of genotypes included |
|---|---|---|
| I | 3 | SMJ 33, SMJ 34 and SMJ 50 |
| II | 1 | SMJ 3 |
| III | 3 | SMJ 37, SME 9 and SME 5 |
| IV | 8 | KUL 18, SMJ 15, SMJ 21, SMJ 18, SME 2, SME 6, SME 16 and SME 28 |
| V | 1 | KUL 25 |
| VI | 1 | SMJ 25 |
| VII | 5 | SME 21, SME 24, SME 26, SME 29 and VPS 12 |
| VIII | 5 | KUL 2, SMJ 4, SMJ 10, VPS 13 and VPS 15 |
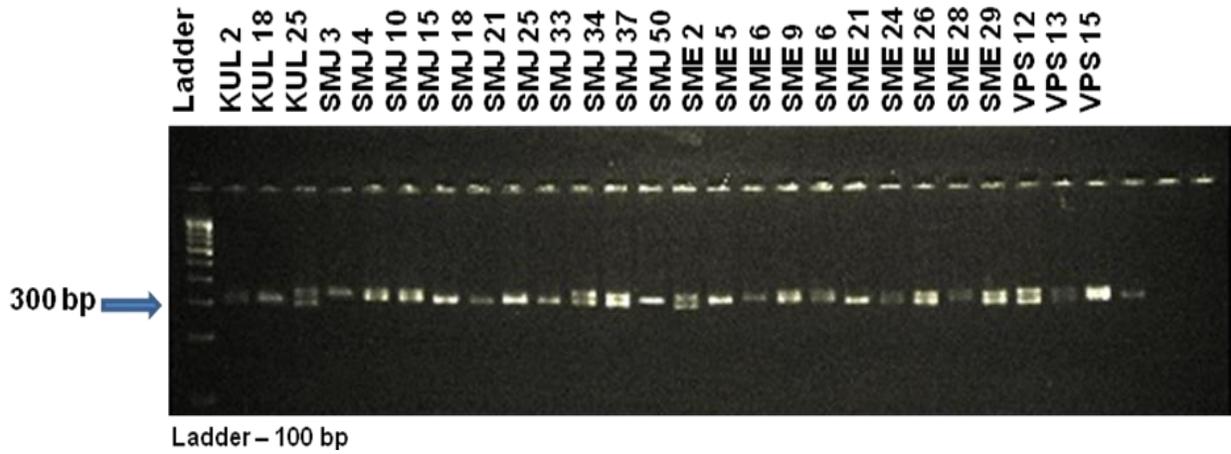
**Figure.2.** SSR marker profile of cocoa candidate plus trees generated by the primer mTcCIR11.
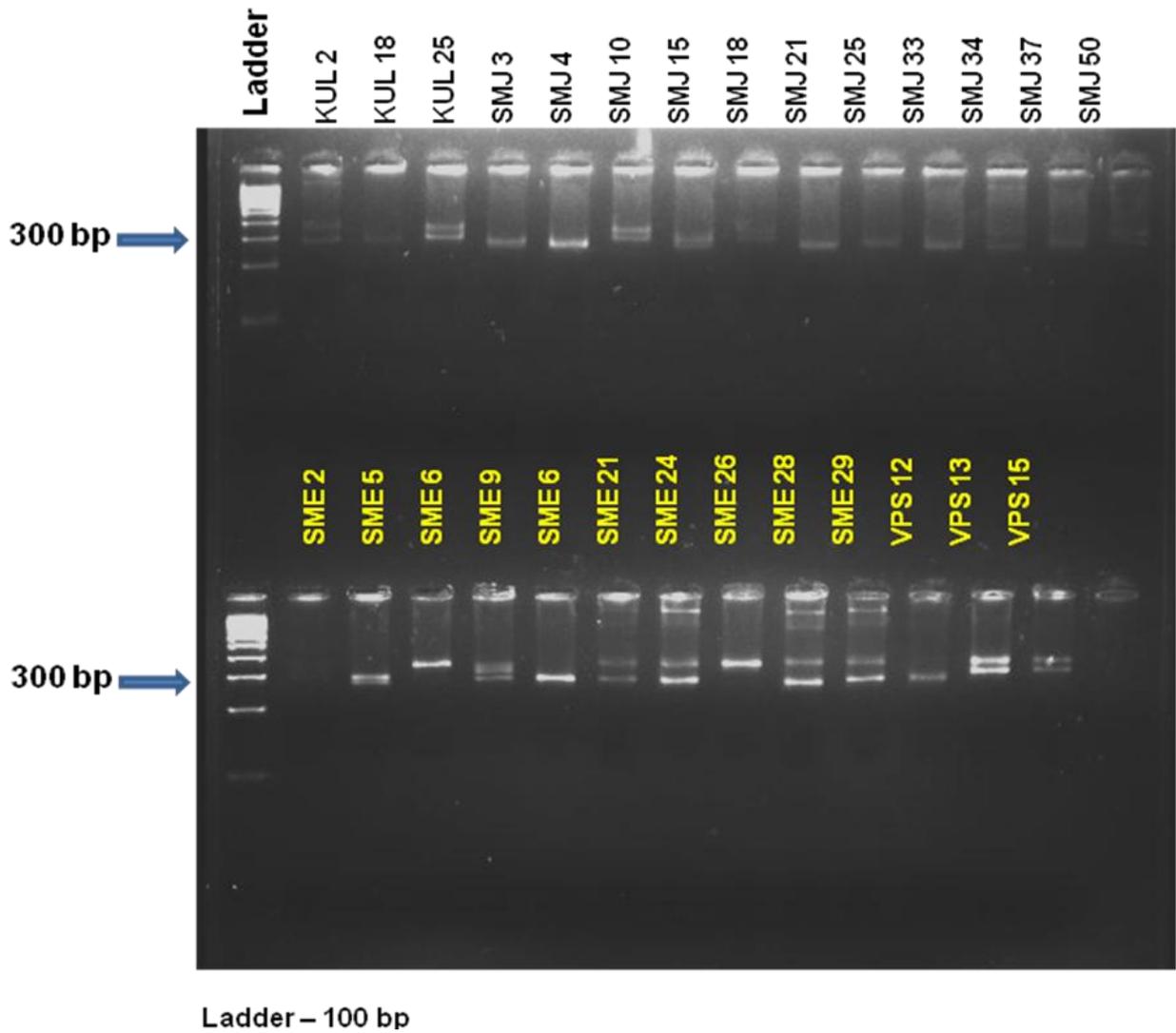


**Figure 3.** SSR marker profile of cocoa candidate plus trees generated by the primer mTcCIR33.

in breeding programmes.

## REFERENCES

Aikpokpodion PO, Motamayor JC, Adetimirin VO, Adu-Ampomah Y, Ingelbrecht I, Eskes AB, Schnell RJ, Kolesnikova-Allen M (2009). Genetic diversity assessment of sub-samples of cacao (*Theobroma cacao* L.) collections in West Africa using simple sequence repeats marker. Tree Genet. Genomes. 5: 699–711.

Ali ML, Rajewski JF, Baenziger PS, Gill KS, Eskridge KM, Dweikat M (2008). Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. Mol. Breed. 21: 497–509.

Alverson WS, Whitlock BA, Nyffler R, Bayer, C. Baum DA (1999). Phylogeny of the core Malvales: evidence from ndhF sequence data. Am. J. Bot. 86: 1474-1486.

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

Anonymous (2010). Information on Cocoa. http://www.unctad.org/infocomm/ anglais/cocoa/ characteristics. htm#descript.

DCCD 2012. Directorate of Cashew and Cocoa Development. (http://dccd.gov.in/stat2.htm).

Dellaporta L, Wood J, Hicks JB (1983). A plant DNA minipreparation: version II. Plant Mol. Biol., Rep. 1: 19–21.

Echevarria-Machado I, Sanchez-Cach LA, Hernandez-Zepeda C, Rivera-Madrid R Moreno-Valenzuela OA (2005). A simple and efficient method for isolation of DNA in high mucilaginous plant tissues. Mol. Biotech. 31: 129-135.

Efombagn MIB, Sounigo O, Nyasse S, Manzanares-Dauleux M, Cilas C, Eskes MAB Kolesnikova-Allen (2006). Genetic diversity in cocoa germplasm of southern Cameroon revealed by simple sequences repeat (SSR) markers. African J. Biotech. 5 (16): 1441-1449.

Johnson ES, Bekele FL, Brown SJ, Song Q, Zhang D, Meinhardt LW, Schnell RJ (2009). Population Structure and Genetic Diversity of the Trinitario Cacao (*Theobroma cacao* L.) from Trinidad and Tobago. Crop Sci. 49:564–572.

Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A, Lagoda PJL (1999). Isolation and characterization of microsatellite in *Theobroma cacao* L. Molecular Ecology. 8: 2141-2143.

Motamayor JC (2001). Etude de la diversité génétique et de la domestication des cacaoyers du groupe Criollo (*Theobroma cacao* L.) à l'aide de marqueurs moléculaires. Ph. D. Thesis. Université Paris XI. p. 179.

Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A, Lanaud C (2002). Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity. 89: 380-386.

Nair RV, Appaiah GN, Nampoothiri KUK (1990). Variability in some exotic accessions of cocoa in India. Ind. Cocoa Arecanut and Spices J. 14: 49-51.

Rohlf FJ (1998). NTSYS-pc. Numerical taxanomy and multi-variate analysis system, Version 2.02, Exeter software. State University of New York at Stony Brook, New York.

Rohlf FJ (2000). NTSYS-pc: Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket. 119.

Sneath PHA, Sokal RR (1973). Taxonomic structure. In: Numerical Taxonomy: The principles and practice of numerical classification. W.H. Freeman, San Francisco. pp. 188-308.

Zhang D, Boccara M, Motilal M, Mischke S, Jhonson ES, Butler DR, Bailey B, Mainhardt L (2009). Molecular characterization of an earliest cocoa collection from upper Amazon using microsatellite DNA markers. Tree Genetics and Genomics. 5: 595-607.