DNA FINGERPRINTING OF PROMISING COCOA (*Theobroma cacao* L.) VARIETIES OF KAU

By

SUJITH S.S.

(2014-11-226)

THESIS

Submitted in partial fulfilment of the requirement for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680656

KERALA, INDIA

2016

DECLARATION

I hereby declare that this thesis entitled **"DNA fingerprinting of promising cocoa (*Theobroma cacao* L.) varieties of KAU"** is a bonafide record of research done by me during the course of research and that the thesis has not previously formed the basis for the award of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara
Date: 09/09/2016

Sujith S.S.
2014-11-226

CERTIFICATE

Certified that this thesis entitled **"DNA fingerprinting of promising cocoa (*Theobroma cacao* L.) varieties of KAU"** is a record of research work done independently by **Sujith S.S.** (2014-11-226) under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to him.

Vellanikkara
Date: 09/09/2016

Dr. Minimol J.S.
(Chairman, Advisory Committee)
Assistant Professor
Plant Breeding and Genetics
Cocoa Research Centre
Vellanikkara, Thrissur

CERTIFICATE

We, the undersigned members of the advisory committee of **Mr. Sujith S.S.** (2014-11-226) a candidate for the degree of **Master of Science in Agriculture** with major field in **Plant Biotechnology,** agree that this thesis entitled **"DNA fingerprinting of promising cocoa (*Theobroma cacao* L.) varieties of KAU"** may be submitted by **Mr. Sujith S.S.,** in partial fulfilment of the requirement for the degree.

Dr. Minimol J.S.
(Chairman, Advisory committee)
Assistant Professor, Plant Breeding and Genetics,
Cocoa Research Centre, Vellanikkara, Thrissur

9/9/2016

Dr. P. A. Valsala
(Member, Advisory Committee)
Professor & Head, CPBMB,
College of Horticulture,
Vellanikkara, Thrissur

Dr. Rose Mary Francies
(Member, Advisory Committee)
Professor, Seed Technology Unit,
College of Horticulture,
Vellanikkara, Thrissur

Dr. Deepu Mathew
(Member, Advisory Committee)
Assistant Professor,
CPBMB, College of Horticulture,
Vellanikkara, Thrissur

Dr. Rekha K.
(External Examiner)
Senior Scientist, Biotechnology
Division, Rubber Research Institute
of India, Kottayam

ACKNOWLEDGEMENT

First of all, I bow my head before the Great Truth, the ***Almighty****, whose blessings were always with me which enabled me to undertake this endeavor successfully.*

My words cannot express the deep sense of gratitude and indebtedness to my mentor ***Dr. Minimol J. S.****, Assistant Professor, Cocoa Research Centre, Kerala Agricultural University and chairman of my advisory committee. With great respect, I wish to place my heartfelt thanks to her for the inspiring guidance, untiring interest, unstinted co-operation, esteemed advice, constructive criticism, valuable suggestion, and immense help rendered during the investigation and preparation of the thesis and thus kindling within me the sacred flame of social commitment.*

I am extremely thankful to ***Dr. P. A. Valsala,*** *Professor and Head, Centre for Plant Biotechnology and Molecular Biology, College of Horticulture and member of my advisory committee for her unfailing support, valuable suggestions and guidance rendered to me for the completion of the research programme and preparation of the thesis. I am genuinely indebted to her for the constant encouragement and affectionate advice rendered throughout the academic programme.*

I am deeply obliged to ***Dr. Deepu Mathew,*** *Assistant Professor, CPBMB, College of Horticulture and member of my advisory committee for his regular guidance, support and enthusiasm, relevant suggestions and constructive criticism rendered throughout the period of investigation and preparation of thesis.*

I am highly thankful to ***Dr. Rose Mary Francies,*** *Professor, Seed Technology Unit, College of Horticulture and member of my advisory committee for the valuable help rendered for interpreting the results and formulating the thesis.*

I have immense pleasure to express my sincere thanks to ***Dr. Suma B. S.****, Professor and Head, Cocoa Research Center for the avail of facilities at the center during the course of my research.*

I am highly indebted to ***Dr. Pradeep Kumar,*** *Professor, Department of Olericulture,* ***Shri. Krishnan,*** *Professor, Department of Agricultural Statistics,* ***Dr. Abida P. S.,*** *Professor and Head, Department of Genetics and Plant Breeding, RARS, Pattambi and* ***Dr. M. R. Shylaja,*** *Professor, CPBMB for their valuable and constructive suggestions and guidance rendered throughout the period of pursuance of my post-graduation.*

My sincere thanks to my dear friends Ajmal, Praveen, Manju, Sameer, Mangalam, Sandesh, Manjesh, Reddy, Darshan, Naresh, Anju, Rose, Aswini, Vinu, Brinda, Saha and all other batchmates who helped me in several ways for the completion of the venture. I wish to express my sincere thanks to all non-teaching staff members and labourers of CPBMB for their whole-hearted cooperation and timely assistance. I also appreciate and acknowledge the facilities provided CPBMB for the successful completion of my research work.

I have immense pleasure to express my sincere thanks to all staff members of Cocoa Research Centre whose helping hands, love and affection fetched a remarkable place during my research days in the university.

I take this opportunity to extend my profound gratitude to ***Dr. A. T. Francis,*** *Librarian, Central Library & College of Horticulture, and other library staff for the facilities offered.*

The award of ***KAU fellowship*** *is thankfully acknowledged.*

Above all, I am forever beholden to my beloved Pappa, Amma, my elder brother, little sister and my grandparents for their unfathomable love, bondless affection, personal sacrifice, incessant inspiration and constant prayers, which supported me to stay at hard times.

Sujith S. S.

Dedicated to my beloved
Parents and Teachers

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE NO.
1	INTRODUCTION	1
2	REVIEW OF LITERATURE	3
3	MATERIALS AND METHODS	23
4	RESULTS AND DISCUSSION	39
5	SUMMARY	83
	REFERENCE	I-XXIV
	ANNEXURES	
	ABSTRACT	

LIST OF TABLES

TABLE NO.	TITLE	PAGE NO.
1	Details of ISSR primers used in this study	33
2	Details of SSR primers used in this study	36
3	Potential yield of cocoa varieties	41
4	Pod qualitative characters of popular cocoa varieties of KAU	42
5	Mean values of yield contributing characters of cocoa varieties	43
6	DNA quantification result of eight cocoa varieties	46
7	Details of amplification pattern obtained with 35 ISSR primers	49
8	Details of ISSR primers selected for fingerprinting	50
9	Details of amplification pattern obtained with 37 SSR primers	56
10	Details of SSR primers selected for fingerprinting	58
11	Details of amplification with selected 10 ISSR primers	73
12	Details of amplification with selected 11 SSR primers	74
13	Clustering based on qualitative characters of cocoa varieties	77
14	Clustering based on quantitative characters of cocoa varieties	77
15	Homology between qualitative and quantitative clusters	77
16	Clustering of cocoa varieties based on D^2 statistics	79
17	Inter cluster distance between cocoa varieties	79
18	Inter varietal distance between eight cocoa varieties	79
19	Clustering based on ISSR data	81
20	Clustering based on SSR data	82
21	Clustering based on ISSR and SSR data	82
22	Homology between quantitative and molecular clustering data	82

LIST OF PLATES

PLATE NO.	TITLE	BETWEEN PAGES
1	Field view of the cocoa varieties used in the study	23-24
2	Isolated DNA used for the study	46-47
3	Amplification pattern generated with primer UBC 810	54-55
4	Amplification pattern generated with primer UBC 815	54-55
5	Amplification pattern generated with primer UBC 826	54-55
6	Amplification pattern generated with primer UBC 827	54-55
7	Amplification pattern generated with primer UBC 835	54-55
8	Amplification pattern generated with primer UBC 841	54-55
9	Amplification pattern generated with primer UBC 842	54-55
10	Amplification pattern generated with primer UBC 846	54-55
11	Amplification pattern generated with primer UBC 855	54-55
12	Amplification pattern generated with primer UBC 866	54-55
13	Amplification pattern of eight cocoa varieties using primers mTcCIR 8 and mTcCIR 11	62-63
14	Amplification pattern of eight cocoa varieties using primers mTcCIR 12 and mTcCIR 18	62-63
15	Amplification pattern of eight cocoa varieties using primers mTcCIR 24 and mTcCIR 33	62-63
16	Amplification pattern of eight cocoa varieties using primers mTcCIR 40 and mTcCIR 42	62-63
17	Amplification pattern of eight cocoa varieties using primers SHRSTc 49 and SHRSTc 51	62-63
18	Amplification pattern of eight cocoa varieties using primers SHRSTc 64	62-63

LIST OF FIGURES

FIGURE NO.	TITLE	BETWEEN PAGES
1	Colour chart developed using UBC 810	54-55
2	Colour chart developed using UBC 815	54-55
3	Colour chart developed using UBC 826	54-55
4	Colour chart developed using UBC 827	54-55
5	Colour chart developed using UBC 835	54-55
6	Colour chart developed using UBC 841	54-55
7	Colour chart developed using UBC 842	54-55
8	Colour chart developed using UBC 846	54-55
9	Colour chart developed using UBC 855	54-55
10	Colour chart developed using UBC 866	54-55
11	Colour chart developed using mTcCIR 8 and mTcCIR 11	62-63
12	Colour chart developed using mTcCIR 12 and mTcCIR 18	62-63
13	Colour chart developed using mTcCIR 24 and mTcCIR 33	62-63
14	Colour chart developed using mTcCIR 40 and mTcCIR 42	62-63
15	Colour chart developed using SHRSTc 49 and SHRSTc 51	62-63
16	Colour chart developed using SHRSTc 64	62-63
17	Fingerprint developed using selected ISSR and SSR primers - CCRP1	68-69
18	Fingerprint developed using selected ISSR and SSR primers - CCRP2	68-69
19	Fingerprint developed using selected ISSR and SSR primers - CCRP4	68-69
20	Fingerprint developed using selected ISSR and SSR primers - CCRP5	68-69
21	Fingerprint developed using selected ISSR and SSR primers - CCRP6	68-69
22	Fingerprint developed using selected ISSR and SSR primers - CCRP7	68-69

23	Fingerprint developed using selected ISSR and SSR primers - CCRP8	68-69
24	Fingerprint developed using selected ISSR and SSR primers - CCRP9	68-69
25	Overall DNA fingerprinting data	70-71
26	Polymorphic information content of 10 ISSR and 11 SSR primers	74-75
27	Resolving power of 10 ISSR and 11 SSR primers	74-75
28	Dendrogram based on similarity coefficient for qualitative and quantitative characters of cocoa varieties	76-77
29	Cluster diagram and proposed cocoa garden lay out	79-80
30	Dendrogram based on similarity coefficient for ISSR and SSR analysis of cocoa varieties	81-82
31	Combined dendrogram based on similarity coefficient for ISSR and SSR analysis of cocoa varieties	81-82

ABBREVIATIONS

%	Percentage
>	Greater than
β	Beta
µg	Microgram
µl	Microliter
AFLP	Amplified Fragment Length Polymorphism
bp	Base pair
CCRP	KAU- Cadbury Co-oprative Cocoa Research Project
CRC	Cocoa Research Centre
cm	Centimetre
CPBMB	Centre for Plant Biotechnology and Molecular Biology
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTPs	Deoxyribo Nucleoside Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
EST	Expressed Sequence Tags
g	Gram
IBPGR	International Board of Plant Genetic Resources
ICTG	International Cocoa Genebank
IPR	Intellectual Property Rights
ISSR	Inter Simple Sequence Repeat
KAU	Kerala Agricultural University
kb	Kilo base pairs
L	Litre
M	Molar

MAS	Marker Assisted Selection
mg	Milligram
ml	Millilitre
mM	Milli Molar
ng	Nano gram
°C	Degree Celsius
OD	Optical Density
PBR	Plant Breeders Right
PCR	Polymerase Chain Reaction
pH	Hidrogen ion concentration
PIC	Polymorphic Information Content
PVP	Poly Vinyl Pyrrolidone
QTL	Quantitative Trait Loci
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
SSR	Simple Sequence Repeat
TAE	Tris Acetate EDTA
TE	Tris EDTA
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
UV	Ultra Violet
V	Volts

Introduction

1. INTRODUCTION

Theobroma cacao L. is a major cash crop of tropical countries and belongs to the family Malvaceae. The family is further divided into six sections Andropetalum, Glossopetalum, Oreanthes, Rhytidocarpus, Telmatocarpus and Theobroma (Alverson *et al.,* 1999). Cocoa which is native of Andes in South America (Wood and Lass, 1985) was introduced to India during 1798 (Ratnam, 1961). *T. cacao* has a rich history of human use, beginning with the Olmec and Mayan peoples of the eastern Mexican gulf who believed that cacao was delivered to them directly by God (Dillinger *et al.,* 2000).

Africa is the main cocoa producing continent, with approximately 70 per cent of the world's production. Cameroon alone produces around 4 per cent of the world's cocoa (Bruno *et al.,* 2008). Among tropical commodities traded in the world market, cocoa is third largest product after sugar and coffee. About 40 to 50 million people worldwide are reported to depend on cocoa for their livelihood. About 3.5 million tons of cocoa are produced annually and the demand for cocoa is increasing. Cocoa products are mainly consumed by the population of developed countries (Fouet *et al.,* 2011).

Cocoa is a diploid fruit tree species (2n = 20) (Aikpokpodion, 2012) and the haploid size of *T. cacao* genome is 0.43 pico gram (equivalent to 415,000 kb) *i.e.*, the genome size is only three times as large as that of the *Arabidopsis thaliana* L. which is considered to have the smallest genome among higher plants (Figueira *et al.,* 1992).

In India, the genetic base of cocoa is widened by the systematic introduction of germplasm from University of Reading, England, through Cocoa Research Centre (CRC), Kerala Agricultural University (KAU) (Minimol and Prasnnakumari, 2013). Exploitation of these germplasm has resulted in the release of 15 cocoa varieties from KAU. In India, cocoa is grown as a mixed crop with palm based cropping systems *i.e.*, with coconut and arecanut in traditional zones

of Kerala and Karnataka. In Tamil Nadu and Andhra Pradesh, cocoa is an inter crop under coconut and to some extend in oil palm gardens (Alban *et al.*, 2016).

Starting point of any breeding programme is a genebank. Characterization of available genetic resources is necessary for any breeding programme for crop improvement (Spooner *et al.*, 2005). During the early history of plant breeding, the markers used for characterization were mainly depended on morphological traits. It is evident that, genetic diversity can be evaluated based on morphological markers. But it is difficult to classify the accessions solely based on their morphological characters as they are highly influenced by environmental factors (Ganesan *et al.*, 2014).

DNA markers, that are not subjected to environmental influences act as an efficient tool to identify and differentiate accessions and cultivars which are similar in morphological characteristics and with indistinct traits. DNA finger printing is successfully applied for cultivar identification, controlling seed purity of hybrids and checking the genetic similarity between cultivars (Aikpokpodion *et al.*, 2009). In this scenario, DNA fingerprinting is very much helpful to protect plant varieties and to solve Intellectual Property Right (IPR) issues.

Central sub-committee on crop standards, notification and release of varieties for agricultural crops has made it mandatory to provide DNA finger printing data along with the varieties where ever the proposal for national release/ notification is submitted. Thus, the study will help to characterize the most popular eight cocoa varieties at morphological and molecular levels. This is highly essential to maintain their identity and to protect the IPR.

In this context, the present study was taken up with the following specific objectives.

1. Characterization of eight cocoa varieties using molecular markers Inter Simple Sequence Repeats (ISSR) and Simple Sequence Repeats (SSR).
2. Development of DNA fingerprint specific to each varieties.

2. REVIEW OF LITERATURE

2.1 General background

Cocoa is a native of Andes in South America and was introduced to India during 1798 (Ratnam, 1961), but its commercial cultivation started only during 1970's (Prasannakumari *et al.*, 2009). Africa is the main cocoa producing continent, with approximately 70 per cent of the world's production. Cameroon alone produces around 4 per cent of the world's cocoa (Bruno *et al.*, 2008). Among tropical commodities traded in the world market, cocoa is third largest product after sugar and coffee. About 40 to 50 million people worldwide are reported to depend on cocoa for their livelihood. Around 3.5 million tons of cocoa are produced annually and the demand for cocoa is increasing (Fouet *et al.*, 2011). Cocoa products are mainly consumed by the population of developed countries. Demand towards such products are progressively increasing and in order to meet these needs by the consumers, a sustainable agriculture for producing countries is necessary for a more productive and high quality cocoa (Chia *et al.*, 2011).

Cocoa is a diploid fruit tree species (2n = 20) (Aikpokpodion, 2012) and the haploid size of *T. cacao* genome is 0.43 pico gram (equivalent to 415,000 kb) *i.e.*, the genome size is only three times as large as that of the *Arabidopsis thaliana* (L.) which is considered to have the smallest genome among higher plants (Figueira *et al.*, 1992).

Starting point of any breeding programme is a genebank. The available genetic materials can be effectively utilized only on its appropriate characterization (Spooner *et al.*, 2005). Characterization can be done based on phenotypic characters or at genome level using molecular markers. Polymerase Chain Reaction (PCR) based DNA markers are being successfully used in genotyping, genome mapping and phylogenetic studies in cocoa (Guiltinan, 2007).

In India, the genetic base of cocoa is widened by the systematic introduction of germplasm from University of Reading, England, through Cocoa

Research Centre (CRC), Kerala Agricultural University (KAU) (Minimol and Prasannakumari, 2013). Exploitation of these germplasm has resulted in the release of 10 cocoa varieties from KAU. Central sub-committee on crop standards, notification and release of varieties for agricultural crops had made it mandatory to provide DNA finger printing data along with the varieties where ever the proposal for national release/ notification is submitted. Thus, the study will help to characterize the most popular eight cocoa varieties at morphological and molecular levels. This is highly essential to maintain their identity and to protect the Plant Breeders Right (PBR).

DNA markers are not subjected to environmental influences and therefore act as an efficient tool to identify and differentiate accessions and cultivars which are similar in morphological characteristics with indistinguishable traits.

2.2 Usefulness of genetic markers in plant

Genetic markers are the biological features that are determined by allelic forms of genes or genetic loci and are able to transmit from one generation to another. Thus they can be used as experimental probes or tags to track individuals, tissues, cells, nuclei, chromosomes or genes. There are different kinds of genetic markers employed for different purpose and with different properties, each having its own advantages and disadvantages to assess the genetic variation among natural populations. However, markers used in plant breeding and genetics can be broadly classified into three: morphological markers (based on visually accessible traits), biochemical markers (based on gene product) and molecular markers (based on DNA assay). Ganesan *et al.* (2014) reported that, genetic diversity pattern, genetic makeup and agronomical parameters of a plant are needed to be studied for successful breeding, crop improvement, domestication and large scale cultivation.

2.2.1 Morphological markers in plant characterization

During the early history of plant breeding, the markers used were mainly depended on morphological traits, such as leaf shape, flower colour, pubescence

colour, pod colour, seed colour, seed shape, hilum colour, awn type and length, fruit shape, rind (exocarp) colour and stripe, flesh colour, stem length, etc. It is evident that, genetic diversity can be evaluated based on morphological markers. But it is difficult to classify the accessions solely based on their morphological characters (Ganesan *et al.,* 2014). Morphological markers generally represent polymorphisms based on visible traits, which are easily identified and manipulated. Investigations so far have revealed that morphological markers are useful for classifying the diversity of cocoa populations and germplasm collections (Efombagn *et al.,* 2009). Smith and Smith (1989) concluded that, morphological characterization is the first step in the description and classification of germplasm. But the disadvantage of such markers is that, the expression of phenotype is highly influenced by environmental conditions.

2.2.1.1 Morphological scoring of genotypes

Morphological scoring is simple to conduct and is based on visible traits of the plant. A morphological descriptor should be easily observable, stable under a wide range of environmental conditions and should be highly discriminating (Martinez *et al.*, 2015). The International Board of Plant Genetic Resources (IBPGR, 1981) recommended a list of 65 morphologic descriptors to characterize cocoa germplasm (Lerceteau *et al.*, 1997). From which, Engels (1986) selected ten qualitative and quantitative descriptors of flowers, leaves, beans and fruits to differentiate between the cocoa groups such as Criollo, Forastero and the hybrid Trinitario. Bekele *et al.* (1994) reported that, in cocoa, flower traits are probably the suitable character for differentiating various types. Santos *et al* (2012) conducted a study to analyze the morphological variability of four wild species, one semi-cultivated and one cultivated species of cocoa, indigenous to Brazil. The characterization was done based on 35 quantitative and 13 qualitative traits comprises of leaf, flower, fruit and seed variability.

With the objective to elucidate the relationships between wild and cultivated germplasm, another study was conducted with 600 cocoa accessions from The International Cocoa Genebank, Trinidad (ICGT) (Bekele *et al.,* 2006).

2.2.1.1.1 Visual markers at vegetative stage

Morphological characterization can help cocoa breeders to select the most suitable accessions for further breeding programs (Engels *et al.,* 1980). Usually, phenotypic characterization of the species are conducted with gene bank resources, which involves the use of leaf, flower, fruit and seed descriptors (Engels *et al.,* 1980; Bekele and Bekele 1996). Among these descriptors, fruit phenotypes play an important role in the characterization of cocoa types and populations (Efombagn *et al.,* 2009). Minimol *et al.* (2011) reported that the fruit apex form plays an important role in determining fruit shape. So far, several studies of morphological diversity have been conducted on flowers, fruits and leaves of cocoa germplasm accessions (Asna *et al.,* 2014).

Domesticated *T. cacao* had a wide diversity in plant morphology and the three cocoa groups such as Criollo, Forastero and Trinitario were mainly classified based on their morphology especially the fruit and seed characters (Wood and Lass, 1985). They were widely divers in morphological characters such as shape of pod, ridge colour, pod basal constriction, shape of pod apex, pod rugosity, cotyledon colour etc.

2.2.1.1.2 Visual markers at reproductive stage

Cocoa is one among the important perennial plantation crops, and the reproductive characters of the plant is quite unique. The inflorescence is a compressed cyme. Cauliflorous flowers are produced on the leafless portion of woody trunk called 'cushion'. Each cushion will bear up to 50 flowers per season. Individual flowers are pedicellate, pentamerous regular and hermaphrodite in nature. Calyx consist of five sepals (polysepalous) and corolla contains five petals (polypetalous). Androecium comprise of 10 stamens in two whorls of which, outer five represents aborted stamens or staminodes and inner five are fertile. Gynoecium has five carpels and syncarpous in nature. Ovary is superior in nature and is with short style and pentafid stigma. Number of ovules in ovary ranges from 40-60 and are arranged in axial placentation (Prasannakumari *et al.,* 2009).

Cocoa is highly cross pollinated on account of its unique flower structure and existence of self-incompatibility (Minimol and Prasannakumari, 2013). Natural pollination is facilitated by small crawling insects especially, by female midges. Cocoa plants starts flowering after 1-2 years of planting. Anther dehiscence starts by 4-5 am and extends up to 8-10 am. The fruit of cocoa is referred to as pod but botanically it is drupe. Pod shows high variability in size, colour, shape and texture. Seeds of cocoa are called 'beans'. Each pod contains 20-60 beans.

2.2.1.2 Cluster analysis

Cluster analysis was carried out for 294 genotypes using principal component analysis and the distribution of cultivars were in agreement with traditional classification of Criollo, Forestero and their sub divisions (Engels, 1986).

In cocoa, hierarchical average linkage cluster analysis was done for determining morphological diversity associated with morphological descriptors (Bekele and Bekele, 1996). For the study, 100 accessions from the germplasm maintained at International Cocoa Gene bank, Trinidad were used. The analysis indicated that, at 75 per cent level of similarity, nine accessions remained ungrouped and remaining were grouped in to 11 clusters.

Stability of the genetic divergence among five non-commercial cocoa cultivars for a period of five years was investigated (Dias *et al.*, 1997). Based on the data obtained, cluster analysis was performed on five yield components on year basis and further pooled the data. Comparison of year wise D^2 value and pooled D^2 value showed a stable clustering pattern in the favourable years. Similarly in the same year, quantification of multivariate phonetic divergence among germplasm SIC and SIAL was performed by using cluster and principal component analysis (Santos *et al.*, 1997).

Asna (2013) conducted cluster analysis of fifty cocoa germplasm (40 exotic and 10 indigenous) maintained at cocoa research centre, KAU. Thirteen

qualitative and 26 quantitative characters were considered for the study and resulted in 9 clusters for exotic accessions and 5 for indigenous ones at 70 percent similarity level. Clustering using D^2 statistics developed by Mahalanobis (1936) for quantitative characters were also done and resulted in development of five clusters for exotic and three for indigenous group.

2.2.2 Biochemical markers in plant characterization

A biochemical marker is one which is developed from the study of chemical products of gene expression. They are also known as isozyme/allozyme markers or protein markers. The limitations of morphological markers have resulted in the deployment of biochemical techniques such as protein electrophoresis and isozyme makers. Karp *et al.* (1997) reported that, characterization of plants based on combination of morphological and biochemical descriptions, provided more detailed and accurate data than classical phenotypic data. Isozyme technique is cost effective simple and could be used as supplementary markers on molecular genetic map construction. The markers representing the dominant proteins could also be useful in breeding as markers of important genes (Dudnikov, 2003). Isozyme markers in plants resulted in differences in pattern and band intensity by tissue used for the analysis and developmental stage of the plant (Montarroyos *et al.*, 2003). The major drawback of biochemical marker system is the lack of sufficient polymorphism among closely related cultivars (El-Nahas *et al.*, 2011). Isozyme markers have been used to validate the conventional classification, but absence of distinct polymorphism does not allow discrimination between cocoa types (Ronning and Schnell, 1994; Lanaud, 1999).

2.2.3 Molecular markers in plant characterization

In genetics, a molecular marker is a fragment of nuclear, mitochondrial or chloroplast DNA that is associated with specific sequences such as RFLP, RAPD, AFLP, ISSR, SSR etc. Molecular markers are used in molecular biology and biotechnology to identify a particular sequence of DNA from a pool of unknown

DNA. Naturally occurring DNA polymorphism is used in molecular markers to differentiate one individual from the other. DNA markers can be effectively used to detect the presence of allelic variation in the genes corresponding the traits. Application of DNA markers in plant breeding, increases the efficiency and precision of breeding programme. The use of DNA markers in plant breeding is called marker-assisted selection (MAS) and is a crucial component of the new discipline of 'molecular breeding' (Collard and Mackill, 2008).

To be informative, molecular marker must be polymorphic and it should present in different forms among the accessions. Mutant genes present in the genome of an individual can be distinguished from the normal genes by a molecular marker. DNA markers are best candidates for evaluation and selection of plant material with varying characters. Different molecular marker system hold varying properties, each with its own advantages and disadvantages (Karp *et al.*, 1997; Weising *et al.*, 2005). Molecular markers segregate as single genes to progenies and the environmental factors do not affect them. It helps to shorten the process of conventional selection, evaluation and identifying genotypes. Kumar *et al.* (2009) concluded that, the degree of genetic similarity can be determined and even in selecting mutant plants, the one with desirable characters can also be selected by using molecular marker. Initially, RFLP and RAPD markers were developed and have proven their ability to characterize germplasm for many species including *Theobroma cacao* (Figueira *et al.*, 1994; N'goran *et al.*, 1994).

Molecular markers were introduced in 90's and proven more useful for evaluation of genetic diversity, because they are more informative and stable than morphological and agronomic traits (Cohen *et al.*, 1991). Markers can be used for quantitative trait locus (QTL) analysis to observe complex traits in cereals (Hodges, 1991). Markers tightly linked to resistance genes effectively helped for marker assisted gene pyramiding in rice and it reduced the duration of the breeding program (Yoshima *et al.*, 1995).

Smith and Smith (1989) reported that the use of morphological traits alone is not always the best way to evaluate varieties. Morphological analysis along

with molecular marker analysis can be a best option. Because the degree of divergence between genotypes at the phenotypic level may not be necessarily correlated with a similar degree of genetic difference (Hamrick and Godt, 1989). Based on the findings by these authors, molecular markers provide a better coverage of the genome, resulting in a better estimate of relationships or divergence at molecular level. Molecular markers are an effective genomics approach for fingerprinting, mapping, gene tagging, as well as for determining the genetic structure and relationships among cocoa genetic groups (Laurent *et al.*, 1993; Motamayor *et al.*, 2003). Molecular markers are located in the specific regions of a chromosome near to the target gene and will be effective flags to breeders whether specific traits have been inherited or not. Further, such markers are very much suitable for early identification of the most desirable genotypes from a diverse population (Figueira, 2004).

Collard and Mackill (2008) opinioned that, in MAS, generally for line development, DNA markers have been integrated in place of conventional schemes, and they are utilized widely to substitute conventional phenotypic selection. Janila and Sharma (2004) reported that, marker-assisted selection improves the efficiency of plant breeding. Selection of tillering in wheat was done by marker-assisted selection (Li *et al.*, 2010). MAS also have potential role in resistance gene pyramiding in peanuts (Mishra *et al.*, 2009; Pandey *et al.*, 2012; Varshney *et al.*, 2014).

The DNA based marker systems are classified into hybridization based (non PCR) markers and PCR based markers (Joshi *et al.*, 1999). The value of PCR based molecular marker is influenced by several factors such as assessment speed, cost and technical simplicity etc. It must be sufficiently informative to distinguish between the individuals (Charters and Wilkinson, 2000).

2.3 Polymerase chain reaction (PCR)

The development of PCR was a technological breakthrough in genome analysis as it enabled the amplification of specific fragments from the total

genomic DNA. The principle of PCR is simple and is based typically on the function of a copying enzyme, DNA polymerase. The enzyme is able to synthesise a duplicate molecule of DNA from a DNA template *in vitro*. DNA polymerase extension product of the original template DNA, when denatured, can serve as a template for the other reaction and vice versa (Mullis and Faloona, 1987). Repeated chain reaction resulted in duplications, leading to an exponential increase in DNA product accumulation. Primers are short stretches of DNA sequence which are complementary to the genome which is supposed to be analysed. The DNA target to be amplified is defined by primer annealing sites. By selecting specific sequences as primers, amplification from virtually any region of a DNA molecule is possible (Karp *et al.*, 1997), but for direct-targeted PCR the sequence of the flanking region of the gene must be known.

One of the drawbacks of PCR based method, was the thermal instability of the Klenow fragment of *Escherichia coli* which was initially used as DNA polymerase I to catalyze the extension of the annealed primers. Denaturation of template as well as newly synthesized strands requires a heating step (above 92°C). Prior to the discovery of Taq polymerase enzyme, the thermal instability of the DNA polymerase enzyme from *E. coli* was a problem since the initial denaturation of DNA template at 92 °C lead to the enzyme degradation. This problem was effectively solved by the replacement of the *E. coli* DNA polymerase I with a thermo stable DNA polymerase (Taq polymerase) purified from the thermophilic bacterium, *Thermus aquaticus* that can survive extended incubation at 95°C (Saiki *et al.*, 1988).

2.4 PCR based molecular marker technique

Several techniques could be uses for the detection of variations (polymorphisms) among individuals at molecular level. Few of these are based upon the initial restriction digestion of the DNA (non PCR based), while others working on the principle of different enzymatic reaction, known as the polymerase chain reaction (PCR). DNA fingerprinting is a PCR based molecular marker technique which is widely adopted for characterizing individual plants and

to study genetic diversity within breeding populations of a crop (Saunders *et al.*, 2004). Agarwal *et al.* (2008) subdivided PCR-based techniques into arbitrarily primed PCR-based techniques and sequence targeted PCR-based techniques. Arbitrarily primed PCR technique or sequence non-specific techniques do not requires prior sequence knowledge to perform. RFLP and RAPD are widely used sequence non-specific PCR techniques. Whereas, microsatellite based marker technique and single nucleotide polymorphism (SNPs) markers are sequence specific and much informative. According to Agarwal *et al.* (2008), an ideal molecular marker technique should have specific features such as: (1) should be polymorphic and evenly distributed throughout the genome, (2) should provide adequate resolution of genetic differences, (3) should be independent and reliable, (4) should be simple, quick and inexpensive (5) amount of tissue and DNA samples required should be minimum and (6) has linkage to distinct phenotypes.

2.5 Molecular markers for varietal characterization

DNA based molecular marker techniques such as RFLP (restriction fragment length polymorphism), RAPD (random amplified polymorphic DNA), SSR (simple sequence repeats) and AFLP (amplified fragment length polymorphism) are widely used in ecological, taxonomical, evolutionary, phylogenic and genetic studies of plant sciences (Agarwal *et al.*, 2008).

A number of molecular tools and procedures are being employed to establish DNA fingerprint profiles and each of these procedures has its strength and weakness. So, a combination of marker system is recommended for concrete and confirmatory results (Karp *et al.*, 1996). Saunders *et al.* (2001) pointed out that, in plants, DNA fingerprinting can be adapted for numerous applications including characterizing individual plants to correct the errors in the identification of accessions and cultivars. Degani *et al.* (2001) reported that, once the identity of individuals has been established, the technique is also effective to study breeding populations of a crop based on its genetic diversity. DNA fingerprinting is successfully applied for cultivar identification, ensuring seed purity of hybrids and checking the genetic similarity between cultivars (Aikpokpodion *et al.*, 2010).

Because of the ease of scoring and analysing data from large numbers of samples, simple sequence repeat (SSR) analysis is particularly useful and commonly used. In addition, SSR analysis has the ability to detect heterozygotes (Bredemeijer *et al.*, 1998).

2.5.1 Inter simple sequence repeats (ISSR) markers in plants

DNA fragments of about 100-3000 bp located between adjacent, oppositely oriented microsatellite regions are referred to as inter simple sequence repeats. Spooner (2005) had reported that, ISSRs are amplified by PCR using microsatellite core sequences as primers with a few selective nucleotides as anchors into the non-repeat adjacent regions (16-18 bp). ISSR analysis will result in 10–60 fragments from multiple loci which are generated simultaneously, separated by gel electrophoresis and scored as the presence or absence of fragments of particular size. The major advantage of this marker technique is that, it doesn't requires any prior sequence data and the analysis requires only low quantity of template DNA (5-50 ng per reaction). And also, ISSRs are abundant and are randomly distributed throughout the genome.

Because of these advantageous features of ISSRs, the marker system emerged as an alternative system with the reliability and advantages of microsatellites (SSR) along with the broad taxonomic applicability of RAPDs. Studies conducted with ISSR markers so far indicated that, they are potentially useful for cultivar differentiation and for genetic diversity analysis (Dongre and Parkhi, 2005; Behera *et al.*, 2008; Luo *et al.*, 2011; Singh *et al.*, 2012; Thul *et al.*, 2012). Polymorphism in inter-microsatellite DNA sequence is detected by ISSR marker system and, the marker produces large number of fragments with high reproducibility and low expenditure.

In a study conducted at KAU, using 11 ISSR and 6 SSR primers, the molecular markers linked to the resistance of bacterial streak dieback (VSD) disease of cocoa were identified (Chandrakant, 2014). Distinct polymorphism was observed using the ISSR primer UBC857 and the 246bp long sequence derived

from the marker was analyzed using BLASTn. Sequence was proven similar to that of the SSR primer mTcCIR42. Subsequently, this primer was found to develop specific markers for VSD resistance.

2.5.1.1 ISSR markers for genetic diversity analysis in plants

With the objective to identify and make use of efficient molecular marker system for genetic analysis of traditional and evolved basmati and non-basmati rice varieties, Nagaraju *et al.* (2001) have used fluorescence based ISSR-PCR and SSR markers. Three rice groups, traditional basmati, evolved basmati and semi dwarf non-basmati were analyzed using 19 simple sequence repeat (SSR) loci and 12 ISSR-PCR primers. A total of 481 ISSR markers and 70 SSR markers were developed in 24 varieties from the three group. The genetic diversity analysis had shown the lowest genetic diversity among traditional basmati varieties. Evolved basmati varieties had shown the highest genetic diversity using both the marker assay and could clearly describe the traditional basmati and semi dwarf non-basmati rice.

Dongre and Parkhi (2005) have conducted a study to identify cotton hybrids through the combination of PCR based RAPD, ISSR and microsatellite marker systems. Hybrid cotton H6 and its parents G.Cot.10 (male) and G.Cot.100 (female) were used for conducting the study. Twenty RAPD primers, nineteen ISSR primers and twenty five reported polymorphic microsatellites were used for the experiment. RAPD primer, OPA 11 was most useful in differentiating parents and hybrids based on polymorphism. ISSR primers IS4 and IS7 developed two female specific amplicons of size 500bp and 1200bp in the hybrid H6. Microsatellites JESPR-2 and JESPR-17 were heteroallelic for parents. The study concluded that, for identification of cotton hybrids, utilization of all three markers - RAPD, ISSR and SSR in combination is reliable and faster than using them individually.

Using 29 RAPD and 15 ISSR markers, genetic analysis of 38 different diverse Indian bitter gourd (*Momordica charantia* var. *charantia*, and *Momordica*

charantia var. *muricata*) accessions was performed (Behera *et al.*, 2008). Cluster analysis performed after grouping the bitter gourd accessions indicated that, RAPD and ISSR based diversity analysis in these germplasm were consistent. Distinct genetic differences while examining *M. charantia* var. *charantia* (domesticated) and var. *muricata* (wild, free-living) accessions provided new strategy for genetic analysis and crop improvement of bitter gourd.

Genetic diversity of 23 mango germplasm accessions collected within Guangxi province of China was estimated using start codon tagged primers (SCoT) and ISSR markers (Luo *et al.*, 2011). For conducting the study, 18 SCoT and 18 ISSR primers were used. Amplification using 18 ISSR primers yielded 15 bands with 55 per cent polymorphism and SCoT primers yielded 65 per cent polymorphism. Then the 23 cultivars were clustered in to three major groups based on ISSR and SCoT analysis with UPGMA. Result indicated that, SCoT marker system is better represents the actual relationship than ISSR.

Singh *et al.* (2012) conducted a study to evaluate the genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR marker systems. A total number of 17 polymorphic primers were used, of which 11 RAPDs and 6 ISSRs were included. When 60 genotypes were analyzed, ISSR primers produced 66 bands and 52 of them were polymorphic with an average of 8.6 polymorphic fragments per primer. Number of bands per primer varied from 1 to 14 with amplicon size ranges from 200 to 2000bp. Average percentage of polymorphism was found to be 95.4 per cent. Using the data, Nei's dendogram was generated and analysis had shown low genetic diversity among the accessions.

Genetic variability of 22 *Capsicum* accessions belonging to C. *annuum*, *C. baccatum*, *C. chinense*, *C. eximium*, *C. frutescens*, and *C. luteum* were investigated using 8 ISSR and 27 RAPD primers (Thul *et al.*, 2012). Comparison of genetic diversity data with morphology had shown that, *C. chinense* accession shared flower characters with *C. frutescens* and found closer regarding genotype. Similarly, *C. luteum* was found to be close to *C. baccatum* both phenotypically and genotypically. The floral characteristics of Capsicum species and the

molecular markers Such as ISSR and RAPD are found to be useful toward the determination of the species specificity and identification of genetic stock.

A study based on morpho-agronomic traits and ISSR markers was conducted by Baliyan *et al.* (2014) to analyze the genetic diversity in chrysanthemum. To survey genetic diversity in a set of 24 chrysanthemum genotypes, nine quantitative (morpho-agronomic) traits, five qualitative traits and 10 ISSR markers were used. UPGMA analysis on the ISSR data has clearly separated the genotypes into distinct groups.

2.5.1.2 ISSR markers for DNA fingerprinting in plants

Blair *et al.* (1999) have conducted ISSR amplification for analysis of microsatellite motif frequency and fingerprinting in rice. In this study, 32 ISSR primers containing different SSR motifs were used for amplification of 59 varieties, which is the representative of the diversity of cultivated rice. The amplification pattern obtained by ISSR was used to group the rice genotypes by cluster analysis. The results revealed that, ISSR fingerprint obtained could be used to differentiate genotypes belongs to *Japonica* and *Indica.* Therefore the study concluded that, ISSR marker technique can be a suitable method for determining genetic diversity among varieties and for rapid identification of cultivars.

McGregor *et al.* (2000) have conducted a comparative assessment of DNA fingerprinting in 39 tetraploid potato (*Solanum tuberosum* L.) germplasm using markers such as RAPD, ISSR, AFLP and SSR. The PCR based DNA fingerprinting of the germplasm was done using 20 RAPD primers, 6 ISSRs, 2 AFLPs and 5 SSRs. With respect to the distinguishing capability of markers between cultivars, all four marker systems could successfully distinguish 39 potato cultivars. The results generated using SSR and AFLP markers were found to be highly reproducible *i.e.*, 100 and 99.6 per cent respectively. For ISSR, it was comparatively lesser (87 per cent). The study concluded that, the AFLP technique is powerful technique for DNA fingerprinting in potato. Genotype index value was significantly higher for SSRs.

Mogalayi (2011) performed DNA fingerprinting of black pepper varieties released by Kerala Agricultural University at Centre for Plant Biotechnology and Molecular Biology. RAPD (30 nos.), ISSR (34 nos.) and SSR (29 nos.) marker systems were used to characterize seven black pepper varieties. Resolving power of each marker system were analyzed and found that, it has ranged from 7.42 to 9.42 in RAPD and 5.42 to 12.28 in ISSR. Polymorphism information content (PIC) value found was ranges from 0.86 to 0.90 in RAPD and 0.80 to 0.89 in case of ISSR. The high resolving power of the markers indicate its usefulness in distinguishing varieties and the observed PIC value indicate the variability among genotypes.

Using 10 each of ISSR and SSR primers, DNA fingerprinting of four KAU brinjal varieties along with three superior accessions and two wild relatives was conducted at CPBMB (Laxman, 2013). According to the banding pattern on gel, resolving power (Rp) of each marker system was calculated. Rp value of ISSR primers ranged from 9.90 to 28.44 and PIC value was ranges from 0.83 to 0.96. Thus, the selected ISSR primers can distinguish the genotypes as well as can be the suitable primers for detecting polymorphism in brinjal.

DNA fingerprinting of five cashew varieties released by Kerala Agricultural University was conducted using eleven each of SSR and ISSR primers at CPBMB (Meena, 2014). Resolving power of each marker technique was analyzed and for ISSRs, it was ranged from 1.4 to 7.0 implied the capability of the selected markers to distinguish the varieties.

2.5.1.3 ISSR markers for genetic diversity analysis in cocoa

Charters and Wilkinson (2000) evaluated the potential of ISSR analysis to genetically evaluate and identify the mislabelled cocoa accessions. For genetic characterization and proper maintenance, 62 accessions were studied. Selected 6 ISSR primers were found sufficient to distinguish and differentiate the accessions. An UPGMA dendogram developed based on the ISSR data was used to provide genetic variability data. Genetic diversity data obtained helped to find out several

documentation errors and provides platform data for locating unwanted material from the collection.

ISSR molecular markers were used to characterize *Theobroma cacao* L. collections at Tingo Maria, Peru (Chia *et al.*, 2011). The study was conducted to differentiate 46 cocoa accessions and to identify the similarity between the accessions. ISSR primers $(GA)_8YG$ and $HBH(AG)_7$ were identified to have more polymorphism determination potential. The results showed that apart from ISSR's dominance nature, these markers have the capacity to group Trinitario accessions in to a common cluster.

2.5.2 Simple sequence repeats (SSR) or microsatellite markers in plants

Microsatellite or simple sequences repeats (SSRs) are short (1-6 bp) tandemly repeated DNA motifs (mono to hexa-nucleotides) that usually occurs abundantly in eukaryotic genomes (Beckman and Soller, 1986). They have been successfully used as distinct molecular markers by targeting the repeat region with unique flanking primer pairs (Asari *et al.*, 2014). SSR length variation is mainly due to the slippage of DNA polymerase or unequal crossover leading to the insertion or deletion of tandem repeats.

Allelic variation in SSR loci can readily be detected by PCR using specific primers flanking the SSR motif. Among the different classes of molecular markers, SSRs have been extensively used in many crop species for numerous applications in plant genetics and breeding because of their simplicity, reproducibility, high degree of polymorphism, codominant inheritance, relative abundance, and good genome coverage (Powell *et al.*, 1996). Moreover, SSRs are technically easy to analyze because of the small amount of DNA required and their suitability for multiplexing on automated systems. SSR markers are widely used in cultivar fingerprinting, genetic diversity assessment, molecular mapping, QTL detection, and marker-assisted selection. Ease of analysing genetic data and the ability to detect heterozygotes by less complicated scoring methods made SSR marker system advantageous over others.

Microsatellites are known to be very efficient marker system for assessing genetic diversity due to their high discriminatory power and polymorphism information content among individuals, which arising from their multi-allelic nature and co-dominant segregation of alleles (Powell *et al.* 1996). SSRs are derived from ESTs, corresponding to the transcribed component of a gene unit and, they have been shown to possess a high potential for inter-specific transferability (Cordeiro *et al.*, 2001; Thiel *et al.*, 2003; Gupta *et al.*, 2003). They are also being widely used in many studies on genotype identification and genetic mapping (Ramsay *et al.*, 2000) and linkage map construction (Pugh *et al.*, 2004; Lu *et al.*, 2005). Another important feature of genomic SSRs is their locus-specificity, which is an important advantage when working with species such as wheat (allopolyploid), in which only one of three homoeologous loci should ideally be tagged (Varshney *et al.*, 2000). The assay shows extensive inter-individual length polymorphisms during PCR analysis of unique loci using discriminatory primers sets (Agarwal *et al.*, 2008).

2.5.2.1 SSR markers for genetic diversity analysis in plants

Varshney *et al.* (2004) reported that, SSRs which are derived from ESTs/genes (ESTSSRs), exhibit a higher potential for transfer through cross-amplification in related species than SSR markers usually generated from genomic DNA libraries. The study was conducted to analyze the interspecific transferability of barley EST-SSR markers in wheat, rye and rice. In this study, a sub-set of 165 EST-SSR markers of barley was examined for its transferability to wheat, rye and rice. Results had shown that, 78.2 per cent of barley markers show amplification in wheat followed by 75.2 per cent and 42.4 per cent in rye and rice respectively.

To differentiate *Capsicum annum* and *Capcicum pubescens* based on its inter and intraspecific characters, 8 ISSR and 19 SSR markers were used to analyse the individuals (Torres *et al.*, 2015). Among 8 ISSRs studied, two primers were given polymorphic data with a PIC value of 0.77 and resolving power 16.08.

Average PIC value observed for SSR was 0.5. Both the markers were found efficient for differentiating the varieties.

Genetic diversity analysis of Orchidaceae species was done using *Dendrobium* SSR markers (Kang *et al.*, 2015). A set of 320 *Dendrobium* SSR markers were utilized to test their transferability and polymorphism across 44 species of 15 Orchidaceae genera. As much as 109 SSRs had shown high transferability and polymorphism. Randomly selected SSR loci were sequenced. The alignment analysis revealed variation in SSR repeat length and sequence of the motifs among the taxa.

2.5.2.2 SSR markers for DNA fingerprinting in plants

Ghislain *et al.* (2009) developed microsatellite based genetic identity kit for potato. Fingerprinting of 742 potato landraces were conducted using 51 SSRs. Using the result obtained, 24 most informative SSRs were identified. The new kit was found ideal for discriminating 93.5 per cent of the 742 potato landraces. Further, a marker-specific set of allele size standards were developed that conveniently provide accurate sizing of all alleles of the 24 SSR markers across laboratories and platforms.

In order to find out the pedigree relationship between clonal tea cultivars, SSR based fingerprinting was conducted (Tan *et al.*, 2015). In the study, 128 Chinese clonal cultivars were analyzed using 30 SSR markers to identify the parent offspring relationship. Fingerprint data identified 47 pairs of parent offspring relationships of which, 33 were found similar with known pedigree information and remaining 14 were identified newly.

2.5.2.3 SSR markers for genetic diversity analysis in cocoa

Lopez *et al.* (2004) have detected and quantified *in vitro* culture induced chimerism in cocoa using 15 SSR markers. Differential amplification of alternate SSR alleles at heterozygous loci was exploited to conduct the assay. Capillary electrophoresis with calibrated reference data was further used to analyze 233 cocoa somatic embryo regenerants.

Saunders *et al.* (2004) have identified 13 SSR primers as international molecular standards for DNA fingerprinting in cocoa. Zhang *et al.* (2006) recommended microsatellite markers for defining genetic identity of varieties. Zhang *et al.* (2009) have characterized international cocoa collection maintained in Costa Rica using microsatellite markers.

The selected international cocoa microsatellite data was further extended by the standardization of allelic size determination (Cryer *et al.*, 2006). Fingerprinting of 429 cocoa trees comprising 345 accessions were analyzed using reported 15 SSR primers. Fingerprinting data was produced for an international quarantine collection and it will further ensure the efficient exchange of quality breeding materials across the world.

Using 12 microsatellite markers, genetic diversity among 574 cocoa accessions covering parental populations in West Africa gene bank and farmers' accessions from Nigeria was evaluated (Aikpokpodion, 2009). The study revealed that, an appreciable genetic diversity was present in on-farm land field genebank collections of cocoa. SSR analysis had shown homozygous pattern which revealed the impact of inbreeding in local parent population. The results concluded with genetic diversity data showing significant difference with in on farm and field genebank.

2.5.2.4 SSR markers for DNA fingerprinting in cocoa

International molecular standards for cocoa DNA fingerprinting was identified and described (Saunders *et al.*, 2004). DNA analysis based on SSR marker was identified as ideal molecular tool for the study. Approximately 690 cocoa accessions were evaluated using 25 SSRs. Based on the results obtained, 15 SSRs with high reproducibility, reliability and differentiation of divergent genotypes were selected. The selected 15 SSR primers were reported as international molecular standards for DNA fingerprinting of *T. cocoa*.

Microsatellite fingerprinting of 924 cocoa germplasm was performed using high throughput genotyping system with 15 microsatellite loci (Irish *et al.*,

2010). The objective of the particular study was to identify the mislabelled and duplicated genotypes in the USDA-ARS germplasm. For the study, the forward primers were labelled with a fluorescent dye and capillary electrophoresis was done on an ABI Prism 3730 genetic analyser platform. Further DNA fingerprinting was done using software GeneMapper 3.0. The observed bands were analysed and condensed to 174 unique fingerprint profiles. Thus, the developed fingerprint was utilized for identifying mislabelled as well as duplicated accessions among the germplasm.

Materials and methods

3. MATERIALS AND METHODS

The study on 'DNA fingerprinting of promising cocoa (*Theobroma cacao* L.) varieties of KAU' was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Vellanikkara, 2014 - 2016. The materials used and methods adopted during the study are described in this chapter.

3.1 Materials

3.1.1 Plant materials

The varieties selected for the study included six promising selections (CCRP 1, CCRP 2, CCRP 4, CCRP 5, CCRP 6 and CCRP 7) and two hybrids (CCRP 8 and CCRP 9) of cocoa which are popular among the farmers (Plate 1.). Twelve year old budded trees of each variety maintained at Cocoa Research Centre (CRC), Kerala Agricultural University (KAU), were used for conducting morphological studies as well as for DNA isolation.

3.1.2 Laboratory chemicals, plastic ware and glassware

Analytical reagent (AR) grade chemicals procured from HIMEDIA, Labcare and Merk India Ltd. were used for conducting the study. PCR master mix components such as Buffer B, $MgCl_2$ and *Taq* DNA polymerase were supplied by GeNei™ and primers used were provided by SIGMA. Gel loading dye used was from Thermo Scientific and ladders were from Invitrogen (100bp to 2000bp), GeNei™ (100bp to 3000bp) and from Thermo Scientific (250bp to 12,000bp). Glassware utilized were supplied by Borosil and plastic ware by Tarson India Ltd. and Axygen.

3.1.3 Laboratory equipments

The DNA fingerprinting of cocoa was carried out using the facilities and equipments available at CPBMB, KAU. Micropipettes used for the study were supplied and calibrated by Eppendrof. Samples were centrifuged using KUBOTA 6500. Spectrophotometer Nanodrop® ND-1000 was used for estimation of quality

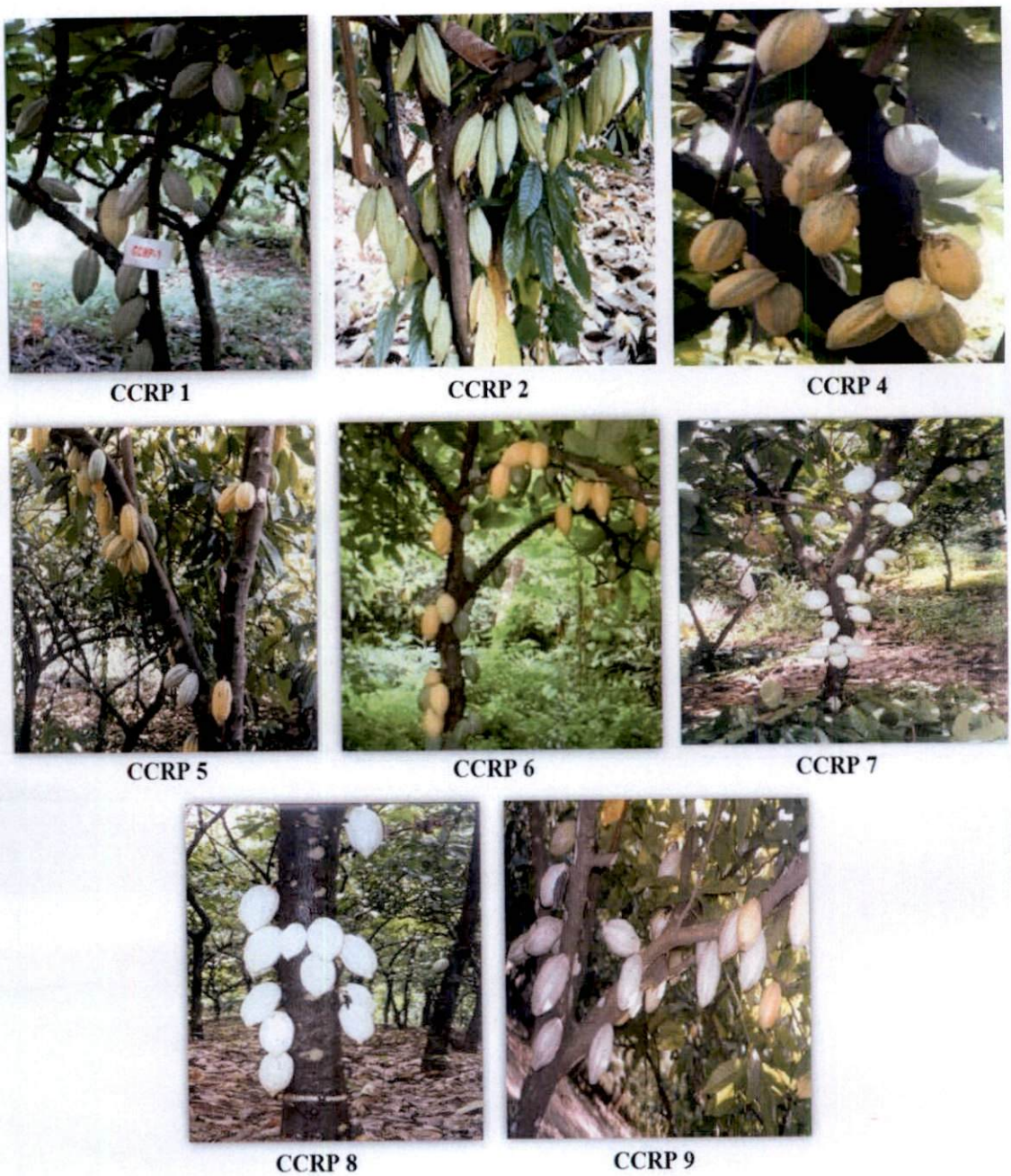


Plate 1. Field view of the cocoa varieties used in the study showing morphological variability

and quantity of isolated DNA. Polymerase chain reaction was performed using thermal cycler Proflex. Agarose gel electrophoresis was performed using wide mini-sub unit and power pac make by BIO-RAD. Gel documentation was done using Gel Doc™ XR+ System from BIO-RAD. Detailed list of equipments used in this study are given in Annexure I.

3.2 Methods

3.2.1 Morphological analysis

Morphological observations on distinguishable quantitative and qualitative characters were recorded on five pods collected at random from each variety as per the descriptor suggested by Bekele and Butler (2000). The descriptor and descriptor states are presented in Annexure II.

3.2.1.1 Analysis of qualitative traits

Six qualitative characters *viz.* colour of ripe pods, pod shape, pod apex form, basal constriction, pod rugosity and colour of bean were recorded on five pods in each variety collected at random. The genetic associations among the varieties were estimated through Jaccard's similarity coefficients (Jaccard, 1908) using NTSYS pc version 2.1 (Rohlf, 1992). Cluster analysis was done on the similarity matrix and a dendrogram was constructed using Unweighted Pair - Group Method (UPGMA) (Sneath and Sokal, 1973).

3.2.1.2 Analysis of quantitative traits

Eight quantitative characters *viz.* pod weight (g), pod length (cm), pod breadth (cm), husk thickness (cm), wet bean weight (g), number of beans per pod, single wet bean weight (g) and single dry bean weight (g) were recorded. Further analysis was done using completely randomized design (CRD) and computed the data for all the individual characters observed. Clustering based on quantitative characters was done using NTSYS pc version 2.1 and a dendrogram was constructed using UPGMA.

3.2.2 Molecular analysis

Molecular characterization and DNA fingerprinting of the eight varieties were accomplished respectively using marker systems inter simple sequence repeats (ISSR) and simple sequence repeats (SSR).

3.2.2.1 Genomic DNA extraction

Tender leaves from the actively growing tips of plants were collected early in the morning. The collected leaves were labeled, covered with aluminium foil and kept in ice box to prevent phenolic oxidation. Samples were brought to the laboratory under cool environment inside ice box. Leaf surface was cleaned by washing with sterile distilled water followed by wiping with 70 per cent ethanol. The samples were weighed and stored at -20 °C until used for DNA extraction. Modified CTAB method reported by Chandrakant (2014) with slight modification was used for the extraction of good quality DNA.

Reagents used

a. CTAB buffer:

- 5 per cent CTAB (W/V)
- 100mM Tris base (pH 8.0).
- 20mM EDTA (pH 8.0)
- 1.4 M NaCl
- 1 per cent polyvinyl pyrrolidin (PVP)
- 0.2 per cent 2-β mercaptoethanol

b. Chloroform: isoamyl alcohol (24:1 v/v)

c. Chilled isopropanol

d. Ethanol 70 and 100 per cent

e. TE buffer:

- 10 mM Tris (pH 8)

- 1 mM EDTA (pH 8)

f. Sterile distilled water

Reagent a. and e. were autoclaved and stored at room temperature.
Further details on the chemicals used and their composition are provided in Annexure III.

Procedure

The freshly prepared CTAB (5 per cent) solution was preheated to a temperature of 60 °C by keeping it in a hot water bath. Freshly collected leaf sample was cleaned by washing and dried with the help of sterile tissue paper. From the collected leaf sample, 0.08-0.10 gm of the tissue was weighed avoiding the mid rib region. The weighed sample then kept in a mortar. A pinch of polyvinyl pyrrolidone (PVP) and 50µl 2-β mercaptoethanol was added to the tissue to prevent poly phenol oxidation. Then for powdering the sample, liquid nitrogen was added to the mortar and quickly ground the tissue using a pestle. The powdered sample was immediately transferred to a 2 ml eppendrof tube containing a pinch of PVP. To the powdered sample, 1.2 ml preheated CTAB was added and the samples were incubated at 65 °C for 30 minutes with intermittent shaking. Then the mixture was centrifuged (KUBOTA 6500) at 12000 rpm for 15 minutes at a temperature of 4 °C. The top clear extraction buffer containing DNA was transferred carefully to another 2 ml tube and equal volume chloroform: isoamyl alcohol was added. The tube contents were properly mixed by gentle inversion and then centrifuged at 12000 rpm for 15 minutes at 4 °C.

Centrifugation resulted in the separation of tube contents in to three distinct phases. Top aqueous phase comprises of DNA with small quantity of RNA and middle phase was formed by protein and other cell debris. Bottom clear phase was formed by chloroform, isoamyl alcohol and some leaf pigments etc. The top aqueous layer was carefully transferred to a sterile 1.5ml centrifuge tube and 0.6 volume of chilled isopropanol was added. The contents in the tubes were mixed by gentle inversion until DNA precipitates by forming thread like

structures. For complete precipitation, the tubes were kept at -20 °C for 30-60 minutes. Tubes were again centrifuged at 12,000 for 15 minutes at 4 °C for condensation and pellet formation of DNA. The supernatant was carefully poured off taking care that only the pellets get retained in the tubes. 50 µl ethanol (70 per cent) was added to this pellet and spun for 5 minutes at 10,000 rpm. The ethanol was decanted and the pellets were air dried until ethanol was evaporated. Air dried pellets were dissolved in 50 µl sterilized distilled water and stored at -20 °C. Quality and quantity of isolated DNA was initially observed by electrophoresis with 0.8 per cent agarose gel and then by using Nanodrop.

3.2.2.2 Purification of DNA

After checking the quality and quantity of isolated DNA, the samples those found contaminated with RNA were further purified by RNase treatment.

Reagents used

a. RNase (1.0 per cent)
b. Chloroform: isoamyl alcohol (24:1 v/v)
c. Chilled isopropanol (100 per cent)
d. Ethanol (70 per cent)

One per cent RNase was prepared by dissolving RNase (SIGMA) in TAE buffer. Inactivation of residual DNase in the reagent was done by heating at 100 °C for 15 minutes. The solution was then cooled to room temperature and stored at -20 °C.

Procedure

RNA contamination was removed by adding 1µl of RNase (1 per cent) to 50 µl of DNA sample, and kept for incubation at 37 °C for 40 minutes in dry bath.

Protein contamination was removed by following the procedure detailed below.

Total volume of DNA was made up to 250 µl using sterile distilled water. Equal volume of chloroform: isoamyl alcohol (24:1) was added to this and mixed by gentle inversion. The mixture was centrifuged for 15 minutes at 12,000 rpm (4

°C). Top aqueous phase that contained DNA was transferred to a sterile centrifuge tube to which equal volume of chloroform: isoamyl alcohol (24:1) was added. The mixture was again centrifuged at 12,000 rpm for 15 minutes at 4 °C. Top aqueous phase containing DNA was transferred to a new sterile centrifuge tube to which 0.6 volume of chilled isopropanol was added. The contents in the tubes were mixed by gentle inversion and kept at -20 °C for complete precipitation. Precipitated DNA was centrifuged at 12,000 rpm for 15 minutes at 4 °C. DNA pellet thus formed was carefully retained in the tube and washed with 70 per cent ethanol. Centrifuged tubes at 10,000 rpm for 10 minutes at 4 °C. The pellet was air dried to ensure removal of ethanol. The pellet then dissolved in 50 µl sterile distilled water and stored at -20 °C. The DNA thus obtained was found to be devoid of RNA/ protein and confirmation was done using electrophoresis with 0.8 per cent agarose gel (at 70 V).

3.2.2.3 DNA quantity assessment by gel electrophoresis

Agarose gel electrophoresis technique can be used to visualize and quantify the isolated DNA samples under UV light with the help of a ladder.

Reagents used

a. Agarose - 0.8 per cent (w/v)
b. 50X TAE buffer (pH 8.0)
c. Running buffer (1X TAE)
d. 6X gel loading dye
e. Ethidium bromide (0.5 µg/ml)

Equipments used

a. Micropipettes
b. Gel casting system (BIO-RAD)
c. Wide mini-sub electrophoresis unit (BIO-RAD)
d. Power pac (BIO-RAD)
e. Gel documentation unit - Gel Doc™ XR+ System (BIO-RAD)

Composition of reagents is provided in Annexure IV

Procedure

Gel was cast in BIO-RAD gel casting system. Gel casting tray and comb were wiped with 70.per cent ethanol. The tray was kept tight inside the gel casting system with properly placed comb. For the preparation of agarose gel (0.8 per cent), 0.8gm agarose was dissolved in 100 ml of 1X TAE buffer taken in a conical flask. To ensure complete dissolution of the gel it was melted by keeping in a microwave oven for 45 to 60 seconds. The molten gel was then allowed to cool down to about 45 to 50 °C to which 2-3 µl diluted Ethidium bromide (0.5 µg/ml) was added and mixed gently. The warm gel was poured in to the gel casting tray and kept undisturbed for 30 to 45 minutes at room temperature for solidification.

After solidification, the comb was carefully removed from the gel. The gel casting tray along with the gel was placed in the electrophoresis unit containing running buffer (1X TAE). 6 µl of samples (5 µl DNA and 1 µl 6X gel loading dye) were loaded in to the wells. A suitable marker of molecular weight (250-12,000bp) was also loaded in one of the wells. Electrophoresis (80V) was carried out until the tracking dye reached 2/3rd of the gel. Electrophoresed gel was carefully transferred to gel documentation unit (Gel Doc^{TM} XR+ System) and observed under UV exposure. Presence of thick bands near the wells indicated the presence of DNA. RNA contamination can be observed as presence of thick bands around 100 bp region. And a thick white patch observed inside the well indicated the presence of protein. Quantity one software provided by BIO-RAD was used to analyze the electrophoresed agarose gel.

3.2.2.4 DNA quantity assessment by spectrophotometer

Further confirmation of the DNA quality and quantity was done using spectrophotometer (Nanodrop ND-1000). The maximum absorbance of nucleic acid was found at 260 nm and that of protein was 280 nm. Thus the absorbance was recorded at both 260 and 280 nm and the purity of the DNA was indicated using the ratio, OD_{260}/OD_{280}. Good quality DNA is found to have an OD_{260}/OD_{280}

value of 1.8 to 2.0. The value below 1.8 indicated that the sample is contaminated with proteins and value above 2.0 indicated RNA contamination.

Procedure

Spectrophotometer (Nanodrop-ND 1000) was connected to the computer in which the software ND-1000 is installed. For initiating quantification, the option nucleic acid from software window was chosen. The sampling arm was opened and loaded with 1µl distilled water into the lower measurement pedestal. The sampling arm was then closed upon which the instrument automatically begins the measurement process with the help of the software.

The instrument was calibrated using distilled water (blank). The reading was set to zero. Thereafter 1µl of each sample was loaded into the measurement pedestal and measured the OD value by choosing the option 'measure' in the software. The OD value was recorded along with quantity of DNA in nano gram per micro litre (ng/µl). After measuring each sample, the sampling arm was opened and both upper and lower pedestals were wiped using tissue paper. The recorded OD values and DNA quantity (ng/µl) values were used for further dilution of the DNA.

3.3 DNA fingerprinting

The study was carried out using two different types of markers, ISSRs (Inter Simple Sequence Repeats) and SSRs (Simple Sequence Repeats). Under each marker system, DNA from all the eight selected varieties of cocoa was amplified using selected primers. The polymorphic bands amplified for each primer in both the systems were identified. DNA amplification pattern thus obtained was represented in the form of colour charts for better interpretation of results.

3.3.1 Conditions for DNA amplification

The concentration of the PCR components as well as the PCR conditions used were standardized to yield effective amplification using both the marker

systems. The standardized reaction mixture consisted of template DNA, assay buffer B, $MgCl_2$, dNTPs, primer, *Taq* DNA polymerase and autoclaved distilled water. Master mix devoid of DNA template was prepared and transferred to 0.2ml PCR tubes and to this, the template DNA was added. Polymerase chain reaction using standardized PCR conditions was performed using thermal cycler (Model: Proflex, Make: Life). PCR conditions were standardized in such a way that, the temperature and time used (denaturation, annealing, elongation) and cycle numbers were optimum for effective amplification. Blank was prepared by adding all PCR components which is devoid of DNA sample.

3.3.2 Inter Simple Sequence Repeats (ISSR) analysis

Good quality cocoa DNA isolated from cocoa leaf samples were diluted to a concentration of 10 ng/µl and were used for ISSR analysis. Initial screening was done using bulked DNA and those primers with good resolving power were chosen for further analysis. Annealing temperature was also standardized using gradient PCR for better results.

Genomic DNA amplification was performed using 20 µl reaction mixture in 0.2 ml PCR tubes and the mixture consisted of the following components.

a.	Taq assay buffer B (10X)	- 2.0 µl
b.	$MgCl_2$	- 2.0 µl
c.	dNTP mix (10mM each)	- 1.5 µl
d.	Primer (10pM)	- 1.5 µl
e.	Autoclaved distilled water	- 10.6 µl
f.	*Taq* DNA polymerase (3U)	- 0.4 µl
g.	Genomic DNA (10 ng)	- 2.0 µl
	Total volume	- 20.0 µl

Standardized PCR condition used for development of ISSR pattern was as following.

a. Initial denaturation (hot start) - 94 °C for 4 minutes
b. Denaturation - 94 °C for 45 seconds
c. Primer annealing - 43 to 55 °C for 1 minute } 36 cycles
d. Primer elongation - 72 °C for 2 minutes
e. Final élongation - 72 °C for 8 minutes
f. Incubation - 4 °C for infinity

3.3.2.1 Screening and analysis of ISSR primers

Based on the previous marker study reports in cocoa (Saunders *et al.*, 2004; Chia *et al.*, 2011; Chandrakant, 2014), thirty six ISSR primers were chosen for the analysis. List of ISSR primers chosen for the analysis are given in Table 1.

The amplified products after PCR were electrophoresed along with 1kb+ DNA ladder (Thermo Scientific) or 100-3000 bp ladder (GeNei™) on 1.8 per cent agarose gel stained with ethidium bromide. Banding pattern obtained was visualized using gel documentation unit (Gel Doc™ XR+ System) under UV exposure. The gel pictures were saved in image format for further scoring and detection of polymorphism among the amplicons.

Table 1. Details of ISSR primers used in this study

Sl. No.	Primer	Nucleotide Sequence
1.	HB 10	5'-GAGAGAGAGAGACC-3'
2.	HB 12	5'-CACCACCACGC-3'
3.	Oligo ISSR 04	5'-ACACACACACACACACC-3'
4.	Oligo ISSR 05	5'-CTCTCTCTCTCTCTTG-3'
5.	Oligo ISSR 06	5'-GAGAGAGAGAGAGAGAC-3'
6.	Oligo ISSR 07	5'-CTCTCTCTCTCTCTTG-3'
7.	Oligo ISSR 08	5'-GAGAGAGAGAGAGAGAT-3'
8.	ISSR 2	5'-ATTATTATTATTATTCAT-3'
9.	ISSR 3	5'-TTATTATTATTATTACTT-3'
10.	ISSR 4	5'-ATTATTATTATTATTGTT-3'
11.	ISSR 5	5'-ATTATTGTTGTTGTTTTC-3'
12.	ISSR 6	5'-TTATTATTATTATTATAA-3'
13.	ISSR 7	5'-ATTATTGTTGTTGTTGTA-3'
14.	ISSR 8	5'-ATTATTATTATTATTGTA-3'
15.	ISSR 9	5'-TTATTATTATTATTATTACT-3'
16.	UBC 808	5'-AGAGAGAGAGAGAGAGC-3'
17.	UBC 810	5'- GAGAGAGAGAGAGAGAT-3'
18.	UBC 811	5'-GAGAGAGAGAGAGAGAC-3'
19.	UBC 815	5'-CTCTCTCTCTCTCTCTG-3'
20.	UBC 818	5'-CACACACACACACACAG-3'
21.	UBC 826	5'-ACACACACACACACACC-3'
22.	UBC 827	5'-ACACACACACACACACG-3'
23.	UBC 835	5'-AGAGAGAGAGAGAGAGYC-3'
24.	UBC 841	5'-GAGAGAGAGAGAGAGAY-3'
25.	UBC 844	5'-CTCTCTCTCTCTCTCTRC- 3'
26.	UBC 846	5'-CACACACACACACACART-3'
27.	UBC 847	5'-CACACACACACACACARC-3'
28.	UBC 848	5'-CACACACACACACACARG-3'
29.	UBC 850	5'-GTGTGTGTGTGTGTGTYC-3'
30.	UBC 854	5'-TCTCTCTCTCTCTCTCRG-3'
31.	UBC 855	5'-ACACACACACACACACYT-3'
32.	UBC 857	5'-ACACACACACACACACYG-3'
33.	UBC 865	5'-CCGCCGCCGCCGCCGCCG-3'
34.	UBC 866	5'-CTCCTCCTCCTCCTCCTC-3'
35.	UBC 873	5'-GACAGACAGACAGACA-3'

3.3.3 Simple Sequence Repeats (SSR) analysis

Good quality cocoa DNA was diluted to a concentration of 10 ng/µl and the diluted samples were used for SSR analysis. Initial screening was done using bulked DNA and those primer pairs with good resolving power were chosen for further analysis. Annealing temperature was standardized using gradient PCR.

Genomic DNA amplification was performed using 20 µl reaction mixture in 0.2 ml PCR tubes and the mixture consisted of the following components.

a.	Taq assay buffer B (10X)	- 2.00 µl
b.	$MgCl_2$	- 2.00 µl
c.	dNTP mix (10mM each)	- 1.50 µl
d.	Forward primer (10pM)	- 0.75 µl
e.	Reverse primer (10pM)	- 0.75 µl
f.	Autoclaved distilled water	- 10.60 µl
g.	*Taq* DNA polymerase (3U)	- 0.40 µl
h.	Genomic DNA (10 ng)	- 2.00 µl
	Total volume	- 20.00 µl

Standardized PCR conditions for the development of SSR pattern were the following.

a.	Initial denaturation (hot start)	- 94 °C for 4 minutes	
b.	Denaturation	- 94 °C for 45 seconds	40cycles
c.	Primer annealing	- 46 to 64 °C for 1 minute	
d.	Primer elongation	- 72 °C for 2 minutes	
e.	Final elongation	- 72 °C for 8 minutes	
f.	Incubation	- 4 °C for infinity	

3.3.3.1 Screening and analysis of SSR primers

Primers for the analysis were chosen based on previous reports (Lanaud *et al.*, 1999; Saunders *et al.*, 2004; Lopez *et al.*, 2004; Takrama *et al.*, 2005; Cryer *et al.*, 2006; Aikpokpodion, 2009; Fouet *et al.*, 2011 and Chandrakant, 2014). Thirty

seven SSR primers were screened by thermal cycling for selecting the ones with useful molecular data. List of SSR primers chosen for the analysis along with their annealing temperature is given in Table 2.

The PCR amplified products were electrophoresed along with 100 bp - 2000 bp DNA ladder (Invitrogen) on 2.0 per cent agarose gel stained with ethidium bromide. Banding pattern was visualized using gel documentation unit under UV exposure and saved in image format for further scoring and detection of fingerprints.

Table 2. Details of SSR primers used in this study

Sl. No.	Primer	Nucleotide Sequence
1.	mTcCIR1	F 5'-GCAGGGCAGGCTCAGTGAAGCA-3' R 5'-TGGGCAACCAGAAAACGAT-3'
2.	mTcCIR6	F 5'-TTCCCTCTAAACTACCCTAAAT-3' R 5'-TAAAGCAAAGCAATCTAACATA-3'
3.	mTcCIR7	F 5'-ATGCGAATGACAACTGGT-3' R 5'-GCTTTCAGTCCTTTGCTT-3'
4.	mTcCIR8	F 5'-CTAGTTTCCCATTTACCA-3' R 5'-TCCTCAGCATTTTCTTTC-3'
5.	mTcCIR10	F 5'-CCGAATTGACAGATGGCCTA-3' R 5'-CCCAAGCAAGCCTCATACTC-3'
6.	mTcCIR11	F 5'-TTTGGTGATTATTAGCAG-3' R 5'-GATTCGATTTGATGTGAG-3'
7.	mTcCIR12	F 5'-TCTGACCCCAAACCTGTA-3' R 5'-ATTCCAGTTAAAGCACAT-3'
8.	mTcCIR15	F 5'-CAGCCGCCTCTTGTTAG-3' R 5'-TATTTGGGATTCTTGATG-3'
9.	mTcCIR18	F 5'-GATAGCTAAGGGGATTGAGGA-3' R 5'-GGTAATTCAATCATTTGAGGATA-3'
10.	mTcCIR22	F 5'-ATTCTCGCAAAAACTTAG-3' R 5'-GATGGAAGGAGTGTAAATAG-3'
11.	mTcCIR24	F 5'-TTTGGGGTGATTTCTTCTGA-3' R 5'-TCTGTCTCGTCTTTTGGTGA-3'
12.	mTcCIR25	F 5'-CTTCGTAGTGAATGTAGGAG-3' R 5'-TTAGGTAGGTAGGGTTATCT-3'
13.	mTcCIR26	F 5'-GCATTCATCAATACATTC-3' R 5'-GCACTCAAAGTTCATACTAC-3'
14.	mTcCIR33	F 5'-TGGGTTGAAGATTTGGT-3' R 5'-CAACAATGAAAATAGGCA-3'
15.	mTcCIR37	F 5'-CTGGGTGCTGATAGATAA-3' R 5'-AATACCCTCCACACAAAT-3'
16.	mTcCIR40	F 5'-AATCCGACAGTCTTTAATC-3' R 5'-CCTAGGCCAGAGAATTGA-3'
17.	mTcCIR42	F 5'-TTGCTGAAGTATCTTTTGAC-3' R 5'-GCTCCACCCCTATTTG-3'
18.	mTcCIR60	F 5'-CGCTACTAACAAACATCAAA-3' R 5'-AGAGCAACCATCACTAATCA-3'

19.	mTcCIR102	F 5'-TTGTGAAAAGATTGCGA-3' R 5'-TTGCTTGTTATTGCTACTAT-3'
20.	mTcCIR121	F 5'-CATGTGCATTTAGGTGTC-3' R 5'-TCTGGCTTCTTAGTGATAC-3'
21.	mTcCIR146	F 5'-GCAAGGTCTTTTTACGAT-3' R 5'-ATGGACACGTCTAAGTTG-3'
22.	SHRSTc49	F 5'-ATCGCAGCAAACTCCCTCTC-3' R 5'-TTCTCTTCCCACCAAGTCCC-3'
23.	SHRSTc51	F 5'-CTGTTTTTGCCTCCCTTGTTCT-3' R 5'-ATTGCTGGTTGTTCTCCATCCT-3'
24.	SHRSTc52	F 5'-TTTTAGAGCATCCACTTCCCT-3' R 5'-CCATTCTTTCCACACTGAGAG-3'
25.	SHRSTc53	F 5'-TTCCCTTTCTTTCTCTCTCTC-3' R 5'-AGTCGTTGCTACTGCTGG-3'
26.	SHRSTc64	F 5'-TCCTACATTCCTGCACCC-3' R 5'-TCGAGAGAAAAGCTCTTACACT-3'
27.	SHRSTc66	F 5'-ACAGGAATCCCCATCAGCGA-3' R 5'-GCAATCGACAGGCATGAGAGAG-3'
28.	SSRKAU11	F 5'-TGTTGGTTGGAGAAACTCCC-3' R 5'-AGGCATTTAAACCAATAGGTAGC-3'
29.	SSRKAU12	F 5'-TCCTCAAGAAATGAAGCTCTGA-3' R 5'-CCTTGGAGATAACAACCACAA-3'
30.	SSRKAU13	F 5'-GGAATAACCTCTAACTGCGGG-3' R 5'-CGATGCCTTCATTTGGACTT-3'
31.	SSRKAU15	F 5'-TGTTGCTCGAACTCTCCAAA-3' R 5'-CATAGGAGAGGTAACCCGCA-3'
32.	SSRKAU19	F 5'-ATTGTACAAAGACCCGTGGC-3' R 5'-GTTGCACACTGGATCAATGC-3'
33.	SSRKAU20	F 5'-AGGGTCCTTCGTTTGGAACT-3' R 5'-GCATTCCACTTGTGAAGCAT-3'
34.	SSRKAU21	F 5'-GGTCCAGTTCAATCAACCGA-3' R 5'-TGAAGTCGTCTCATGGTTCG-3'
35.	SSRKAU22	F 5'-GCAGAGGATATTGCATTCGC-3' R 5'-CAAACCGAACTCATCAAGGG-3'
36.	SSRKAU24	F 5'-CCGAGGCGAATCTTGAATAC-3' R 5'-GCACCATCTCTTGTGCCTCT-3'
37.	SSRKAU25	F 5'-CTCGTCTTTAGGTATCAATGGAGAT-3' R 5'-TCAATGCTACTCAATGGCTCA-3'

3.4 Data analysis

The selected eight cocoa varieties were further analyzed by using selected 10 ISSR and 11 SSR markers (from Table 1 and Table 2).

Distinct bands generated through electrophoresis were visually scored based on molecular weight and used for DNA fingerprinting. Different colour codes were given to the amplicons of same size for easy detection and interpretation of results. Identification of polymorphic amplicons were done by visual scoring of gel pictures obtained by individual primers wits eight varieties. Polymorphic bands which were found in maximum of three varieties were considered as fingerprints. Based on the visual scoring, colour charts were developed for effective data analysis.

The polymorphism obtained were further processed by Jaccards coefficient of similarities and dendogram based on similarity coefficient was generated using Unweighted Pair Group Method with Arithmetic Mean (UPGMA). Well resolved and distinct amplicons were scored.

The primers with excellent distinguishing capacity among selected varieties were identified by calculating resolving power (Rp) of the individual primers (Prevost and Wilkinson, 1999). The Resolving Power (Rp) of a primer is: $Rp = \Sigma\ Ib$ where, Ib indicates band informativeness which is equal to $1 - [2 \times (0.5 - p)]$, p denotes the proportion of the 8 varieties of cocoa containing the band.

Hollman *et al.* (2005) reported that, polymorphism information content (PIC) value represents the ability of a marker to detect polymorphism within a population based on detectable alleles and allele frequency. The PIC of the selected primers was calculated as $1-\Sigma\ pi^2$ where, *pi* is the frequency of i^{th} allele. Calculation of PIC value has further helped to confirm the reliability and usefulness of the particular primer for DNA fingerprinting.

Results and discussion

4. RESULTS AND DSCUSSION

Eight popular varieties of cocoa released from Kerala Agricultural University were used for the study. The study was conducted at Centre for Plant Biotechnology and Molecular Biology (CPBMB), KAU during 2014-2016. The results of the present investigation on 'DNA fingerprinting of promising cocoa (*Theobroma cacao* L.) varieties of KAU' are presented below.

4.1 Morphological characterisation

Morphological observations on distinguishable quantitative and qualitative characters were recorded on five pods collected from each variety using reported descriptors (Bekele and Butler, 2000). Selected eight varieties included six selections (CCRP 1, CCRP 2, CCRP 4, CCRP 5, CCRP 6, and CCRP 7) and two hybrids (CCRP 8 and CCRP 9).

Variations in important qualitative characters such as pod shape, colour of ripe pod, pod apex form, pod basal constriction, pod rugosity and cotyledon colour were recorded (Table 3). Similarly, important quantitative characters such as pod weight, pod length, pod breadth, husk thickness, number of beans, number of flat beans, wet bean weight/pod, single wet bean weight and single dry bean weight were also recorded and depicted in Table 4.

4.1.1 Qualitative characters

The varieties evaluated grouped under three different pod shapes *viz.* cundeamor, amelonado and angoleta. Varieties CCRP 4, CCRP 5, CCRP 6 and CCRP 7 were of angoleta (oval) type and CCRP 2 produced amelonado (melon shaped) shaped pods. Clone CCRP 1 and hybrids CCRP 8 and CCRP 9 produced cundeamor (ridged and with bottle neck) shaped pods. All varieties with Cundeamor had acute apex. Out of four varieties with Angoleta three expressed obtuse end indicating fruit shape can be identified by its apex form. The results were in tune with the early study by Minimol *et al.* (2011) stating that fruit shape is influenced by fruit apex.

Greenish yellow pods were produced by CCRP 7, CCRP 8 and CCRP 9, whereas the pods produced by CCRP 4 and CCRP 5 were yellow in colour. CCRP 1, CCRP 2 and CCRP 6 yielded yellowish green coloured pods (Table 4). Although there are slight variations in fruit colour, all varieties were grouped under the category Forastero. The other characters like rugosity and colour of cotyledon also support the fact that all the varieties were forastero types. Rugosity was medium and colour of cotyledon ranged from pink to dark pink. The same features were described by Wood and Lass, (1985) for forastero types.

4.1.2 Quantitative characters

4.1.2.1 Yield (number of pods/ tree/year)

Potential yield (number of pods/tree/year) of the selected varieties is explicated in table 3. It was evident that hybrids showed more yield when compared to selections. However CCRP 6 a selection from seedlings of exotic variety IMC 10 excelled over the hybrids. The parent IMC 10 had been reported to be a good performer with high yield and wet bean weight (Iwaro *et al.*, 2003).

4.1.2.2 Yield contributing characters

The eight varieties differed significantly with respect to all the nine quantitative characters *viz.* pod weight, pod length, pod breadth, husk thickness, number of beans, number of flat beans, wet bean weight/pod, single wet bean weight and single dry bean weight (Table 5).

A significant difference was expressed for pod weight among the varieties. The highest pod weight was recorded in CCRP 6 (608.84 g) followed by CCRP 4 (583.20 g) and they were at par. The least pod weight recorded was 378.56 g by CCRP 2 (Table 5). However all the varieties expressed pod weight more than 350 g which was the selection criteria recommended by Francies *et al.* (2002). CCRP 6 was found to yield pods with an average of 18.82 cm length and was found maximum. Whereas, hybrid CCRP 9 was having small pod length of 12.83 cm (Table 5). The highest pod breadth of 9.12 cm was recorded for CCRP 6 (9.12 cm) and minimum in CCRP 9 (6.93 cm) (Table 5). Length and breadth of the pods

were found to be proportional. Husk thickness had a significant role in deciding pod weight (Rubeena, 2015). Husk thickness of one cm or less than one cm is the desirable character (Enriquez and Soria, 1966) and almost all varieties fell within the range. Only one selection (CCRP 4) expressed a value higher than one (1.02 cm) but the increase is negligible.

The average number of beans per pod was recorded and found diverse among the varieties. However there is no significant difference among the varieties (CCRP 1, CCRP 4, CCRP 5, CCRP 7 and CCRP 8) with highest bean count. Large number of beans alone cannot consider as a selection criteria. More number of beans with less weight may yield less butter content (Mossu, 1992). Yield expressed as wet or dry bean weight is highly variable (Pound 1932; Enriquez and Soria, 1966). Maximum wet bean weight was expressed by CCRP 6 (142.86g) and the least by CCRP 2 (67. 42 g) (Table 5).

The most important economic part of cocoa is beans. Size of bean is considered as one of the important component of yield in cocoa (Soria 1978). Bean characters expressed variations with in the species (Adewale *et al.*, 2010). However in the present study the peeled dry weight of single bean of all the varieties was more than 0.8 g which is the accepted standard (Wood and Lass, 1985).

Table 3. Potential yield of cocoa varieties

Sl. No.	Varieties	No. of pods/tree/year
1	CCRP 1	72
2	CCRP 2	90
4	CCRP 4	93
5	CCRP 5	75
6	CCRP 6	180
7	CCRP 7	95
8	CCRP 8	131
9	CCRP 9	105

Table 4. Pod qualitative characters of popular cocoa varieties of KAU

Characters/ Varieties	Pod shape	Colour of ripe pod	Pod apex form	Pod basal constriction	Pod rugosity	Colour of bean
CCRP 1	Cundeamor	Yellowish green	Acute	Slight	Intermediate	Dark purple
CCRP 2	Amelonado	Yellowish green	Rounded	Slight	Intermediate	Light purple
CCRP 4	Angoleta	Yellow	Obtuse	Slight	Intermediate	Dark purple
CCRP 5	Angoleta	Yellow	Obtuse	Slight	Intermediate	Dark purple
CCRP 6	Angoleta	Yellowish green	Obtuse	Slight	Intermediate	Light purple
CCRP 7	Angoleta	Greenish yellow	Acute	Slight	Intermediate	Dark purple
CCRP 8	Cundeamor	Greenish yellow	Acute	Slight	Intermediate	Dark purple
CCRP 9	Cundeamor	Greenish yellow	Acute	Intermediate	Intermediate	Dark purple

Table 5. Mean values of yield contributing characters of cocoa varieties

CCRP clones/ hybrids	Pod weight (g)	Pod length (cm)	Pod breadth (cm)	Husk thickness (cm)	Number of beans/pod	Number of flat beans/pod	Wet bean weight/pod (g)	Wet bean weight/ single seed (g)	Dry bean weight/ single seed (g)
CCRP 1	439.06^{cd}	15.74^{b}	7.62^{b}	0.84^{cd}	43.60^{ab}	0.60	109.92^{bc}	2.53^{bc}	0.94^{b}
CCRP 2	378.56^{d}	14.36^{bc}	8.60^{a}	0.93^{abc}	33.40^{cd}	0.60	67.42^{d}	2.01^{d}	0.78^{cde}
CCRP 4	583.20^{ab}	16.00^{b}	9.08^{a}	1.02^{a}	47.80^{a}	0.60	132.78^{a}	2.96^{ab}	1.05^{a}
CCRP 5	486.20^{c}	16.36^{ab}	8.78^{a}	0.91^{bcd}	42.80^{ab}	1.60	116.10^{bc}	2.76^{bc}	0.86^{bc}
CCRP 6	608.84^{a}	18.82^{a}	9.12^{a}	0.95^{ab}	39.60^{bc}	1.80	142.86^{ab}	3.35^{a}	1.09^{a}
CCRP 7	447.48^{cd}	15.82^{b}	7.76^{b}	0.94^{abc}	42.40^{ab}	0.60	102.48^{c}	2.43^{cd}	0.75^{c}
CCRP 8	475.91^{c}	14.63^{bc}	7.64^{b}	0.81^{d}	44.40^{ab}	0.60	126.03^{abc}	2.84^{bc}	0.84^{cd}
CCRP 9	510.50^{bc}	12.83^{c}	6.93^{c}	0.82^{d}	31.80^{d}	1.40	107.35^{c}	3.38^{a}	0.77^{de}
CD (0.05)	**91.189**	**2.8**	**0.622**	**0.102**	**6.443**	**NS**	**25.184**	**0.475**	**0.085**
CV (%)	**14.409**	**13.957**	**5.891**	**8.772**	**12.281**	**6.443**	**17.281**	**13.248**	**7.404**

4.2 Molecular characterization

4.2.1 Isolation and quantification of DNA

For the conduct of any molecular analysis, the basic requirement is good quality DNA. A number of DNA extraction protocols were reported for molecular marker analysis in various crops (Murray and Thompson, 1980; Dellaporta *et al.*, 1983; Couch and Fritz, 1990; Lanaud *et al.*, 1995; Rogers and Bendich, 1994; Perry *et al.*, 1998; Bhattacharjee *et al.*, 2004). Young leaves (second or third leaf from the shoot tip) were reported to be the desirable part for extraction of good quality DNA (Charters and Wilkinson, 2000; Bhattacharjee *et al.*, 2004; Kumar and Sharma, 2011). Tender pale yellow or reddish colored leaves were also reported to yield good quality DNA (Chandrakant, 2014).

In the present study tender leaves from the actively growing tips of cocoa trees were collected during early morning for extraction of genomic DNA. The modified Doyle and Doyle (1987) method of extraction protocol was found to yield good quality DNA (Chandrakant, 2014). However when employed here, the DNA extracted was found to be dominated by RNA contamination and high amount of mucilage content as the plant belongs to family Malvaceae. So, slight modification were made in the protocol and isolated good quality DNA (Plate 2). Initially, sample quantity was reduced to 0.08-0.10 gm from 1 gm. Liquid nitrogen was used for homogenization and complete disruption of the tissue which further resulted in isolation of "nick"-free DNA (Blin and Stafford, 1976) and also for DNA isolation from dehydrated plant tissue (Tai and Tanksley, 1990). Inhibition of polyphenol was done by addition of anti-oxidant β-mercaptoethanol and PVP in excess quantity. Further, the addition of chloroform: isoamyl alcohol (24:1) during the first centrifugation step was avoided and instead of 2 per cent CTAB, 5 per cent was used. The protocol used in the study is detailed in 3.2.2.1.

Bayer *et al.* (1999) reported that leaf samples from many Malvaceae samples contain high amounts of mucilage, and it is difficult to remove the mucilaginous supernatant after centrifugation without losing most of the DNA. In studies conducted by many scientist it is revelled that in cocoa, a plant belongs to

Malvaceae family, the leaf tissue used for DNA isolation were characteristically high in polysaccharides, polyphenols, tannins and other secondary metabolites, which interfere with effective DNA extractions and polymerase chain reaction (PCR) amplification (Machado *et al.*, 2005; Chandrakant, 2014). Polyphenols have viscous, glue-like texture and make the DNA unmanageable in pipetting and unsuitable for PCR by inhibiting Taq polymerase activity (Fang *et al.*, 1992). Upon cell lysis, these compounds bind to DNA and cannot be removed by conventional extraction procedures (John, 1992).

Jose and Usha (2000) described a protocol for the extraction of genomic DNA from *Abelmoscus esculentus*, a Malvaceae plant. This method involved a combination of Dellaporta *et al.* (1983), the CTAB method (Doyle and Doyle, 1990), and alkali lysis. Protocols developed for DNA extraction from Malvaceae plant species rich in polysaccharides, polyphenols, and latex were found to be time consuming or expensive (Jeske *et al.*, 2001). Other techniques for DNA extraction from different highly mucilaginous plants, which included modifications of the CTAB-based protocols, had been reported (Barnwell *et al.*, 1998; Mansoor *et al.*, 1999). Expensive commercial kits are also available and recommended for the extraction of DNA form cocoa (Haymes *et al.*, 2004).

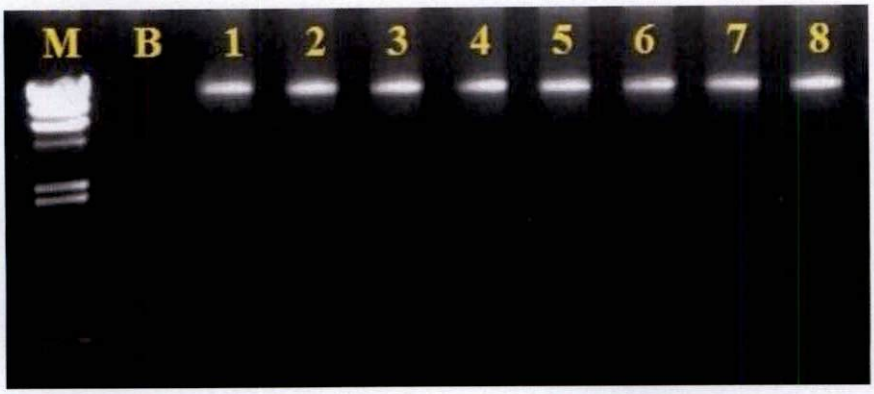
Quantification of isolated DNA can be done in many ways. Flow cytometry (FCM) using DNA-selective fluorochromes is one of the method for the measurement of nuclear DNA content in plants (Dolezel *et al.*, 2007). Similarly, in a study conducted in small grains, Matrix Assisted Laser Desorption Ionisation-Time of Flight Mass Spectrometry (MALDITOF-MS), was used to provide accurate molecular weight determination (Terzi *et al.*, 2005).

In the present study, spectrophotometer analysis was performed using nanodrop ND-1000® and in the quantification, optical density at 260 and 280 nm was taken in to consideration. Absorbance value was calculated as OD 260/280 and the ratio was found to be in between 1.80 to 2.00 (Table 6) which was considered as good quality DNA (Meena, 2012; Laxman, 2013; Thakur *et al.*, 2014).

Thus, the quantified DNA samples were diluted to a final concentration of 40 ng/µl as per the previous reports (Lanaud *et al.*, 1995). But due to the presence of polyphenols, amplification rate was found so poor and then the concentration was reduced to 10 ng/µl. Further, the amplification was found excellent and the same concentration (10 ng/µl) was used for both ISSR as well as SSR analysis. Reproducibility of a molecular marker is very important (Nayak *et al.*, 2003) and in this study, reproducibility of selected ISSR and SSR primers were tested by repeating the PCR reaction for at least three times under same PCR conditions and using same chemicals.

Table 6. DNA Quantification result of eight cocoa varieties

Varieties	UV absorbance at 260nm (A_{260})	UV absorbance at 280nm (A_{280})	Optical Density ($A_{260/280}$)	Quantity (ng/µl)
CCRP 1	17.2	9.10	1.89	205.02
CCRP 2	9.57	5.04	1.90	480.86
CCRP 4	11.79	6.48	1.82	597.12
CCRP 5	6.17	3.07	2.01	393.33
CCRP 6	13.34	6.95	1.92	143.34
CCRP 7	11.48	6.14	1.87	190.10
CCRP 8	9.53	5.21	1.83	284.70
CCRP 9	6.27	3.23	1.94	451.12



M- Molecular weight marker, B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 2. Good quality DNA isolated from eight cocoa varieties

4.3 Molecular marker analysis

Isolated DNA was bulked and screening was done to find out suitable primers with desirable banding pattern. Thirty five ISSR and thirty seven SSR primers were validated with bulked DNA and selected primers were further utilized to amplify all the eight samples.

4.3.1 Inter Simple Sequence Repeat (ISSR) analysis

4.3.1.1 Primer screening for ISSR assay

Primers based microsatellite regions in the genome are designed and utilized to amplify inter simple sequence repeat regions in the genome (Zietkiewicz *et al.,* 1994). Thus, the designed primers which are successfully annealed to two microsatellite regions within an amplifiable stretch will generate a band of particular molecular weight. ISSR is a PCR based molecular marker assay that assess variation in the numerous microsatellite regions dispersed throughout the genome (Chandrakant, 2014).

Thirty five primers were used for ISSR analysis (Table 7) with reaction mixture composition and thermal conditions mentioned earlier (3.3.2). ISSR is highly reproducible, sensitive and dominant marker system and had been successfully used for evolutionary and genetic diversity study of many crops including fingermillet (Salimath *et al.,* 1995), wheat (Nagaoka and Ogihara, 1997), diploid banana (Godwin *et al.,* 1997), citrus (Fang and Roose, 1997), rice (Joshi *et al.,* 2000), Vigna (Ajibade *et al.,* 2000), litchi (Clyde *et al.,* 2005), mango (Pandit *et al.,* 2007) etc. ISSR primers are designed form SSR motifs and can be undertaken for any plant species containing a sufficient number and distribution of SSR motifs in the genome (Gupta *et al.,* 1996). ISSR targets the highly variable microsatellite regions of the nuclear genome providing a large number of polymorphic fragments (Gupta *et al.,* 1994). Therefore this technique has been used extensively to evaluate genetic diversity both within and between plant populations in angiosperms and gymnosperms (Osborn *et al.,* 2005). The high level reproducibility of ISSR markers is already well established

in many crops and hence these markers offer the potential for direct usage in MAS (Bornet and Branchard, 2001; Reddy *et al.,* 2002).

In the present study based on the amplification pattern with different primers, ten ISSR primers were selected. Selected primers and their annealing temperature details are provided in Table 8. ISSR markers are better than the others since they do not require prior information of the DNA sequence, have low development costs, and have high transferability to other plant species, and generate a number of polymorphic fragments (Barth *et al.*, 2002; Brandao *et al.,* 2011). This marker also have high reproducibility and potential to be used in species without detailed DNA sequence information (Rossi *et al.,* 2009; Almeida *et al.,* 2009; Brandao *et al.,* 2011).

Table 7. Details of amplification pattern obtained with 35 ISSR primers

Sl. No.	Primer	Amplification pattern			
		No. of amplicon	Types of amplicons		Remarks
			Distinct	Faint	
1.	HB 10	0	0	0	Not selected
2.	HB 12	6	5	1	Not selected
3.	Oligo ISSR 04	0	0	0	Not selected
4.	Oligo ISSR 05	0	0	0	Not selected
5.	Oligo ISSR 06	0	0	0	Not selected
6.	Oligo ISSR 07	0	0	0	Not selected
7.	Oligo ISSR 08	0	0	0	Not selected
8.	ISSR 2	0	0	0	Not selected
9.	ISSR 3	0	0	0	Not selected
10.	ISSR 4	0	0	0	Not selected
11.	ISSR 5	0	0	0	Not selected
12.	ISSR 6	0	0	0	Not selected
13.	ISSR 7	2	2	0	Not selected
14.	ISSR 8	0	0	0	Not selected
15.	ISSR 9	0	0	0	Not selected
16.	UBC 808	4	1	3	Not selected
17.	UBC 810	11	7	4	Selected
18.	UBC 811	3	2	1	Not selected
19.	UBC 815	10	7	3	Selected
20.	UBC 818	2	1	1	Not selected
21.	UBC 826	10	6	4	Selected
22.	UBC 827	8	7	1	Selected
23.	UBC 835	9	7	2	Selected
24.	UBC 841	8	7	1	Selected
25.	UBC 842	12	7	5	Selected
26.	UBC 846	6	5	1	Selected
27.	UBC 847	3	2	1	Not selected
28.	UBC 848	0	0	0	Not selected
29.	UBC 850	2	1	1	Not selected
30.	UBC 854	0	0	0	Not selected
31.	UBC 855	9	7	2	Selected
32.	UBC 857	5	4	1	Not selected
33.	UBC 865	0	0	0	Not selected
34.	UBC 866	6	6	0	Selected
35.	UBC 873	2	1	1	Not selected

Table 8. Details of ISSR primers selected for fingerprinting

Sl. No.	Primer	Annealing temperature (°C)	Nucleotide sequence
1.	UBC 810	50.5	5'- GAGAGAGAGAGAGAGAT-3'
2.	UBC 815	44.9	5'-CTCTCTCTCTCTCTCTG-3'
3.	UBC 826	53.3	5'-ACACACACACACACACC-3'
4.	UBC 827	54.9	5'-ACACACACACACACACG-3'
5.	UBC 835	45.6	5'-AGAGAGAGAGAGAGAGYC-3'
6.	UBC 841	46.0	5'-GAGAGAGAGAGAGAGAY-3'
7.	UBC 842	46.5	5'-CTCTCTCTCTCTCTCTRC- 3'
8.	UBC 846	53.7	5'-CACACACACACACACART-3'
9.	UBC 855	52.1	5'-ACACACACACACACACYT-3'
10.	UBC 866	55.0	5'-CTCCTCCTCCTCCTCCTC-3'

4.3.1.1 DNA fingerprinting with selected ISSR primers

In order to develop ISSR fingerprint of selected eight varieties, amplification of respective DNA samples using selected 10 ISSR primers (Table 8) were performed. Amplification pattern obtained with each primers were observed thoroughly to identify the polymorphic bands. The details of amplification with the ten selected ISSR primers are detailed below.

4.3.1.1.1 UBC 810

Eight cocoa DNA samples were amplified using primer UBC 810. (Plate 3). The primer generated an average of eight amplicons in each accessions. The pattern of amplification is shown (Fig. 1). Molecular size of the amplicons generated were varied from 400 bp to 1300 bp. On careful analysis, amplicon obtained at 850 bp was found to be distinct polymorphic band which was observed only in CCRP 6. This primer was found to be a good candidate to identify CCRP 6 and the amplicon observed at 850 bp can be a fingerprint of the same variety.

4.3.1.1.2 UBC 815

ISSR analysis of the DNA samples using the primer UBC 815 generated an average of eight amplicons per accession on agarose gel (Plate 4). The molecular size of the amplicons varied from 250 to 1600 bp. The pattern of amplification is shown in Fig. 2. Through amplicon scoring revealed the presence of a band at 600 bp in CCRP 2, CCRP 4 and CCRP 5. Thus, the primer UBC 815 can be used for distinguishing varieties CCRP 2, CCRP 4 and CCRP 5 from a mixed population.

4.3.1.1.3 UBC 826

On an average of nine amplicons were obtained when the DNA samples were amplified with the primer UBC 826 (Plate 5). The molecular size of the bands ranged from 400 to 1800 bp. On careful observation of the gel picture,

distinct polymorphic band was found in CCRP 6 at 1100 bp (Fig. 3). Thus the primer UBC 826 can be used for identifying CCRP 6.

4.3.1.1.4 UBC 827

Amplification pattern obtained using the primer UBC 826 consisted of an average of seven amplicons per genotype (Plate 6). Molecular size of the amplicons generated varied from 400 to 1950 bp. At 740 bp length, a distinct polymorphic amplicon was generated for CCRP 9. Further amplicon scoring revealed the presence of a band at 400 bp in CCRP 6, CCRP 7 and CCRP 9 (Fig. 4). Thus UBC 827 can be effectively utilized for identifying CCRP 6, CCRP 7 and CCRP 9.

4.3.1.1.5 UBC 835

UBC 835 generated an average of twelve amplicons in each accession. The pattern of amplification is shown in plate 7. Amplicons obtained with this primer on genotypes CCRP 8 and CCRP 9 were found polymorphic. The molecular size of generated bands using the primer varied from 280 to 1600 bp. In hybrids CCRP 8 and CCRP 9, polymorphic bands were obtained at 550 bp, 1400 bp and 1600 bp length. The primer UBC 835 can be an ideal primer for identifying the hybrids CCRP 8 and CCRP 9 (Fig. 5).

4.3.1.1.6 UBC 841

UBC 841 had generated an average of eight clear amplicons in all eight genotypes (Plate 8). The molecular size of the band varied from 300 to 1200 bp. Amplicons generated from this primer were polymorphic, produced a unique band at 650 bp in CCRP 5. Also UBC 841 produced polymorphic bands in CCRP 2, CCRP 5 and CCRP 6 at 850 bp and at 700 bp. The primer yielded polymorphic band in varieties CCRP 2, CCRP 4 and CCRP 5 (Fig. 6). Thus, primer UBC 841 can be an ideal primer for identification of CCRP 2, CCRP 4, CCRP 5 and CCRP 6.

4.3.1.1.7 UBC 842

Amplification pattern developed using the ISSR primer UBC 842 produced an average of ten bands in all genotypes (Plate 9). Molecular size of the bands developed varied from 300 to 1500 bp. Amplicons generated at 550 bp was found to be polymorphic and the band was present in CCRP 1, CCRP 2 and CCRP 5 (Fig. 7). Thus the primer UBC 842 can be used for identifying these three varieties via, the presence of polymorphic band.

4.3.1.1.8 UBC 846

On an average of four amplicons were obtained when the DNA samples were amplified with the primer UBC 846 (Plate 10). The molecular size of the bands developed ranged from 250 to 1600 bp. On careful observation of the gel picture, distinct polymorphic band was found in CCRP 5 at three sites *i.e.*, 300 bp, 1300 bp and 1600 bp (Fig. 8). The hybrid CCRP 9 also produced a polymorphic distinct band at 350 bp. Another polymorphism was observed in CCRP 5 and CCRP 8, and the band was observed at 500 bp length. Thus the primer UBC 846 was found ideal for distinguishing CCRP 5, CCRP 8 and CCRP 9.

4.3.1.1.9 UBC 855

The ISSR primer UBC 855 had generated an average of seven clear amplicons in all eight genotypes (Plate 11). The molecular size of the band varied from 600 to 1700 bp. Amplicons generated from this primer were polymorphic, produced a band at 1700 bp in CCRP 2, CCRP 8 and CCRP 9 (Fig. 9).

4.3.1.1.10 UBC 866

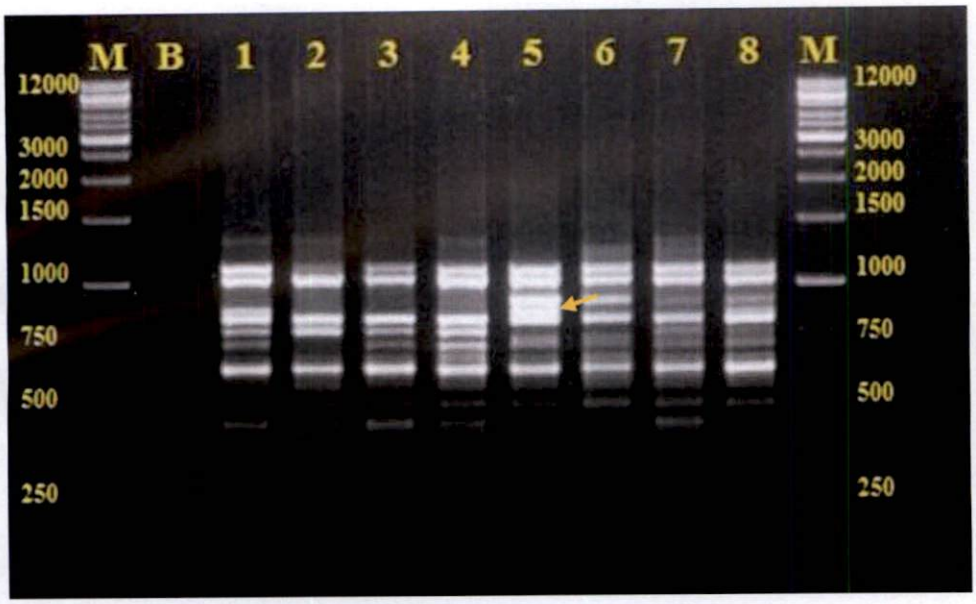
ISSR analysis of the DNA samples using the primer UBC 866 generated an average of six amplicons per accession on agarose gel (Plate 12). The molecular size of the amplicons varied from 350 to 1600 bp. The pattern of amplification is shown in Fig. 7b. Through amplicon scoring revealed the presence of a unique band at 600 bp in CCRP 9 which serves as a fingerprint for the variety. Similarly, in varieties CCRP 8 and CCRP 9, polymorphic bands were

obtained at two locations *i.e.*, 800 bp and 1200 bp (Fig. 10). Thus, the primer UBC 866 can be used for distinguishing the hybrids CCRP 8 and CCRP 9 from a mixed population.

In the present study all the 10 ISSR primers (Table 8) used were found to yield at least single useful polymorphic bands among the eight selected varieties. All the selected ten primers belonged to the series UBC (University of British Columbia). UBC 810 was found to yield distinct polymorphic band at 850 bp which was observed only in CCRP 6. UBC 826 developed a unique polymorphic band at 1100 bp in CCRP 6. The hybrid CCRP 9 generated a distinct polymorphic band at 740 bp. Primer UBC 846 yielded three distinct polymorphic bands in CCRP 5 (300 bp, 1300 bp and 1600 bp) and one distinct band in CCRP 9 at 350 bp. UBC 866 also generated a polymorphic band in CCRP 9 at 600 bp was also unique and distinct.

Among the polymorphic bands selected for construction of final DNA fingerprint, five distinct polymorphic bands were shared only by the hybrids CCRP 8 and CCRP 9. Primer UBC 835 generated bands at three different sites (550 bp, 1400 bp and 1600 bp) and primer UBC 866 produced bands at 800 bp and 1200 bp in both the hybrids. Similarly, at 500 bp, UBC 846 generated an amplicon which was shared by CCRP 5 and CCRP 8

In cocoa, ISSR markers were effectively used in the characterization of germplasm maintained at Tingo Maria (Julio *et al.,* 2011). Similarly, population structure and genetic diversity in natural populations of *Theobroma subincanum* was studied using ISSR markers (Rivas *et al.,* 2013). Giustina *et al.* (2013) also used ISSR markers to screen natural populations of *Theobroma speciosum* and able to detect polymorphism in the populations analyzed and proved to be reproducible for the species.



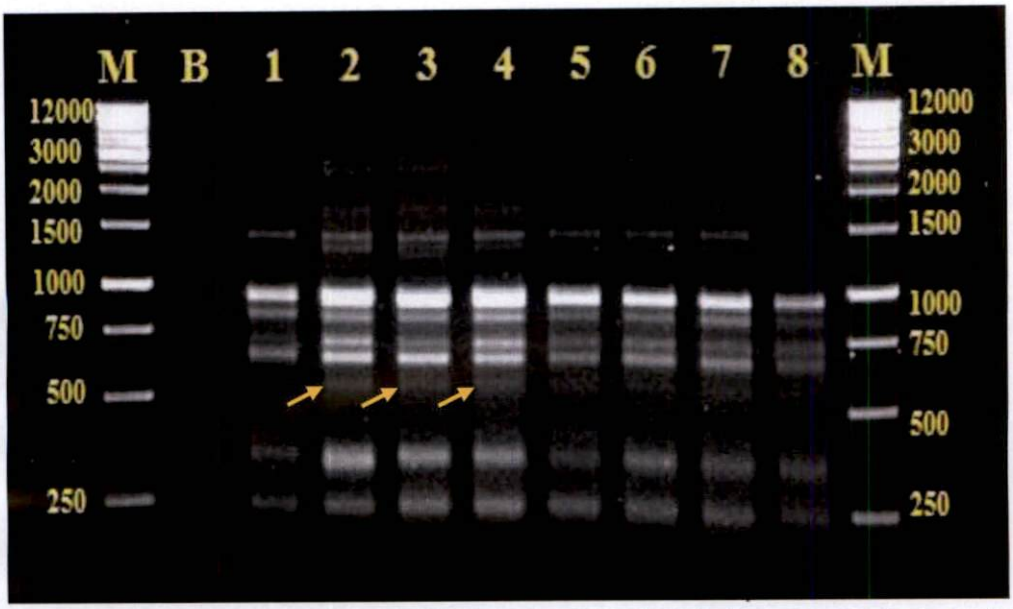
M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 3. Amplification pattern generated with primer UBC 810

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1300								
1200								
1000								
900								
850								
800								
750								
600								
500								
400								

Colour code for sharing of bands among varieties

Fig. 1 Colour chart developed using UBC 810



M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

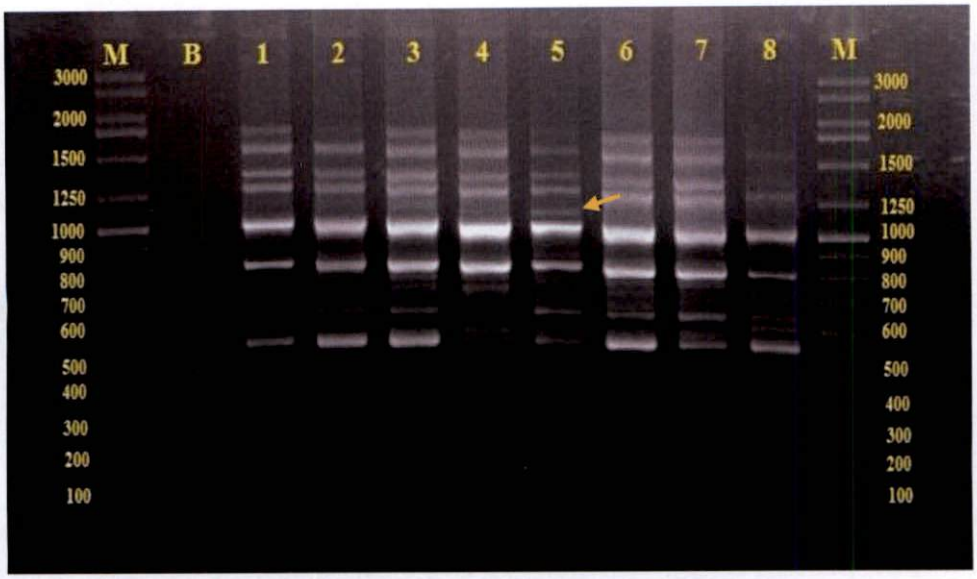
Plate 4. Amplification pattern generated with primer UBC 815

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1600								
1500								
1400								
1000								
900								
750								
700								
600								
400								
250								

Colour code for sharing of bands among varieties

1	2	3	4	5	6	7	8

Fig. 2 Colour chart developed using UBC 815



M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

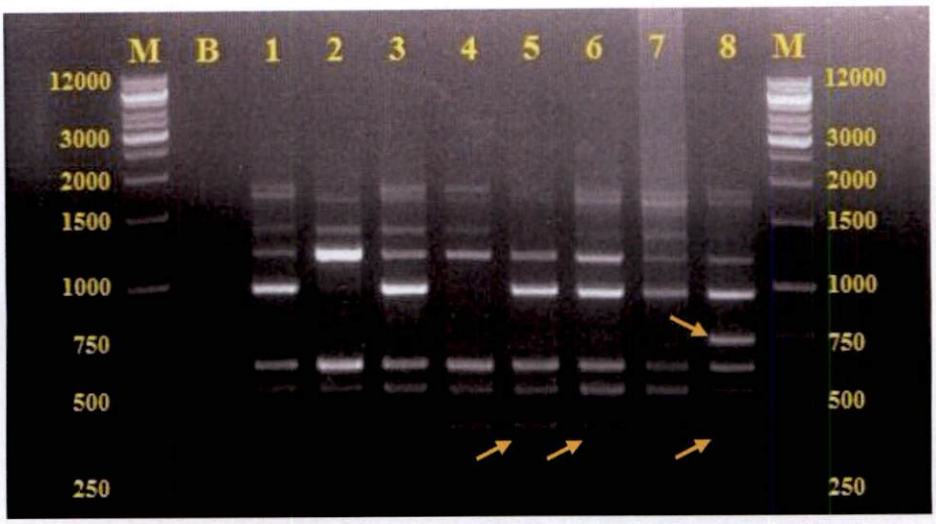
Plate 5. Amplification generated with primer UBC 826

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1800								
1600								
1300								
1200								
1100								
1000								
800								
600								
550								
450								
400								

Colour code for sharing of bands among varieties

1	2	3	4	5	6	7	8

Fig. 3 Colour chart developed using UBC 826



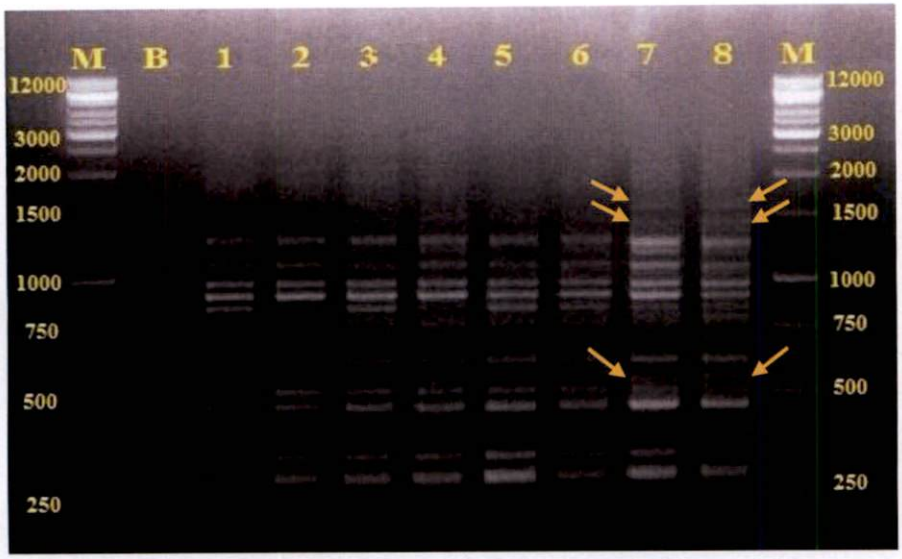
M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 6. Amplification pattern generated with primer UBC 827

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1950								
1900								
1400								
1200								
950								
740								
600								
550								
450								
400								

Colour code for sharing of bands among varieties

Fig. 4 Colour chart developed using UBC 827



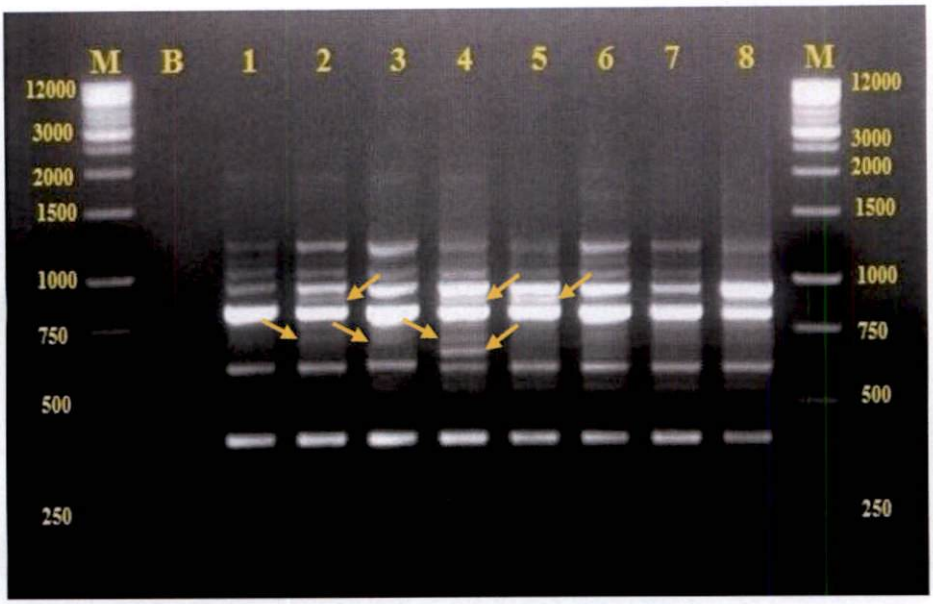
M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 7. Amplification pattern generated with primer UBC 835

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1600								
1400								
1250								
1200								
1100								
950								
900								
850								
800								
600								
550								
500								
480								
300								
280								

Colour code for sharing of bands among varieties

Fig. 5 Colour chart developed using UBC 835



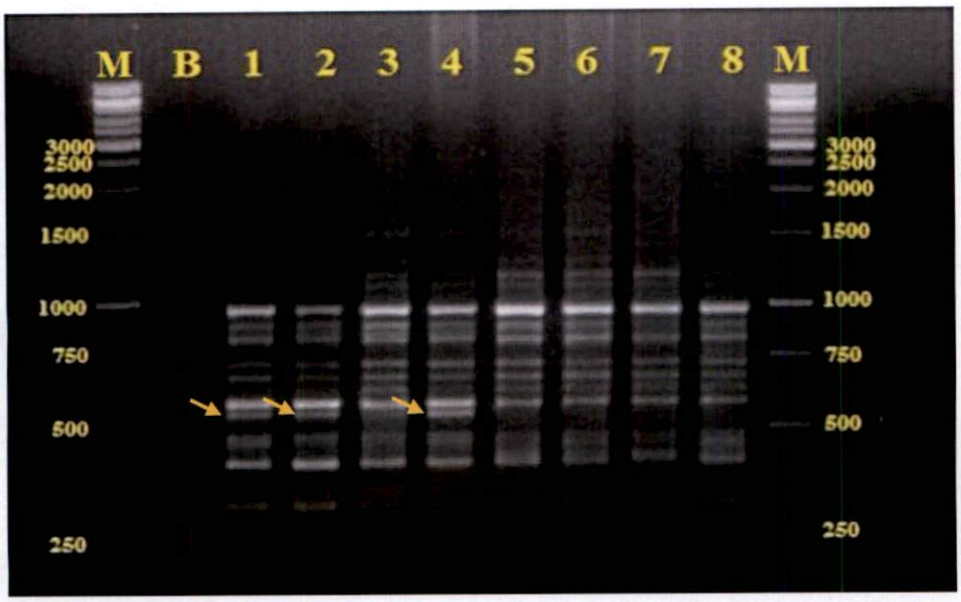
M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 8. Amplification pattern generated with primer UBC 841

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1200								
1100								
1000								
900								
850								
800								
700								
650								
600								
550								
400								
300								

Colour code for sharing of bands among varieties

Fig. 6 Colour chart developed using UBC 841



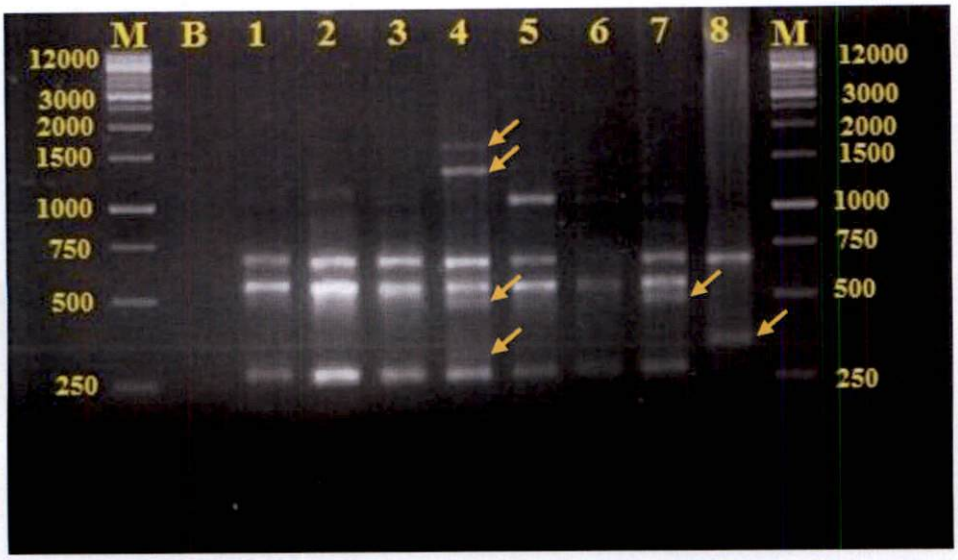
M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 9. Amplification pattern generated with primer UBC 842

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1500								
1300								
1250								
1200								
900								
850								
800								
700								
650								
600								
550								
450								
400								
350								
300								

Colour code for sharing of bands among varieties

Fig. 7 Colour chart developed using UBC 842



M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

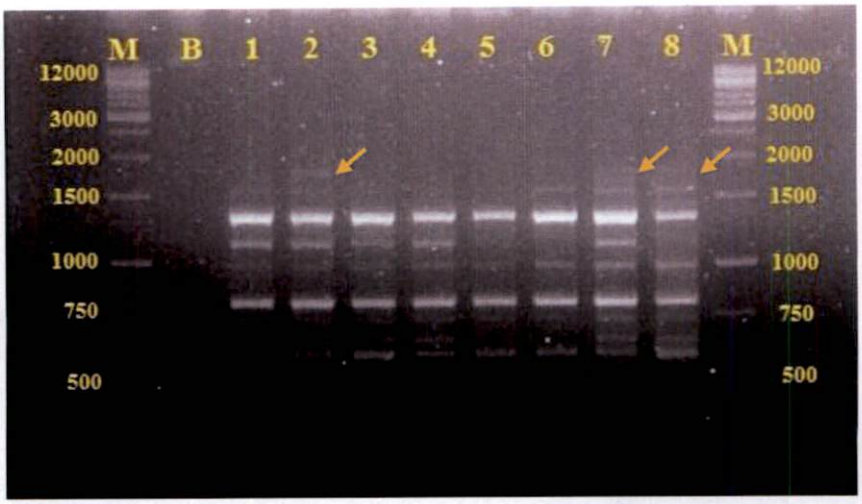
Plate 10. Amplification pattern generated with primer UBC 846

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1600								
1300								
1100								
700								
550								
500								
350								
300								
250								

Colour code for sharing of bands among varieties

1	2	3	4	5	6	7	8

Fig. 8 Colour chart developed using UBC 846



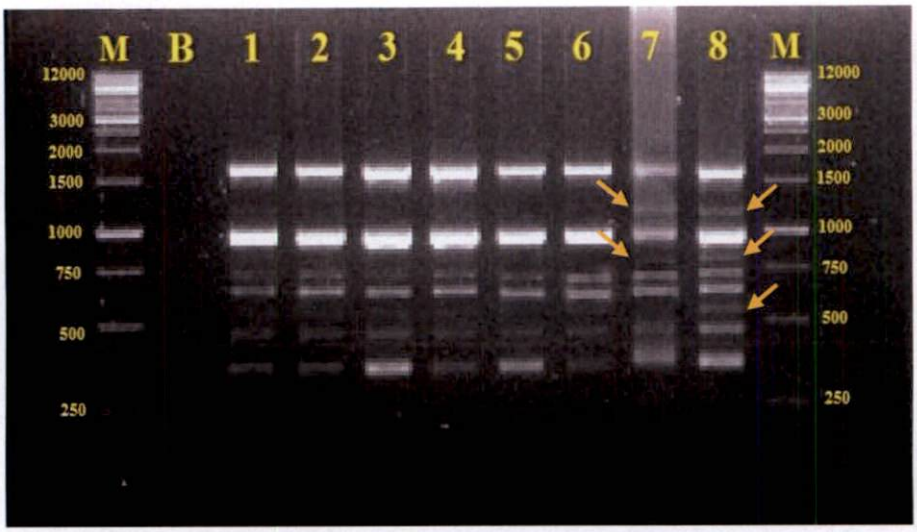
M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 11. Amplification pattern generated with primer UBC 855

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1700								
1550								
1400								
1200								
1000								
800								
750								
650								
600								

Colour code for sharing of bands among varieties

Fig. 9 Colour chart developed using UBC 855



M- Molecular weight marker (12kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 12. Amplification pattern generated with primer UBC 866

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1600								
1200								
950								
800								
700								
650								
600								
450								
350								

Colour code for sharing of bands among varieties

Fig. 10 Colour chart developed using UBC 866

4.3.2 Simple Sequence Repeat (SSR) analysis

Microsatellites are known to be very efficient marker system for assessing genetic diversity due to their high discriminatory power and polymorphism information content among individuals, which arising from their multi-allelic nature and co-dominant segregation of alleles (Powell *et al.*, 1996). SSRs are derived from ESTs, corresponding to the transcribed component of a gene unit and, they have been shown to possess a high potential for inter-specific transferability (Cordeiro *et al.*, 2001; Thiel *et al.*, 2003; Gupta *et al.*, 2003). They were also being widely used in many studies on genotype identification and genetic mapping (Ramsay *et al.*, 2000) and linkage map construction (Pugh *et al.*, 2004; Lu *et al.*, 2005). Due to their codominant nature, single-locus behaviour and high polymorphism, microsatellites constitute a set of useful markers that are transferable across different mapping populations. SSRs also allowing QTL position comparison, and are easily transferable to laboratories (Pugh *et al.*, 2004).

4.3.2.1 Primer screening for SSR assay

Thirty seven primer sets were screened for SSR analysis (Table 9) with reaction mixture composition and thermal conditions as mentioned in previous chapter (3.3.3). Based on the amplification pattern obtained with different primer sets, eleven SSR primer sets were chosen for development of fingerprints. The selected primer sets and their annealing temperature details are depicted in table 10. Eight out of eleven primers belonged to the series mTcCIR and three belonged to SHRSTc.

SSR markers are widely used due to their high polymorphic information content, ease of genotyping through PCR, codominant nature, allelic nature and high discriminating power (Russell *et al.*, 1997). Similarly, Rallo *et al.* (2000) reported that, presence of easily scorable, unique alleles or allele combination made the primer system an ideal tool for cultivar identification. Moreover, only small quantity of DNA is required and the quality of DNA need not to be high which is necessary for other advanced DNA assays (Refalski *et al.*, 1996)

Table 9. Details of amplification pattern obtained with 37 SSR primers

Sl. No.	Primer	Amplification pattern			
		No. of bands	Types of bands		Remarks
			Distinct	Faint	
1.	mTcCIR1	0	0	0	Not selected
2.	mTcCIR6	0	0	0	Not selected
3.	mTcCIR7	0	0	0	Not selected
4.	mTcCIR8	3	1	2	Selected
5.	mTcCIR10	0	0	0	Not selected
6.	mTcCIR11	1	1	0	Selected
7.	mTcCIR12	3	1	2	Selected
8.	mTcCIR15	0	0	0	Not selected
9.	mTcCIR18	1	1	0	Selected
10.	mTcCIR22	0	0	0	Not selected
11.	mTcCIR24	1	1	0	Selected
12.	mTcCIR25	0	0	0	Not selected
13.	mTcCIR26	0	0	0	Not selected
14.	mTcCIR33	3	2	1	Selected
15.	mTcCIR37	0	0	0	Not selected
16.	mTcCIR40	2	1	1	Selected
17.	mTcCIR42	2	2	0	Selected
18.	mTcCIR60	0	0	0	Not selected
19.	mTcCIR102	0	0	0	Not selected
20.	mTcCIR121	0	0	0	Not selected
21.	mTcCIR146	0	0	0	Not selected
22.	SHRSTc49	4	2	2	Selected

23.	SHRSTc51	3	2	1	Selected
24.	SHRSTc52	0	0	0	Not selected
25.	SHRSTc53	0	0	0	Not selected
26.	SHRSTc64	3	2	1	Selected
27.	SHRSTc66	1	0	1	Not selected
28.	SSRKAU11	0	0	0	Not selected
29.	SSRKAU12	0	0	0	Not selected
30.	SSRKAU13	0	0	0	Not selected
31.	SSRKAU15	0	0	0	Not selected
32.	SSRKAU19	0	0	0	Not selected
33.	SSRKAU20	0	0	0	Not selected
34.	SSRKAU21	0	0	0	Not selected
35.	SSRKAU22	0	0	0	Not selected
36.	SSRKAU24	0	0	0	Not selected
37.	SSRKAU25	0	0	0	Not selected

Table 10. Details of SSR primers selected for fingerprinting

Sl. No.	Primer	Annealing temperature (°C)	Nucleotide sequence
1.	mTcCIR8	46.0	F 5'-CTAGTTTCCCATTTACCA-3' R 5'-TCCTCAGCATTTTCTTTC-3'
2.	mTcCIR11	46.0	F 5'-TTTGGTGATTATTAGCAG-3' R 5'-GATTCGATTTGATGTGAG-3'
3.	mTcCIR12	46.0	F 5'-TCTGACCCCAAACCTGTA-3' R 5'-ATTCCAGTTAAAGCACAT-3'
4.	mTcCIR18	51.0	F 5'-GATAGCTAAGGGGATTGAGGA-3' R 5'-GGTAATTCAATCATTTGAGGATA-3'
5.	mTcCIR24	46.0	F 5'-TTTGGGGTGATTTCTTCTGA-3' R 5'-TCTGTCTCGTCTTTTGGTGA-3'
6.	mTcCIR33	51.0	F 5'-TGGGTTGAAGATTTGGT-3' R 5'-CAACAATGAAAATAGGCA-3'
7.	mTcCIR40	51.0	F 5'-AATCCGACAGTCTTTAATC-3' R 5'-CCTAGGCCAGAGAATTGA-3'
8.	mTcCIR42	55.0	F 5'-TTGCTGAAGTATCTTTTGAC-3' R 5'-GCTCCACCCCTATTTG-3'
9.	SHRSTc49	58.0	F 5'-ATCGCAGCAAACTCCCTCTC-3' R 5'-TTCTCTTCCCACCAAGTCCC-3'
10.	SHRSTc51	57.0	F 5'-CTGTTTTTGCCTCCCTTGTTCT-3' R 5'-ATTGCTGGTTGTTCTCCATCCT-3'
11.	SHRSTc64	54.0	F 5'-TCCTACATTCCTGCACCC-3' R 5'-TCGAGAGAAAAGCTCTTACACT-3'

4.3.2.1 DNA fingerprinting with selected SSR primers

For the development of SSR fingerprints, amplification of DNA samples using selected 11 SSR primer sets (Table 10) were performed. Amplification pattern obtained with each primers were observed to identify the polymorphic bands. The details of amplification with the eleven selected SSR primers were as follows

4.3.2.1.1 mTcCIR 8

Amplification of DNA samples with the primer mTcCIR 8 yielded bands at 3 different lengths among the varieties studied (Plate 13a). A unique band was observed in the variety CCRP 1 and CCRP 4 at 280 bp and 420 bp respectively. Similarly, at 350 bp, polymorphic bands were obtained in hybrids CCRP 8 and CCRP 9. The amplicon observed at 300 bp was polymorphic but was present in all varieties except CCRP1 (Fig. 11a). Hence, mTcCIR 8 can be effectively used for identifying CCRP 1 and CCRP 4 from a mixed population. Similarly, the band shared between CCRP 8 and CCRP 9 at 350 bp can be a fingerprint of the hybrids.

4.3.2.1.2 mTcCIR 11

Amplification pattern of eight genotypes with SSR primer mTcCIR 11 was performed (Plate 13b). The amplicons obtained with this primer set was found polymorphic by generating a shared band between varieties CCRP 4, CCRP 5 and CCRP 6 at 280 bp. Molecular size of another amplicon was 300 bp (Fig. 11b) which was shared by all other varieties. The amplicon obtained at 280 bp was further utilized for consolidated DNA fingerprint.

4.3.2.1.3 mTcCIR 12

SSR analysis of the DNA samples using the primer mTcCIR 12 generated polymorphic amplicons among the varieties (Plate 14a). The varieties CCRP 2 and CCRP 5 were found to share polymorphic band at 250 bp. (Fig. 12a). At 200 bp, varieties CCRP 1, CCRP 2 and CCRP 4 were found to share a band among them and the band information was useful for generation of final DNA

fingerprint. Thus, for the identification of varieties CCRP 2 and CCRP 5, mTcCIR 12 can be used.

4.3.2.1.4 mTcCIR 18

Amplification pattern of eight genotypes with SSR primer mTcCIR 18 was performed (Plate 14b). The amplicons obtained with this primer set was found to be polymorphic. But the amplicon obtained was shared between more than three varieties and found undesirable for final DNA fingerprint. Molecular size of obtained amplicons were 320 bp and 350 bp (Fig. 12b).

4.3.2.1.5 mTcCIR 24

Amplification pattern of eight genotypes with SSR primer mTcCIR 24 was performed (Plate 15a). The amplicons obtained with this primer set was found polymorphic. The amplicons obtained at 200 bp were found to be shared among varieties CCRP 6, CCRP 8 and CCRP 9 (Fig. 13a). This particular amplicon information was incorporated to final DNA fingerprint.

4.3.2.1.6 mTcCIR 33

Amplification of DNA samples with the primer mTcCIR 33 yielded bands at four different molecular size among the varieties studied (Plate 15b). A unique polymorphic band was observed in the variety CCRP 2 at 320 bp which was a specific fingerprint of the variety. Similarly, at 340 bp, polymorphic bands were shared by varieties CCRP 4, and CCRP 5. Another amplicon observed at 280 bp was found to be present in varieties CCRP 1, CCRP 2 and CCRP 5 (Fig. 13b). Therefore, the primer mTcCIR 33 can be specifically used to identify CCRP 2, CCRP 4 and CCRP 5. Such useful amplicon information were utilized for development of final DNA fingerprint.

4.3.2.1.7 mTcCIR 40

SSR analysis of the DNA samples using the primer mTcCIR 40 generated amplicons at two different lengths (Plate 16a). The obtained amplicons were found to be shared between more than three varieties and couldn't consider for the

final DNA fingerprint (Fig. 14a). Amplicons were generated at 260 bp and 280 bp.

4.3.2.1.8 mTcCIR 42

SSR analysis of the DNA samples using the primer mTcCIR 42 generated amplicons at three different locations (Plate 16b). Polymorphic bands observed at 250 bp were present in all the varieties studied except CCRP 9. And in genotype CCRP 4, a unique amplicon was observed at 200 bp. Similarly, in hybrid CCRP 9, unique amplicon was generated at 220 bp (Fig. 14b) which implied the ability of the primer to distinguish respective varieties.

4.3.2.1.9 SHRSTc 49

Out of the four amplicons developed using the primer SHRSTc 49 (Plate 17a), polymorphic bands obtained at 260 bp were shared between varieties CCRP 1, CCRP 4 and CCRP 7. The hybrids CCRP 8 and CCRP 9 were found to share polymorphic amplicon at 100 bp (Fig. 15a) and found to be the DNA fingerprint of respective varieties. Another band obtained at 250 bp was found to be shared between five varieties. Useful amplicon pattern were incorporated into final DNA fingerprint.

4.3.2.1.10 SHRSTc 51

Amplification of DNA samples with the primer SHRSTc 51 yielded bands at 3 different lengths among the varieties studied (Plate 17b). Polymorphic bands observed in the varieties CCRP 5, CCRP 7 and CCRP 8 at 180 bp were useful polymorphic amplicons. Another bands obtained at 200 bp and 380 bp were shared by more than three varieties and couldn't consider for final DNA fingerprint (Fig. 15b). Thus, the primer SHRSTc 51 can be used to identify the varieties CCRP 5, CCRP 7 and CCRP 8.

4.3.2.1.11 SHRSTc 64

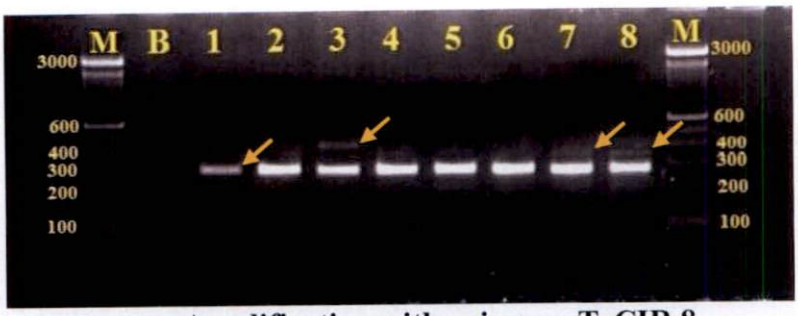
Amplification of DNA samples with the primer SHRSTc 64 yielded bands at 2 different lengths among the varieties studied (Plate 18). Polymorphic band

was observed at 450 bp in the varieties CCRP 1, CCRP 2 and CCRP 9. Another monomorphic band was observed at 300 bp in all the genotypes (Fig. 16). Therefore the amplicon generated at 450 bp was useful polymorphism and considered for final DNA fingerprint.

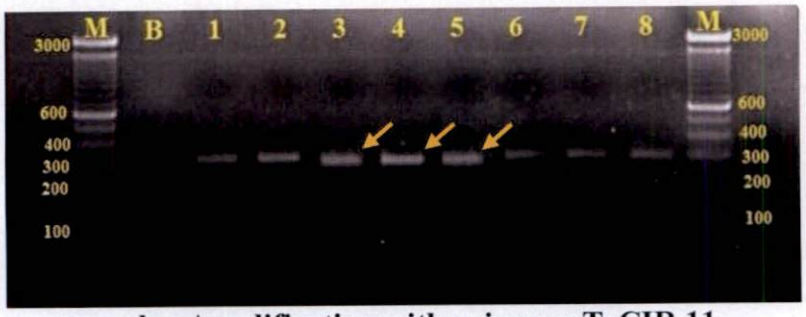
In the present study, out of eleven primers selected, three of them generated unique bands which can act as variety specific DNA fingerprints. Except primers mTcCIR 18 and mTcCIR 40, all other selected primers yielded at least single distinct polymorphic bands. In CCRP 4, two unique bands were generated by primers mTcCIR 42 (200 bp) and mTcCIR 8 (420 bp) which were specific DNA fingerprint of the particular variety. Primer mTcCIR 42 generated a unique amplicon in hybrid CCRP 9 at 220 bp which act as the fingerprint of the variety. In varieties CCRP 1 and CCRP 2, such polymorphic bands were obtained at 280 bp (mTcCIR 8) and 320 bp (mTcCIR 33) respectively.

Hybrids CCRP 8 and CCRP 9 developed shared bands at 350 bp (mTcCIR 8) and 100 bp (SHRSTc 49) respectively. mTcCIR 12 generated polymorphic band at 250 bp and was found only in varieties CCRP 2 and CCRP 5. Another band shared between varieties CCRP 4 and CCRP 5 were generated by primer mTcCIR 33 at 340 bp (Fig. 13 b). DNA fingerprints obtained were helped to characterize the genotypes and the data can be used to detect infringement of breeder's right.

A number of scientist had exploited the potential of SSR primers in cocoa (Lanaud *et al.*, 1999; Pugh *et al.*, 2004; Saunders *et al.*, 2004; Cryer *et al.*, 2006; Aikpokpodion *et al.*, 2009). SSRs as they are codominant markers, shows high level of polymorphism than other markers and they are widely used in various genetic studies of cocoa including QTL mapping and genetic map construction (Brown *et al.*, 2005; Faleiro *et al.*, 2006). DNA fingerprinting of potato (McGregor *et al.*, 2000) and tea (Tan *et al.*, 2015) were also performed using SSR marker analysis.



a. Amplification with primer mTcCIR 8



b. Amplification with primer mTcCIR 11

M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 13. Amplification pattern of eight cocoa varieties using primers mTcCIR 8 and mTcCIR 11

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
420								
350								
300								
280								

a. Colour chart - mTcCIR 8

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
300								
280								

b. Colour chart - mTcCIR 11

Colour code for sharing of bands among varieties

1 2 3 4 5 6 7 8

Fig. 11 Colour chart developed using mTcCIR 8 and mTcCIR 11

a. Amplification with primer mTcCIR 12

b. Amplification with primer mTcCIR 18

M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 14. Amplification pattern of eight cocoa varieties using primers mTcCIR 12 and mTcCIR 18

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
250								
200								
180								

a. Colour chart - mTcCIR 12

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
350								
320								

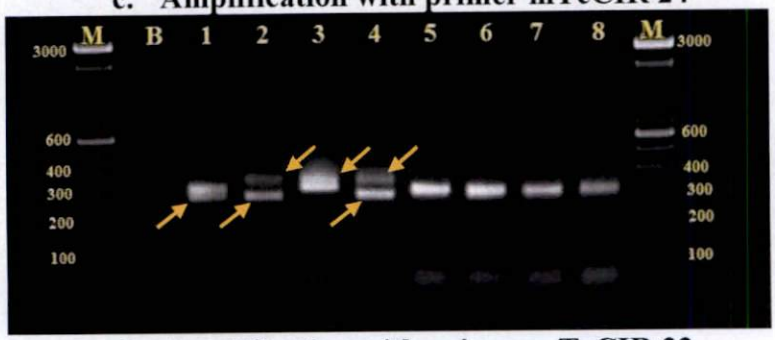
b. Colour chart - mTcCIR 18

Colour code for sharing of bands among varieties

1	2	3	4	5	6	7	8

Fig. 12 Colour chart developed using mTcCIR 12 and mTcCIR 18

c. Amplification with primer mTcCIR 24



d. Amplification with primer mTcCIR 33

M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 15. Amplification pattern of eight cocoa varieties using primers mTcCIR 24 and mTcCIR 33

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP1	CCRP2	CCRP4	CCRP5	CCRP6	CCRP7	CCRP8	CCRP9
200								
180								

a. Colour chart - mTcCIR 24

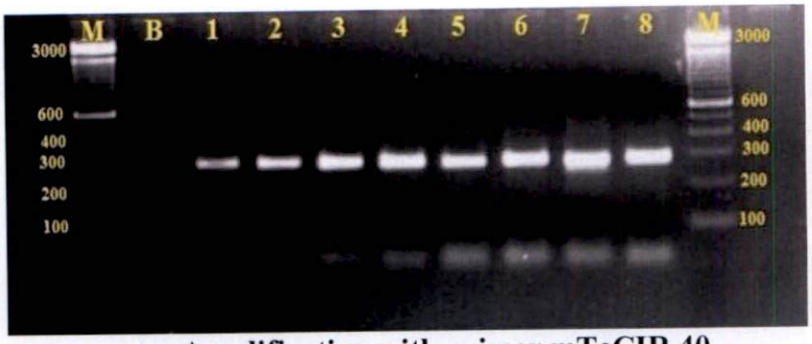
	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP1	CCRP2	CCRP4	CCRP5	CCRP6	CCRP7	CCRP8	CCRP9
340								
320								
300								
280								

c. Colour chart - mTcCIR 33

Colour code for sharing of bands among genotype

1	2	3	4	5	6	7	8

Fig. 13 Colour chart developed using mTcCIR 24 and mTcCIR 33



a. Amplification with primer mTcCIR 40

b. Amplification with primer mTcCIR 42

M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 16. Amplification pattern of eight cocoa varieties using primers mTcCIR 40 and mTcCIR 42

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
280								
260								

a. Colour chart - mTcCIR 40

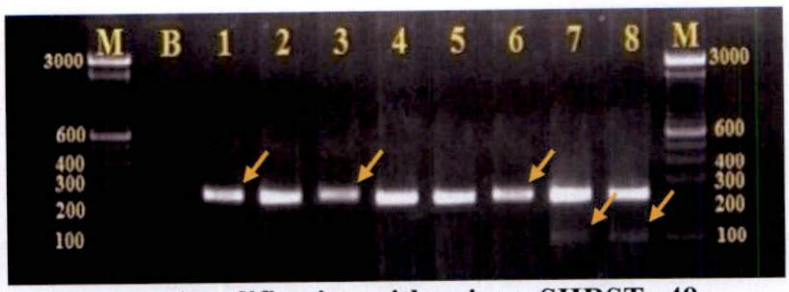
	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
250								
220								
200								

b. Colour chart - mTcCIR 42

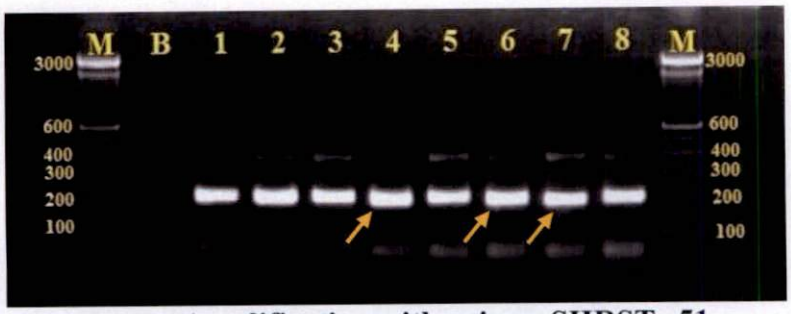
Colour code for sharing of bands among genotype

1	2	3	4	5	6	7	8

Fig. 14 Colour chart developed using mTcCIR 40 and mTcCIR 42



a. Amplification with primer SHRSTc 49



a. Amplification with primer SHRSTc 51

M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 17. Amplification pattern of eight cocoa varieties using primers SHRSTc 49 and SHRSTc 51

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
260								
250								
100								

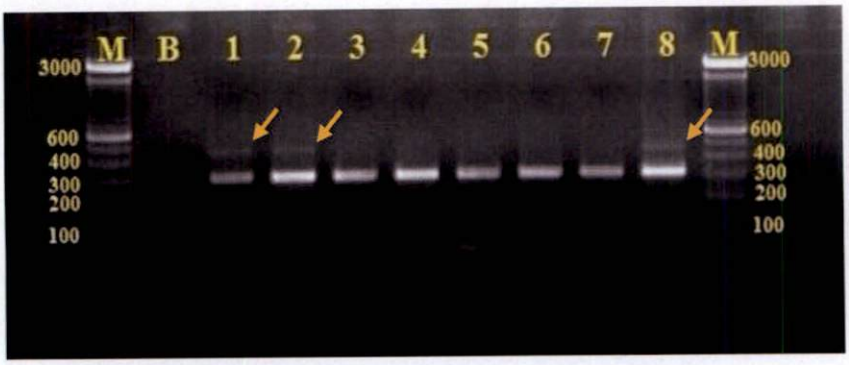
a. Colour chart - SHRSTc 49

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
380								
200								
180								

b. Colour chart - SHRSTc 51

Colour code for sharing of bands among genotype

Fig. 15 Colour chart developed using SHRSTc 49 and SHRSTc 51



M- Molecular weight marker (3kb), B- Blank, 1-CCRP 1, 2- CCRP 2, 3- CCRP 4, 4- CCRP 5, 5- CCRP 6, 6- CCRP 7, 7- CCRP 8, 8- CCRP 9

Plate 18. Amplification with primer SHRSTc 64

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
450								
300								

Colour code for sharing of bands among genotype

Fig. 16 Colour chart developed using SHRSTc 64

4.4 DNA fingerprinting of individual cocoa varieties

Individual primer data obtained using selected 10 ISSR primers were further utilized for locating useful polymorphic amplicons in each variety. The consolidated polymorphic amplicon colour chart was developed. Similarly for SSR DNA fingerprint data, the individual primer set data generated using 11 selected primer sets were used. Colour charts thus developed for individual varieties in such a way were utilized for further interpretation. To give polymorphism coverage to all eight varieties, amplicons which were distinct and shared only among maximum of three varieties were identified and selected. Variety wise DNA fingerprint details are described below.

4.4.1 CCRP 1

4.4.1.1 ISSR profile

Using selected 10 ISSR primers, 78 distinct loci were observed in CCRP 1. The obtained amplicon sizes ranged between 250 bp to 1950 bp (Fig. 17a). Primers UBC 835 and UBC 866 yielded only monomorphic bands. A maximum of 13 bands were observed with the primer UBC 842 and for UBC 846, it was only 3. Total number of polymorphic amplicons were 19 and the amplicon developed using UBC 842 at 550 bp was found useful and considered for final DNA fingerprint of the variety.

4.4.1.2 SSR profile

Selected 11 SSR primer sets were amplified the DNA and yielded 14 amplicons with the genotype CCRP 1. Obtained amplicon size ranged from 180 bp to 450 bp. Total number of polymorphic amplicons were 13 of which, five were useful polymorphism (Fig. 17b). Among the five polymorphic amplicons, the unique band which obtained at 280 bp using primer mTcCIR 8 was found to be the specific fingerprint of the variety. Other useful polymorphic amplicons also can be utilized to identify the genotypes.

4.4.2 CCRP 2

4.4.2.1 ISSR profile

In genotype CCRP 2, amplification of DNA using 11 selected SSR primer sets yielded a total of 84 amplicons of which, 25 were polymorphic. The obtained amplicon size ranged from 250 bp to 1900 bp (Fig. 18a). Primers UBC 835 and UBC 841 produced 11 amplicons each and the least number of 4 bands were developed by primer UBC 846. Polymorphic amplicons developed by UBC 815 (600 bp), UBC 841 (700 and 850 bp) UBC 842 (550 bp) and UBC 855 (1700 bp) were found to be useful and utilized for final fingerprint data.

4.4.2.2 SSR profile

SSR analysis of genotype CCRP 2 using selected 11 primer sets formed 15 amplicons of which, 14 were polymorphic (Fig. 18b). Molecular size of the obtained amplicons ranged from 180 bp to 450 bp. Among the 14 polymorphic amplicons, five were found useful and chosen for development of final SSR fingerprint. Primer mTcCIR 33 generated a unique amplicon at 320 bp which can act as the specific DNA fingerprint of the variety. Similarly at 200 bp (mTcCIR 12), 250 bp (mTcCIR 12), 280 bp (mTcCIR 33) and at 450 bp (SHRSTc 64) useful polymorphic amplicons were generated. Thus the useful amplicon information were incorporated in to the final consolidated DNA fingerprint data.

4.4.3 CCRP 4

4.4.3.1 ISSR profile

Selected 10 ISSR primers together generated 84 amplicons with molecular size ranged from 250 bp to 1950 bp (Fig. 19a). Among 84 amplicons generated, 25 were polymorphic of which, two were useful polymorphism. Maximum number of amplicons were generated by UBC 842 with 14 bands and least was 3 amplicons which is produced by UBC 846. Primer UBC 866 yielded monomorphic bands alone. The distinct polymorphic band obtained by UBC 815

at 600 bp and UBC 841 at 700 bp were chosen and utilized for the development of final DNA fingerprint.

4.4.3.2 SSR profile

Analysis of CCRP 4 using selected 11 SSR primers yielded amplicons with molecular size ranged between 180 bp to 420 bp. Total number of amplicons generated were 15 and of them, six were useful polymorphism (Fig. 19b). Primer mTcCIR 8 yielded a polymorphic band at 420 bp which was a specific fingerprint of the variety. Similarly, mTcCIR 42 generated another fingerprint by developing a unique amplicon at 200 bp. Other useful amplicons were obtained at 200 bp (mTcCIR 12), 260 bp (SHRSTc 49), 280 bp (mTcCIR 11) and at 340 bp (mTcCIR 33). Thus, six such useful amplicons were chosen for the construction of final SSR fingerprint of variety CCRP 4.

4.4.4 CCRP 5

4.4.4.1 ISSR profile

ISSR analysis of genotype CCRP 5 using selected 10 primer generated 90 amplicons of which, 31 were polymorphic (Fig. 20a). Nine of 31 polymorphic amplicons were found with useful polymorphism. The obtained amplicon sizes ranged between 250 bp to 1950 bp. Maximum number of amplicons were generated by UBC 842 with 14 bands and least was 6 amplicons which is produced by UBC 866. Primer UBC 866 yielded monomorphic bands alone. Amplicons generated using primers UBC 815 (600 bp), UBC 841 (650 bp, 700 bp, 850 bp), UBC 842 (550 bp) and UBC 846 (300 bp, 500 bp, 1300 bp, 1600 bp) were found distinct and polymorphic. Thus, those nine distinct polymorphic amplicons were chosen for the construction of final ISSR fingerprint of CCRP 5. Primers UBC 841 (650 bp) and UBC 846 (300 bp, 1300 bp and 1600 bp) generated specific DNA fingerprints of the particular variety by generating four such unique amplicons.

4.4.4.2 SSR profile

SSR analysis of genotype CCRP 5 using 11 selected primers formed 13 amplicons of which, 12 were polymorphic (Fig. 20b). Molecular size of obtained amplicons ranged from 180 bp to 340 bp. Among the 12 polymorphic amplicons, five were found useful and chosen for development of final SSR fingerprint. Such five amplicons were generated by primers mTcCIR 11 (280 bp), mTcCIR 12 (250 bp), mTcCIR 33 (280 bp and 340 bp) and SHRSTc 51 (180 bp). Thus details of those five useful amplicons were utilized for the construction of final DNA fingerprint.

4.4.5 CCRP 6

4.4.5.1 ISSR profile

In CCRP 6, selected 10 ISSR primers generated 83 amplicons with molecular size ranged from 250 bp to 1950 bp (Fig. 21a). Among 83 amplicons, 23 were polymorphic of which, four were useful polymorphic amplicons. Maximum number of amplicons were generated by UBC 842 with 13 bands and least was 4 amplicons produced by primers UBC 846 and UBC 855 respectively. Primers UBC 855 and UBC 866 yielded monomorphic bands alone. Amplicons generated using primers UBC 810 (850 bp), UBC 826 (1100 bp) were unique and thus recorded as the DNA fingerprint of the variety CCRP 6. Other bands obtained using primers UBC 827 (400 bp) and UBC 841 (850 bp) were also selected for development of ISSR fingerprint.

4.4.5.2 SSR profile

Analysis of CCRP 6 using selected 11 SSR primers yielded total of twelve amplicons with molecular size ranged between 180 bp to 380 bp. Total number of polymorphic amplicons generated were 11 and of them, two were useful polymorphism (Fig. 21b). The useful polymorphic amplicons were generated by primers mTcCIR 11 and mTcCIR 24 at 280 bp and 200 bp respectively.

4.4.6 CCRP 7

4.4.6.1 ISSR profile

Analysis of genotype CCRP 7 using 10 selected ISSR primers generated 84 amplicons of which, 25 were polymorphic (Fig. 22a). Out of 25 polymorphic amplicons, only one band was useful. The obtained amplicon sizes ranged from 250 bp to 1950 bp. Maximum numbers of amplicons were generated by UBC 842 with 14 bands and least was 4 amplicons which was produced by UBC 846. Primer UBC 866 yielded monomorphic bands alone. Amplicons generated using primers UBC 827 (400 bp) was found distinct and polymorphic. Thus, the amplicon was utilized for constructing final ISSR fingerprint.

4.4.6.2 SSR profile

SSR analysis of genotype CCRP 7 using 11 selected primers generated 12 amplicons of which, 11 were polymorphic (Fig. 22b). Molecular size of the obtained amplicons ranged from 180 bp to 380 bp. Primers SHRSTc 49 (260 bp) and SHRSTc 51 (180 bp) generated a useful polymorphic amplicons at respective molecular size and the band information were incorporated in to the final SSR fingerprint.

4.4.7 CCRP 8

4.4.7.1 ISSR profile

ISSR analysis of hybrid CCRP 8 using selected 10 primer generated 89 amplicons of which, 30 were polymorphic (Fig. 23a). The obtained amplicon sizes ranged between 250 bp to 1950 bp. Maximum number of amplicons were generated by UBC 835 with 15 bands and least was 5 amplicons produced by UBC 846. Amplicons generated using primers UBC 835 (550 bp, 1400 bp, 1600 bp), UBC 846 (500 bp), UBC 855 (1700 bp) and UBC 866 (800 and 1200 bp) were found distinct and polymorphic. Thus, those seven polymorphic amplicons were utilized for final ISSR fingerprint.

4.4.7.2 SSR profile

Analysis of CCRP 8 using selected 11 SSR primers generated amplicons with molecular size ranged between 100 bp to 380 bp. Total number of amplicons

generated were 14 and of them, 13 were polymorphic (Fig. 23b). Useful polymorphic bands were four. Amplicons generated using primer sets mTcCIR 8 (350 bp), mTcCIR 24 (200 bp), SHRSTc 49 (100 bp) and SHRSTc 51 (180 bp) were found distinct and polymorphic. Thus, those four band information were utilized for constructing final SSR fingerprint. The amplicons obtained at 100 bp (SHRSTc 49) and 350 bp (mTcCIR 8) were shared between two varieties and could be used for differentiating the genotypes from others.

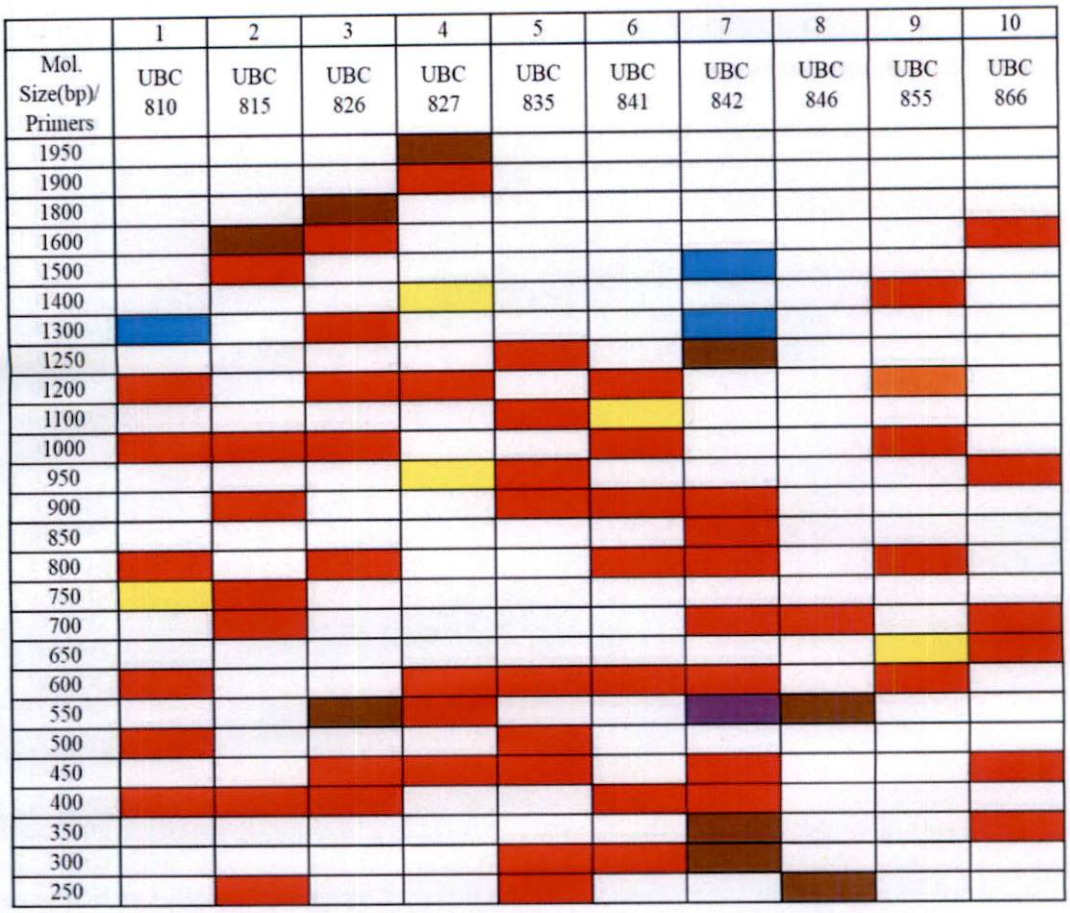
4.4.8 CCRP 9

4.4.8.1 ISSR profile

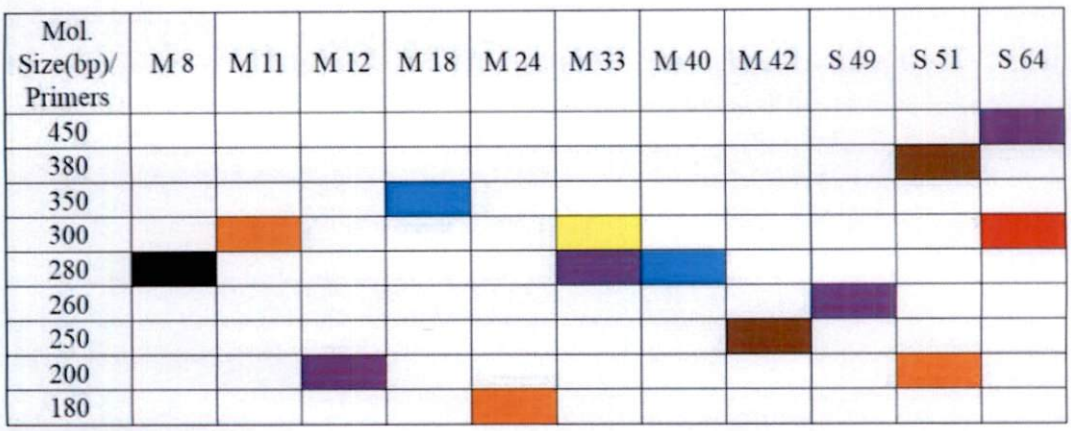
ISSR analysis of hybrid CCRP 9 using selected 10 primer generated 87 amplicons of which, 28 were polymorphic (Fig. 24a). Genome of the hybrid CCRP 9 generated ten useful polymorphic amplicons which was the highest. The obtained amplicon sizes ranged between 250 bp to 1950 bp. Maximum number of amplicons were generated by UBC 835 with 15 bands and least number of 2 amplicons were produced by UBC 846. Unique DNA fingerprints were generated by primers UBC 827 (750 bp), UBC 846 (350 bp) and UBC 866 (600 bp) which were the specific identity of the hybrid. Amplicons generated using primers UBC 827 (400 bp), UBC 835 (550 bp, 1400 bp, 1600 bp), UBC 855 (1700 bp) and UBC 866 (800 bp and 1200 bp) were also found to be distinct and polymorphic. Thus, those ten polymorphic amplicons were utilized for final ISSR fingerprint.

4.4.8.2 SSR profile

SSR analysis of hybrid CCRP 9 using 11 selected primer sets generated 15 amplicons of which, 14 were polymorphic (Fig. 24b). Molecular size of the obtained amplicons ranged from 100 bp to 450 bp. Amplicons generated using primers mTcCIR 8 (350 bp), mTcCIR 24 (200 bp), SHRSTc 49 (100 bp) and SHRSTc 64 (450 bp) were found distinct and polymorphic. Unique amplicon generated at 220 bp using the primer mTcCIR 42 was a specific fingerprint of the hybrid. Thus, those five amplicons were further utilized for constructing final SSR fingerprint.



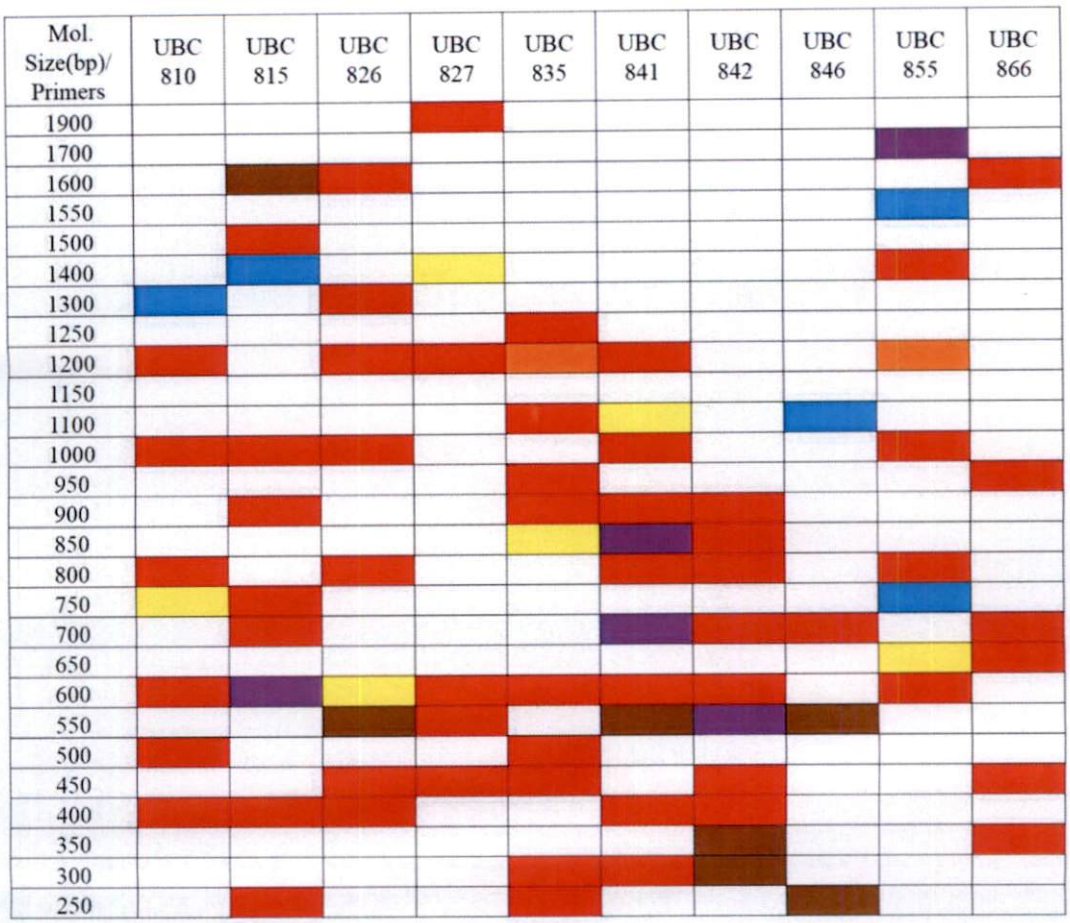
a. ISSR fingerprint



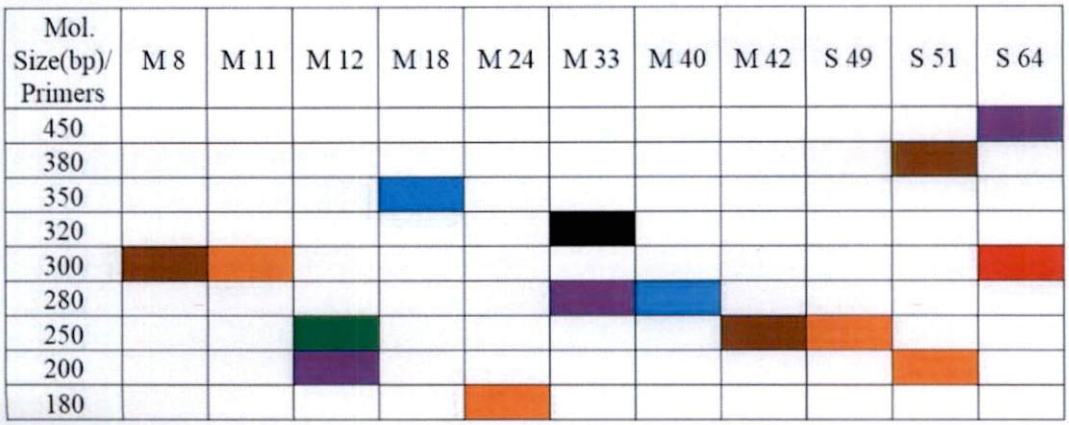
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 17 Fingerprint developed using selected ISSR and SSR primers - CCRP1



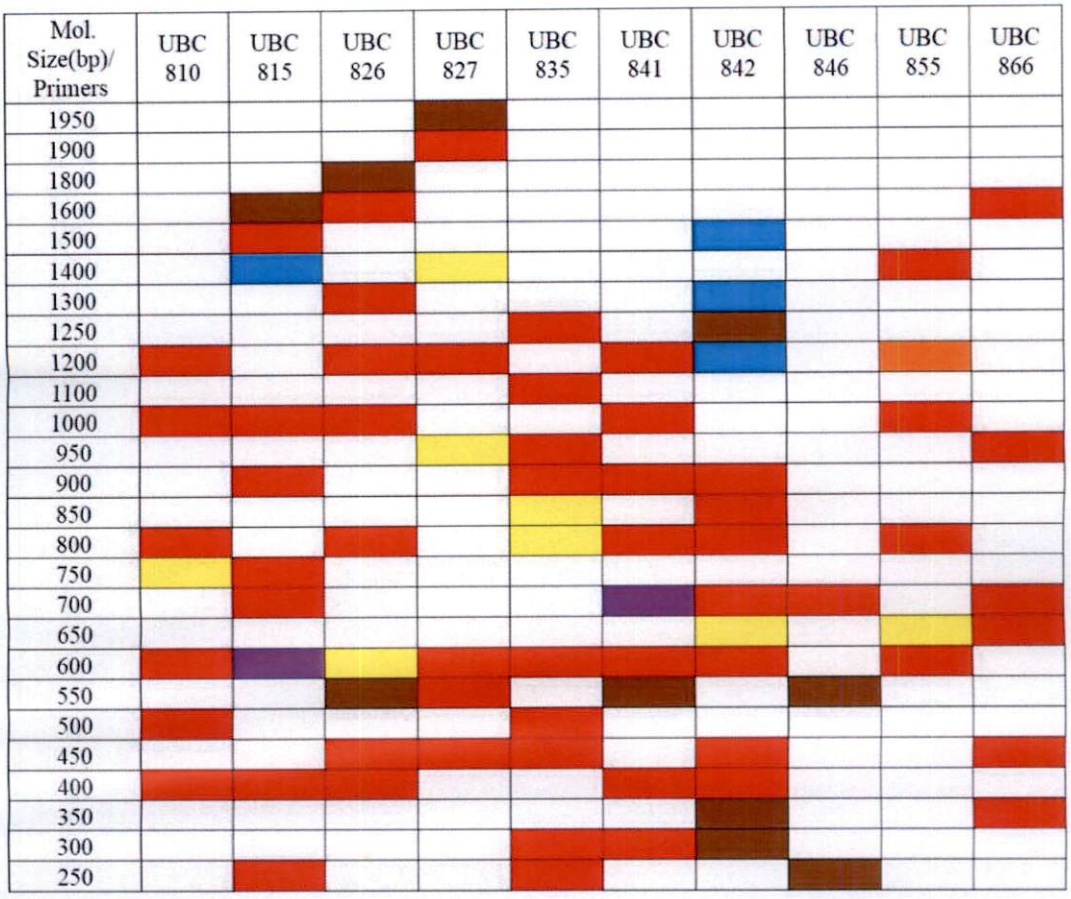
a. ISSR fingerprint



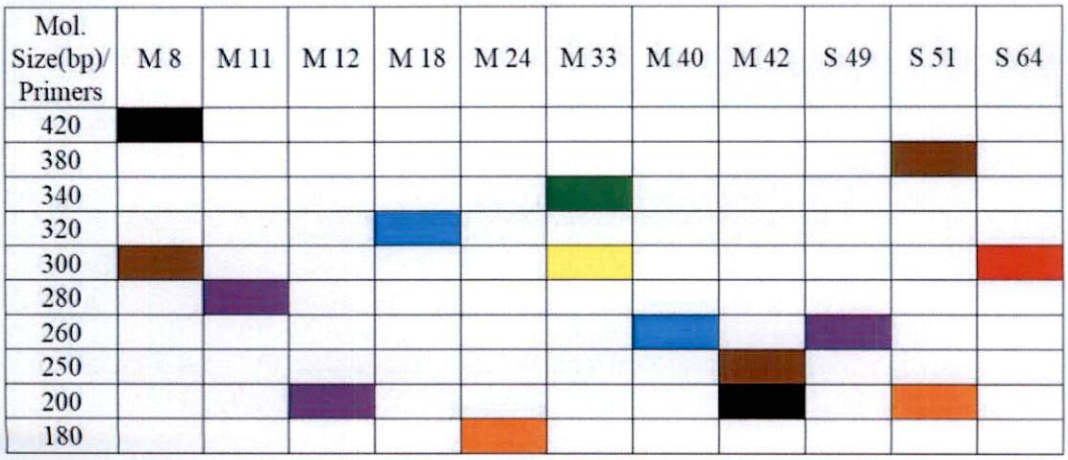
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 18 Fingerprint developed using selected ISSR and SSR primers - CCRP2



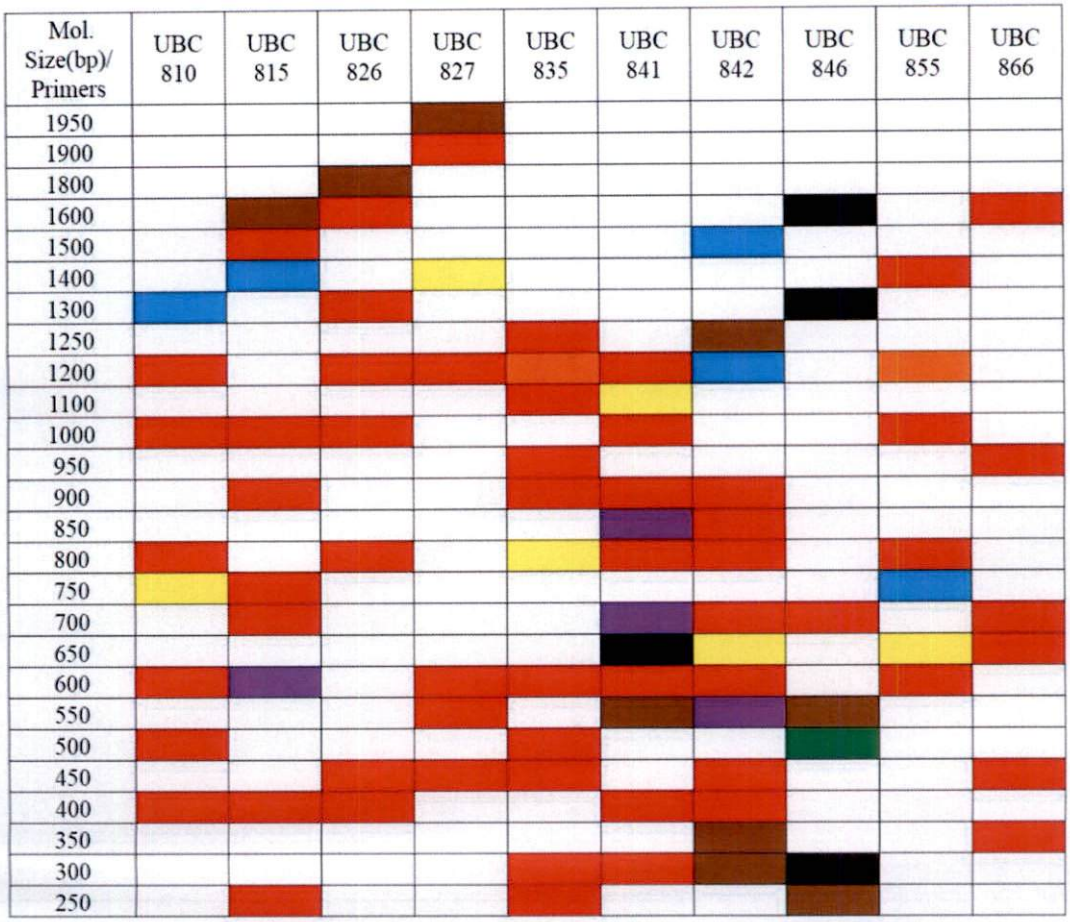
a. ISSR fingerprint



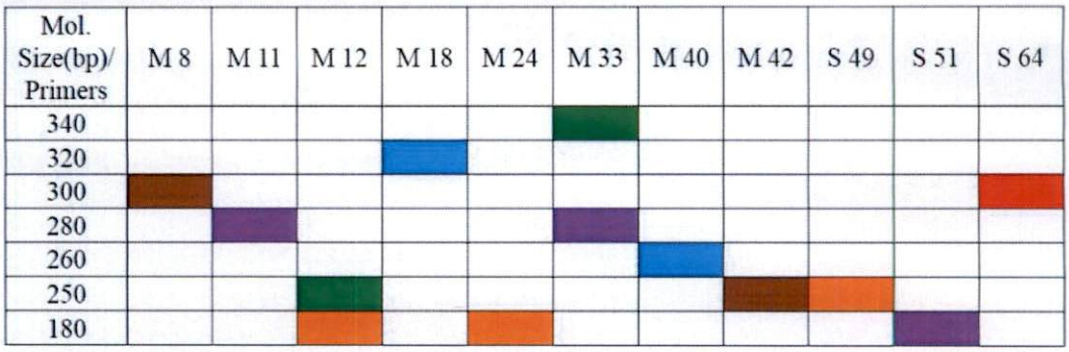
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 19 Fingerprint developed using selected ISSR and SSR primers - CCRP4



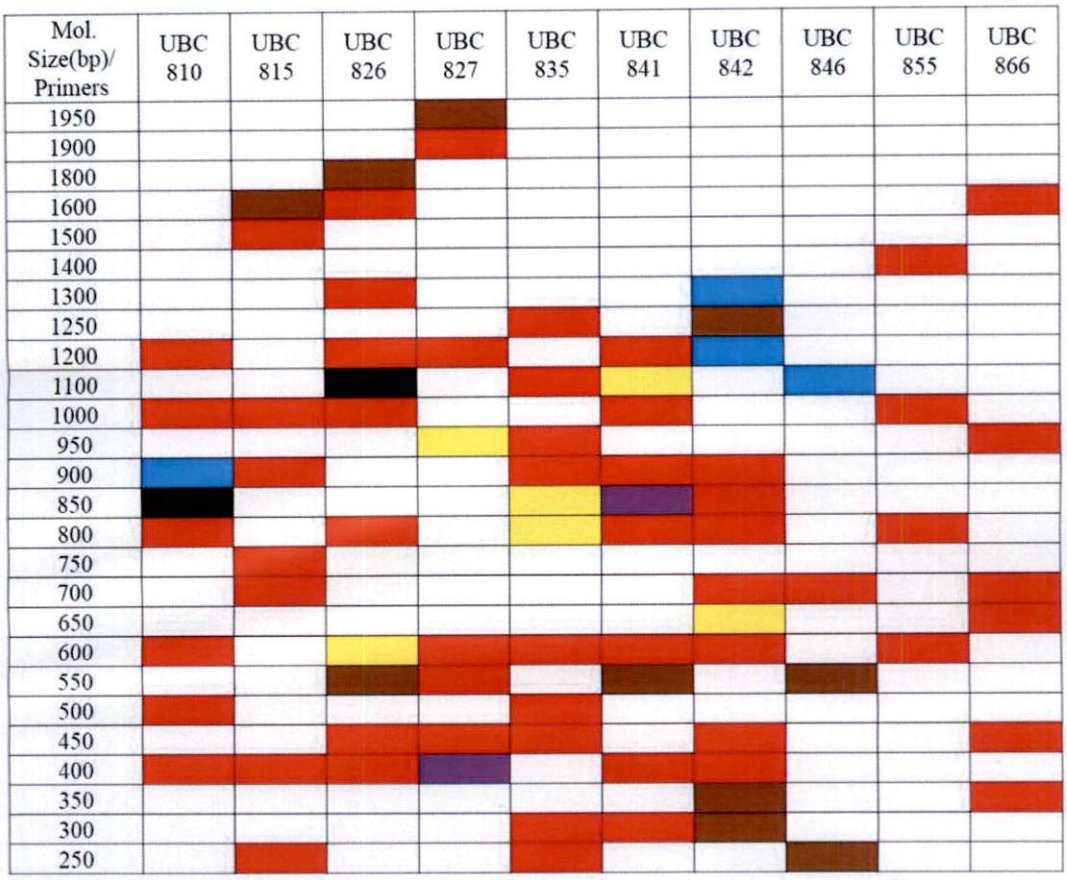
a. ISSR fingerprint



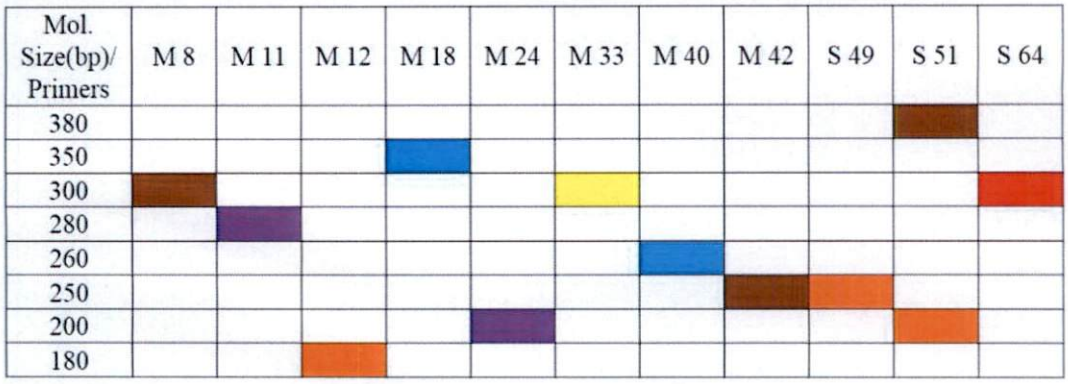
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 20 Fingerprint developed using selected ISSR and SSR primers - CCRP5



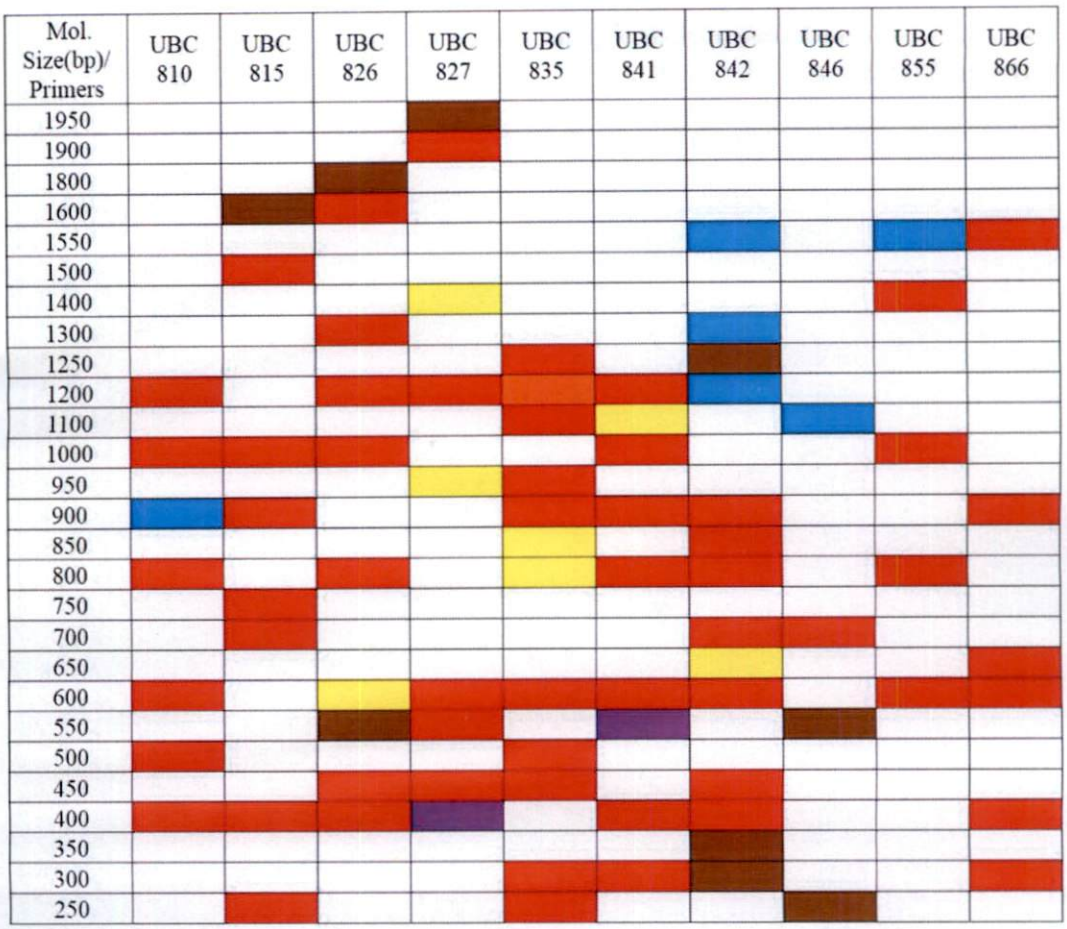
a. ISSR fingerprint



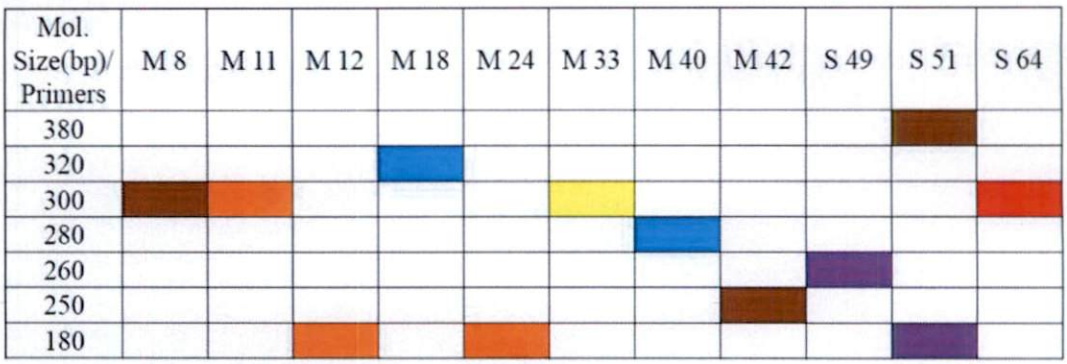
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 21 Fingerprint developed using selected ISSR and SSR primers - CCRP6



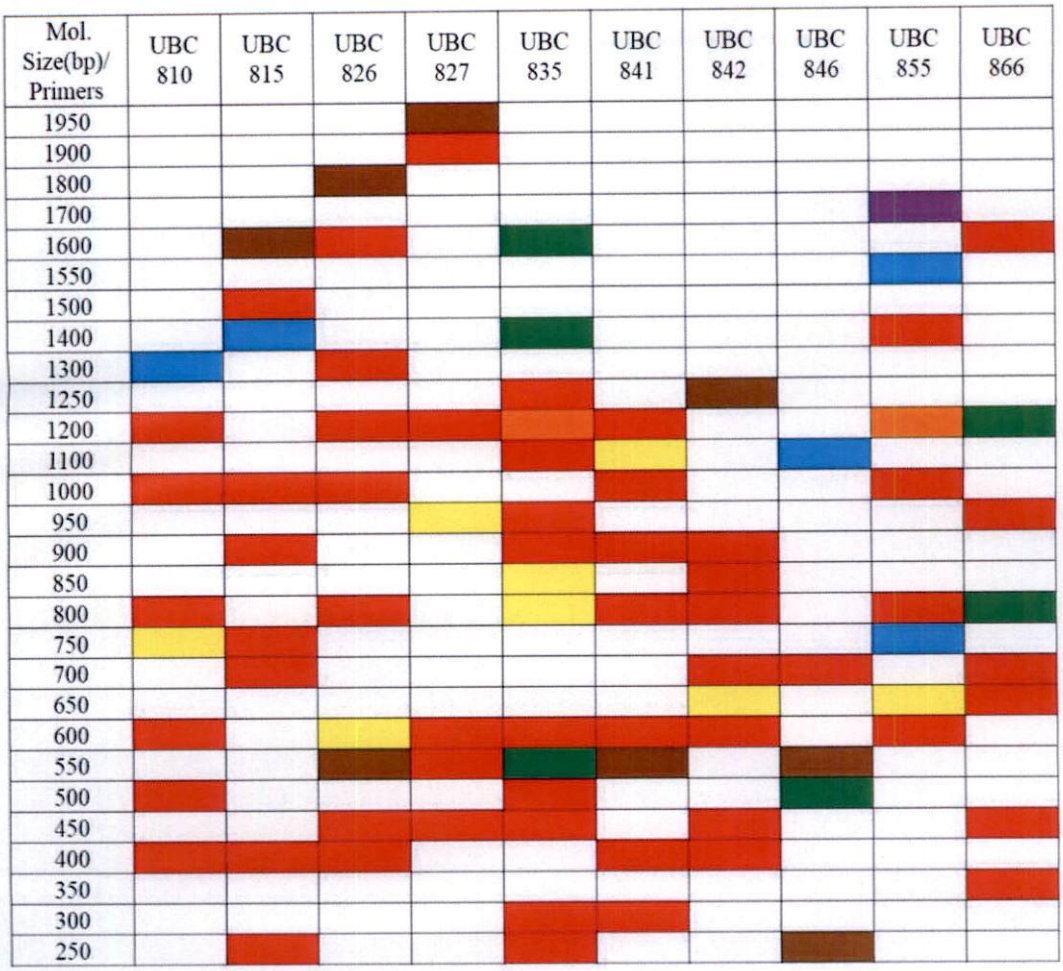
a. ISSR fingerprint



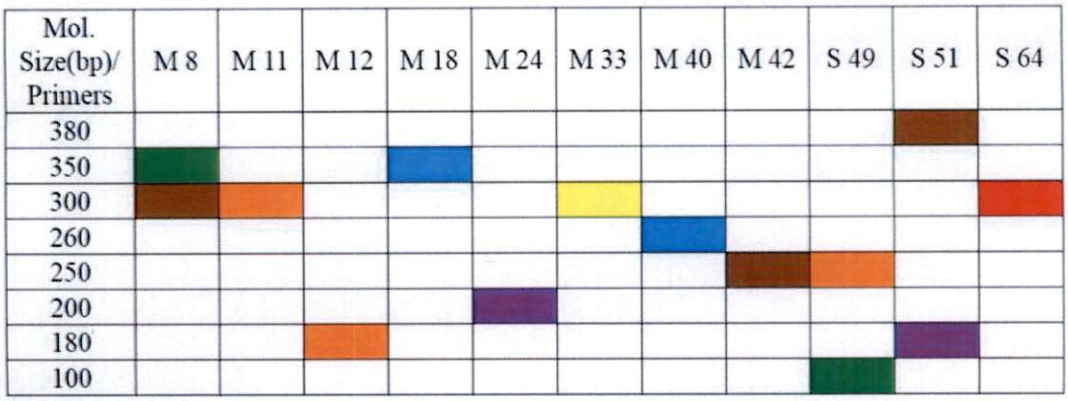
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 22 Fingerprint developed using selected ISSR and SSR primers - CCRP7



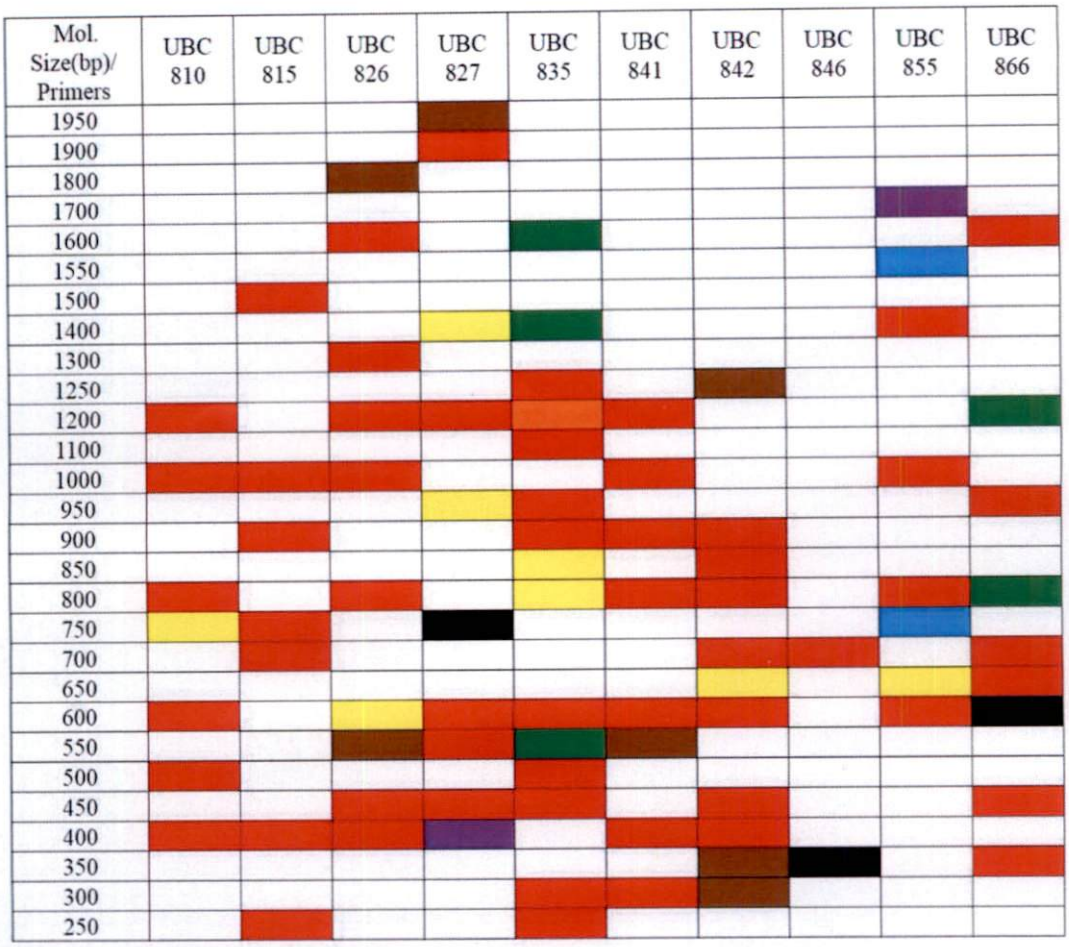
a. ISSR fingerprint



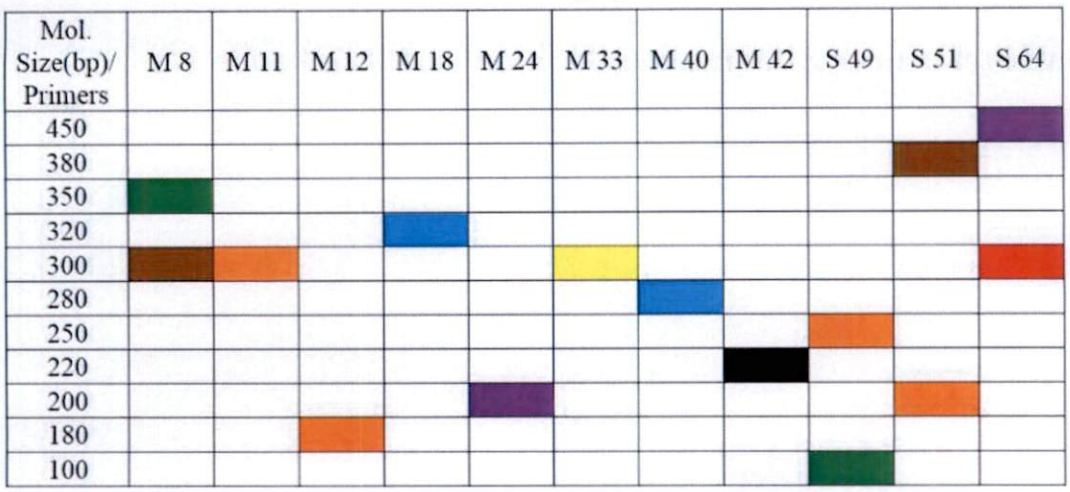
b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 23 Fingerprint developed using selected ISSR and SSR primers - CCRP8



a. ISSR fingerprint



b. SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 24 Fingerprint developed using selected ISSR and SSR primers - CCRP9

4.5 Overall DNA fingerprinting data

4.5.1 ISSR fingerprint

Amplification pattern of the selected 10 ISSR primers were further analysed and located useful distinct polymorphic amplicons using colour chart. All the 10 selected primers generated at least one useful polymorphic amplicons. Amplicons which were found only in maximum of three genotypes were taken in consideration for the development of final ISSR fingerprint chart (Fig. 25a). Using the 10 ISSR primers, 39 such polymorphic amplicons were obtained over eight genotypes studied. Molecular size of the ISSR fingerprints ranged from 300 bp to 1700 bp.

Maximum number of ten fingerprints were observed in hybrid CCRP 9, whereas, CCRP 1 and CCRP 7 were found to have only one such amplicons. Nine among 39 polymorphic amplicons were unique, six of them were shared by two varieties and bands which shared by three varieties were six in number. Maximum number of seven fingerprints were generated by the primer UBC 841. Whereas, UBC 810 and UBC 826 generated only one such amplicons. The result thus obtained is highly useful for the identification of popular cocoa varieties studied and this information can be effectively utilized to protect our IPR.

4.5.2 SSR fingerprint

Consolidated SSR fingerprint generated among eight varieties using 11 SSR primer set is depicted in the form of colour chart (Fig. 25b). Out of the eleven selected primers, nine primers generated at least one useful polymorphic amplicons among the studied varieties. Distinct amplicons which were found in maximum of three varieties were chosen. In that way, 34 such polymorphic amplicons were identified. The molecular size of DNA fingerprints thus generated ranged between 100 bp and 450 bp.

Maximum number of six fingerprints were identified in variety CCPR 4 followed by CCRP 1, CCRP 2, CCRP 5 and CCRP 9 with five fingerprints each. Five among 24 polymorphic amplicons were unique, four of them were shared by

two varieties and bands which shared by three varieties were seven in number. Unique DNA fingerprints generated by the primers mTcCIR 8 at 280 bp (CCRP 1) and 420 bp (CCRP 4) were specific fingerprint of the respective varieties. Primer mTcCIR 33 generated fingerprint at 320 bp in CCRP 2. Primer mTcCIR 42 generated unique fingerprints for varieties CCRP 4 (200 bp) and CCRP 9 (220 bp). The SSR result thus obtained is highly useful for the identification of studied varieties and this information can be effectively utilized to protect our IPR.

The current study has helped to characterize the promising cocoa varieties and the data generated was found useful for the variety registration, detection of infringement of breeder's right and biopiracy. An ISSR based DNA fingerprinting study was conducted on thirty four major cultivar strains of *Auricularia auricula* in China (Li-Hua *et al.,* 2008). The study concluded with the positive remarks of ISSR markers on DNA fingerprinting of the studied cultivars. Similarly, SSR primer sets were made use of for developing DNA fingerprints of 96 soybean cultivars (Rongwen *et al.,* 1995). The study implies the ability of the marker system to differentiate among cultivars. In case of cocoa DNA fingerprinting, a set of 15 internationally accepted microsatellite primers were identified (Saunders *et al.,* 2004).

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
1700		UBC 855					UBC 855	UBC 855
1600				UBC 846			UBC 835	UBC 835
1400							UBC 835	UBC 835
1300				UBC 846				
1200							UBC 866	UBC 866
1100					UBC 826			
850		UBC 841		UBC 841	UBC 841 UBC 810			
800							UBC 866	UBC 866
750								UBC 827
700		UBC 841	UBC 841	UBC 841				
650				UBC 841				
600		UBC 815	UBC 815	UBC 815				UBC 866
550	UBC 842	UBC 842		UBC 842			UBC 835	UBC 835
500				UBC 846			UBC 846	
400					UBC 827	UBC 827		UBC 827
350								UBC 846
300				UBC 846				

a. Final ISSR fingerprint

	1	2	3	4	5	6	7	8
Mol. Size(bp)/ Varieties	CCRP 1	CCRP 2	CCRP 4	CCRP 5	CCRP 6	CCRP 7	CCRP 8	CCRP 9
450	S 64	S 64						S 64
420			M 8					
350							M 8	M 8
340			M 33	M 33				
320		M 33						
280	M 8 M 33	M 33	M 11	M 11 M 33	M 11			
260	S 49		S 49			S 49		
250		M 12		M 12				
220								M 42
200	M 12	M 12	M 12 M 42		M 24		M 24	M 24
180				S 51		S 51	S 51	
100							S 49	S 49

b. Final SSR fingerprint

Colour code for sharing of bands among genotype

Fig. 25 Overall DNA fingerprinting data

4.6 Polymorphic information content of selected ISSR and SSR primers

The polymorphic information content (PIC) value for ten ISSR and eleven SSR primers were calculated for analysing the suitability of primers to yield polymorphic amplicons (Fig. 26a). Calculated PIC value for selected ten ISSR primers were ranged from 0.17 (UBC 826) to 0.68 (UBC 846). Botstein *et al.* (1980) reported that, quality of a genetic marker is measured by its heterozygosity in the population of interest, and for a molecular marker, its polymorphism information content. The PIC value was used as an indicator for understanding the usefulness of a particular primer for characterizing a variety (Hollman *et al.*, 2005). Average PIC value for selected primers in this study was found to be 0.32 which indicated their usefulness (Table 11). Polymorphism percentage for selected ISSR primers ranged from 9.43 to 71.19. The highest polymorphism percentage was recorded by primer UBC 846 (71.19 per cent) and lowest was by the primer UBC 866 (9.43 per cent).

Similarly, PIC value for selected eleven SSR primers were also calculated (Fig. 26b). Out of 11 selected primer sets, all of them were found to have 100 per cent polymorphism except SHRSTc 64 (27.27 per cent). Calculated PIC value for selected eleven SSR primers were ranged from 0.42 (SHRSTc 64) to 0.80 (mTcCIR 12, mTcCIR 33 and SHRSTc 49). Average PIC value for selected primers were found to be 0.72 which indicated their usefulness (Table 12).

4.8 Resolving power of selected ISSR and SSR primers

Resolving power is the ability of an informative primer to differentiate between varieties among a studied population (Ikegami *et al.*, 2009). In the present study, resolving power of selected ten ISSR and eleven SSR primers were calculated using the formula given (3.4.1). Rp value was reported for variability analysis in many crops and the range was found between 2.59 to 12.50 (Prevost and Wilkinson, 1999). The result obtained on Rp value indicated the effective distinguishing capacity of each marker system over the studied eight varieties.

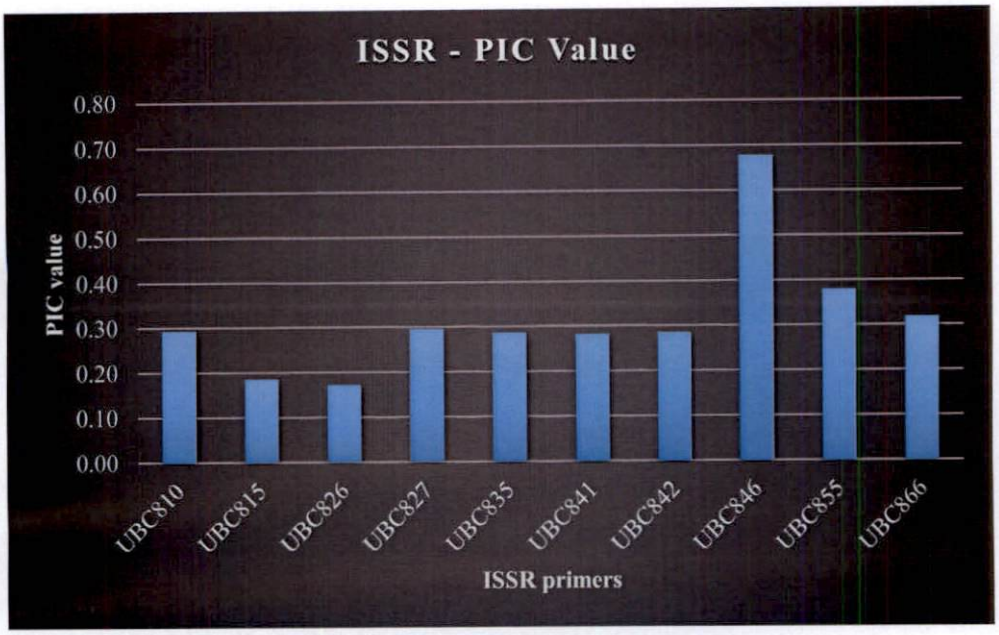
Rp value for selected ten ISSR primers were calculated (Fig. 27a). With an average value of 14.75, the Rp value was found to be ranged from 8.00 (UBC 846) to 17.50 (UBC 815) (Table 11). Similarly, Rp value for selected eleven SSR primers were also calculated and the range was found to be from 2.00 to 3.75 (Fig. 27b). Ikegami *et al.* (2009) reported that, microsatellites offer higher resolving power in genetic studies than other previous molecular markers. Highest resolving power of 3.75 was obtained by the primer SHRSTc 51 and least Rp was for the primers mTcCIR 11, mTcCIR 18, mTcCIR 24 and mTcCIR 40 with a value of 2.00 (Table 12). Average Rp value for SSR primers were found to be 2.50, indicated the usefulness of selected primer for the development of DNA fingerprint.

Table 11. Details of amplification with selected 10 ISSR primers

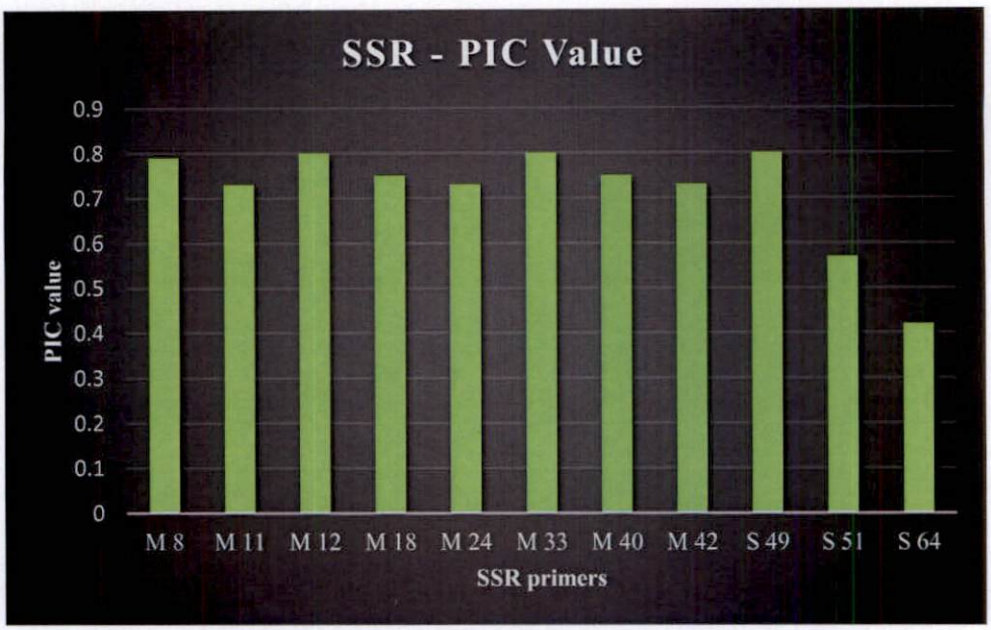
Sl. No.	ISSR Primer	Total no of amplicons	No of polymorphic amplicons	Polymorphism (%)	Polymorphic information content (PIC)	Resolving power (Rp)
1	UBC 810	63	15	23.81	0.29	15.75
2	UBC 815	70	14	20.00	0.18	17.50
3	UBC 826	77	21	27.27	0.17	17.25
4	UBC 827	63	22	34.92	0.30	15.75
5	UBC 835	95	23	24.21	0.29	15.25
6	UBC 841	76	20	26.32	0.28	15.00
7	UBC 842	98	41	41.84	0.28	16.25
8	UBC 846	31	23	74.19	0.68	8.00
9	UBC 855	54	22	40.74	0.38	13.50
10	UBC 866	53	5	9.43	0.32	13.25
Total		**680**	**206**	**322.73**	**3.17**	**147.5**
Average		**68**	**20.6**	**32.27**	**0.32**	**14.75**

Table 12. Details of amplification with selected 11 SSR primers

Sl. No.	SSR Primer	Total no of amplicons	No of polymorphic amplicons	Polymorphism (%)	Polymorphic information content (PIC)	Resolving power (Rp)
1.	mTcCIR 8	11	3	100	0.79	2.75
2.	mTcCIR 11	8	0	100	0.73	2.00
3.	mTcCIR 12	14	4	100	0.80	2.50
4.	mTcCIR 18	8	0	100	0.75	2.00
5.	mTcCIR 24	8	0	100	0.73	2.00
6.	mTcCIR 33	11	11	100	0.80	3.00
7.	mTcCIR 40	9	1	100	0.75	2.00
8.	mTcCIR 42	9	1	100	0.73	2.25
9.	SHRSTc 49	20	4	100	0.80	2.50
10.	SHRSTc 51	18	10	100	0.57	3.75
11.	SHRSTc 64	12	4	27.27	0.42	2.75
Total		**136**	**38**	**1027.27**	**7.87**	**27.50**
Average		**11.33**	**3.17**	**93.39**	**0.72**	**2.50**

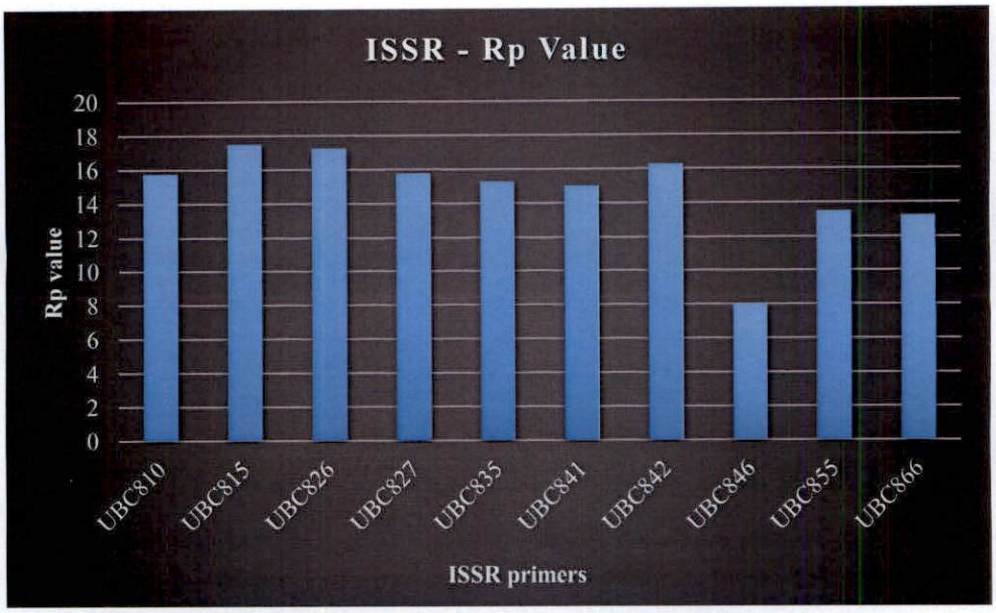


a. PIC value of 10 selected ISSR primers

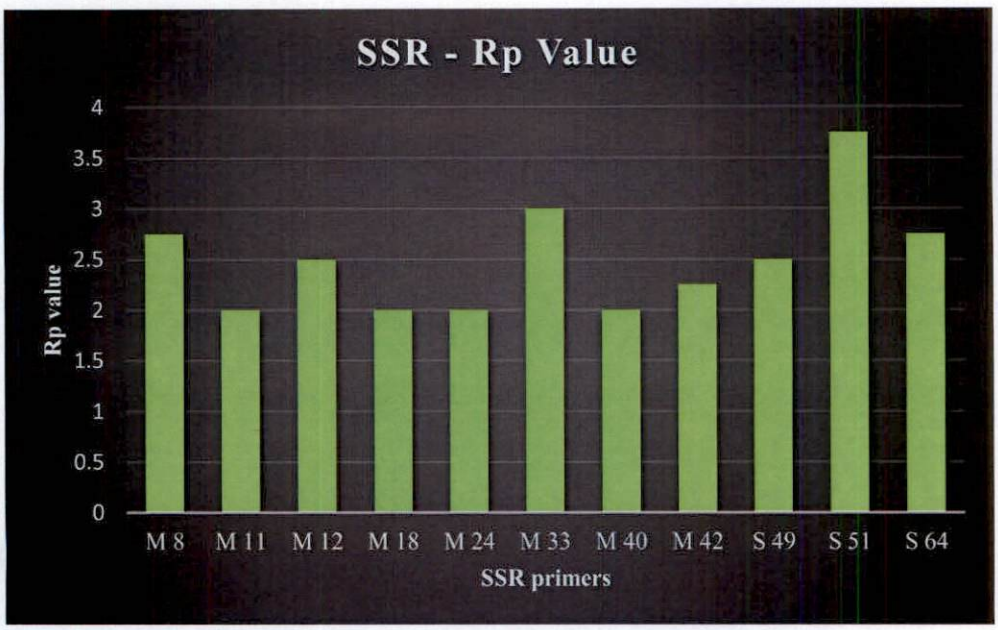


b. PIC value of 11 selected SSR primers

Fig. 26 Polymorphic information content of 10 ISSR and 11 SSR primers



a. Rp value of 10 selected ISSR primers



b. Rp value of 11 selected SSR primers

Fig.27 Polymorphic information content of 10 ISSR and 11 SSR primers

4.9 Diversity analysis

Cluster analysis was done for determining the diversity associated with morphological and molecular markers.

4.9.1 Morphological characters

4.9.1.1 Cluster analysis based on qualitative traits

Agglomerative hierarchical clustering based on Jaccard's similarity coefficient was done using the UPGMA method with 6 qualitative characters. Dendrogram was constructed and presented in Fig 28a.

Eight varieties grouped into four clusters at 66 percent similarity level (Table 13). Cluster along with the varieties are presented in table 6. Cluster III and IV had two varieties each. Members of cluster I included CCRP 1, CCRP 8 and CCRP 9. Cluster II was monogenic with CCRP 7 alone. Cluster IV comprised of CCRP 2 and CCRP 6 and cluster V included CCRP 4 and CCRP 5.

All the members of cluster I CCRP 1, CCRP 8 and CCRP 9 possessed pods of cundeamor shape, medium rougosity and dark purple beans. Cluster III had CCRP 2 and CCRP 6. Though they differed in pod shape, they were of same colour, medium rougosity and had dark purple seeds. Members of Cluster IV CCRP 4 and 5 were similar in morphological characters.

4.9.1.2 Cluster analysis based on quantitative traits

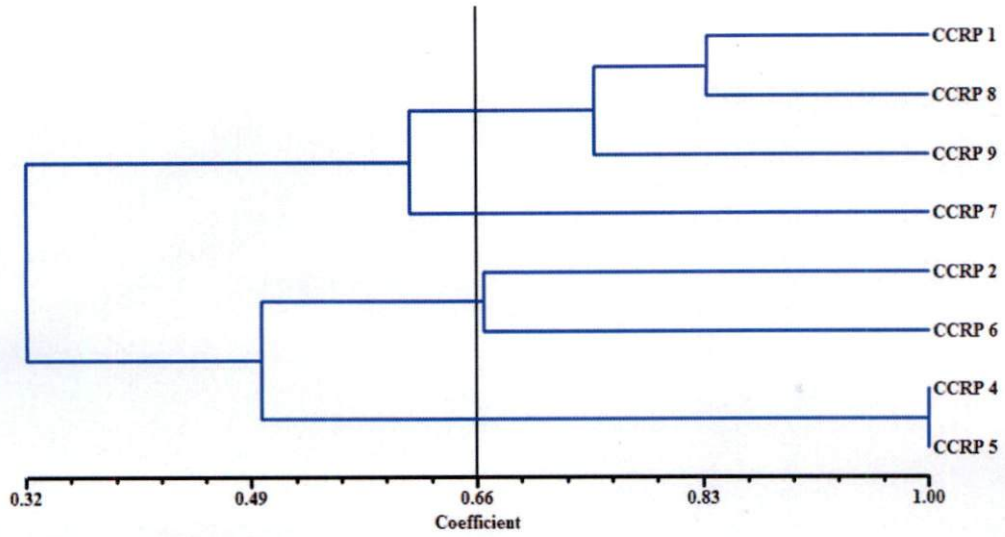
Cluster analysis was done based on Jaccard's similarity coefficient using UPGMA method with quantitative data and result obtained in such a way is represented in the form of dendrogram (Fig. 28b).

Eight varieties evaluated grouped into four clusters at 50 per cent similarity level (Table 14). These clusters along with their cluster members are presented in table 7. Cluster I possessed the maximum members of genotypes (4 no.s). The members of this cluster CCRP 1, CCRP 2, CCRP 7 and CCRP 8 were exactly similar. Cluster II had two members *viz.* CCRP 5 and CCRP 6. They were

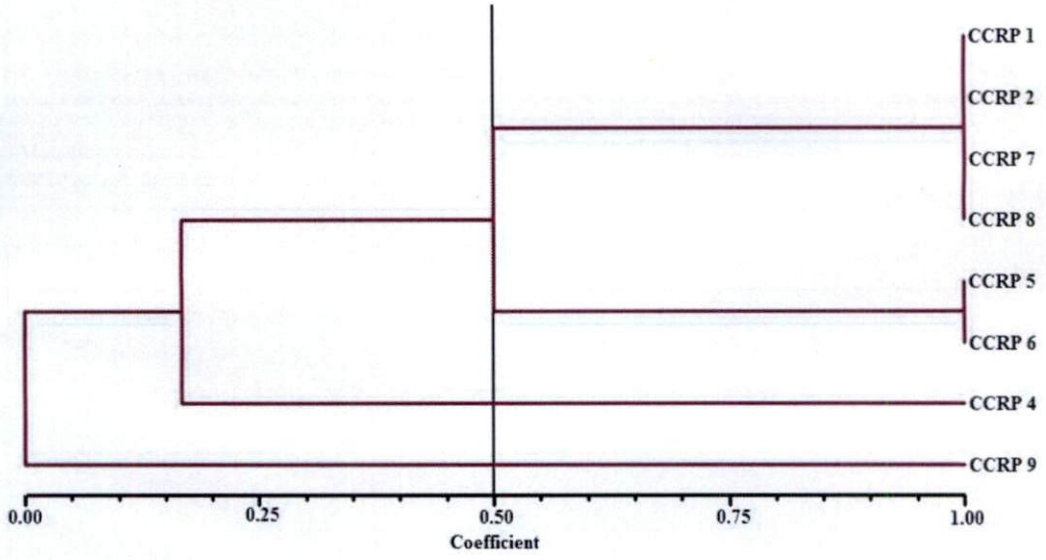
also identical in nature. Cluster III and IV are with individual members. CCRP 4 fell in cluster III and CCRP 9 in cluster IV respectively. They were distinct from other varieties.

4.9.1.3 Comparison of qualitative and quantitative clustering patterns

Homology between qualitative and quantitative clustering pattern was worked out for the varieties studied and presented (Table 15). Distribution pattern of varieties based on qualitative and quantitative clustering varied. In cluster I, CCRP 1 and CCRP 8 were common. Hybrids CCRP 8 and CCRP 9 were grouped in a single cluster with respect to qualitative data, but in case of quantitative clustering they were grouped into cluster I and cluster II respectively.



a. Dendrogram based on qualitative characters of cocoa varieties



b. Dendrogram based on quantitative characters of cocoa varieties

Fig. 28 Dendrogram based on similarity coefficient for qualitative and quantitative characters of cocoa varieties

Table 13. Clustering based on qualitative characters for cocoa varieties

Cluster no	No. of varieties	Name of varieties
I	3	CCRP 1 CCRP 8 CCRP 9
II	1	CCRP 7
III	2	CCRP 2 CCRP 6
IV	2	CCRP 4 CCRP 5

Table 14. Clustering based on quantitative characters of cocoa varieties

Cluster no	No. of cluster members	Cluster members
I	4	CCRP 1 CCRP 2 CCRP 7 CCRP 8
II	2	CCRP 5 CCRP 6
III	1	CCRP 4
IV	1	CCRP 9

Table 15. Homology between qualitative and quantitative clusters

Qualitative cluster	No. of varieties	Quantitative cluster			
		I	II	III	IV
I	3	66.66 (CCRP1, CCRP 8)	-	-	33.33 (CCRP 9)
II	1	100 (CCRP 7)	-	-	-
III	2	50 (CCRP 6)	50 (CCRP 2)	-	-
IV	2	-	50 (CCRP 5)	50 (CCRP 4)	-

4.9.1.4 Clustering of cocoa varieties based on D^2 statistics

Cluster analysis of eight cocoa varieties using D^2 statistics (Mahalanobis, 1936) based on 9 quantitative characters resulted in four clusters. This has been successfully exploited in biology to determine divergence among populations in terms of 'generalised group distance' (Murty and Arunachalam, 1966; Murthy *et al.,* 1967; Chandrasekhariah *et al.,* 1969; Ram and Panwar, 1970). In cocoa D^2 statistics was used by many scientist like Engels (1986) and Asna (2013) for cluster analysis.

The varieties included in each cluster is presented are table format (Table 16). Maximum members fall in cluster III (CCRP1, CCRP 5, CCRP7 and CCRP 8). The hybrid CCRP 9 was clearly distinct from other varieties as indicated by its position alone in cluster IV. Similarly, cluster II was also found to have only one member CCRP 2 which showed its distinctiveness from other varieties.

The intra and inter cluster distances are presented in table (Table 17 and Table 18) and the distance between different cluster centroids are presented in Fig. 2a. Cluster I and III showed minimum distance (23.77) indicated the similarity between the cluster members. Maximum inter cluster distance was observed between cluster II and IV (140.70) indicating that CCRP 2 and CCRP 9 were highly divergent. This is further proved by inter varietal distance (Table 18).Maximum inter varietal distance is expressed between CCRP 2 and CCRP 9 (140.69).Maximum similarity was observed between CCRP 1 and 8 (7.97) which fall together in Cluster III, even though one is a selection and other is a hybrid.

Genotypes of same genetic constitution will show cross incompatibility and finally will result in reduction in yield (Mallika *et al.,* 2002). Hence while designing a poly clonal garden or planting bud woods for commercial cultivation care should be taken not to plant varieties falling in same cluster (Fig. 29a). This will facilitate maximum pod set. A proposed lay out is designed in fig 29b when cocoa is planted as intercrop in coconut garden. Following the same pattern, modified designs can be employed.

Table 16. Clustering of cocoa varieties based on D^2 statistics

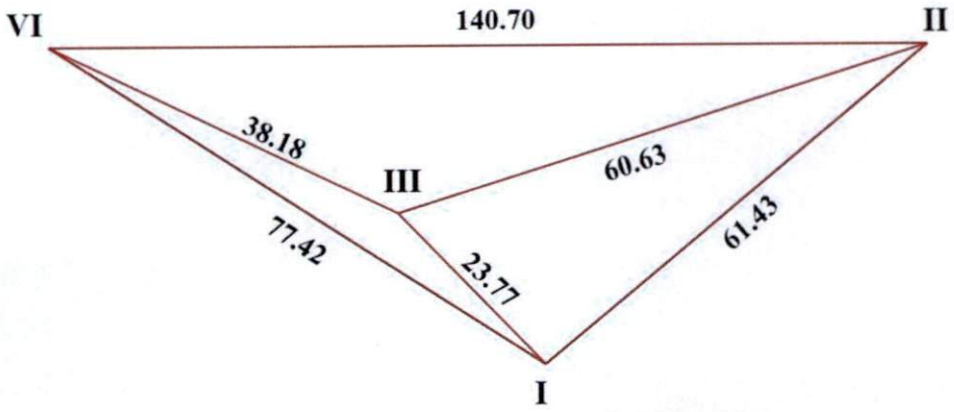
Cluster ID	No. of cluster members	Cluster members
Cluster 1	2	CCRP 4 CCRP 6
Cluster 2	1	CCRP 2
Cluster 3	4	CCRP 1 CCRP 5 CCRP 7 CCRP 8
Cluster 4	1	CCRP 9

Table 17. Inter cluster distance between cocoa varieties

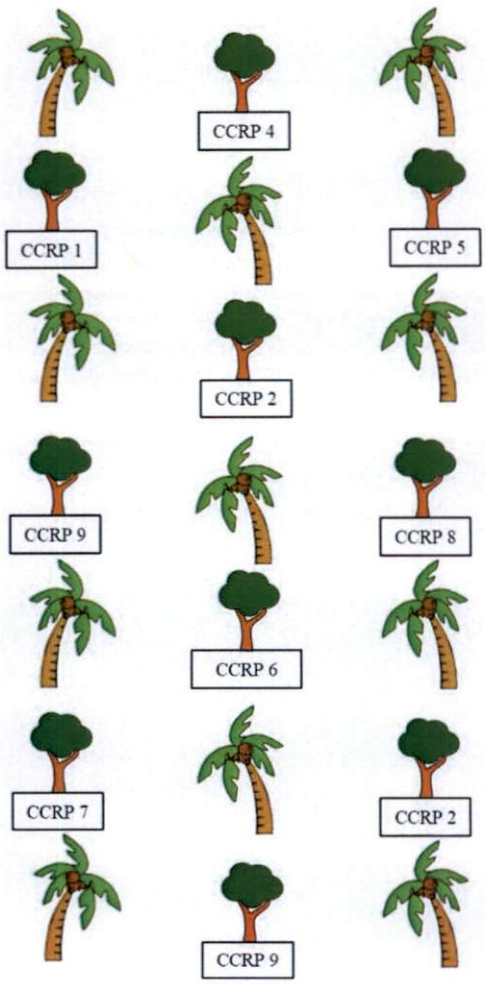
	Cluster I	Cluster II	Cluster III	Cluster IV
Cluster I	11.62			
Cluster II	61.43	0.00		
Cluster III	23.77	60.63	11.28	
Cluster IV	77.42	140.70	38.18	0.00

Table 18. Inter varietal distance between eight cocoa varieties

	CCRP 9	CCRP 8	CCRP 7	CCRP 6	CCRP 5	CCRP 4	CCRP 2	CCRP 1
CCRP 1	43.97	7.97	16.63	20.89	13.26	14.99	69.85	0
CCRP 2	140.69	58.66	70.19	49.95	43.81	72.92	0	
CCRP 4	65.19	23.87	35.64	11.62	16.07	0		
CCRP 5	51.45	9.51	11.08	15.82	0			
CCRP 6	89.65	25.52	37.35	0				
CCRP 7	25.99	9.02	0					
CCRP 8	31.32	0						
CCRP 9	0							



a. Cluster diagram based on cluster distance



b. Proposed lay out of cocoa as inter crop with coconut

Fig. 29 Cluster diagram and proposed cocoa garden lay out

4.9.2 Molecular characters

4.9.2.1 Cluster analysis based on ISSR data

Using the amplification pattern data generated by selected ISSR primers, dendrogram was constructed (Fig. 30a). At 85 per cent similarity coefficient. The eight varieties were grouped in to five clusters (Table 19). First cluster consisted of four varieties CCRP 1, CCRP 4, CCRP 6 and CCRP 7. Other four clusters were with only one member *i.e.*, Cluster II, III, IV and V with CCRP 2, CCRP 5, CCRP 8 and CCRP 9 respectively.

4.9.2.2 Clustering analysis based on SSR data

Polymorphism generated by the varieties with selected SSR markers were used to group them. At 85 per cent similarity, the eight varieties were grouped in to five clusters (Fig. 30b). Cluster II and III were found to have single cluster members in it *ie.*, CCRP 4 and CCRP 5 respectively (Table 20). Other three clusters were formed with two members each *ie.*, cluster I (CCRP 1 and CCRP 2), IV (CCRP 6 and CCRP 8) and V (CCRP 7 and CCRP 9).

4.9.2.3 Clustering based on combined ISSR and SSR data

Polymorphism obtained from ISSR and SSR marker analysis was combined and dendrogram was constructed at 85 per cent similarity coefficient (Fig. 31). Cluster I was found to have maximum number of four individuals (CCRP 1, CCRP 4, CCRP 6 and CCRP 7). Cluster II (CCRP 2) and cluster III (CCRP 5) were formed with unique members. Hybrids CCRP 8 and CCRP 9 were fell in single cluster shown certain similarities among them (Table 21).

CCRP 1, CCRP 4, CCRP 6 and CCRP 7 grouped together based on both ISSR markers and combined ISSR-SSR clustering. But based on SSR marker data, the members were split in to four different clusters indicating their diversity. However, all these four varieties were grouped together in to single cluster by ISSR and combined molecular analysis. This indicated that, there exist some

amount of similarity at molecular level. CCRP 2 was present alone in cluster II of ISSR, but when SSR was considered, it was grouped with cluster I.

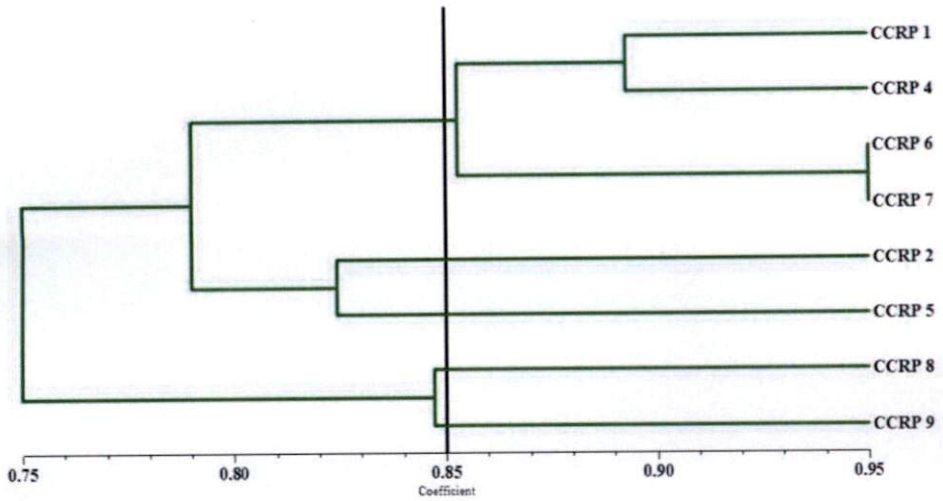
Based on ISSR analysis, CCRP 6 and CCRP 8 were exactly similar but when SSR clustering was done, CCRP 6 shown more similarity with CCRP 7 and CCRP 8 with CCRP 9.

4.9.3 Comparative clustering based on morphological and molecular characters

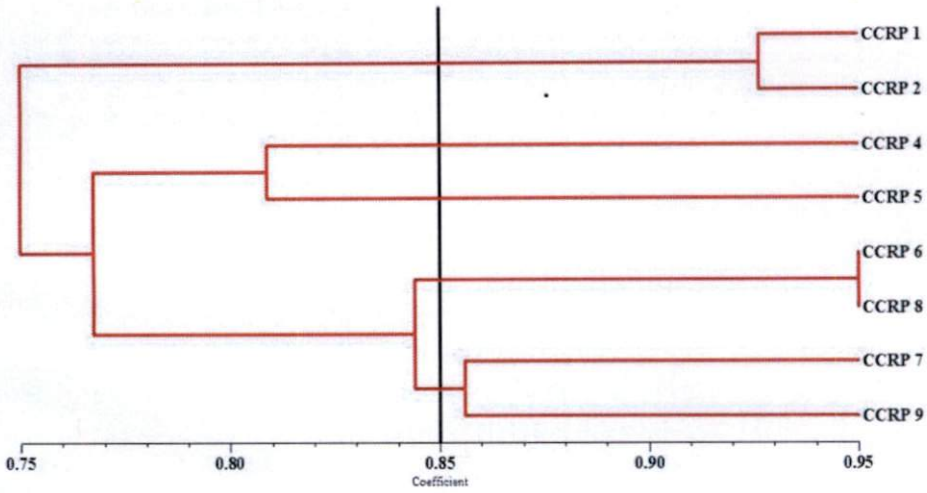
Quantitative clustering pattern of eight cocoa varieties were compared with molecular cluster data (combined ISSR-SSR data) and the result is presented below (Table 22). In both quantitative and molecular clustering, four clusters were formed but distribution of varieties were different. This may be because molecular clustering was not done based on actual coding sequence for morphological traits, but it depends on non-coding sequence. For more precise result, homology has to be worked out between expressed gene sequence and morphological data.

Table 19. Clustering based on ISSR data

Cluster no	No. of cluster members	Cluster members
I	4	CCRP 1 CCRP 4 CCRP 6 CCRP 7
II	1	CCRP 2
III	1	CCRP 5
IV	1	CCRP 8
V	1	CCRP 9



a. ISSR Dendrogram using NTsys



b. SSR Dendrogram using NTsys

Fig. 30 Dendrogram based on similarity coefficient for ISSR and SSR analysis of cocoa varieties

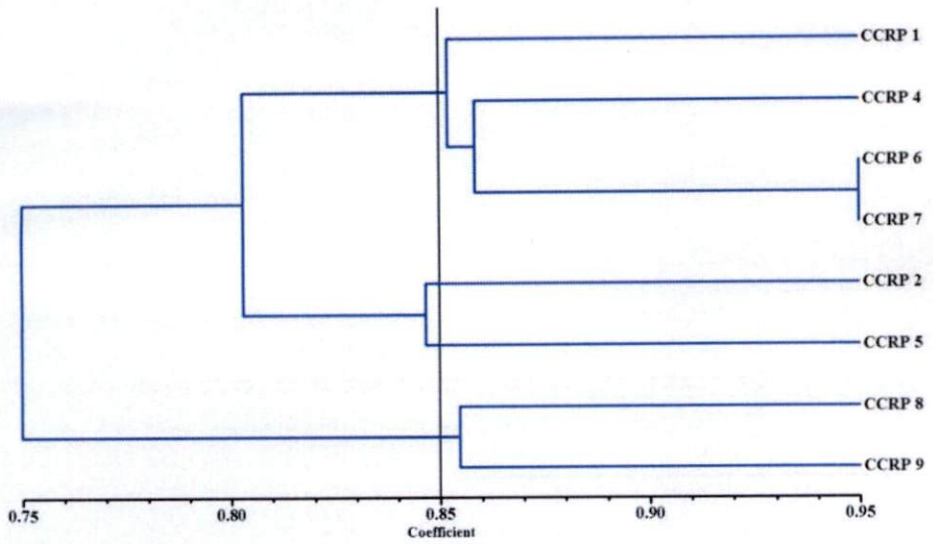


Fig 31. Combined dendrogram based on similarity coefficient for ISSR and SSR analysis of cocoa varieties

Table 20. Clustering based on SSR data

Cluster no	No. of cluster members	Cluster members
I	2	CCRP 1 CCRP 2
II	1	CCRP 4
III	1	CCRP 5
IV	2	CCRP 6 CCRP 8
V	2	CCRP 7 CCRP 9

Table 21. Clustering based on ISSR and SSR combined data

Cluster no	No. of cluster members	Cluster members
I	4	CCRP 1 CCRP 4 CCRP 6 CCRP 7
II	1	CCRP 2
III	1	CCRP 5
IV	2	CCRP 8 CCRP 9

Table 22. Homology between quantitative and molecular clustering data

Quantitative cluster	No. of varieties	Molecular cluster			
		I	II	III	IV
I	4	50 (CCRP1, CCRP7)	25 (CCRP 2)	-	25 (CCRP 8)
II	2	50 (CCRP 6)	-	50 (CCRP 5)	-
III	1	100 (CCRP 4)	-	-	-
IV	1	-	-	-	100 (CCRP 9)

Summary

5. SUMMARY

The study entitled "DNA fingerprinting of promising cocoa (*Theobroma cacao* L.) varieties of KAU" was carried out at the Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara during the period 2014-2016. The objective of the study was to characterize eight promising cocoa varieties released from KAU, using SSR and ISSR markers and to develop DNA fingerprints. Eight promising cocoa varieties CCRP 1, CCRP 2, CCRP 4, CCRP 5, CCRP 6, CCRP 7, CCRP 8 and CCRP 9 maintained at Cocoa Research Centre, KAU were used for the study.

The salient findings of the study were as follows

1. Morphological characterization of eight cocoa varieties were carried out by recording six qualitative and nine quantitative characters. Characterization was done based on standard descriptors and the variety CCRP 6 and CCRP 8 were found superior.
2. Developed an improved protocol for DNA isolation from cocoa. Chandrakant's (2014) protocol with slight modification was found to be the most appropriate method for DNA isolation.
3. Quantification of isolated DNA was done using spectrophotometer NanoDrop® ND-1000. The UV absorbance ratio ranged from 1.82-2.01 which indicated good quality DNA. Therefore the good quality DNA utilized for molecular characterization.
4. Two marker systems namely, ISSR and SSR were employed for molecular characterization and DNA fingerprinting of the selected eight varieties.
5. Optimum PCR conditions and master mix composition for both ISSR and SSR analysis in cocoa were standardized.
6. A total of 35 ISSR primers and 37 SSR primers were screened for selecting primers with good amplification and polymorphism. Thus, ten ISSR and eleven SSR primers with such characteristics were selected for further analysis.

7. Amplification pattern generated using selected ISSR and SSR primers were scored based on presence or absence of band. For further interpretation, scoring was done using colour codes. Thus, individual variety wise ISSR and SSR fingerprints were developed.
8. Most of the amplicons were found to be shared among varieties. However, amplicons shared by maximum of three varieties were considered as useful polymorphism and incorporating such useful amplicons information, consolidated DNA fingerprint was developed.
9. The DNA fingerprint thus generated was good enough to distinguish between the varieties. Unique amplicons were recorded as specific DNA fingerprints of respective varieties.
10. Variety CCRP 5 generated unique amplicon at 650 bp using primer UBC 841, which was a unique fingerprint of the variety. At 300 bp, 1300 bp and 1600 bp, primer UBC 846 also generated fingerprints which could recognize the variety.
11. Primer UBC 810 and UBC 826 generated unique amplicons at 850 bp and 1100 bp respectively in variety CCRP 6. These were the specific fingerprints of the variety.
12. Amplicons generated using the primer UBC 846 (350 bp), UBC 827 (750 bp) and UBC 866 (600 bp) were specific DNA fingerprints of the hybrid CCRP 9.
13. SSR primer mTcCIR 8 generated unique amplicons at 280 bp and 420 bp in varieties CCRP 1 and CCRP 4 respectively which were the specific fingerprints of those varieties.
14. Unique fingerprint was obtained at 320 bp in CCRP 2.
15. Primer mTcCIR 42 developed DNA fingerprints in varieties CCRP 4 (200 bp) and CCRP 9 (220 bp) which were good enough to distinguish them from other genotypes.
16. Both the marker systems generated shared bands between different varieties at different location. Further investigation is required to correlate certain morphological expressions with the shared amplicon data.

17. Unique amplicons generated using certain primers in some varieties could be further characterized and correlated to specific traits of the particular variety. For that, further studies have to be done.
18. Polymorphic information content and Rp value of selected ten ISSR primers were calculated. With 74.19 per cent polymorphism, primer UBC 846 recorded highest PIC value of 0.68. Average Rp value was found to be 14.75 and range was between 8.00 (UBC 846) to 17.25 (UBC 826).
19. Generating 100 per cent polymorphic amplicons, SSR primers mTcCIR 12, mTcCIR 33 and SHRSTc 49 were recorded highest PIC value of 0.80. Least PIC value of 0.42 was recorded by the primer, SHRSTc 64. Average Rp value was 2.50 and the range was between 2.00 to 3.75.
20. The DNA fingerprint developed was good enough to provide varietal identity and the analysis could effectively represents the relatedness/ variability among the eight studied varieties.

Future line of work includes exploration of the scope of linking polymorphic amplicons with expression of morphological traits among the varieties which will help to develop trait specific markers. Unique ISSR markers can be used for developing their corresponding SSR markers. DNA fingerprinting of other seven cocoa varieties released by KAU also have to be done.

References

REFERENCES

Adewale, B. D., Okonji, C., Oyekanmi, A. A., Akintobi, D. A. C. and Aremu, C. O. 2010. Genetic variability and stability of some grain yield components of cowpea. *Afr. J. Agric. Res.* 5: 874-880.

Agarwal, M., Shrivastava, N. and Padh, H. 2008. Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Rep.* 27: 617-631.

Aikpokpodion, P. O., Kolensnikova-Allen, M., Adetimirin, V. O., Guiltinan, M. J., Eskess, A. B., Motamayor, J. C. and Schnell, R. J. 2010. Population Structure and Molecular Characterization of Nigerian Field Genebank Collections of Cacao, *Theobroma cacao* L. *Silvae Genetica.* 59(3): 273-285.

Aikpokpodion, P.O. 2012. Phenology of flowering in cacao (*Theobroma cacao*) and its related species in Nigeria. *Afr. J. Agric. Res.* 7(23): 3395-3402.

Aikpokpodion, P.O., Motamayor, J. C., Adetimirin, V. O., Adu-Ampomah, Y., Ingelbrecht, I., Eskes, A. B., Schnell, R. J. and 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.

Ajibade, S. R., Weeden, N. F. and Chite, S. M. 2000. Inter simple sequence repeat analysis of genetic relationships in the genus *Vigna*. *Euphytica* 111: 47-55.

Alban, M. K. A., Apshara, S. E., Hebbar, K. B., Mathiyas, T. G. and Severin, A. 2016. Morpho-physiological criteria for assessment of two month old cocoa (*Theobroma cacao* L.) genotypes for drought tolerance. *Ind.J. Plant Physiol.* 21(1): 23-30.

Almeida, C. M. A., Lima, G. S. A., Brito, J. Z. and Donato, V. M. T. S. 2009. Caracterizaçao molecular de cultivares de cana-deaçucar utilizando marcadores ISSR. *Cienc. Agrotecnol.* 33: 1771-1776.

Alverson, W. S., Whitlock, B. A., Nyffler, R., Bayer, C. and Baum, D. A. 1999. Phylogeny of the core Malvales: evidence from ndhF sequence data. *Am. J. Bot.* 86: 1474-1486.

Asari, N. S., Ramaswamy, M., Subbaian, E. A., Palliyarakkal, M. K., Malhothra, S. K. and Karun, A. 2014. Standalone EST microsatellite mining and analysis tool (SEMAT): for automated EST-SSR analysis in plants. *Tree Genet. & Genomes.* 10: 1755-1757.

Asna, A. C. 2013. Performance analysis of selected accessions of cocoa (*Theobroma cacao* L.). M.Sc. thesis, Kerala Agricultural University, Thrissur, 76 p.

Asna, A. C., Prasnnakumari, K. T., Minimol, J. S. and Krishnan, S. 2014. Variability analysis in bean characters of selected accessions of cocoa (Theobroma cacao L.). *J. Plant. Crops.* 42(2): 246-251.

Baliyan, D., Sirohi, A., Kumar, M., Kumar, V., Malik, S., Sharma, S. and Sharma, S. 2014. Comparative genetic diversity analysis in chrysanthemum: A pilot study based on morpho-agronomic traits and ISSR markers. *Scientia Horticulturae.* 167: 164-168.

Barnwell, P., Blanchard, A. N., Bryant, J. A., Smirnoff, N. and Weir, F. 1998. Isolation of DNA from the highly mucilaginous succulent plant *Sedum telephium. Plant Mol. Biol. Rep.* 16: 133-138.

Barth, S., Melchinger, A. E. and Lübberstedt, T. H. 2002. Genetic diversity in *Arabidopsis thaliana* L. Heynh. investigated by cleaved amplified polymorphic sequence (CAPS) and inter-simple sequence repeat (ISSR) markers. *Mol. Ecol.* 11: 495-505.

Bayer, C., Fay, M. F. and Bruijn, A. Y. 1999. Support for an expanded family concept of Malvaceae within a re-circumscribed order Malvales: a combined analysis of plastid atpB and rbcL DNA sequences. *Bot. J. Linnean Soc.* 129: 267-303.

Beckman, J. S. and Soller, M. 1986. Restriction fragment length polymorphisms in plant genetic improvement. In: Miflin, B. J. (ed.), *Oxford Surveys of Plant Molecular and Cell Biology*. Vol III, Oxford Press, Oxford. p. 196-250.

Behera, T. K., Singh, A. K. and Staub, J. E. 2008. Comparative analysis of genetic diversity in Indian bitter gourd (*Momordica charantia* L.) using RAPD and ISSR markers for developing crop improvement strategies. *Scientia Horticulturae.* 115: 209-217.

Bekele, F. and Bekele, I. 1996. A Sampling of the Phenetic Diversity of Cacao in the International Cocoa Gene Bank of Trinidad. *Crop Sci.* 36(1): 57-64.

Bekele, F. L. and Butler, D. R. 2000. Proposed short list of cocoa descriptors for characterization. In: Eskes, A. B., Engels, J. M. M. and Lass, R. A. (Eds.), *Working Procedures for Cocoa Germplasm Evaluation and Selection.* Proceedings of the CFC/ICCO/IPGRI project workshop, Montpiller, France, pp. 41-48.

Bekele, F. L., Bekele, I., Butler, D. R. and Bidaisee, G. G. 2006. Patterns of morphological variation in a sample of cacao (*Theobroma cacao* L.) germplasm from the International Cocoa Genebank, Trinidad. *Genet. Res. and Crop Evol.* 53: 933-948.

Bekele, F. L., Kennedy, A. J., McDavid, C., Lauckner, F. B. and Bekele, I. 1994. Numerical taxonomic studies on cocoa (*Theobroma cacao* L.) in Trinidad. Euphytica 75: 234-240.

Bhattacharjee, R., Allen, K. M., Aikpokpodion, P., Taiwo, S. and Ingelbrecht, I. 2004. An improved semi-automated rapid method of extracting genomic DNA for molecular analysis in cocoa, Theobroma cacao L. *Plant Mol. Biol. Rep.* 22: 428-435.

Blair, M. W., Panaud, O. and McCouch, S. R. 1999. Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 98: 780-792.

Blin, N. and Stafford, D. W. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acid Res.* 3: 2303-2308.

Bornet, B. and Branchard, M. 2001. Nonanchored inter simple sequence repeat (ISSR) markers: reproducible and specific tools for genome fingerprinting. *Plant Mol. Biol. Rep.* 19: 209-215.

Botstein, D., Reymond, W. L., Skolnick, M. and Roland, W. D. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32: 314-331.

Brandao, M. M., Vieira, F. A. and Carvalho, D. 2011. Estrutura genetica em microescala espacial de *Myrcia splendens* (Myrtaceae). *Rev. Arvore.* 35: 957-964.

Bredemeijer, G. M. M., Arens, P., Wouters, D., Visser, D. and Vosman, B. 1998. The use of semi-automated fluorescent microsatellite analysis for tomato cultivar identification. *Theor. Appl. Genet.* 97: 584-590.

Brown, J. S., Schnell, R. J., Motamayor, J. C., Lopes, U., Kuhn, D. N. and Borrone, J. W. 2005. Resistance gene mapping for witches' broom disease in *Theobroma cacao* L. in an F2 population using SSR markers and candidate genes. *J. Am. Soc. Horti. Sci.* 130: 366-373.

Bruno, M. I., Efombagan, I. B. M., Motamayor, J. C., Sounigo, O., Eskes, A. B., Nyasse, S., Cilas, C., Schnell, R., Manzanares-Dauleux, M. J. and Kolensnikova-Allen, M. 2008. Genetic diversity and structure of farm and GenBank accessions of cacao (*Theobroma cacao* L.) in Cameroon revealed by microsatellite markers. *Tree Genet. Genomes*. 4: 821-831.

Chandrakant, E. S. 2014. Identification of molecular marker linked to the resistance for vascular streak dieback disease in cocoa (*Theobroma cacao* L.). M.Sc. thesis, Kerala Agricultural University, Thrissur, 87 p.

Chandrasekhariah, S. R., Murty, B. R. and Arunachalam, V. 1969. Multivariate analysis of genetic divergence in Eu-Sorghum. *Proc. Nat. Sci.* 35: 172-195.

Charters, M. and Wilkinson, M. J. 2000. The use of self-pollinated progenies as 'in-groups' for the genetic characterization of cocoa germplasm. *Theor. Appl. Genet.*100: 160-166.

Chia, W. J., Garcia, C. L., Suni, N. M. and Eskes, B. 2011. Characterization of a *Theobromacacao* L. collection at Tingo Maria using ISSR molecular markers. *Rev. Aporte Santiaguino*. 4(2): 195-202.

Clyde, M. M., Chew, p. C., Normah, M. N., Ramanatha-Rao, V. and Salma, I. 2005. Genetic diversity of *Nephelium ramboutan* L. assessed using RAPD and ISSR. *Acta Hort.* 665: 171-181.

Cohen, J. I., Alcorn, J. B. and Potter, C. S. 1991. Utilization and conservation of genetic resources: international projects for sustainable agriculture. *Economic Botany*. 45:190-199.

Collard, C. C. Y. and Mackill, D. J. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Trans. R. Soc.* 363: 557-572.

Cordeiro, G. M., Casu, R., McIntyre, C. L., Manners, J. M. and Henry, R. J. 2001. Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. *Plant Sci.* 160: 1115-1123.

Couch, J. A. and Fritz, P. J. 1990. Isolation of DNA from plants high in phenolics. *Plant Mol. Biol.* Report. 8: 8-12.

Cryer, N. C., Fenn, M. G. E., Turnbull, C. J. and Wilkinson, M. J. 2006. Allelic size standards and reference genotypes to unify international cocoa (*Theobroma cacao* L.) microsatellite data. *Genet. Resour. and Crop Evol.* 53: 1643-1652.

Degani, C., Rowland, L. J., Saunders, J. A., Hokanson, S. C., Ogden, E. L., Golan-Goldhirsh, A. and Galletta, G. J. 2001. A comparison of genetic relationship measures in strawberry (*Fragaria ananassa* Duch.) based on AFLPs, RAPDs, and pedigree data. *Euphytica.* 117: 1-12.

Dellaporta, S. J., Wood, J. and Hicks, J. B. 1983. A plant DNA mini preparation: version II. *Plant Mol. Biol. Report.* 1(4): 19-21.

Dias, L. A., Dos, S. and Kageyama, P. Y. 1997. Temporal stability of multivariate genetic divergence in cacao (*Theobroma cacao* L.) in southern Bahia conditions. *Euphytica.* 93(2): 181-187.

Dillinger, T. L., Barriga, P., Esca, R. S., Jimenez, M., Lowe, D. S. and Grivetti, L. E. 2000. Foods of the gods: cure for humanity? A cultural history of the medicinal and ritual uses of chocolate. *J. Nutr.* 130: 2057-2072.

Dolezel, J., Greilhuber, j. and Suda, J. 2007. Estimation of DNA content in plants using flow cytometry. *Nat. Protocols.*2: 2233-2244.

Dongre, A. and Parkhi, V. 2005. Identification of cotton hybrid through the combination of PCR based RAPD, ISSR and microsatellite markers. *J. Plant Biochem. Biotechnol.* 14: 53-55.

Doyle, J. J. and Doyle, J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19: 11-15.

Doyle, J. J. and Doyle, J. L. 1990. Isolation of plant DNA from fresh tissue. *Phytochem. Bull.* 12: 13-15.

Dudnikov, A. J. 2003. Allozymes and growth habit of *Aegilops tauschii*: genetic control and linkage patterns. *Euphytica* 129: 89-97.

Efombagn, M. I. B., Sounigo, O., Nyasse, S., Manzanares-Dauleux, M. and Eskes, A. B. 2009. Phenotypic variation of cacao (*Theobroma cacao* L.) on farms and in the gene bank in Cameroon. *J. Plant Breed. Crop Sci.* 1(6): 258-264.

El-Nahas, A. I., El-Shazly, H. H., Ahmed, S. M. and Omran, A. A. A. 2011. Molecular and biochemical markers in some lentil (*Lens culinaris* Medik.) genotypes. *Ann. Agric. Sci.* 56(2): 105-112.

Engels, J. M. M., Bartley, B. G. D. and Enriquez, G. A. 1980. Cacao descriptors, their stage and modus operandi. *Turrialba*. 30(2): 209-218.

Engels, J.M.M., 1986. The identification of cocoa cultivars. *Actae Horticulturae* 182: 195-202.

Enriquez, C. G. A. and Soria, V. J. 1966. Estudio de la variabile de varies characteristicas delar mazereas de cocoa (Theobroma cacao L.). *Fitotecnia Latino - Americana*, 3(1): 99-118.

Faleiro, F. G., Queiroz, V. T., Lopes, U. V., Guimaraes, C. T., Pires, J. L., Yamada, M. M., Araujo, I. S., Pereira, M. G., Souza Filho, G. A., Brown, J. S., Schnell, R., Ferreira, C. F., Barros, E. G. and Moreira, M. A. 2006. Mapping QTLs for witches broom (*Crinipellis perniciosa*) resistance in cacao (*Theobroma cacao* L.). *Euphytica* 149: 227-235.

Fang, D. Q. and Roose, M. L. 1997. Identification of closely related citrus cultivars with inters simple sequence repeat markers. *Theor. Appl. Genet.* 95: 408-417.

Fang, G., Hammar, S. and Grumet, R. 1992. A quick and inexpensive method for removing polysaccharides from plant genomic DNA. Biofeedback 13: 52-53.

Figueira, A. 2004. New methodologies for cocoa breeding. In *Cocoa Futures: A Source Book of Some Important Issues Confronting the Cocoa Industry*, Flood, F. and Murphy, R. (Ed.), CABI Commodities and USDA, Chinchina, Colombia. p. 65-76.

Figueira, A., Janick, J. and Goldsbrough, P. 1992. Genome Size and DNA Polymorphism in *Theobroma cacao*. *J. Am. Soc. Hortic. Sci.* 177(4): 673-677.

Figueira, A., Janick, J., Levy, M. and Goldsbrough, P. 1994. Reexamining the classification of *Theobroma cacao* L. using molecular markers. *J. Am. Soc. Hortic. Sci.* 119(5): 1073-1082.

Fouet, O., Allegre, M., Argout, X., Jeanneau, M., Lemainque, A., Pavek, S., Boland, A., Risterucci, A. M., Loor, G., Tahi, M., Sabau, X., Courtois, B. and Lanaud. 2011. Structural characterization and mapping of functional EST-SSR markers in *Theobroma cacao*. *Tree Genet. Genomes*. 7: 799-817.

Francies, R. M., Oomen, A., Mallika, V. K., and Nair, V. R. 2002. Selection indices in cocoa (*Theobroma cacao* L.). In: Bhat, R., Balasimha, D., and Jayasekhar, S. (Eds.), *Technologies for Enhancing Productivity in Cocoa.* Proceedings of a National Seminar, 29 - 30 November 2002. Vittal, Karnataka. Central Plantation Crops Research Institute, India, pp. 29-30.

Ganesan, S. K., Singh, R., Choudhury, D. R., Bharadwaj, J., Guptha, V. and Singode, A. 2014. Genetic diversity and population structure study of drumstick (*Moringa oleifera* Lam.) using morphological and SSR markers. *Ind. Crops Products*. 60: 316-325.

Ghislain, M., Nunez, J., Herrera, M. R., Pignataro, J., Guzman, F., Bonierbale, M. and Spooner, D. M. 2009. Robust and highly informative microsatellite-based genetic identity kit for potato. *Mol. Breeding*. 23: 377-388.

Giustina, L. D., Luz, L. N., Vieira, F. S., Rossi, F. S., Soares-Lopes, C. R. A., Pereira, T. N. S. and Rossi, A. A. B. 2013. Population structure and genetic diversity in natural populations of *Theobroma speciosum* Wild. Ex Spreng (Malvaceae). *Genet. Mol. Res.* 13(2): 3510-3519.

Godwin, I. D., Aitken, E. A. and Smith, L. W. 1997. Application of Inter simple sequence repeats (ISSR) markers to plant genetics. *Electrophoresis* 18(9): 1524-1528.

Guiltinan, M. J. 2007. Cacao. In: Pua, E. C. and Davey, M. R. (Eds.) Biotechnology in Agriculture and Forestry. Transgenic Crops V. Section III.2 pp. 497-518.

Gupta, M., Chyi, Y. S., Romero-Sevenson, J. and Owen, J. L. 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple sequence repeats. *Theor. Appl. Genet.* 8: 998-1006.

Gupta, P. K., Balyan, H. S., Sharma, P. C. and Ramesh, B. 1996. Microsatellites in plants: a new class of molecular markers. *Curr. Sci.* 70: 45-54.

Gupta, P. K., Rustgi, S., Sharma, S., Singh, R., Kumar, N. and Balyan, H. S. 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat, *Mol. Gen. and Genom.* 270: 315-323.

Hamrick, J. L. and Godt, M. J. 1989. Allozyme diversity in plant species. In: Brown, A. D. H., Clegg, M. T., Kahler, A. L. and Weir, B. S. (Eds.), Plant Population Genetics, Breeding, and Genetic Resources. Sinauer Association, Sunderland, pp. 43-63.

Haymes, K. M., Ibraham, I. A., Mischke, S., Scott, D. L. and Saunders, J. A. 2004. Rapid isolation of DNA from chocolate and date palm tree crops. *J. Agri. Food. Chem.* 52(17): 5456-5462.

Hodges, T. 1991. Predicting crop phenology. CRC Press, Boca Raton. p. 33-51.

Hollman, A. B., Stier, J. C., Casler, M. D., Jung, G. and Brilman, L. A. 2005. Identification of putative velvet bengrass clones using RAPD marker. *Crop Sci.* 45: 923-930.

IBPGR Secretariat. 1981. Report: IBPGR Working Group on genetic resources of cocoa. International Board for Plant Genetic Resources, Rome, pp. 28.

Ikegami, H., Nogata, H., Hirashima, K., Awamura, M. and Nakahara, T. 2009. Analysis of genetic diversity among European and Asian fig varieties (*Ficus carica* L.) using ISSR, RAPD, and SSR markers. *Genet. Resour. Crop. Evol.* 56: 201-209.

Irish, B. M., Gonenaga, R., Zhang, D., Schnell, R., Brown, S. and Motamayor, J. C. 2010. Microsatellite fingerprinting of the USDA-ARS Tropical Agricultural Research Station cocoa (*Theobroma cacao* L.) germplasm collection. *Crop Sci.* 50: 656-667.

Iwaro, A. D., Bekele, F. L., and Butler, D. R. 2003. Evaluation and utilization of cacao (*Theobroma cacao* L.) germplasm at the International Cocoa Genebank, Trinidad. Euphytica. 130: 207- 221.

Jaccard, P. 1908. Nourelles recherché sur la distribution florale. *Bull. Vaud. Sci. Nat.* 44: 223-270.

Janila, P. and Sharma, B., 2004. RAPD and SCAR markers for powdery mildew resistance gene in pea. *Plant Breed.* 123: 271-274.

Jeske, H., Lutgemeier, M. and Preib, W. 2001. DNA forms indicate rolling circle and recombination-dependent replication of Abutilon mosaic virus. *EMBO J.* 20: 6158-6167.

John, M. E. 1992. An efficient method for isolation of RNA and DNA from plants containing polyphenolics. *Nucleic Acids Res.* 20(9): p2381.

Jose, J. and Usha, R. 2000. Extraction of geminiviral DNA from a highly mucilaginous plant (*Abelmoscus esculentus*). *Plant Mol. Biol. Rep.* 18: 349-355.

Joshi, S., Ranjekar, P. and Gupta, V. 1999. Molecular markers in plant genome analysis. *Curr. Sci.* 77: 230-240.

Julio, C. W., Luis, G. C., Mery, S. N. and Bertus, E. 2011. Characterization of a *Theobroma cacao* L. collection at Tingo Maria using ISSR molecular markers. *Rev. Aporte Santiaguino* 4(2): 195-202.

Kang, J., Lu, J., Qiu, S., Chen, Z., Liu, J. and Wang, H. 2015. Dendrobium SSR markers play a good role in genetic diversity and phylogenetic analysis of Orchidaceae species. *Scientia Horticulturae.* 183: 160-166.

Karp, A., Kresovich, S., Bhat, K. V., Ayad, W. G. and Hodgkin, T. 1997. Molecular tools in plant genetic resources conservation: a guide to the technologies. IRGRI Technical bulletin No.2. International Plant Genetic resources Institute, Rome, Italy.

Karp, A., Seberg, O. and Buiatti, M. 1996. Molecular techniques in the assessment of botanical diversity. *Ann. Bot.* 78: 143-149.

Kumar, P., Gupta, V. K., Misra, A. K., Modi, D. R. and Pandey, B. K. 2009. Potential of Molecular Markers in Plant Biotechnology. *Plant Omics J.* 2(4): 141-162.

Kumar, V. and Sharma, S. N. 2011. Comparative Potential of Phenotypic, ISSR and SSR Markers for Characterization of Sesame (*Sesamum indicum* L.) Varieties from India. *J. Crop Sci. Biotech.* 14: 163-171.

Lanaud, C., Risterucci, A. M., N'goran, J. A. K., Clement, D., Flament, M. H., Laurent, V. and Falque, M. 1995. A genetic linkage map of *Theobroma cacao* L. *Theor. Appl. Genet.* 91: 987-993.

Lanaud, C., Risterucci, A. M., Pieretti, I., Falque, M., Bouet, A. and Lagoda, P. J. L. 1999. Isolation and characterization of microsatellites in *Theobroma cacao* L. *Mol. Ecol.* 8: 2141-2152.

Laurent, V., Risterucci. A. M. and Lanaud, E.1993. Chloroplast and mitochondrial DNA diversity in *Theobroma cacao. Theor. Appi. Genet.* 87: 81-88.

Laxman, P. V. 2013. DNA fingerprinting of brinjal (*Solanum melongena* L.) varieties and related species. M.Sc. thesis, Kerala Agricultural University, Thrissur, 98 p.

Lerceteau, E., Quiroz, J., Soria, J., Flipo, S., Petiard, V. and Crouzilat, D. 1997. Genetic differentiation among Ecuadorian *Theobroma cacao* L. accessions using DNA and morphological analyses. *Euphytica.* 95: 77-87.

Li, Z., Peng, T., Xie, Q., Han, S. and Tian, J. 2010. Mapping of QTL for tiller number at different stages of growth in wheat using double haploid and immortalized F_2 populations. *J. Genet.* 89: 409-415.

Li-Hua, T., Yang, X. and Yin-Bing. 2008. ISSR fingerprint analysis and SCAR marker of major cultivated strains of *Auricularia auricular* in China. *Mycosystema.* 27(2): 243-251.

Lopez, C. M. R., Wetten, A. C. and Wilkinson, M. J. 2004. Detection and quantification of in vitro-culture induced chimerism using simple sequence repeat (SSR) analysis in *Theobroma cacao* (L.). *Theor. Appl. Genet.* 110: 157-166.

Lu, H., Redus, M.A., Coburn, J. R., Rutger, J. N., McCouch, S. R. and Tai, T. H. 2005. Population structure and breeding patterns of 145 U.S. Rice cultivars based on SSR marker analysis. *Crop Sci.* 45: 66-76.

Luo, C., He, X., Chen, H., Ou, S., Gao, M., Brown, J. S., Tondo, C. T. and Schnell, R. J. 2011. Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. Biochemical Syst. Ecol. 39: 676-684.

Machado, I. E., Cach, L. A. S., Zepeda, C. H., Madrid. R. R. and Moreno, O. A. 2005. A Simple and Efficient Method for Isolation of DNA in High Mucilaginous Plant Tissues. Mol. Biotechnol. 31: 129-136.

Mahalanobis, P. C. 1936. On the generalized distance in statistics. *Proc. Nat. Ins. Sci.* India. 2: 49-55.

Mallika, V. K., Amma, S. P., Nair, R. V. and Namboothiri, R. 2002. Cross-compatibility relationship within selected clones of cocoa. In: Proceedings of seminar on strategies for enhancing productivity of cocoa, 29-30 November 2002. Ravi, B., Balasimha, D., and Jayasankar, S. (Eds.), Central Plantation Crops Research Institute, Regional station, Vittal, Karnataka, pp. 19-27.

Mansoor, S., Khan, S. H. and Bashir, A. 1999. Identification of a novel circular single-stranded DNA associated with cotton leaf curl disease in Pakistan. *Virology.* 259: 190-199.

Martinez, I. B., Cruz, M. V., Nelson, M. R. and Bertin, P. 2015. Morphological characterization of traditional cacao (*Theobroma cacao* L.) plants in Cuba. *Genet. Resour. Crop Evol.* pp. 1-27.

McGregor, C. E., Lambert, C. A., Greyling, M. M., Louw, J. H. and Warnich, L. 2000. A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. *Euphytica.* 113: 135-144.

Meena, P. 2014. Molecular characterization and DNA fingerprinting of selected cashew (*Anacardium occidentale* L.) varieties of KAU. M.Sc. thesis, Kerala Agricultural University, Thrissur, 120 p.

Minimol, J. S. and Prasannakumari, A. S. 2013. Self incompatibility assessment in vascular streak dieback (VSD) disease resistant hybrids of cocoa (*Theobrama cocoa* L.). *Asian J. Hortic.* 8(1): 114-117.

Minimol, J. S., Amma, S. P., Krishnan, S. and Vasudevan, N. R. 2011. Influence of fruit apex on fruit shape in selected accessions of cocoa. In: Proceedings of seminar on strategies for enhancing productivity of cocoa, 28-29 January 2011, Central Plantation Crops Research Institute, Regional station, Vittal, Karnataka. pp. 23-25.

Mishra, R. K., Kumar, A., Chaudhary, S. and Kumar, S. 2009. Mapping of the multifoliate pinna (*mfp*) leaf blade morphology mutation in grain pea *Pisum sativum. J. Genet.* 88: 227-232.

Mogalayi, M. 2011. DNA fingerprinting of selected black pepper (*Pepper nigrm* L.) varieties. M.Sc. thesis, Kerala Agricultural University, Thrissur, 109 p.

Montarroyos, A. V. V., de Lima, M. A. G., dos Santos, E. O. and de Franca, J. G. E. 2003. Isozyme analysis of an active cassava germplasm bank collection. *Euphytica.* 13: 101-106.

Moosu, G. 1992. *Cocoa*. The Macmillan Press Ltd., London, 103p.

Motamayor, J. C., Risterucci, A. M. Heath, M. and C. Lanaud. 2003. Cocoa domestication II: progenitor germplasm of the Trinitario cocoa cultivar. *Heredity*. 91: 322-330.

Mullis, K. B. and Faloona, F. A. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enymology*. 155: 335-350.

Murray, M. G. and Thompson, W. F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acid Res*. 8: 4321-4325.

Murty, B. R. and Arunachalam, V. 1966. The nature of divergence in relation to breeding system in some crop plants. *Indian J. Genet*. 26: 188-198.

Murty, B. R., Arunachalam, V. and Saxena, M. B. I. 1967. Catalogue of world collection of Sorghum and *Pennisetum*. *Indian J. Genet*. 27: 390-394.

N'goran, J. A. K., Laurent, V., Risterucci, A. M. and Lanaud, C. 1994. Comparative genetic diversity studies of *Theobroma cacao* L. using RFLP and RAPD markers, Heredity 73: 589-597.

Nagaoka, T. and Ogihara, Y. 1997. Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theor. Apppl. Genet*. 94: 597-602.

Nagaraju, J., Kathirvel, M., Kumar, R. R., Siddiq, E. A. and Hasnain, E. 2001. Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc. Natl. Acad. Sci. USA. 99: 5836-5841.

Nayak, S., Debata, B. K., Shrivastava, V. K. and Sangwan, N. S. 2003. Evaluation of agronomically useful somaclonal varients in Jamrosa (a hybrid

Cymbopogon) and detection of genetic changes through RAPD. *Plant Sci.* 164: 1029-1035.

Osborn, B., Lightfoot, D. R. and Hill, A. 2005. Assessment of genetic polymorphism of Pinus ponderosa by Inter Simple Sequence Repeats (ISSR). *J. Sci.* 1: 46-51.

Pandey, M. K., Monyo, E., Ozias-Akins, P., Liang, X., Guimarães, P. and Nigam, S. N. 2012. Advances in *Arachis* genomics for peanut improvement. *Biotechnol. Adv.* 30:639-651.

Pandit, S. S., Mitra, S., Giri, A. P., Pujari, K. H., Patil, B. P., Jambhale, N. D. and Gupta, V. S. 2007. Genetic diversity of mango cultivars using inter simple sequence repeats markers. *Curr. Sci.* 93: 1135-1141.

Perry, M. D., Davey, M. R., Power, J. B., Lowel, K. C., Frances, H., Bligh, J., Roach, P. S. and Jones, C. 1998. DNA isolation and AFLP genetic fingerprinting of *Theobroma cacao* (L.). Plant Mol. Biol. Report. 16: 49-59.

Pound, F.J. 1932. The genetic constitution of the cocoa crop. In: *First Annual Report on Cocoa Research for 1931.* Trinidad and Tobago, pp. 25-28.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* 2(3): 225-238.

Prasannakumari A. S., Nair, R. V., Lalithabai, E. K., Mallika, V. K., Minimol, J. S., and Abraham, K. 2009. Cocoa in India. Kerala Agricultural University, India, 72.

Prevost, A. and Wilkinson, M. J. 1999. A new system for the comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theor. Appl. Genet.* 98: 107-112.

Pugh, T., Fouet, O., Risterucci, A. M., Brottier, P., Abouladze, M., Deletrez, C., Courtois, B., Clement, D., Larmande, P. N.,'Goran, J. A. K. and Lanaud, C. 2004. A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theor. Appl.Genet.* 108: 1151-1161.

Rafalski, J. A., Vogel, J. M., Morgante, M., Powell, W., Andre, C. and Tingey, S. V. 1996. Generating and using DNA markers in plants. In: Birren, B. and Lai, E. (eds.) Non mammalian genomic analysis. *A practical guide.* Academic Press, San Diego, pp 75-134.

Rallo, P., Dorado, G. and Martin, A. 2000. Development of simple sequence repeats (SSRs) in olive tree (*Olea europeae* L.). *Theor. Appl. Genet.* 101: 984-989.

Ram, J. and Panwar, D. V. S. 1970. Intraspecific divergence in rice. *Indian J. Genet.* 30: 1-11.

Ramsay, L., Macaulay, M., Degli, I. S., MacLean, K., Cardle, L., Fuller, J., Edwards, K. J., Tuvesson, S., Morgante, M., Massari, A., Maestri, E., Marmiroli, N., Sjaskte, T., Ganal, M., Powell, W. and Waugh, R. 2000. A simple sequence repeat-based linkage map of barley. *Genet.* 156: 1997-2005.

Ratnam, R. 1961. Introduction of criollo cacao into Madras State South Indian Hort., 9: 24-29.

Reddy, M. P., Sarla, N., Siddiq, E. A. 2002. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* 128: 9-17.

Rivas, L. H., Giustina, L. D., Luz, N. L., Karsburg, I. V., Pereira, T. N. S. and Ros, A. A. B. 2013. Genetic diversity in natural populations of *Theobroma subincanum* Mart. in the Brazilian Amazon. *Genet. Mol. Res.* 12(4): 4998-4500.

Rogers, S. O. and Benedich, A. J. 1994. Extraction of total cellular DNA from plants, algae and fungi. Plant Mol. Biol. Manual. Eds. Gelvin, S. and Schilperoort, R. A.,3rd Edition, Kluwer Academic Publisher, Bostan, Mars, pp.1-18.

Rohlf, F. J. 1992. *NTSYSpc. Numerical Taxonomy and Multivariate Analysis System Version 2.0.* Department of Ecology and Evolution, State University of New York. 38p.

Rongwen, J., Akkaya, M. S., Bhagwat, A. A., Lavi, U. and Cregan, P.B. 1995. The use of microsatellite DNA markers for soybean genotype identification. *Theor. Appl. Genet.* 90: 43-48.

Ronning, C.M. and R.J. Schnell, 1994. Allozyme diversity in a germplasm collection of *Theobroma cacao* L. J Heredity 85(4): 291–295.

Rossi, A. A., de Oliveira, L. O., Venturini, B. A. and dos Santos, S. R. 2009. Genetic diversity and geographic differentiation of distinct Atlantic and Amazonian populations of *Psychotria ipecacuanha* (Rubiaceae). *Genetica* 136: 57-67.

Rubeena, M. 2015. Analysis of bean characters in cocoa (Theobroma cacao L.) hybrids bred for bold beans. MSc Thesis. Kerala Agricultural University, Thrissur, Kerala, 63p.

Russell, J. R., Fuller, J. D., Macaulay, M. Hatz, B. G., Jahoor, A., Powell, W. and Waugh, R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theor. Appl. Genet.* 95: 714-722.

Saiki, R.K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Sci.* 239: 487-491.

Salimath, S. S., de Oliveira, A. C., Godwin, I. D. and Bennetzen, J. L. 1995. Assessment of genome origins and genetic diversity in the genus Eleusine with DNA markers. *Genome* 38: 757-763.

Santos, R. C., Dias, L. A., Kageyama, P. Y. and Castro, G. C. T. 1997. Multivariate phonetic divergence in germplasm preservation of cacao (*Theobroma cacao* L.). *Agrotropica.* 9(1): 20-49.

Santos, R. C., Pires, J. L. and Correa, R. X. 2012. Morphological characterization of leaf, flower, fruit and seed traits among Brazilian *Theobroma* L. species. *Genet. Resour. Crop Evol.* 59: 327-345.

Saunders, J. A., Hemeida, A. A. and Mischke, S. 2001. USDA DNA fingerprinting programme for identification of Theobroma cacao accessions. In: Proceedings of the international workshop on new technologies and cocoa breeding, 16–17 October 2000, Kota Kinabalu, Malaysia. pp. 108-114.

Saunders, J., Mischke, S., Leamy, E. and Hemeida, A. 2004. Selection of international molecular standards for DNA fingerprinting of *Theobroma cacao*. *Theor. Appl. Genet.*110 (1): 41-70.

Singh, S., Panda, M. K. and Nayak, S. 2012. Evaluation of genetic diversity in turmeric (*Curcuma longa* L.) using RAPD and ISSR markers. *Ind. Crops Products*. 37: 284-291.

Smith, J. S. C. and Smith, O. S. 1989. The description and assessment of distances between inbred lines of maize: The utility of morphological, biochemical,

and genetic descriptors and a scheme for the testing of distinctiveness between inbred lines. *Maydica.* 34: 151-161.

Sneath, P. H. and Sokal, R. R. 1973. Numerical taxonomy: the principles and practice of numerical classification. San Francisco, Freeman, 573 p.

Soria, V. J. 1978. The breeding of cacao (*Theobroma cacao* L.) *Trop. Agric. Res. Ser.* 11: 161-168.

Spooner, D., Treuren, R. V. and Vicente, M. C. 2005. Molecular markers for genebank management. IPGRI Technical Bulletin No. 10. International Plant Genetic Resources Institute, Rome, Italy. 136pp.

Tai, T. H. and Tanksley, S. D. 1990. A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant Mol. Bio. Rep.* 8: 297-303.

Takrama, J. F., Martinez, C. C., Mora, W. P., Brown, J. S., Motamayor, J. C. and Schnell, R. J. 2005. Determination of Off-Types in a Cocoa Breeding Programme using Microsatellites. *Ingenic News Lett.* 10: 2-8.

Tan, L., Peng, M., Xu, L., Wang, L., Chen, S., Xou, Y., Qi, G. and Cheng, H. 2015. Fingerprinting 128 Chinese clonal tea cultivars using SSR markers provides new insights into their pedigree relationships. *Tree Genet. Genomes.* 90: 1-12.

Terzi, V., Morcia, C., Gorrini, A., Stanca, A. M., Shewry, P. R. and Faccioli, P. 2005. DNA-based methods for identification and quantification of small grain cereal mixtures and fingerprinting of varieties. *J. Cereal Sci.* 41: 213-220.

Thakur, P. P., Mathew, D., Nazeem, P. A., Abida, P. S., Indira, P., Girija, D., Shylaja, M. R., and Valsala, P. A. 2014. Identification of allele specific AFLP markers linked with bacterial wilt [*Ralstonia solanacearum*

(Smith) Yabuuchi, et al.] resistance in hot peppers (*Capsicum annuum* L.). *Physiol. Mol. Plant Pathol.* 87: 19-24.

Thiel, T., Michalek, W., Varshney, R. K. and Graner, A. 2003. Exploiting EST databases for the development of cDNA derived microsatellite markers in barley (*Hordeum vulgare* L.), *Theor. Appl. Genet.* 106: 411-422.

Thul, S. T., Darokar, M. P., Shasany, A. K. and Khanuja, S. P. S. 2012. Molecular profiling for genetic variability in *Capsicum* species based on ISSR and RAPD markers. *Mol. Biotechnol.* 51: 137-147.

Torres, P. I., Moctezuma, E. V., Grajales, M. P., Campos, J. R. and Flores, M. E. J. 2015. Inter- and intraspecific differentiation of *Capsicum annuum* and *Capsicum pubescens* using ISSR and SSR markers. *Scientia Horticulturae.* 181: 137-146.

Varshney, R. K., Kumar, A., Balyan, H. S., Roy, J. K., Prasad, M. and Gupta, P. K. 2000. Characterization of microsatellites and development of chromosome specific STMS markers in bread wheat, *Plant Mol. Biol. Rep.* 18: 1-12.

Varshney, R. K., Pandey, M. K., Janila, P., Nigam, S. N., Sudini, H. K. and Gowda, M. V. C. 2014. Marker-assisted introgression of a QTL region to improve rust resistance in three elite and popular varieties of peanut (*Arachis hypogaea* L.). *Theor. Appl. Genet.* 127: 1771-1781.

Varshney, R. K., Sigmund, R., Borner, A., Korzun, V., Stein, N., Sorrells, M. E., Langridge, P. and Graner, A. 2004. Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Sci.* 168: 195-202.

Weising, K., Nybom, H., Wolff, K. and Kahl, G. 2005. *DNA Fingerprinting in Plants: Principles, Methods and Applications* (2nd Ed.), CRC Press, Taylor and Francis Group, pp. 472.

Wood, G. A. R and Lass, R. A. 1985. *Cocoa*. 4th edn. Tropical Agriculture Series, Longman Publications, New York, 620p.

Yoshima, S., Yoshima, A., Iwata, N. and Mcouch, S. R. 1995. Tagging and combining bacterial blight resistance genes in rice using RAPD and RFLP markers. *Mole. Breed.* 1: 375-387.

Zhang, D., Boccara, M., Motilal, L., Mischke, S., Johnson, E. S., Butler, D. R., Bailey, B. and Meinhardt, L. 2009. Molecular characterization of an earliest cacao (*Theobroma cacao* L.) collection from Upper Amazon using microsatellite DNA markers. *Tree Genet. Genomes*. 5: 595-607.

Zhang, D., Gardini, E. A., Mischke, S., Cernades, L. Z., Chavez, A. B. and Aguila, J. A. D. 2006. Genetic Diversity and Structure of Managed and Semi-natural Populations of Cocoa (*Theobroma cacao*) in the Huallaga and Ucayali Valleys of Peru. *Ann. of Bot.* 98: 647-655.

Zietkiewicz, E., Rafalski, A. and Labuda, D. 1994. Genomic fingerprinting by simple sequence repeats (SSR)-anchored polymerase chain reaction amplification. *Genomics*. 20:176-183.

Annexures

ANNEXURE I

List of laboratory equipments used for the study

Refrigerated centrifuge	:	Kubota, Japan
Horizontal electrophoresis system	:	BIO-RAD, USA
Thermal cycler	:	Proflex
Gel documentation system	:	Gel Doc™ XR+ BIO-RAD
Micropipettes	:	Eppendrof
Spectrophotometer	:	Nanodrop® ND-1000, USA

ANNEXURE II

Descriptors and descriptor states used for qualitative characters

Sl. No.	Descriptor	Descriptor state	Description
1.	Pod Shape	1	Cundeamor
		2	Angoleta
		3	Amelonado
		4	Calabacillo
		5	Criollo
2.	Colour of ripe pod (Ridge & furrow colour)	0	Absent (green)
		3	Slight (greenish yellow)
		5	Intermediate (yellowish green)
		7	Intense (yellow)
3.	Pod apex form	1	Attenuate
		2	Acute
		3	Obtuse
		4	Rounded
		5	Mammelate
4.	Pod basal constriction	0	Absent
		1	Slight
		2	Intermediate
		3	Strong
5.	Pod rugosity	0	Absent
		3	Slight
		5	Intermediate
		7	Intense
6.	Cotyledon colour (Bean colour)	1	White
		2	Grey
		3	Light purple
		4	Medium purple
		5	Dark purple
		6	Mottled
		7	Mixed

ANNEXURE III

Reagents required for DNA isolation

1. 2x CTAB extraction buffer (100 ml)

CTAB	:	2g
(Cetyl trimethyl ammonium bromide)		
Tris base	:	1.21 g
EDTA	:	0.745 g
NaCl	:	8.18 g
PVP	:	1.0 g

Adjusted the pH to 8 and made up final volume up to 100 ml.

2. Chloroform- Isoamyl alcohol (24:1 v/v)

To chloroform (24 parts), isoamyl alcohol (1 part) was added and mixed properly.

3. Chilled isopropanol

Isopropanol was stored in refrigerator at 0°C and was used for the study.

4. Wash buffer

Ethyl alcohol 76 ml and distilled water 24 ml.

Ammonium acetate 0.077 g

Add 0.077 gm ammonium acetate in 100 ml 76% ethyl alcohol and mixed well.

5. Ethanol (70 %)

To the 70 parts of absolute ethanol (100 %), 30 parts of sterile distilled water was added to make 70 per cent ethanol.

6. TE buffer (pH 8, 100 ml)

Tris HCl (10 mM)	:	0.1576 g
EDTA (1 mM)	:	0.0372 g

(The solution was prepared, autoclaved and stored at room temperature).

ANNEXURE IV

Composition of buffers and dyes used for gel electrophoresis

1. **TAE Buffer 50X**

Tris base	:	242 g
Glacial acetic acid	:	57.1 ml
0.5M EDTA (pH 8.0)	:	100 ml

2. **Loading dye 6X**

   0.25 per cent bromophenol blue

   0.25 per cent xylene cyanol

   30 per cent glycerol in water

3. **Ethidium bromide**

   The dye was prepared as a stock solution of 10 mg/ml in water and was stored at room temperature in a dark bottle.

Abstract

DNA FINGERPRINTING OF PROMISING COCOA (*Theobroma cacao* L.) VARIETIES OF KAU

By

SUJITH S.S.

(2014-11-226)

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the requirement for the degree of

Master of Science in Agriculture

(PLANT BIOTECHNOLOGY)

Faculty of Agriculture

Kerala Agricultural University, Thrissur

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY

COLLEGE OF HORTICULTURE

VELLANIKKARA, THRISSUR - 680656

KERALA, INDIA

2016

ABSTRACT

Cocoa (*Theobroma cacao* L.) also known as 'chocolate tree' is a major cash crop of tropical countries and belongs to the family Malvaceae. The plant is a native of Andes, South America and was introduced to India during 1970s. In India, cocoa is cultivated as an important intercrop with in coconut, arecanut, rubber, oil palm etc. The statistics shows that, 80 per cent of cocoa plantations in India is established with planting materials distributed from Kerala Agricultural University (KAU).

In India, the genetic base of cocoa is widened by the systematic introduction of germplasm from University of Reading, UK. Cocoa Research Centre (CRC), KAU holds Asia's largest germplasm with 640 accessions. Exploitation of these germplasm has resulted in the release of 15 cocoa varieties from KAU.

Central sub-committee on crop standards, notification and release of varieties for agricultural crops has made it mandatory to provide DNA finger printing data along with the varieties where ever the proposal for national release/ notification is submitted. DNA markers, that are not subjected to environmental influences act as an efficient tool to identify and differentiate accessions and cultivars which are similar in morphological characteristics and with indistinct traits. DNA finger printing is successfully applied for cultivar identification, controlling seed purity of hybrids and checking the genetic similarity between cultivars. Hence, the technique act as a powerful tool to protect Plant Breeder's Right (PBR).

In the present investigation, eight promising cocoa varieties CCRP 1, CCRP 2, CCRP 4, CCRP 5, CCRP 6, CCRP 7 (selections), and CCRP 8 and CCRP 9 (hybrids) released from KAU were characterized using morphological and molecular markers. For morphological characterization, six qualitative and nine quantitative characters were recorded. And it was observed that, CCRP 6 and CCRP 8 were superior based on the performance of their yield contributing characters.

Molecular characterization was performed with genomic DNA isolated using modified Chandrakant's (2014) protocol. Based on the polymorphism, ten ISSR (inter simple sequence repeats) and eleven SSR (simple sequence repeats) primers were selected. These selected primers were used for developing DNA fingerprints of varieties under study. In the further analysis, amplicon generation pattern were carefully scored to locate polymorphism. Individual

ISSR and SSR amplification pattern further converted into variety wise fingerprints and thus consolidated DNA fingerprints on each marker system were developed.

ISSR primers UBC 810 and UBC 826 were found to differentiate CCRP 6 from other genotypes. Primers UBC 827, UBC 846 and UBC 866 were generated unique amplicons in CCRP 9. UBC 841 and UBC 846 were capable of distinguishing CCRP 5 from other genotypes. Primers UBC 835 and UBC 866 were generated amplicons in hybrids CCRP 8 and CCRP 9 alone and the markers can be used for differentiating these hybrids.

SSR marker analysis was performed using selected eleven primers. Selected primers generated polymorphic amplicons and were capable of distinguishing between varieties. Primer mTcCIR 8 generated unique amplicon at 220 and 420 base pairs (bp) in CCRP 1 and CCRP 4 respectively which was a fingerprint for those varieties. The unique amplicon generated by mTcCIR 33 at 320 bp was a fingerprint of CCRP 2. Polymorphic amplicons generated by mTcCIR 42 at 200 bp (CCRP 4) and at 220 bp (CCRP 9) were fingerprints of respective varieties.

In future, more morphological characters have to be screened and correlated with the markers shared by two or three varieties. This will help to know whether the shared bands are responsible for expression of some distinct traits. DNA fingerprints for remaining seven released cocoa varieties also have to be designed.