CHARACTERIZATION AND VALIDATION OF MICROSATELLITE MARKERS FOR RESISTANCE TO VASCULAR STREAK DIEBACK DISEASE IN COCOA (*Theobroma cacao* L.)

By

WAGHMARE SANDESH TULSHIRAM

(2014-11-101)

THESIS

Submitted in partial fulfillment of the requirement

for the degree of

Master of Science in Agriculture

(Plant Biotechnology)

Faculty of Agriculture

Kerala Agricultural University

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY
COLLEGE OF HORTICULTURE
VELLANIKKARA, THRISSUR – 680656
KERALA, INDIA

2016

DECLARATION

I hereby declare that the thesis entitled **"Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)"** is a bonafide record of research work done by me during the course of research and that it has not been previously formed the basis for the award to me of any degree, diploma, fellowship or other similar title, of any other university or society.

Place -Vellanikkara
Date: 09-09-2016

WAGHMARE SANDESH TULSHIRAM
(2014-11-104)

CERTIFICATE

Certified that the thesis entitled **"Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)** is a record of research work done independently by **Mr. Waghmare Sandesh Thulsiram (2014-11-101)** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, associateship, fellowship to him.

Place- Vellanikkara

Date- 09.09.2016

Dr. Deepu Mathew
(Chairperson, Advisory Committee)
Assistant Professor
Centre for Plant Biotechnology and Molecular Biology

CERTIFICATE

We, the undersigned members of the advisory committee of **Mr. Waghmare Sandesh Thulsiram (2014-11-101)**, a candidate for the degree of **Master of Science in Agriculture** with major field in **Plant Biotechnology**, agree that the thesis entitled **Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)** may be submitted by **Mr. Waghmare Sandesh Thulsiram (2014-11-101)** in partial fulfillment of the requirement for the degree.

9/9/16

Dr. Deepu Mathew
(Chairperson, Advisory Committee)
Assistant Professor
Centre for Plant Biotechnology and Molecular Biology
College of Horticulture, Vellanikkara

Dr. P. S. Abida
(Member, Advisory Committee)
Professor
Department of Plant Breeding
and Genetics
RARS, Pattambi

Dr. J. S. Minimol
(Member, Advisory Committee)
Assistant Professor
Dept. of Plant Breeding and Genetics
Cocoa Research Centre
College of Horticulture, Vellanikkara

9/9/16

Dr. P. A. Valsala
(Member, Advisory Committee)
Professor and Head
Centre for Plant Biotechnology
and Molecular Biology
College Of Horticulture,
Vellanikkara

9/9/16

(External Examiner)
Dr. T. PARIMELAZHAGAN
Professor
Department of Botany
Bharathiar University
Coimbatore - 641 046

ACKNOWLEDGEMENT

*First and foremost, I exalt and praise the "**Almighty**" for having blessed me and who finally made this humble effort a reality.*

*I appraise my deep sense of gratitude and indebtness to my benevolent and perspicacious Professor, Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, Chairperson of my advisory committee, **Dr. Deepu Mathew** for her erudite counselling, untiring interest and constructive ideas which helped in completing this thesis and whose encouragement, constant guidance and support from the initial to the final level enabled me to develop an understanding of the subject. My sense of gratitude and indebtedness to her will remain till eternity.*

*My sincere thanks are due to **Dr. P.A.Valasala.** Professor and Head, Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, member of my advisory committee, for the valuable guidance and help rendered in the conduct of the experiment.*

*It is a gratifying moment to express my indebtness with devotion to **Dr. Abida P. S.** Professor, Department of Plant Breeding and Genetics RARS, Pattambi, member of my advisory committee, for her inspiring guidance, support and encouragement rendered in the conduct of the experiment.*

*My sincere thanks are due to convey my sincere gratitude to the **Dr. Minimol. J.S.** Assistant professor, Department of Plat Breeding and Genetics, COH, Vellanikkara, member of the advisory committee, for his valuable guidance rendered in the conduct of the experiment.*

*It is a great pleasure to record my sincere thanks to **Dr. M. R. Shylaja.** Associate professor Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, for their untiring support and help during my course of study.*

My sincere thanks to ***Sandhya , Shiny , Shylaja, shruthi,, kavya , Remya , Danasree and Misha, Shabnam, milsha, Sherin,Asha and research fellow of*** *CPBMB, College of Horticulture, Vellanikkara, who helped me in several ways for the completion of this venture.*

I wish to express my sincere thanks to all the non-teaching staff members and labourers of CPBMB for their whole-hearted cooperation and timely assistance. I also appreciate and acknowledge the facilities provided at CPBMB and Bioinformatics Centre for the successful completion of my research work.

My heartfelt gratitude cannot be captured in words for the unflinching support, constant encouragement, warm concern, patience and valuable advice of my friends ***Anju, Andrew, Ashwini, Brinda, Naresh, Vinusree, Kalavati, Rosemole, Manglam, Sujith,Darshan, Saurav, Manjesh, Narasimha*** *whose prayers, love and affection rendered me a successful path which propped up my career all along. My duty is incomplete if I forget my Senior friends* ***Rohini, Dolagovinda, Sreeja, Ajinkya , Sachin and Aoundumbar*** *and my junior friends* ***Ashwin, Basil, Debashish, Swapnil, Kishor, Pramod, Priya , Deepali, Geethu, Aruneema and Giri*** *whose helping hands, love and affection fetched a remarkable place in my days in Kerala.*

I owe special thanks to Librarian, College of Horticulture, ***Dr. A. T. Francis*** *and all other staff members of library, who guided me in several ways, which immensely helped for collection of literature for writing my thesis.*

I take this opportunity to express my deep sense of gratitude to Kerala Agricultural University and Department of Biotechnology for all the support rendered for the present study.

The award of DBT fellowship is greatly acknowledged.

Above all, I am forever beholden to my beloved parents, my brothers, my sisters and my family members for their unfathomable love, boundless affection, personal sacrifice, incessant inspiration and constant prayers, which supported me to stay at tough tracks.

Waghmare sandesh Tulshiram


Dedicated to
my beloved
parents and
teachers

TABLE OF CONTENTS

LIST OF TABLES

LIST OF PLATES

LIST OF FIGURES

LIST OF ANNEXURES

ABBREVIATIONS

%	Percentage
@	At the rate
<	Less than
=	Equal to
>	Greater than
µg	Microgram
µl	Microlitre
AFLP	Amplified Fragment Length Polymorphism
BLAST	Basic Local Alignment Search Tool
bp	Base pair
cm	Centimetre
cM	Centimorgan
CPBMB	Centre for Plant Biotechnology and Molecular Biology
CTAB	Cetyl Trimethyl Ammonium Bromide
DNA	Deoxyribose Nucleic Acid
dNTPs	Deoxyribo Nucleoside Triphosphate
EDTA	Ethylene Diamine Tetra Acetic acid
g	Gram
ha	Hectare
ISSR	Inter Simple Sequence Repeat
Kb	Kilo basepairs
L	Litre
M	Molar
Mb	Mega bytes
mg	Milligram
ml	Millilitre
mM	Milli molar

NCBI	National Centre for Biotechnology Information
ng	Nanogram
°C	Degree Celsius
OD	Optical Density
PCR	Polymerase Chain Reaction
pH	Hydrogen ion concentration
PVP	Poly vinyl pyrolidone
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
SCAR	Sequence Characterized Amplified Region
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeats
TAE	Tris Acetate EDTA
TE	Tris EDTA
U	Unit
UV	Ultra violet
V	Volts
β	Beta

Introduction

1. INTRODUCTION

The cocoa (*Theobroma cacao* L.) (2n = 20) is a main tropical crop belonging to the family *Malvaceae* and is known as the chocolate tree (Alverson *et al.*, 1999). Cocoa has a small genome with the size of 390 Mb (Couch *et al.*, 1993). This crop native to the Amazon region of South America and provides sustainable economic benefits to the farmers. Beans are used for the production of liquor and butter whereas cocoa powder is used in the chocolate industry. The global production and consumption of cocoa is around 27.00 lakh MT. In India, cocoa is widely cultivated in Kerala, Karnataka, Tamil Nadu and Andhra Pradesh, with the total area of 78,000 ha, production of 16,050 MT and average productivity of 475kg/ha (DCCD, 2015). Kerala is the leading cocoa producing state.

Vascular Streak Dieback (VSD) disease is one of the major constraints in cocoa cultivation, causing heavy losses in mature trees as well as seedlings. VSD is caused by a basidiomycete (Ceratobasidi ales) fungus *Ceratobasidium theobromae* (Keane &Prior 1991). Cocoa is the main host of *Ceratobasidium theobromae* and the fungus sporulates on cracks by the side of the main vein of the leaves, with necrotic blotch (Purwantara *et al.*, 2009). The initial symptoms of VSD are chlorosis of a single leaf, generally on the second or third flush behind the shoot apex, with spotted islets of green tissue with 2 to 5 mm in diameter. In India, VSD was noticed in Kerala and initial record of this disease was made by Abraham (1981) and Chandramohanan and Kaveriappa (1981). The incidence was found in Kozhikode, Kottayam, Idukki and Pathanamthitta districts.

Conventional breeding relies on the field screening for the confirmation of the transfer of desired genes in the offspring and this process is highly influenced by the environmental factors. Marker assisted selection is the most reliable alternative where the tightly linked molecular markers will be employed to confirm the presence of the gene of interest in the selected plants (Salimath *et al.*, 1995). Molecular markers were introduced in 1990s and proven more useful for gene tagging, because they are more informative and stable than morphological and agronomic traits (Cohen *et al.*, 1991).

Use of molecular markers is an effective genomic approach for fingerprinting, mapping, gene tagging, as well as for determining the genetic structure and relationships (Laurent *et al.,* 1993; Motamayor *et al.,* 2003). In Marker Assisted Selection, molecular markers can be used for indirect selection of traits (Kelly, 1995).

Simple sequence repeats (SSRs) are stretches of DNA dispersed throughout the genome, containing tandem repeats of nucleotide units. SSRs are generated from unequal crossing over and replication errors that leads to the formation of a typical secondary structure like hair pins or slipped strands (Pearson *et al.* 1998). Inter Simple Sequence Repeats (ISSRs) are the regions lying between two SSR regions. Most of the ISSR regions represent the coding regions. ISSR marker system does not require prior sequence information since the primers used are of random in nature (Reddy *et al.*, 2002). High degree of polymorphism, simplicity and low cost makes ISSR as an efficient marker system (Salimath *et al.*, 1995; Borba *et al.*, 2005).

Cocoa Research Centre, Kerala agricultural University, India has developed as much as 5921 hybrids of which a large number are VSD resistant and their field performance is being evaluated. The resistance hybrids are maintained in Cacao Research Centre (Minimol and Prasannakumari, 2013). In the inoculation test for screening, 566 seedlings survived and these hybrids were field planted during June 1998. Efficiency of resistance was compared between mother plant and their progenies it was seen that G VI 126 (Scavin 6) shown maximum resistance and male parent G VI 55 indicated highest recovery. A total of 267 hybrids were found free from the disease and 51 recorded satisfactory yield levels (Minimol *et al.*, 2016). Five best performing VSD resistant hybrids S IV1.26, S IV2.29, S IV4.29, S IV6.18 and VSD I31.8 are identified to be released for commercial cultivation (Amma, 2013). Through hybridization programme, CCRP 15 (VSD I 31.8), a VSD resistant cocoa variety was developed and released by Ker 3 ،gricultural University for commercial cultivation in 2015 (Minimol *et al.*, 2015).

Two marker systems ISSR and SSR were used for characterization of the genotypes for VSD resistance. Genotypes of cocoa *viz.*, VSD I- 4.6, VSD I- 4.11,

VSD I- 5.8, VSD I- 6.9, G VI- 50, G VI- 52, G VI- 82, G VI- 100, G VI- 25, G VI- 53, G VI- 67, G VI- 144 with varying response to the VSD were used for marker development. Five ISSR viz., UBC811, UBC815, UBC826, UBC857 and UBC866 and one SSR markers mTcCIR42, linked to VSD resistance were identified at Kerala Agricultural University (Chandrakant, 2014). These identified markers have to be validated by studying the segregation of markers in a large population. Twenty hybrids showing resistance to VSD and four susceptible clones have been selected from Cocoa Research Centre, College of Horticulture, Kerala Agricultural University and employed in the study. Further, the characterization of these markers by sequencing can throw light on the genes as well as the proteins involved in the resistance mechanism in cocoa and also for the development of gene specific markers for the use in MAS.

The objective of this study was 'to validate and characterize identified microsatellite SSR and ISSR markers for the genes, offering resistance to vascular streak dieback disease of cocoa (*Theobroma cacao* L.) and to characterize the ISSR markers to develop corresponding SSR markers.'

Review of Literature

2. REVIEW OF LITERATURE

The present study on "Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)" has been executed through the amplification of ISSR and SSR region in different cacao accession. The relevant literatures available on various aspects of this study were collected and are reviewed in this chapter, under different heads.

Theobroma cacao L. (Family - Malvaceae) is an important tropical crop, known as the chocolate tree. The annual world cocoa production is about 3 million tons, among this two third is used for cocoa powder and cocoa butter, and one third portion is used for cocoa liquor production (Wood and Lass, 1985).

Cacao is perennial tree crop having a cropping cycle of more than 50 years; cultivation helps to conserve soil, watershed, biodiversity, and buffer zones near endangered areas (Rice and Greenberg, 2003).

2.1. Botany of cocoa

Theobroma cacao is a diploid with ten pairs of chromosome (2n = 20) and the genome size of cacao varies from 390 Mb to 415 Mb (Figueira *et al.,* 1992; Couch *et al.,* 1993). Flowers of cacao are borne on the trunk and branches, a habit referred to as cauliflory or truncate. Flowers are produced on wood of a certain minimum physiological age, usually two or three years old under favorable growing conditions. Flowers are borne on long pedicels and have five free petals, ten stamens and ovary of five united carpels. The ten stamens which form the androecium of the flower are in two whorls consist of five long non fertile staminodes, while the inner whorl has five fertile stamens. The stamens bear two anthers which lie in the pouch of corresponding petal. The ovary has five parts containing many ovules arranged around a central axis. Flowers are pink with darker tissue in the staminodes and petals, but there is a considerable variation between cultivars in the size and color of the flowers (Wood and Lass, 1985).

The sustainability of cocoa is under increasing threat from both coevolved and newly- encountered diseases, which constitutes the most serious constraint to production in the Neotropics (Evans, 2007). Annually cacao diseases reduce the crop's potential globally by almost 30 per- cent, with some farms experiencing 100 per- cent losses (Keane, 1992; Bowers *et al.*, 2001).

2.2. Vascular Streak Dieback disease of cocoa

Turner (1967) recorded the widespread occurrence of vascular streak dieback (VSD) disease in 37 cocoa growing countries in Asia, Africa, the Caribbean, Central America and South America. Investigations in these countries attributed the dieback to one or a combination of factors which included environmental, nutritional, tree age, vigor, varietal response, pests and diseases.

VSD was first reported in Papua New Guinea in 1960 and it reached to epidemic levels posing severe threat to plantations (Shaw, 1963; Bridgland *et al.*, 1966; Keane, 1981, Dennis, 1991). VSD infection occurs in vegetative shoots of the plants is transferred frequently to cocoa from an unidentified host in Southeast Asia (Keane 1992).

In India for the first time Abraham (1981) and Chandramohan and Kaveriappa (1982) have reported the presence of VSD in Kottayam district of Kerala. The maximum spread of VSD was in Kottayam, followed by Thiruvananthpurm, Kozhikode, Idduki and Pathanamthitta districts of Kerala (Abraham and Ravi, 1991). Earlier, VSD was not present in Thrissur and Palakkad district, but a further survey conducted in 1993 has revealed the incidence of the disease in Thrissur District of Kerala (KAU, 1993, 1995).

2.3. Causal organism of VSD

VSD is caused by the fungus *Ceratobasidium theobromae* (Talbot and Keane, 1971; Keane and Prior, 1991). Cacao is the main host of *Ceratobasidium theobromae* (Anderson, 1989; Dennis, 1991). The fungus *Ceratobasidium theobromae* belongs to the family Ceratobasidiaceae. The fungus cannot sporulate on dead branches,

although the fungus readily grows from infected tissue onto water agar. *Ceratobasidium theobromae* cannot be maintained in pure culture, but the sexual stage has only been induced in tissue culture (Lam *et al.*, 1988, Dennis, 1991).

Gonzalez *et al.* (2001) observed *Ceratobasidium* is more diverse than Thanatephorus, paraphyletic *Ceratobasidium* and a monophyletic Thanatephoru. Wind dispersed basidiospores of the fungus released during periods of continuous rainfall and high humidity are the sources of inoculum, and that spore infects to very young, unhardened leaves. *Ceratobasidium theobromae* is closely related to species in Ceratobasidium and Thanatephorus (Maddison and Maddison, 2005).

Ceratobasidium theobromae enters in leaf, xylem vessels and causes vascular browning in veins of the lamina. The fungus transfers into the midrib, petiole and it spread in the branch. The pathogen of VSD grows on the main stem and cause death of a whole tree. The fungal hyphae emerge from the vascular traces in the leaf wound and form white corticoid basidiomata in and around the scar. The fungus sporulates in the main veins of leaves with necrotic blotches (Purwantara *et al.*, 2009).

2.4. Symptoms of Vascular Streak Dieback disease of cocoa

Shaw (1963) has first reported the symptoms of VSD. Subsequently the detailed symptoms were given by Keane *et al.* (1972), Keane and Prior, (1991), and Guest and Keane, (2007). Symptoms of VSD, includes chlorosis of leaves with green spots and necrotic blotches appear 3 to 5 months after infection. After the full development of spotted chlorotic symptoms on the leaf, leaf abscission occurs within a few days, even if leaves with the new necrotic symptoms cling much more persistently to the stem.

The VSD spreads to the lateral branches on the infected stem and on such branches; the leaves will turn chlorotic and drop off in succession from the base. Leaves in the latest flush of the infected seedlings or branches shows interveinal necrosis, characteristic to the calcium deficiency. The leaf scars developed due to the fall of chlorotic leaves will be covered by white, effused, adherent fruiting bodies of

the fungus. These fruiting bodies will be seen on the leaf scar and adjacent bark. Infected stem has brown cambium and discolored xylem with brown streaks.

Prior (1980) has reported that the maximum disease occurrence was seen in 3-5 months after seasonal rainfall. Seedlings inoculated with spores of the fungus develop typical symptoms of the VSD (Keane, 1981).

Wood and Lass (1985) noted that in Malaysia, occurrence of interveinal necrosis was more common compared to yellow leaves with green isolated patches as that Papua New Guinea. A study conducted at KAU, under the Cadbury Cocoa Research Project has revealed many variations in symptoms of VSD (KAU, 1995).

2.5. Yield losses in cocoa due to VSD

VSD in Malaysia is reported to have which caused significant crop losses (Guest and Keane, 2007). Byrne (1976) has estimated that overall loss due to VSD was between 25 to 40 per-cent of total production. In areas severely affected by VSD, in Malaysia yield has reduced considerably in new plantations (Shepherd *et al.,* 1977). Severe death of cacao seedlings in the nursery and immature field planting has also been reported (Ahmad, 1982).

2.6. VSD resistance in cocoa

Host resistance to vascular streak dieback disease was observed at first in Papua New Guinea in the 1960s. Later in Malaysia, this resistance was crucial to overcome epidemics that destroyed the cocoa industries (Zainal Abidin *et al.*, 1984; Keane and Prior, 1992). Generally, the resistance to VSD is found to be is partial, limiting infection to branch tips thus preventing the fungus from reaching larger branches and killing trees (Van der Vossen, 1997).

In Papua New Guinea, breeding with surviving genotypes was successful, and has relegated VSD to minor importance in most years. Cacao in Sulawesi Island were highly resistant to VSD, most likely as natural selection from diverse cacao germplasm (Tan and Tan, 1988)

Screenings of resistance in cacao seedlings and clonal cuttings had been attempted by inoculating spore suspensions of *Ceratobasidium theobromae* in the

upper surface of a young unexpanded leaf and onto the stipules of the apical bud. Considerable variation in resistance and susceptibility had been observed by this method. No cultivars have been observed to be completely resistant to VSD (Prior, 1978).

2.7. Gene action in VSD resistance

According to Tan and Tan (1988) in VSD resistance, gene effects were mostly additive for nearly all characters of the cocoa. Resistance to VSD is in the form of horizontal resistance. It is polygenic and largely inherited as additive genes. VSD resistance is durable and is inherited quantitatively. The action of the gene is additive when each additional gene increases the expression of the trait by the same increments.

2.8. Breeding for VSD resistance in cocoa

In Kerala Agricultural University during 1995 to1998, breeding for resistance to vascular streak dieback (VSD) disease was performed. During this period explorative crosses using 298 resistant parents were made and 953 hybrid seedlings shown various levels of resistance (CCRP report, 1998).

Five best performing VSD resistant hybrids S IV1.26, S IV2.29, S IV4.29, S IV6.18 and VSD I31.8 were identified to be released for commercial cultivation (Amma, 2013). Through hybridization programme CCRP 15 (VSD I 31.8) VSD resistant cocoa variety was developed and released by Kerala Agricultural University for commercial cultivation in 2015 (Minimol *et al.*, 2015) .

Breeding programme for VSD resistance was initiated in Kerala Agricultural University and it had produced 5921 hybrid seedlings. In screening of inoculation test 566 seedlings survived and these hybrids were field planted during June 1998. Efficiency of resistance was compared between mother plant and their progenies, it was seen that G VI 126 (Scavin 6) shown maximum resistance and male parent G VI 55 indicated highest recovery (168 nos). In total 267 hybrids were found free from the disease and 51 recorded satisfactory yield levels (Minimol *et al.*, 2016).

2.9. Molecular markers in plant

Molecular marker is a DNA sequences that is tightly linked with the gene of our interest and readily detected and is inherited in the similar pattern as gene does. Molecular markers are not influenced by the environmental factors; shorten the process of evaluation and identifying genotypes (Kumar *et al.*, 2009)

Molecular markers were introduced in the 90s and proven more useful for evaluation of genetic diversity, because they are more informative and stable than morphological and agronomic trait (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 help to marker assisted gene pyramiding in rice and it reduces the duration of breeding programme (Yoshima *et al.*, 1995).

The DNA based marker systems are classified in 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 speed, cost and technical simplicity, but must be sufficiently informative to distinguish between the most individuals (Chartes and Wilkinsons, 2000).

2.10. PCR based molecular techniques

The Polymerase Chain Reaction (PCR) technique has been the basis of a growing range of new techniques for genome analysis based on the selective amplification of genomic DNA fragments (Saiki *et al.*, 1988). Williams *et al.* (1990) reported the use of PCR with short oligonucleotide primers of arbitrary (random) sequence to generate markers, the basis of the Random Amplified Polymorphic DNA (RAPD). Welsh and McClelland (1990) reported an Arbitrarily Primed Polymerase Chain Reaction (AP-PCR). DNA Amplification Fingerprinting (DAF) was also reported as another technique of PCR used in various experiments (Caetano-Anolles *et al.*, 1991). The PCR reaction requires deoxynucleotides, DNA polymerase, primer, template and buffer containing magnesium. Typical PCR amplification utilizes

oligonucleotide primers which hybridize to complementary strands. The product of DNA synthesis of one primer serves as a template for another primer. The PCR process requires repeated cycles of DNA denaturation, annealing and extension with DNA polymerase enzyme, leading to amplification of the target sequence. Specific region amplified by the primer produces multiple copies of DNA (Saiki *et al.*, 1988). The technique can be applied to detect polymorphism in various plants, animals, bacterial species and fungi.

The introduction of the PCR technique has revolutionized standard molecular techniques and has allowed for the proliferation of new tools for detecting DNA polymorphism (Hu and Quiros, 1991). The electrophoresis pattern of fragments generated by each primer for one isolate can be used as DNA fingerprints for assaying diversity (Tommerup *et al.*, 1998). Insertion can change the size of a DNA fragment without preventing its amplification (Williams *et al.*, 1990). PCR is simple, fast, specific, sensitive and the main advantage of this technique over others is its inherent simplistic analysis and the ability to amplify extremely, small quantities of DNA (Welsch *et al.*, 1991).

2.10.1. SSR markers in crop plants

The allelic variation at microsatellite loci can easily be detected by PCR using specific flanking primers. Variation in repeated motifs is most likely due to unequal crossing-over or slippage during DNA replication (Levinson and Gutman, 1987).

Microsatellites, or SSRs, are present in the majority of eukaryotic genomes and consist of simple, short tandemly repeated di- to penta-nucleotide sequence motifs (Beckman and Soller, 1986). Microsatellites are abundant, have a high degree of polymorphism, locus specificity, reproducibility, requires low amount of DNA, suitability for multiplexing on automated systems and, their co dominant mode of inheritance. SSR loci has tended to cluster reported in several species, including

sorghum (Bhattramakki *et al.,* 2000), barley (Ramsay *et al.*, 2000), rye grass (Jones *et al.,* 2002) and rice (McCouch *et al.,* 2002).

Microsatellites used in population genetics studies; it ranges from the individual level to closely related species. The high rate of mutation in ISSR makes them unsuitable for higher taxonomic study levels. In gene mapping studies microsatellite are ideal markers (Hearne *et al.*, 1992; Morgante and Olivieri, 1993; Jarne and Lagoda, 1996).

Assessment of genetic variation in germplasm is more effective using molecular markers (Mohammadi and Prasanna, 2003). Genic SSRs are helpful for estimating genetic relationships and to examine functional diversity in relation to adaptive variation (Eujayl *et al.*, 2001).

Screening of microsatellite variation can be automated, the use of automatic sequencers is an option EST-SSR markers are one class of marker that can contribute to 'direct allele selection', if they are completely associated or even responsible for a targeted trait (Sorrells and Wilson, 1997).

Yu *et al.* (2004) reported two EST-SSR markers in wheat, which are tightly linked to the photoperiod response gene (*ppd*). In cereals species EST-SSR loci have been integrated and genome-wide genetic maps have been prepared. Many genic SSRs have been positioned on the genetic maps of wheat (Nicot *et al.*, 2004, Holton, 2002 and Gao *et al.*, 2004).

2.10.2. Inter Simple Sequence Repeats (ISSR)

ISSR is DNA fragments of about 100-3000bp located between adjacent, oppositely oriented microsatellite regions. This technique, reported by Zietkiewicz *et al.,* (1994) primers based on microsatellites is used to amplify ISSR region present in DNA sequences. Primers which anchors into the non-repeat adjacent regions (16-18 bp) is used for ISSR assay. From multiple loci 10-60 fragments are generated these

fragments are separated by gel electrophoresis and used for scoring. Single Primer Amplification Reaction (SPAR) and Directed Amplification of minisatellite region DNA (DAMD) uses a single primer having a core motif of a microsatellite and minisatellite respectively. The main advantage of ISSRs is that no prior sequence data are needed for primer construction. ISSR assay includes PCR and requires low quantities of template DNA (5-50 ng per reaction). Multilocus fingerprinting profiles obtained, through ISSR analysis can be applied in studies involving genetic identity, parentage, clone and strain identification, and taxonomic studies of closely related species. In addition, ISSRs are considered useful in gene mapping studies (Zietkiewicz *et al.*, 1994, Gupta *et al.*, 1994, Godwin *et al.*, 1997).

2.11. Molecular marker assisted plant breeding

Use of molecular markers is an effective genomic 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).

Markers located in the specific regions of a chromosome near the target gene, which effectively flag to breeders whether specific traits have been inherited, early identification of the most desirable genotypes will be possible (Figueira, 2004). Selection of tillering in wheat was done by marker-assisted Selection (Li *et al.*, 2010). Marker-assisted selection improves the efficiency of plant breeding (Janila *et al.,* 2013). This technology was also used to successfully pyramid resistance genes using marker assisted selection process (Mishra *et al.,* 2009; Pandey *et al.,* 2012 ; Varshney *et al.,* 2014).Validation of markers linked to disease resistance markers accelerates the process of introgression of into preferred peanut genotypes (Sujay *et al.,* 2012; Gajjar *et al.,* 2014).

2.12. Characterization of resistance genes using marker

Zhu *et al.* (2006) have reported Simple Sequence Repeat (SSR) markers in tomato associated with the late blight resistance allele. The crossing was made between 5 inbred susceptible lines and resistant line CLN2037E for developing F_2 population. Resistance is dominant and inherited as monogenic trait. Late blight resistance gene *Ph-ROL* has a distance of 5.7 cM from the SSR marker TOM236 and is positioned on chromosome 9.

Shan *et al.* (2010) developed markers in wheat governing resistance to powdery mildew; sequence-tagged sites (STS) marker was developed for *PmLK906* and *Pm4* using gene chip hybridization combined with bulked segregate analysis. Primers p9-7p1 and p9-7p2 resulting from *TaAetPR5*was used to amplify lines possessing *Pm4a*, *Pm4b* and *PmLK906*. *TaAetPR5* was linked to *PmLK906* at a genetic distance of 7.62 cM, and co segregated with *Pm4a*. The marker was found to co segregate with *Pm4a*, alleles at *Pm4* locus.

Zhang *et al.* (2011) developed AFLP markers co-segregated with gene *Lr24* (leaf rust resistance genes) in wheat and validated for marker assisted selection (MAS). Four AFLP markers, *P-AGA/M-CTT289 bp*, *P-AGC/M-CAC188 bp*, *P-AGC/MCAC162 bp*, and *P-ACG/M-CGC239 bp*, were co-segregated with *Lr24*. Markers 5R615, 5R616, 1R13, and 1R17 were identified and validated to contain gene *Lr24*.

Ren *et al.* (2012) reported a resistance gene for the stripe rust resistance designated as *YrC110*. A molecular map was constructed through SSR markers.

Ning *et al.* (2012) have reported that the leaf mold resistance gene *Cf*-10 in tomato has conferred resistant to prevailing physiological races of *Cladosporium fulvum* existing three northeastern provinces of China. Bulked segregate analysis was done and SSR and AFLP markers tightly linked to *Cf*-10 gene were obtained.

Lei *et al.* (2013) have developed the map of 2 genes, *Pi60* (t) and *Pi61*(t), associated to blast resistance in rice with F2 and F3 populations ensuing from a cross between the prone cv. Lijiangxintuanheigu (LTH) and resistant cv. 93-11.

Park *et al.* (2013) reported that in tomato TG328 and TG591 regions are tightly linked to the Ph-3 locus conferring resistance to late blight, by performing a mapping in F_6 families resulting from crossing of LB-resistant accession "L3708" (*Solanum pimpinellifolium*) with the LB-susceptible accession "AV107-4" (*S. lycopersicum*) using SSR marker system.

Jie *et al.* (2014) reported the fine mapping and candidate gene analysis of the resistance to soybean mosaic virus (SMV) in the F_2 population and near isogenic lines (nILs) from the residual heterozygous lines (RhLs) of Qihuang 1×nannong 1138-2. Resistance is controlled by a dominant gene (*RSC3Q*) which is located on chromosome number 13. Two genomic-simple sequence repeat (SSR) markers BARCSOYSSR-13-1114 and BARCSOYSSR-13-1136 with the distance of 651 Kb were found flanking the two sides of the *RSC3Q*.

Resistance to powdery mildew in wheat was governed by a dominant gene, *MlWE4*. Genetic linkage map shown that *MlWE4* is located on chromosome arm 5BL. Genes *MlWE4*, *Pm36* and *Ml3D232* have co-segregated with markers *XBD37670* and *XBD37680*, representing that they are same gene or alleles in the same locus (Dong *et al.*, 2015).

2.13. Molecular markers in cocoa

In cocoa many microsatellite markers are present (Kuhn *et al.*, 2003; Clement *et al.*, 2004 Pugh *et al.*, 2004) which are used in DNA fingerprinting, QTL analysis and Marker Assisted Breeding (MAB) programme. SSR markers in cocoa were initially used to study the genetic variation using various sets of primers (Saunders *et al.*, 2004, Efombagn *et al.*, 2006). Microsatellite markers have been developed and mapped in cacao (Lanaud *et al.*, 1999; Risterucci *et al.*, 2000; Lanaud *et al.*, 2004). The distribution of microsatellite markers within the linkage groups was not random.

2.13.1. RAPD markers in cocoa

The prior sequence information is not needed for RAPD and ISSR markers the survey of plant genomes, but has the problems associated with reproducibility and dominance. RAPD is the simple and cost effective marker system developed by Williams *et al.* (1990), and is proven valuable for a number of purposes.

RAPD analysis was the first multi-locus protocol based on Polymerase Chain Reaction (PCR) to be applied for the genetic characterization of cocoa (Wilde *et al. 1992*).

Genetic relationships among cacao population was determined using RAPD marker by (Russel *et al.*, 1993; Laurent *et al.*, 1993; Lerceteau *et al.*, 1997; Whitkus *et al.*, 1998). N'Goran *et al.* (1994) analyzed the genetic diversity of 106 genotypes in Cote D'Ivoire belonging to the various morphogeographic groups within Criollo, using 49 repeatable polymorphic RAPD products. Markers had shown a lucid structure among Forastero and Criollo groups with clear differences between upper and lower Amazon Forastero.

RAPD analysis was carried out using 10 polymorphic primers to distinguish cocoa accessions (Sane *et al.*, 2002). Accessions were efficiently identified with RAPD markers. Polymorphism was lowest in the Nigerian accessions and highest in the KAU collections. The UPGMA algorithm, based on Jaccard's coefficient, grouped 76 accessions into six groups based on genetic distance, and revealed four highly divergent accessions, *viz.* BE 10, EQX 78, I-56 and SCA 12.

Faleiro *et al.* (2004) have studied the genetic variability among 19 clonal accessions of *Theobroma cacao* L. from the Brazilian, Ecuadorian, and Peruvian Amazons, using RAPD and microsatellite markers. In assessment, 56 RAPD and 45 microsatellite markers were generated and cluster analysis, had shown that the different origin groups separately.

The RAPD markers are also proven capable for the discrimination of cocoa cultivars in relation to yield components. Molecular analysis showed a close

relationship with the genetic distance (GD) and Mahalanobis distance (MD) of yield component data, and between the biplots of GD and MD (Dias *et al.*, 2005).

Thirty clonal cocoa accessions showing resistance to witches broom disease were used for genetic diversity analysis using RAPD marker and pedigree information. The association between genetic similarities was very low (Santos *et al.*, 2005).

2.13.2. RFLP markers in cocoa

Laurent *et al.* (1993) used three types of RFLP probes to assess the genetic diversity of 203 genotypes. Ribosomal nuclear DNA probes distinguished Criollo, American Trinitario and Forastero. Mitochondrial probes revealed considerable variability among Criollo clones and cDNA probes confirmed the original structuring into Forastero and Criollo.

Lerceteau *et al.* (1997) have studid genetic variability in 155 cocoa trees which were belonging to different morphological group's using18 primers and 43 RFLP probes. RFLP variability observed within the species has reflected the hybridization and introgressions between trees of different origins. However, the Nacional type was different from Forastero, Criollo and Trinitario and genetically specific. The low heterozygosity rate was observed in some genotypes and it represented the original Nacional pool.

Risterucci *et al.*, (2000) performed linkage analysis in 181 cacao lines from a cross between Forastero and Trinitario, Linkage map comprised of 424 markers (Five isozymes, six gene loci, 65 genomic RFLPs, 104 cDNA RFLPs, three probes for telomeric region, 30 RAPD markers, 20 microsatellites and 191 AFLPs marker) with an average spacing of 2.1 cM. Markers were dispersed over ten linkage groups and covered 885.4 cM.

2.13.3. AFLP markers in cocoa

A linkage map was developed clones based on amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), and simple sequence repeat markers (SSR) markers in F_2 population resulted from the cross between the Scavina-6 and ICS-1 cocoa. A total of 273 markers, distributed in 14 linkage groups (LGs) were used and the Linkage group 9 has -one -QTL linked to butter content and for butter stiffness, two QTL were present in linkage groups 9 and 7. QTL associated for butter content and hardness mapped at the LG9 had 51.0 and 28.8 per-cent of the phenotypic variation, respectively (Araújo *et al.,* 2009).

2.13.4. SSR Marker in cocoa

Microsatellite markers are proven efficient to fingerprint and to resolve identity issues in cacao collections (Figueira, 1998; Risterucci *et al.,* 2001; Saunders *et al.,* 2004; Cryer *et al.,* 2006; Zhang *et al.,* 2006). Brown *et al.* (2005) have constructed the resistant gene mapping for witches broom disease in cocoa using the 170 SSR markers and they found two quantitative trait loci for resistance to witches broom disease, showing important dominance effect. Efombagn *et al.,*(2006) have analyzed the genetic diversity in cocoa using 13 SSR markers. When 282 alleles were analyzed farm accessions were differentiated according to the geographical area. Accessions from the East province clustered together with local Trinitario accessions from the genebank.

Nagai *et al.* (2009) have fingerprinted 100 cocoa trees using eleven SSR markers. The survey group was found to include Criollo, Trinitario, Forastero and their hybrid types, and 77 trees were genetically unique among the samples. The large genetic variation was observed among cacao trees grown in Hawaii.

SSR variation and individual identity in cocoa was analyzed using a capillary electrophoresis in 612 accessions of cacao collected from the Nanay Marañon, and Ucayali river systems of the Peruvian Amazon. The definite structure of genetic

diversity, stratified by the river systems of the Peruvian Amazon was confirmed by Bayesian clustering method and principal coordinate analyses (Zhang *et al.,* 2009).

Lima *et al.* (2010) have reported EST SSR makers for resistance to witches broom disease, clear polymorphism was obtained using 12 markers, having 47 alleles in total, with an average of 3.92 alleles per locus. The 11 genomic SSR markers have covered a total of 47 alleles, with an average of 5.22 alleles per locus. The relationship of genomic SSR with EST-SSR has improved the analysis of genetic distance between the genotypes.

Yang *et al.* (2011) have reported the *Theobroma cacao* cpSSRs markers for genetic variation within the Trinitario cultivars. Eight haplotypes were identified out of 95 accessions sampled. The pentameric repeat (CaCrSSR1) was the most polymorphic and it has proven as a complementary tool for genetic identification of hybrid cultivars between Forastero and Criollo accessions and among all cacao accessions.

Polymorphism between 27cocoa trees in the major cocoa cultivating regions of Tamilnadu was detected using ten SSR markers. Nearly 54 per-cent genetic similarity was observed and it has resulted in nine clusters (Thondaiman *et al.,* 2013).

Suwastika *et al.* (2015) have studied the genetics and clear pot morphology variation based on SSR marker in numerous clones of cocoa, collected from Central Sulawesi farms. The characterization of clone was done based on pod performance including; shape, size, bean properties. The polymorphism among the samples indicated high evolution rate in cacao trees over Sulawesi Island.

2.13.5. ISSR markers in cocoa

ISSR markers are considered useful in gene mapping studies (Godwin *et al.,* 1997). Genome analysis in cacao was performed using ISSR (Charters and Wilkinson, 2000).

Charters and Wilkinson, (2000) have characterized the cocoa genotypes using ISSR markers. Three pairs of the 62 accessions were distinguished and dendrogram of UPGMA was used to provide a measure of the genetic variability among genotypes.

Chia *et al.* (2011) have reported ISSR markers to differentiate 46 cacao accessions maintained in Tingo Maria -Perú. ISSR marker, in spite of its dominance nature, established clear grouping of Trinitario accessions into a common cluster.

Rivas *et al.* (2013) have used ISSR markers in three three geographically separate natural populations of *T. subincanum* for genetic diversity. Fifty nine individuals were characterized by Nei's genetic distance using 13 polymorphic primers.

2.14. Microsatellite markers for VSD resistance in cocoa

Two marker systems ISSR and SSR were used for characterization of the genotypes for VSD resistance. Cocoa genotypes *viz.*, VSD I- 4.6, VSD I- 4.11, VSD I- 5.8, VSD I- 6.9, G VI- 50, G VI- 52, G VI- 82, G VI- 100, G VI- 25, G VI- 53, G VI- 67, G VI- 144 with varying response to the VSD were selected for marker development. For variability study a total of 71 ISSR primers and 46 SSR primer pairs were used for screening. Thirteen ISSR primers were selected on the basis of amplification pattern. Among the thirteen primers UBC 811, UBC 815, UBC 826, UBC 857 and UBC 866 has given the polymorphism which is linked to VSD resistance in cacao. On amplification of genomic DNA from all genotypes with the primer UBC 811 amplicons obtained were distinctly polymorphic for the resistance to VSD. Marker UBC 811 has shown polymorphic bands of 0.950 kb in resistant, but absent in susceptible lines. This marker was able to differentiate resistant and susceptible lines. Using the ISSR primer UBC 815 amplicons obtained with this primer were found polymorphic for the VSD resistance. The polymorphic band was

observed at 750 bp in resistant lines, but absent in susceptible lines. This primer was candidate with the potential to differentiate between resistant and susceptible lines.

ISSR analysis of the DNA samples using the primer UBC 826 has shown variation. Amplicons obtained with this primer were polymorphic and these amplicon were linked with the VSD resistance. The polymorphic band of size 650 bp was present in all resistant lines and in one susceptible line (S2). This marker was associated with a gene with minor contribution to resistance. The polygenic nature of the resistance was observed in marker UBC 826.The ISSR primer UBC 857 has generated polymorphic amplicon with a band size of 450 bp in all the resistant lines but was absent in all the susceptible lines. This marker has shown clear polymorphism between the resistant and susceptible lines of cocoa. Amplicons obtained with the marker UBC 866 were polymorphic. Two polymorphic bands present at 1.300 kb (marked 'a') and 1.5 kb (marked 'b') in two susceptible lines but absent in all the resistant lines. It was found to be this marker linked with the gene offering resistance to VSD. ISSR marker UBC857 had shown 94 per cent identity with the *Theobroma cacao* microsatellite DNA clone of mTcCIR42 SSR (NCBI accession number AJ271944) (Chandrakant, 2014).

2.15. Proteins and phytocompounds involved in disease resistance

Hang *et al.* (1997) analyzed the inhibitory activity of proteins against aflatoxin. Two fractions from corn seeds were inhibitory to aflatoxin formation. Using a sensitive laboratory assay inhibition of fungal growth and inhibition of aflatoxin biosynthesis was examined. Aqueous extracts from seeds of Tex6, a corn inbred shown to be highly resistant to aflatoxin accumulation in field and laboratory evaluations. Two biologically active fractions were identified. One inhibited growth of *Aspergillus flavus* and, thus, aflatoxin accumulation, and the other inhibited aflatoxin formation with little effect on fungal growth. The compounds responsible for these activities appear to be proteaceous. Plant-derived compounds are inhibitors of aflatoxin biosynthesis and a few may be amenable to pathway engineering for

tissue-specific expression in susceptible host plants as a defense against aflatoxin contamination. Phytocompounds has shown promise as protectants during crop storage. Inhibitors with different modes of action could be used in comparative transcriptional and metabolomic profiling experiments to identify regulatory networks controlling aflatoxin biosynthesis (Holmoes *et al.,* 2008).

Systemin, hydroxyproline-rich glycopeptide systemins (Pearce *et al*, 2001) from solanaceous plants and *At*Pep1 peptide from *Arabidopsis* (Huffaker *et al.,* 2006) are involved in Defense mechanism. These peptides are 18 to 23 amino acids in length and are processed from wound- and JA-inducible precursor proteins, and play roles in the activation of local and systemic responses against wounding and pest attack. Systemin is synthesized from prosystemin and stored in the cytoplasm (Narvaez-Vasquez and Ryan, 2004).

A low potassium status in plant triggers expression of high affinity K1 transporters, up-regulates some K1 channels, and activates signaling cascades, some of which are similar to those involved in wounding and other stress responses. The molecules that signal low K1 status in plants include reactive oxygen species and phytohormones, such as auxin, ethylene and jasmonic acid (Ashley *et al*., 2006).

Plant defense metabolites arise from the main secondary metabolic routes, the phenylpropanoid, the isoprenoid and the alkaloid pathways. plant defense responses against pathogens and environmental pollutants may overlap, leading to the unspecific synthesis such as phenylpropanoids, tropane. Quinine and capthotecin are quinoline alkaloids and lysergic acid diethylamide (LSD) is an ergot alkaloid, all these arising from tryptophan these phytochemicals are involved both in resistance against pathogens and in tolerance towards abiotic stresses, such as atmospheric pollution (Iriti and Faoro, 2009) .

Materials and Methods

3. MATERIALS AND METHODS

The study on "Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)" was carried out at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Thrissur during 2014 to 2016.The research work, including the materials used and the methodologies adopted are discussed in this chapter.

3.1 Materials

3.1.1 Plant Materials

Twenty VSD resistant hybrids and four susceptible lines identified and maintained by, Cocoa Research Centre (CRC), College of Horticulture, Kerala Agricultural University (KAU), Thrissur, Kerala were selected for the present study. The details of the hybrids and clones used in this research are presented in Table 3.1.

Table 3.1 Details of the cocoa hybrids and clones used in the study

Sl. No.	Hybrids	Genotype	Response to vascular streak dieback disease
1	H1	M 13.12 X GVI 55	Resistant
2	H2	GII 19.5 X GVI 55	Resistant
3	H3	GVI 4 X GVI 55	Resistant
4	H4	GVI 126 X GIV 18.5	Resistant
5	H5	GVI 126 X GIV 18.5	Resistant
6	H6	GVI 126 X GIV 18.5	Resistant
7	H7	GVI 126 X GIV 18.5	Resistant
8	H8	GVI 126 X GVI 55	Resistant
9	H9	GVI 126 X GVI 55	Resistant
10	H10	GVI 126 X GVI 55	Resistant

Sl. No.	Hybrids	Genotype	Response to vascular streak dieback disease
11	H11	GVI 126 X GVI 55	Resistant
12	H12	GVI 137 X GVI 55	Resistant
13	H13	GVI 137 X GVI 55	Resistant
14	H14	GVI 140 X GVI 55	Resistant
15	H15	GVI 140 X GVI 55	Resistant
16	H16	GVI 140 X GVI 55	Resistant
17	H17	GVI 143 X GVI 55	Resistant
18	H18	GVI 143 X GV I55	Resistant
19	H19	GVI 143 X GVI 55	Resistant
20	H20	GVI 167 X GIV 18.5	Resistant
21	C1	G VI- 50	Susceptible
22	C2	G VI- 52	Susceptible
23	C3	G VI- 82	Susceptible
24	C4	G VI- 100	Susceptible

Resistant and susceptible hybrids are shown in (Plate :1 to 6).

3.1.2 Laboratory chemicals, glassware and plastic wares

The chemicals used in this study were of AR grade procured from Labcare, HIMEDIA and Merk India Ltd. The *Taq* DNA polymerase, *Taq* buffer and molecular weight marker (λ DNA */Hind III + Eco RI* double digest) were supplied by Invitrogen, RNase was supplied by Sigma, USA. The plastic wares used for the study were purchased from Tarsons India Ltd. and Axygen, USA.

H1-M 13.12 X GVI 55

H2- GII 19.5 X GVI 55

H3-GVI 4 X GVI 55

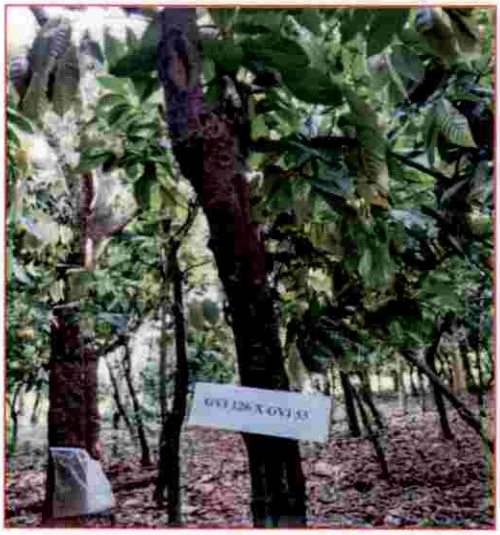
H4-GVI 126 X GIV 18.5

Plate 1: Resistant plants (genotypes 1 to 4)

H5-GVI 126 X GVI 18.5

H6-GVI 126 X GVI 18.5

H7-GVI 126 X GVI 18.5



H8-GVI 126 X GVI 155

Plate 2: Resistant plants (genotypes 5 to 8)

H 9-GVI 126 X GVI 155

H 10-GVI 126 X GVI 155

H 11- GVI 126 X GVI 155

H 12- GVI 137 X GVI 55

Plate 3: Resistant plants (genotypes 9 to 12)

H 13-GVI 126 X GVI 155

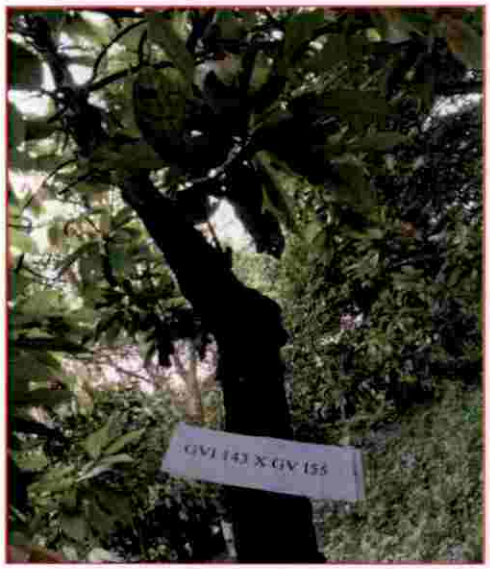
H 14-GVI 140 X GVI 55

H 15-GVI 140 X GVI 55

H 16- GVI 140 X GVI 55

Plate 4: Resistant plants (genotypes 13 to 16)

H 17- GVI 143 X

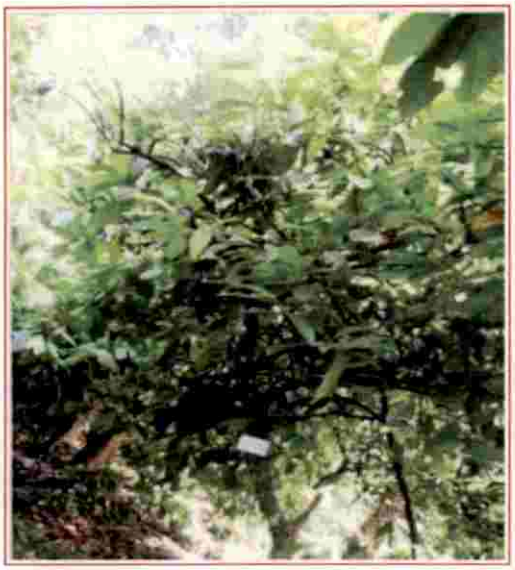


H 18-GVI 143 X

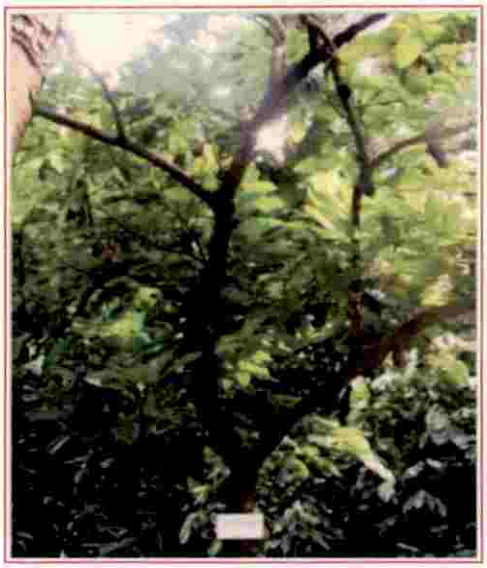
H 19-GVI 143 X

H20-GVI 143 X

Plate 5: Resistant plants (genotypes 17 to 20)



C1-G VI- 50



C2-G VI- 52

C3-G VI- 82

C4-G VI- 100

Plate 6: Susceptible plants (genotypes 21 to 24)

3.1.3 Equipment and machinery

The present research work was carried out using the molecular biology facilities and equipments available at CPBMB. Centrifugation was done in high speed refrigerated centrifuge (KUBOTA 6500, Japan). NanoDrop® ND-1000 spectrophotometer was used for the estimation of quality and quantity of DNA. Thermal cycling was done in the thermal cycler of Life Technologies (Proflex) and Agilent Technologies (super cycler 8800). Horizontal gel electrophoresis system (BIO-RAD, Wide Mini Sub cell GT, USA) was used for agarose gel electrophoresis. Gel Doc (BIO-RAD- Universal Hood) was used for imaging and documenting the agarose gel. Laboratory equipments used for the study are given in annexure I.

3.2 Methods

Validation and characterization of microsatellite markers linked to vascular streak dieback disease of cocoa was carried out in twenty VSD resistant hybrids and four susceptible lines.

3.2.1 DNA isolation

Young needle shaped leaves were selected as the ideal part for extraction of the genomic DNA. Tender, pale green leaves yielded good quality DNA in sufficient quantity. From individual plants, leaves were collected early in the morning. The collected leaves were quickly covered in aluminum foils and brought to the laboratory in ice flask. The surface was cleaned by washing with sterile water and wiping with 70 per cent ethanol and stored at -80°C. The modified Doyle and Doyle method suggested by Chandrakant (2014) was used for the extraction of genomic DNA. Concentration of extraction buffer and reagents is mentioned below, quantity of reagents required for DNA isolation are given in annexure II.

Reagents

I. CTAB Buffer (2X)

-2.0 per cent CTAB (w/v)

-100 mM Tris base (pH-8.0)

-20 mM EDTA (pH-8.8)

-1.4 M Nacl

- 2.0 per cent polyvinyl pyrrolidin (PVP)

- 0.2 per cent β-mercaptoethanol

II. TE buffer:

-10 mM Tris base (pH-8.0)

-1mM EDTA

III.7.5 M Ammonium acetate

IV. Wash buffer

- 70 per cent ethanol

- 10 mM ammonium acetate

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

VI. Chilled isopropanol

VI. 70 and 100 per cent ethanol

VIII. Sterile distilled water

Reagents I and III were autoclaved separately and stored at room temperature.

Procedure for DNA isolation

Using water bath, 1 ml of C-TAB isolation buffer (2X) was preheated in 2 ml centrifuge tube at 60°C. Fresh leaf tissue of 0.1g was ground with a pinch of polyvinyl pyrrolidin (soluble), 50 µl of β-mercaptoethanol and preheated CTAB isolation buffer using mortar and pestle.The crushed sample was incubated at 60°C for 30 minutes with occasional gentle swirling.After thirty minutes of incubation, sample was centrifuged at 12,000 rpm for 15 min at 4°C. Supernatant was taken into a new tube and equal volume of chloroform:isoamyl alcohol (24:1) was added and mixed gently by inversion. The sample was centrifuged (KUBOTA 6500) at 12,000 rpm for 15 minutes at 4°C. The content got separated into three different layers, aqueous top layer having DNA with small quantities of RNA, middle layer with proteins and lower layer containing chloroform, pigments and cell debris.

Top aqueous layer was transferred to a sterile centrifuge tube and equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed by inversion. Further, it was centrifuged at 12,000 rpm for 15 minutes at 4°C. After centrifugation, the aqueous phase was transferred into a clean tube and 0.6^{th} volume (300 µl) of chilled isopropanol was added and mixed by gentle inversions till the DNA was precipitated. These tubes were kept at -20°C for half an hour for complete precipitation. After thirty minutes of incubation, tubes were centrifuged at 12,000 rpm for 10 minutes at 4°C and supernatant was gently poured off. The DNA pellet was washed with 100 µl of wash buffer with centrifugation at 10,000 rpm for 5 minutes. Supernatant was removed and washed with 70 per cent ethanol. Tubes were then spun at 3000 rpm for 1 minute and ethanol was decanted. The DNA pellet was air dried, dissolved in 50µl sterile distilled water and stored at -20°C.

3.2.2 Assessing the quality of DNA by electrophoresis

The quality of isolated DNA was evaluated through agarose gel electrophoresis and gel profile was documented in Gel documentation unit.

Reagents and equipments

1. Agarose
2. 50X TAE buffer (pH 8.0)
3. 6X loading /tracking dye
4. Ethidium bromide (10 mg/ml)
5. Electrophoresis unit, power pack (BIO-RAD), gel casting tray, comb
6. UV transilluminator (Herolab®)
7. Gel documentation and analysis system (BIO-RAD)

Chemical composition of buffers and dyes are given in Annexure III.

Procedure for agarose gel electrophoresis of DNA

The gel casting tray and comb was swabbed with 70 per cent ethanol.The comb was placed in casting tray about 1 inch away from one end and positioned the comb vertically such that the teeth are about 1 to 2 mm above the surface of the tray. Agarose gel of 0.8 per cent was prepared in a conical flask by dissolving 0.8 gm of agarose powder in 100 ml 1X TAE buffer. It was then microwaved for 45 to 60 seconds until agarose was completely dissolved. The solution was allowed to cool under room condition and at this point 5 µl ethidium bromide (10 mg/ml) was added and mixed well. This warm gel solution was poured into the casting tray to a depth of 5 mm and the gel was allowed to solidify for 30-40 minutes at room temperature.

The comb was removed from the gel casting tray and kept in the electrophoresis chamber containing 1X TAE running buffer. Samples for electrophoresis were prepared by adding 1 µl of 6X gel loading dye for every 5µl of DNA sample and mixed thoroughly. Mixture of 6µl DNA and dye was added per well and the molecular weight marker (λ DNA *Eco*RI / *Hind* III double digest- Invitrogen)

was loaded in the first lane. Electrophoresis was carried out at 80 volt till the dye has migrated to two-third portion of the gel.

3.2.3 Gel documentation

The image was documented in gel documentation system (BIO-RAD) (BioRad Gel DOC-It^{TM}). The gel profile was examined for intactness, clarity of DNA band, presence of contamination such as RNA and proteins.

3.2.4 Purification of DNA

The DNA, which had RNA as a contaminant (as observed from the electrophoresis) was purified by RNase treatment and subsequent precipitation.

3.2.4.1. Reagents

I. Chloroform: Isoamyl alcohol (24:1 v/v)

II. Chilled isopropanol (100 per cent)

III. 70 per cent ethanol

IV. 1 per cent RNase (10 mg/ml)

(One per cent RNase solution was prepared by dissolving RNase (Sigma, USA) in TE buffer at 100°C for 15 minutes to inactivate residual DNase. The solution was cooled to room temperature, dispensed into aliquots and stored at -20°C.

Procedure for purification of DNA

To remove RNA contamination from DNA, 1 µl of 1 per cent RNase solution was added in 50 µl DNA sample and incubated at 37° C in a dry bath for 30 minutes. For removal of protein contamination, after 30 minutes of incubation the total volume was made up to 250 µl by adding distilled water. Equal volume of chloroform: isoamyl alcohol (24:1) was added and mixed gently. This mixture was centrifuged at 12000 rpm for 15 minutes at 4°C.The aqueous phase was collected into a new micro

centrifuge tube and equal volume of chloroform: isoamyl alcohol (24: 1) was added. Tubes were centrifuged at 12,000 rpm for 15 minutes at 4°C.

The aqueous phase was transferred into a clean centrifuge tube and 0.6^{th} volume of chilled isopropanol was added and mixed by gentle inversion till the DNA precipitated. For complete precipitation of DNA, tubes were kept at -20° C for half an hour. After centrifugation at 10,000 rpm for 15 minutes at 4° C. DNA pellet was washed with 70 per cent ethanol and centrifuged at 10,000 rpm for 5 minutes at 4° C, the pellet was air dried and dissolved in 50 µl sterile distilled water and stored at - 20° C. The samples were loaded on 0.8 per cent agarose gel at constant voltage of 70V to test the quality of the DNA.

3.2.5. Assessing the quality and quantity of DNA by NanoDrop® method

The quality and quantity of genomic DNA was estimated using NanoDrop® ND-1000 spectrophometer (NanoDrop® Technologies Inc,USA). Before taking the sample readings, NanoDrop was set to zero using 1µl autoclaved distilled water as blank. For quantification, one microlitre of DNA sample was measured at a wavelength of 260 nm and 280 nm and the OD_{260}/OD_{280} ratio were recorded to assess the purity of DNA. A ratio of 1.8 to 2.0 for OD_{260}/OD_{280} indicated good quality of DNA. The quantity of DNA in the pure sample was calculated using the relation 1 OD_{260} equivalent to 50 ng double stranded DNA/ml sample.

1 OD at 260 nm = 50 ng DNA/µl

Therefore OD at 260 × 50 gives the quantity of DNA in µg/ml.

Procedure for quantification of DNA using Nanodrop®

To check the concentration and purity of DNA, the NanoDrop spectrophotometer utilizes the software ND-1000. In the software, the option nucleic acid was selected. Sampling arm was opened and 1µl distilled water loaded onto the lower measurement pedestal. Spectral measurement was initiated after closing the sampling arm. The

sample column is automatically drawn between the upper and lower measurement pedestals and the spectral measurement is made. NanoDrop reading was set to zero with sample blank using distilled water. For taking the sample reading, 1µl of DNA sample was pipetted onto measurement pedestal and the measurement was done using the option 'measure'. After the measurement, the upper and lower pedestals are wiped using a soft laboratory wipe. Sample wiping prevents sample carryover in successive measurements.

3.3. Markers used for validation and characterization study

The markers reported for VSD resistance (Chandrakant, 2014) in cocoa using the ISSR primers UBC811, UBC815, UBC826, UBC857, UBC866 and SSR primer mTcCIR42 were used for validation and characterization as a result of their reproducibility. For every marker system, DNA from all the twenty four genotypes of cocoa was amplified with the reported primers. These markers have helped to obtain the amplification pattern in all genotypes.

3.3.1. DNA amplification conditions

The PCR conditions required for effective amplification in ISSR and SSR analyses included appropriate proportions of the components of the reaction mixture. The reaction mixture includes template DNA, assay buffer A, $MgCl_2$, *Taq* DNA polymerase, dNTPs and primers. The aliquot of this master mix was dispensed into 0.2 ml PCR tubes. The PCR was carried out in Thermal Cycler (Life Technologies, Proflex and Agilent). Temperature profile is an important factor in thermal cycler which affects the amplification pattern. The thermal cycler was programmed for desired number of cycles and temperatures for denaturation, annealing and extension step was adjusted for better amplification.

3.3.2. Inter Simple Sequence Repeat (ISSR) analysis

The good quality genomic DNA (30ng/µl) isolated from different genotypes of cocoa leaf samples were subjected to ISSR analysis. The reported primers were used for ISSR assay. PCR amplification was performed in a 20 µl reaction mixture and the composition of the reaction mixture consisted of,

a) Genomic DNA (30 ng/µl)	- 2 µl
b) 10X *Taq* assay buffer A	- 2.0 µl
c) $MgCl_2$	- 2.0µl
c) dNTP mix (10mM each)	- 1.5 µl
d) *Taq* DNA polymerase (3U)	- 0.4 µl
e) Primer (10 pM)	- 1.5 µl
f) Autoclaved distilled water	- 10.6µl
Total volume	- 20.0 µl

The amplification was carried out with the following program

94^0C for 2 minutes	-	Initial denaturation	
94^0C for 45 seconds	-	Denaturation	35 cycles
43^0C to 55^0C for 1 minutes	-	Primer annealing	
72^0C for 2 minutes	-	Primer extension	
72^0C for 8 minutes	-	Final extension	

4° C for infinity to hold the sample

The ISSR markers used for validation study are listed in Table 3.3.

The amplified product was separated on 1.8 per cent agarose gel using 1X TAE buffer stained with ethidium bromide along with marker (100 bp marker). The

profile was visualized under UV (312nm) transilluminator and documented in gel documentation unit for further analysis. The documented ISSR profiles were carefully examined for amplicon size.

Table 3.2: List of ISSR primers used for validation and characterization study

Sl. No.	Primer name	Sequence	Annealing temp. (°C)	Product length
1	UBC 811	5'GAGAGAGAGAGAGAGAC3'	43.3	950 bp
2	UBC 815	5'CTCTCTCTCTCTCTCTG3'	44.9	750 bp
3	UBC 826	5'ACACACACACACACACC3'	53.3	650 bp
4	UBC 857	5' - ACACACACACACACACYG-3'	52.7	450 bp
5	UBC 866	5'- CTCCTCCTCCTCCTCCTC- 3'	55	1300 bp and 1500 bp

3.3.3. Simple Sequence Repeat (SSR) analysis

The good quality genomic DNA (30 ng/µl) isolated from different genotypes of cocoa leaf samples was subjected to SSR analysis. SSR primers supplied by Sigma, USA was used for amplification of DNA. The amplification was carried out in Thermal cycler of Life Technologies (Proflex, Agilent). PCR amplification was performed in a 20 µl reaction mixture which consist of following components.

a) Genomic DNA (30 ng/µl)	-	2.0µl
b) 10X Taq assay buffer A	-	2.0µl
c) dNTP mix (10mm each)	-	1.5µl
d) *Taq* DNA Polymerase (3U)	-	0.3µl
e) Forward primer (10pM)	-	0.75µl
f) Reverse primer (10pM)	-	0.75µl
g) Autoclaved distilled water	-	12.7µl
Total volume	-	20.0µl

The PCR reaction was carried out with the following programme

94^0C for 3 minutes	-	Initial denaturation	
94^0C for 1 minute	-	Denaturation	35 cycles
40^0C to 64^0C for 1 minute	-	Primer annealing	
72^0C for 1 minute	-	Primer extension	
72^0C for 5 minutes	-	Final extension	

4^0C for infinity to hold the sample

The amplified products were separated on two per cent agarose gel using 1 X TAE buffer stained with ethidium bromide along with marker (1 kb). The profile was documented in gel documentation unit and visualized under UV transilluminator (312 nm). The documented SSR profile was carefully examined for amplification of DNA bands.

Table.3.3. List of SSR primers used for validation and characterization study

Sl. No.	Primer name	Sequence	Annealing temp. (° C)	Product length
1	mTcCIR 42	**F P** 5' TTGCTGAAGTATCTTTTGAC 3' **R P** 5'GCTCCACCCCTATTTG 3'	55	650 bp and 400 bp

3.3.4 Characterization of marker

Characterization of microsatellite marker was done by cloning of polymorphic band generated in resistant genotypes of cacao and After getting the sequence information from cloning result sequence analysis was done.

3.3.4.1. Elution of polymorphic band and cleaning

The polymorphic ISSR band was excised from the gel using a sterile, sharp scalpel avoiding much exposure to UV on a transilluminator. The band was eluted using SIGMA ALDRICH kit protocol. Concentration of eluted DNA was measured in NanoDrop ND 1000 and was checked on 0.8 per cent agarose gel and stored at-20°C for further cloning work.

1. **Excision of band**

The DNA fragment of interest from the agarose gel was excised with a clean, sharp scalpel .The excess amount of gel was trimmed to minimize the amount of agarose.

2. **Weighing of gel slices**

Gel slices were weighed in a tared colourless tube.

3. **Solubilization of gel.**

Volume of gel solubilization solution more than three times that of the gel slice was added to the tube containing gel slice. For every 100 mg of agarose gel, 300 µl of gel solubilization solution was added. The gel mixture was incubated at 50-60 °C for 10 minutes, until the gel slice was completely dissolved. It was further vortexed briefly every 2-3 minutes during incubation to dissolve the gel.

4. **Preparation of binding column**

Placed the GenElute binding column G onto the vacuum manifold. Vacuum was applied and 500 µl of the column preparation solution was added to the column. The column preparation solution was allowed to pass completely through the column.

5. **Colour checking of the mixture**

Once the gel slice was completely dissolved, the colour of the mixture changed to yellow (similar to fresh gel solubilization solution with no gel slice).

6. **Addition of isopropanol**

One gel volume of 100 per-cent isopropanol was added and mixed thoroughly until it became homogenous.

7. **Binding of DNA**

The solubilized gel solution mixture was loaded into the binding column and the mixture was allowed to pass through the column.

8. **Washing of column**

Wash solution of 700 µl was added to the binding column and allowed to pass through.

9. **Transfer of column**

Binding column was removed from the vacuum manifold and transferred to a clean new collection tube. It was further centrifuged for 1 minute at 12,000 to remove excess ethanol.

10. **Elution of DNA**

Binding column was transferred to a fresh collection tube. Elution buffer of 50µl was added to the center of the membrane and incubated for 1 minute and centrifuged for 1 minute.

3.3.4.2. Transformation of *E. coli* cells

Transformation of *E.coli* cells was carried out with the vector pUC 18 for competency checking

3.3.4.2.1. Preparation of competent cells

Competent cells possess altered cell wall with pores on them, through which the foreign DNA can pass through.

A single colony of *E. coli* strain DH5α from master plate was streaked on to a sterile LB plate and kept for incubation at 37°C for overnight. After incubation, on the next day 10-12 bacterial colonies were picked from the plate and streaked on to another LB plate and incubated for 16- 20 hours. Bacterial cells were then transferred the to an oakridge tube containing 50 ml of LB broth and kept for overnight at 37 °c in shaker incubator. After 24 hours of incubation, cells were kept on ice for 15 min and recovered by centrifugation at 3500 rpm for 10 min at 4°c. Supernatant was discarded and 30 ml of $MgCl_2$- $CaCl_2$ solution was added to the pellet. The tube was kept on ice for 15 min and the cells were recovered by centrifugation at 3500 rpm for 10 min at 4°C. Supernatant was removed and the pellet was resuspended in 2 ml of ice cold 0.1M $CaCl_2$. Tubes were further centrifuged at 3500 rpm for 10 min at 4°C. Supernatant was decanted and tubes were kept on ice. In oakridge tube, 2 ml of

glycerol solution was added. Dispensed the suspension as 150 µl aliquots to sterile chilled 1.5 ml tubes and stored at -80° C for further use.

3.3.4.2.2. Screening of competent cell

The prepared competent cells were screened to check their transformation efficiency, by transforming them using a plasmid (pUC18) containing ampicillin resistance marker. The procedure followed for competency checking is as given below.

Procedure

For competency checking, 50 ml of both LB broth and media was prepared. The competent cells were thawed on ice for 10 min. Dilution of pUC18 was made to 1:9 ratio. In thawed competent cells, 2 µl of pUC18 was added. The contents were mixed gently and kept on ice for 30 min. Meanwhile, the water bath was set to 42°C. The tube was rapidly taken from ice and a heat shock at 42°C was given exactly for 90 sec. Without shaking, tube was placed back on ice for 5min. Under aseptic conditions, 250µl of LB broth was added to a vial and the tubes were inverted twice to mix the contents. The tubes were then incubated at 37°C for 1 hour with shaking. The transformed cells were placed on LB agar/ampicillin (5µg/l) overlaid with IPTG and X-gal (80µl). The plates were incubated overnight at 37°C.

3.3.5. Cloning of DNA eluted from polymorphic band of UBC 811 marker

The ISSR polymorphic band eluted from resistant genotypes was cloned into pGEMT vector system supplied by Promega.

A) Preparation of ligation mixture

Appropriate amount of eluted product (insert) required for ligation was calculated by estimating the quantity of eluted DNA using NanoDrop® Spectrophotometer. The amount of DNA required for cloning was calculated by the following relationship.

$$\text{No. of insert} = \frac{\text{50 ng of vector x 0.95 Kb size of insert}}{\text{3kb vector}} \times \frac{3}{1}$$

Table. 3.4. Components of ligation mixture

Components	Quantity required per reaction
2X ligation buffer	5.0 µl
pGEMT vector (50ng)	1.0 µl
Eluted product	3.0 µl
T4 DNA Ligase (3units/µl)	1.0 µl
Total volume	**10 µl**

The reaction mixture was prepared by taking the above components and incubated for 1 hour at room temperature and further was kept at 4°C overnight.

B) Transformation of ligated product

Reagents

1. Ampicillin- 5 mg/ml in water
2. IPTG- 20 mg/ml in water
3. X-gal- 10 mg/ml in DW

The ligated PCR product was added to 100µl of thawed competent cells and kept on ice for 30 min. Heat shock was given at 42°C for 90 seconds in a dry bath and immediately placed back on ice for 5 min. LB broth (250µl) was added to the cells and incubated at 37°C for 1 h in a shaker at 160 rpm. The aliquots of transformed cells were placed on LB agar/ampicillin (5µg/ml)/ IPTG (40 µl)/ X-gal (40 µl) plates and incubated overnight at 37°C.

C) Colony PCR of recombinant cells

The putative transformants which appeared as white colonies were picked from the plate and streaked on grided LB agar plates containing ampicillin and Xgal/ IPTG. The plate was incubated overnight at 37°C. This was done to check whether the white colonies retained their colour and also get more colonies of each transformant. Then it was stored at 4°C for further use.

Procedure for colony PCR

A single white colony was picked out from the plate and mixed with 20 µl of distilled water. It was then heated at 98°C for 3 min and centrifuged at maximum speed for 2 min. Supernatant was then taken in new a tube and was used for PCR reaction. Colony PCR master mix consisted of the following components,

Components	**Quantity**
Genomic DNA	4.0 µl
10 X *Taq* buffer	5.0 µl
$MgCl_2$	2.5 µl
dNTPs	4.0 µl
Forward primer T_7 (10pM)	1.0 µl
Reverse primer Sp 6(10pM)	1.0 µl
Taq polymerase (3U)	0.4 µl

Autoclaved distilled water	32.1 µl
Total volume	50 µl

The thermal cycling was carried out with the following program

Initial denaturation	-	94°C for 2 minutes
Denaturation	-	94°C for 45 seconds
Primer annealing	-	55°C for 1 minute
Primer extension	-	72°C for 1 minute
Final extension	-	72°C for 5 minutes
Hold at 4°C for infinity		

After amplification, the PCR product was separated on 1 per cent agarose gel. In the agarose gel, the product that showed desired single band was sent for sequencing at SciGenom lab, Cochin.

3.3.6. Cloning of the eluted product from the polymorphic band generated from marker UBC 826 in resistant genotypes of cocoa

The polymorphic ISSR band of size 650 bp generated from marker UBC 826 in resistant lines was eluted. Concentration of eluted DNA was measured using NanoDrop ND 1000. Eluted DNA was reamplified to increase the concentration of DNA. Required concentration of DNA was calculated using the equation described in section 3.3.5.2.3. DNA insert was ligated into pGEMT vector and kept for 1hr at 37°C and 24 hr at 4°C for ligation. After complete ligation, it was transformed in competent cells. Cells having ligated product were streaked on LB medium containing ampicillin, X-gal and IPTG. Plates were kept for incubation at 37°C for 16 hr. After 16 hr of incubation blue and white colonies were developed on plates (Fig. 5.6). Colony PCR was done SP 6 and T7 primer, after screening PCR product was sent for sequencing.

3.3.7. Cloning and characterization of polymorphic band generated from marker UBC 857 in resistant genotypes of cocoa

The polymorphic band of size 450 bp generated from marker UBC 857 in resistant lines was eluted using *SIGMA ALDRICH* Gel and PCR clean up kit. The quality and quantity of the eluted DNA was checked. The required concentration of DNA insert was calculated using equation described in section 3.3.5.2.3. The eluted DNA was ligated into pGEMT vector and ligated product was inserted into competent cells. Cells having ligated product were streaked on LB medium containing ampicillin, X-gal and IPTG. Plates were kept for incubation at 37 °C for 16 hr. After 16 hr of incubation blue and white colonies were developed on plates (Fig 5.7). DNA insert was confirmed through colony PCR and PCR product was sent for sequencing

3.3.8. Sequence analysis

A) BLASTn analysis

The nucleotide sequence obtained from marker UBC 811, UBC 826 and UBC 857 was annoted in BLASTn programme to check the sequence homology.

B) ORF finder analysis

The nucleotide sequence obtained from marker UBC 811, UBC 826 and UBC 857, after sequencing was annoted in ORF finder for searching open reading frames (ORFs) present in the DNA sequence.

C) Identification of microsatellites

The frequency and distribution of SSR motifs, dimmers to decamers present in the sequence was determined. Detection and characterization of SSRs motifs between 2 to 10 base pairs was done.

Results

4. RESULTS

The study on "Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)" has been done at the Centre for Plant Biotechnology and Molecular Biology (CPBMB), College of Horticulture, Kerala Agricultural University. The objective of the study was to validate the identified SSR and ISSR markers linked to the gene/s offering resistance to vascular streak dieback disease in cocoa and to characterize the ISSR markers to develop corresponding SSR markers. The results of various aspects of the investigations are presented in this chapter.

4.1. Plant material

Cocoa Research Centre, Kerala Agricultural University have identified and maintained genotypes of cacao. Twenty four cocoa genotypes showing various responses to vascular streak dieback disease were used in the study. The genotypes were selected on the basis of field screening experiments started in 1998-1999. Details on the clones and their disease response are mentioned in Table (3.1).

4.2.1. Isolation, purification and quantification of DNA

Young leaves were selected as the ideal part for extraction of genomic DNA. The modified Doyle and Doyle method of (Chandrakant, 2014) was used for the isolation of genomic DNA. Treatment of β-mercaptoethanol and PVP was given to avoid phenolic oxidation and to remove the brown color of DNA pellet. To remove protein, chloroform: isoamyl alcohol (24:1) treatment was given two times and the DNA pellet was washed with ammonium acetate. The isolated DNA has shown RNA contamination (Fig 4.1 and 4.2). RNase treatment and further precipitation gave sufficient quantity of good quality DNA. The agarose gel electrophoresis has shown clear and discrete bands with no RNA contamination (Fig.4.3 and 4.4) and spectrophotometric analysis shown the acceptable ratio of UV absorbance (A_{260}/A_{280}) between 1.8 and 2.0 (Table.4.1).

Table 4.1 Quality and quantity of genomic DNA isolated from cocoa genotypes using NanoDrop® spectrophotometer

Sl. No.	Hybrids	UV absorbance at 260 nm (A_{260})	UV absorbance at 280 nm (A_{280})	Quantity (ng/µl)	$A_{260}/_{280}$	Quality
1	H1	8.00	4.25	400	1.88	Good
2	H2	19.13	10.09	956	1.89	Good
3	H3	22.29	12.03	1114	1.85	Good
4	H4	22.42	12.01	1121	1.86	Good
5	H5	21.58	11.82	1079	1.82	Good
6	H6	30.50	16.31	1525	1.87	Good
7	H7	17.50	9.40	875	1.86	Good
8	H8	23.34	12.96	1162	1.81	Good
9	H9	16.04	8.74	802	1.83	Good
10	H10	18.06	9.54	903	1.89	Good
11	H11	11.49	6.16	574	1.86	Good
12	H12	17.16	9.15	858	1.87	Good
13	H13	14.01	7.43	803	1.88	Good
14	H14	22.91	12.27	1145	1.86	Good
15	H15	9.21	5.03	460	1.83	Good
16	H16	18.14	10.01	907	1.81	Good
17	H17	10.71	5.82	541	1.84	Good
18	H18	15.09	8.10	754	1.86	Good

19	H19	9.90	5.25	495	1.88	Good
20	H20	28.65	15.64	1432	1.83	Good
21	C1	14.38	7.59	819	1.89	Good
22	C2	17.69	9.55	1054	1.85	Good
23	C3	8.35	4.58	416	1.82	Good
24	C4	19.05	10.09	952	1.88	Good

4.3. Molecular marker analysis

For molecular analysis two marker systems namely ISSR and SSR was used. Genomic DNA of twenty four cocoa genotypes was amplified separately with selected primers. ISSR and SSR assay was done using the valid protocols.

4.3.1 Inter Simple Sequence Repeat (ISSR) analysis

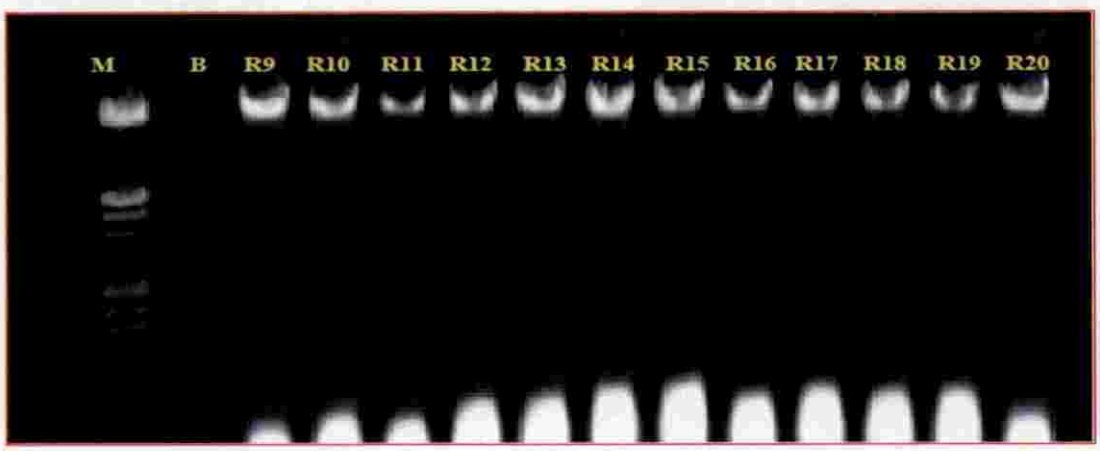
Five ISSR primers were employed for to the amplification of genomic DNA, with the thermal settings mentioned earlier under the material and methods section. Details on the primers are mentioned in Table 3.1.The analysis of result on individual primer is detailed here under.

4.3.1.1 UBC 811

Genomic DNA isolated from twenty four genotypes was diluted to 30 ng/µl. Diluted DNA of twenty four genotypes was amplified with primer UBC 811. Average of nine amplicons was obtained in each genotype, on DNA amplification. The pattern of amplification is shown in Fig. (4.5 and 4.6). The molecular weight of the amplicons varied from 300 bp to 2 kb. Amplicons obtained with this primer were distinctly polymorphic for the resistance to VSD. The polymorphic band of size 0.950

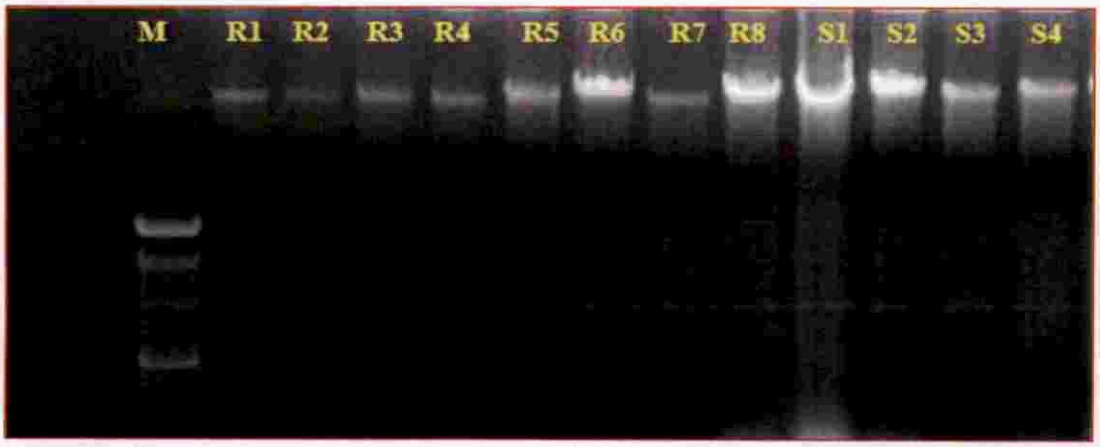
M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-**H7**, **R8**-H8, **S1**-C1, **S2**-C2 , **S3**- C3, **S4**-C4

Fig. 4.1. The gel profile obtained by electrophoresis of DNA samples isolated using modified Doyle and Doyle method, before RNase treatment (Genotypes R1 to S4)



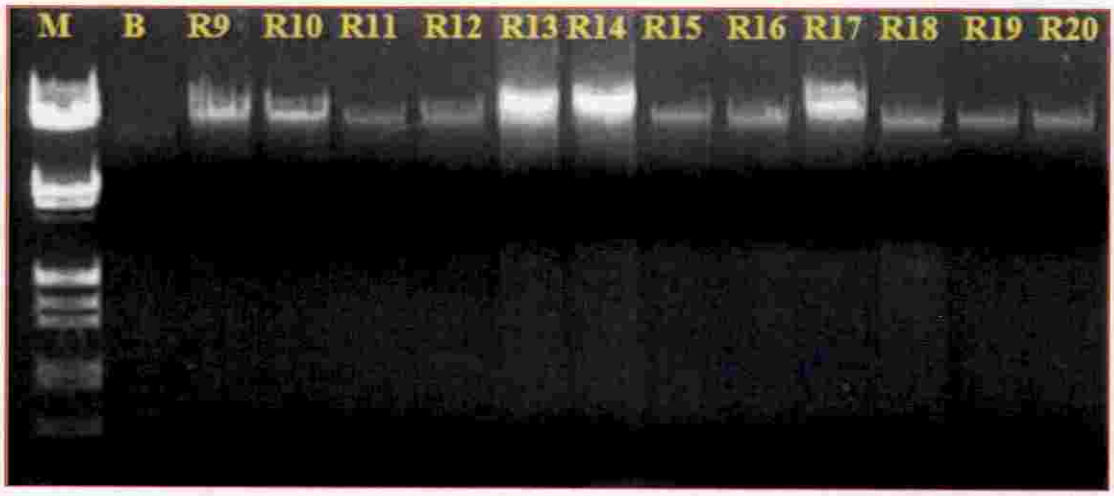
M-marker 100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.2. The gel profile obtained by electrophoresis of DNA samples isolated using modified Doyle and Doyle method, before RNase treatment (Genotypes R9 to R 20)



M-marker 100 bp, **B-** Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7-H7, R8**-H8, **S1**-C1, **S2**-C2 , **S3-** C3, **S4**-C4

Fig. 4. 3. The gel profile obtained by electrophoresis of isolated DNA samples using modified Doyle and Doyle method after RNase treatment (Genotypes R1 to S4)



M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.4. The gel profile obtained by electrophoresis of isolated DNA samples using modified Doyle and Doyle method after RNase treatment (Genotypes R9 to R 20)

kb was present in twenty resistant lines (R1 to R 20) but absent in four susceptible lines (S1 to S4) (Fig 4.5). This marker was able to clearly differentiate resistant and susceptible lines. The marker UBC 811 had shown the same results as previously reported. This ISSR assay was performed repeatedly and this marker was found highly reproducible.

4.3.1.2 UBC 815

DNA samples of twenty resistant and four susceptible lines were amplified with primer UBC 815. On amplification with the primer UBC 815, an average five amplicons were obtained in each genotype. The pattern of amplification is shown in (Fig. 4.7 and 4.8.). The molecular weight of the amplicons varied from 0.4 to 1.2 kb. On careful analysis, amplicons obtained with this primer were found polymorphic for the VSD resistance. The polymorphic band of size 750 bp was observed in all resistant lines (R1 to R 20) but it was absent in four susceptible lines (S1 to S4). This marker was found to be a good candidate with potential to differentiate between resistant and susceptible lines.

4.3.1.3 UBC 826

DNA samples of twenty four cocoa genotypes were amplified with the primer UBC 626. ISSR analysis of the DNA samples using the primer UBC 826 generated an average of ten amplicons per genotypes on 1.5 percent agarose gel. The molecular weight of the amplicons varied from 400 bp to 1.50 kb. The pattern of amplification is shown in (Fig. 4.9 and 4.10). Amplicon obtained from primer UBC 826 was polymorphic and the amplicon is linked with the VSD resistance. The polymorphic band of size 650 bp was present in all twenty resistant lines (R1 to R20) and in one susceptible line (S2) but absent in three susceptible lines (S1, S2 and S3). Marker UBC 826 had shown polygenic nature in ISSR analysis.

4.3.1.4. UBC 857

DNA samples of twenty four cocoa genotypes were amplified with primer UBC 857 for polymorphism study. The ISSR primer UBC 857 had generated an average of nine clear amplicons on 1.8 per cent agarose gel (Fig. 4.11 and 4.12). The molecular weight of the bands varied from 300 bp to 1.5 kb. Amplicon generated from this primer was polymorphic at a band size of 450 bp, the polymorphic band was present in twenty resistant lines (R1 to R 20) but it was absent in four susceptible lines (S1, S2, S3 and S4). This marker had shown clear polymorphism between the resistant and susceptible lines. This marker was highly repeatable and associated with the VSD resistance in cocoa.

4.3.1.5. UBC 866

Twenty four genotypes of cocoa were screened with the ISSR primer UBC 866, on amplification an average eight amplicons were obtained in each genotype. The pattern of amplification is shown in Fig. 4.13 and 4.14. Amplicon obtained with this primer was polymorphic. The molecular weight of the bands varied from 0.400 to 1.500 kb. Two polymorphic bands were present at 1.300 kb (marked 'a') and 1.5 kb (marked 'b') in two susceptible lines (S1 and S2) but absent in all twenty resistant and two susceptible line (S3 and S4). Marker UBC 866 found to be repeatable and linked with the gene offering susceptibility to VSD.

4.3.2 Simple Sequence Repeat (SSR) analysis

SSR primer set mTcCIR 42 was used to amplify the SSR regions in the genomic DNA of twenty four cocoa genotypes, with the thermal settings mentioned earlier, primer details has given in (Table 3.1).

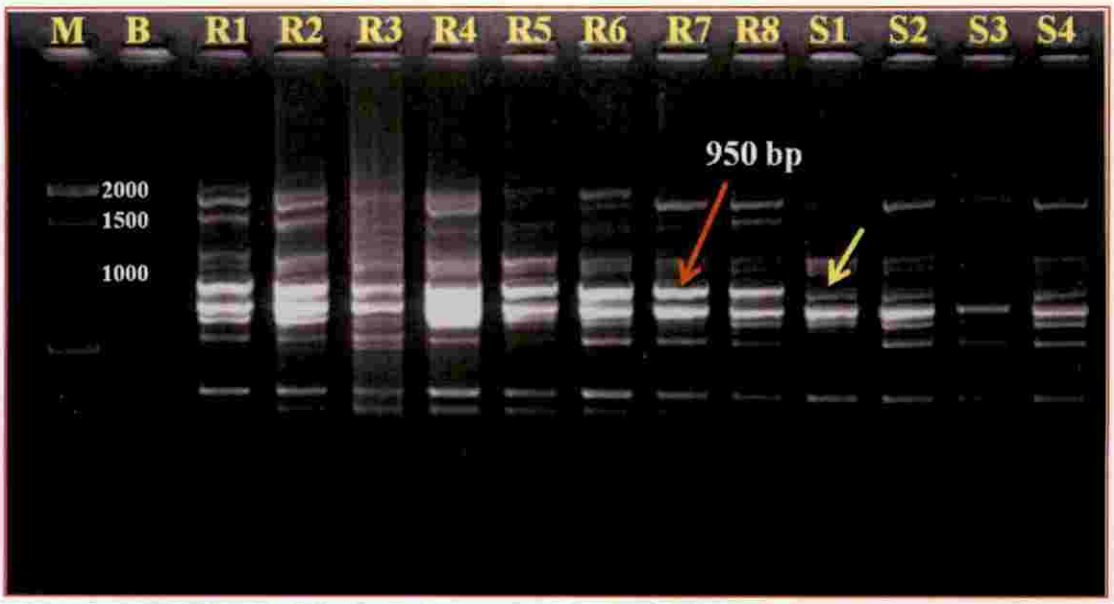
4.3.2.1. mTcCIR 42

SSR region in twenty four genotypes of cacao was amplified using the primer set mTcCIR 42 followed by electrophoresis on two per cent agarose gel. Three different amplicons were generated in gel. The molecular weight of the amplicons varied from size 200 bp to 500 bp (Fig.4.15 and 4.16). The amplicons generated were

polymorphic showing clear differentiation in resistant and susceptible lines. The polymorphic band of size 200 bp is tightly linked with VSD resistance. This polymorphic band was present in 13 resistant lines but it was absent in four susceptible and seven resistant lines. This marker had shown association with VSD resistance.

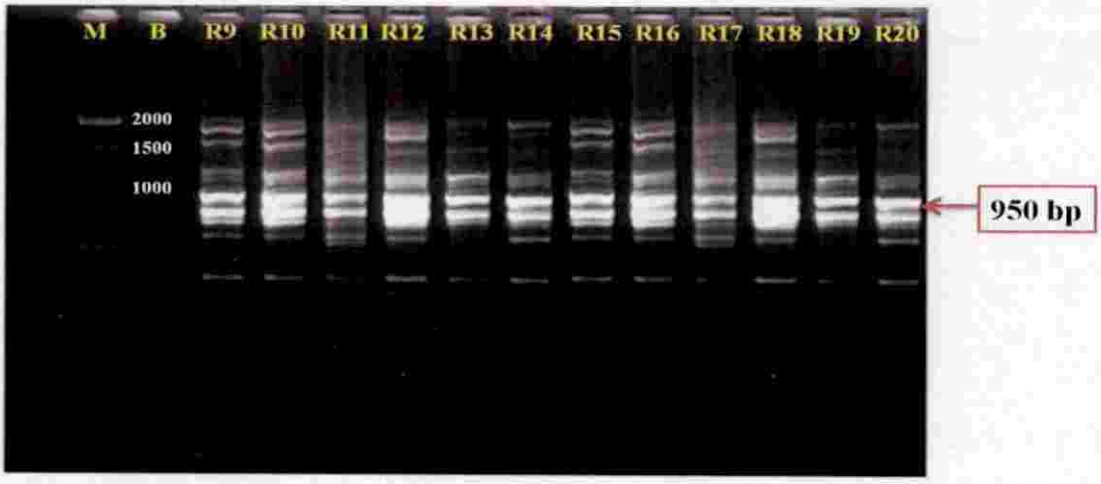
For every ISSR primer, the marker profiles generated through electrophoresis of the PCR products on 1.8 per cent agarose gel were documented and carefully examined. The polymorphic bands were identified in relation to the disease response. The ISSR primers UBC 811, UBC 815, UBC 826, UBC 857 and UBC 866 were repeatable and given reliable polymorphic amplicons. Markers were able to differentiate resistant and susceptible cocoa genotypes against VSD.

Amplification of twenty four cocoa genotypes with the primer UBC 811 has generated an average of ten amplicons. Among them six distinct and four faint band was present. Primer UBC 815 on amplification with cocoa genotypes had produced three distinct and two faint band in each line. When twenty four cocoa genotypes screened with the primer UBC 826 for VSD resistance an average of 10 amplicons were present including seven distinct and three faint band. Primer UBC 857 had generated seven amplicons in each genotype four distinct and three faint band were observed with this primer on agarose gel. On amplification with the primer UBC 866



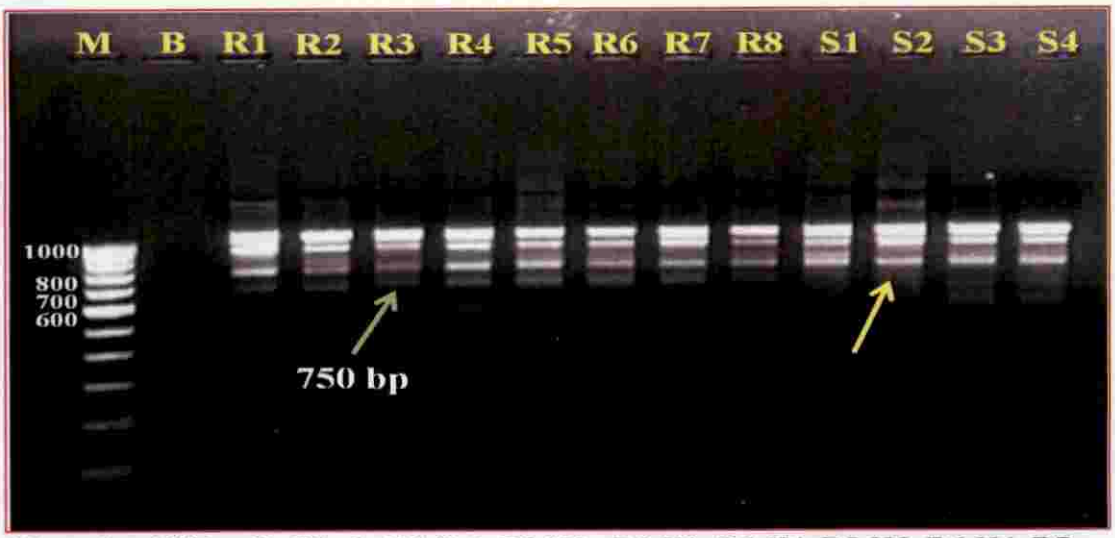
M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-**H7, R8**-H8, **S1**-C1, **S2**-C2 , **S3**- C3, **S4**-C4

Fig. 4.5. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 811 (Genotypes R1 to S4)



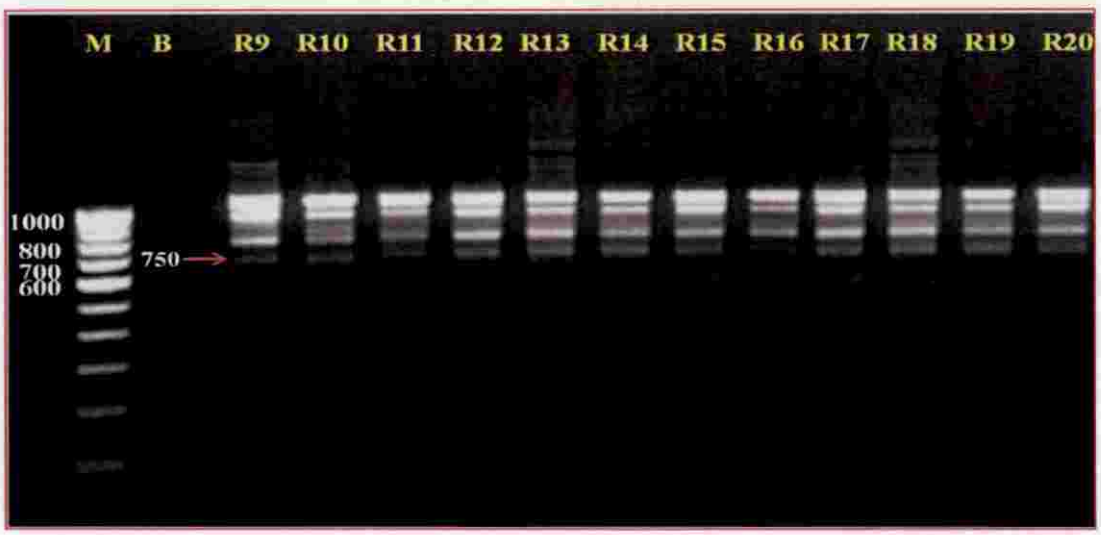
M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.6. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 811 (Genotypes R9 to R 20)



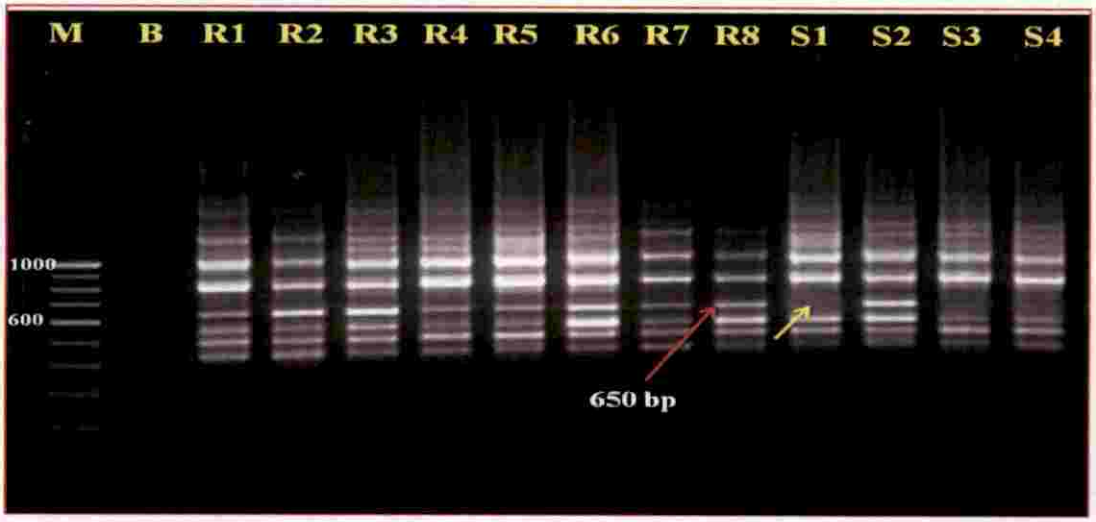
M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-H7, **R8**-H8, **S1**-C1, **S2**-C2 , **S3**- C3, **S4**-C4

Fig. 4.7. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 815 (Genotypes R1 to S4)



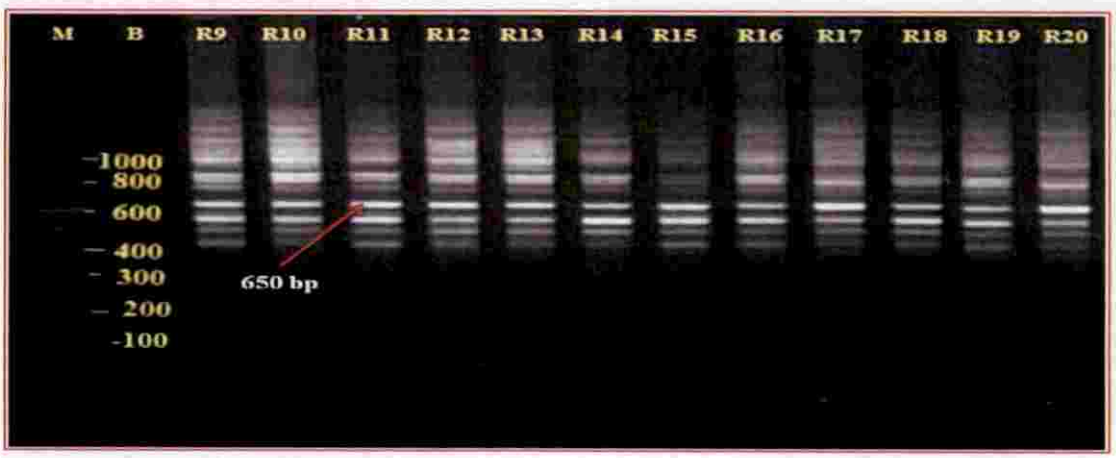
M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.8. Amplification pattern of twenty four cocoa genotypes generated with ISSR primer UBC 815 (Genotypes R9 to R 20)



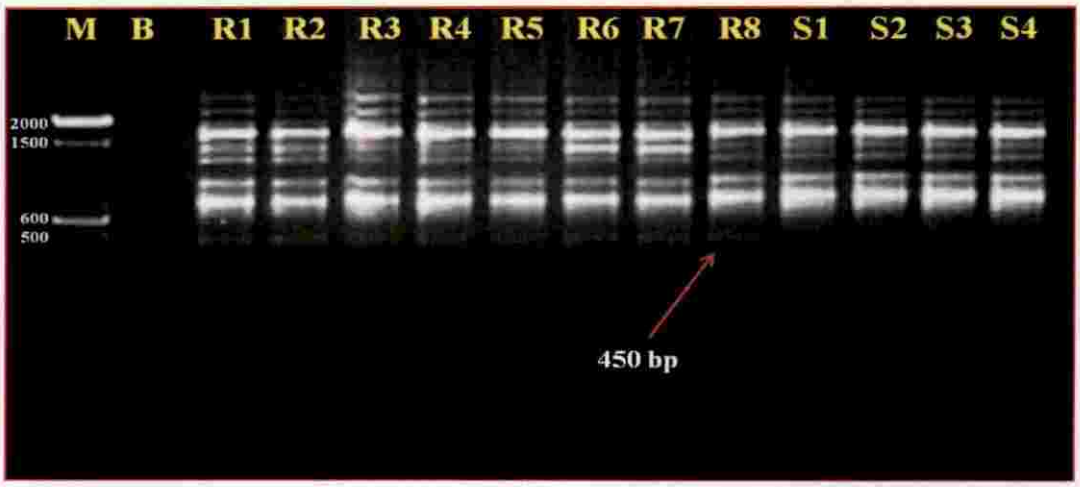
M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-**H7**, **R8**-H8, **S1**-C1, **S2**-C2 , **S3**- C3, **S4**-C4

Fig. 4.9. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 826 (Genotypes R1 to S4)



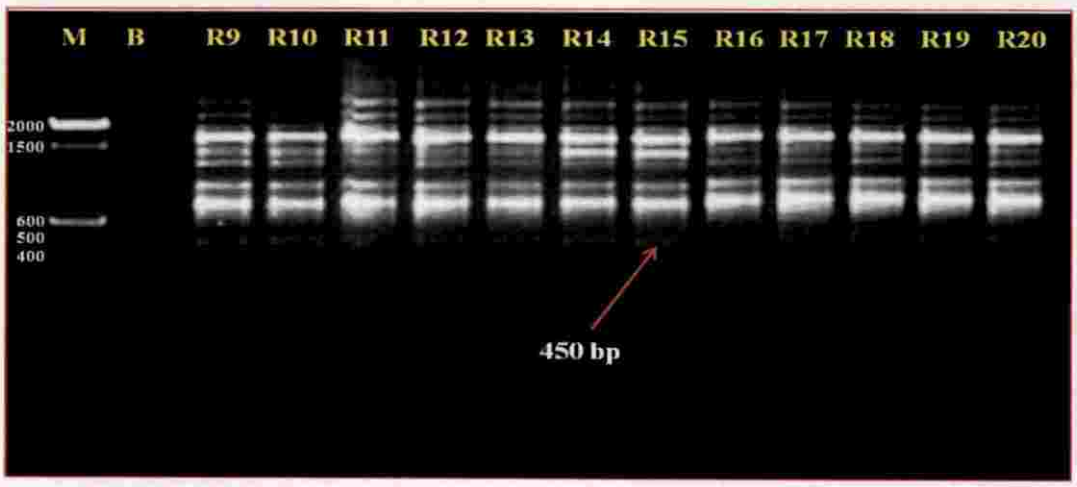
M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.10. Amplification pattern of twenty four cocoa genotypes generated with ISSR primer UBC 826 (Genotypes R9 to R 20)



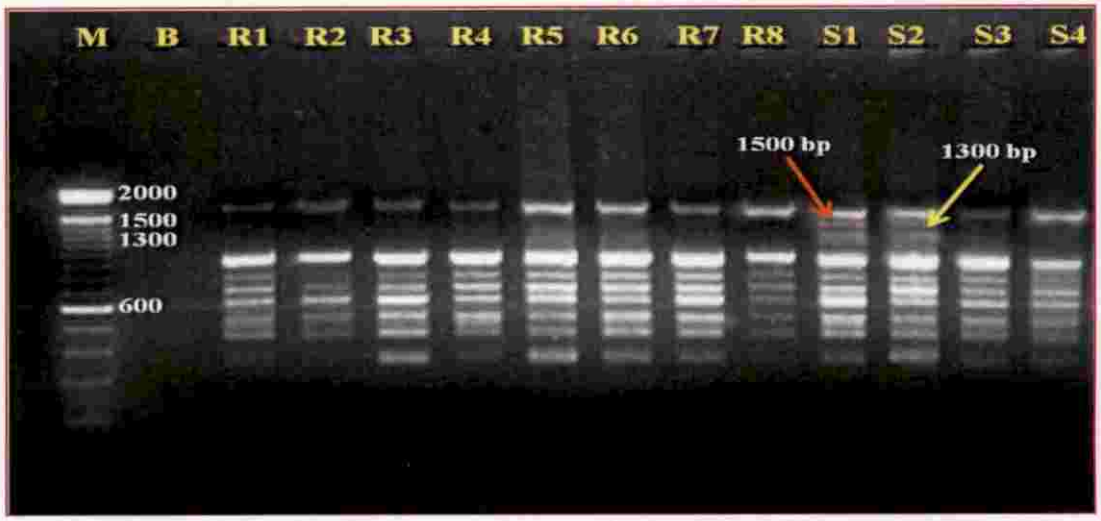
M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-**H7, R8**-H8, **S1**-C1, **S2**-C2 , **S3**- C3, **S4**-C4

Fig. 4.11. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 857 (Genotypes R1 to S4)



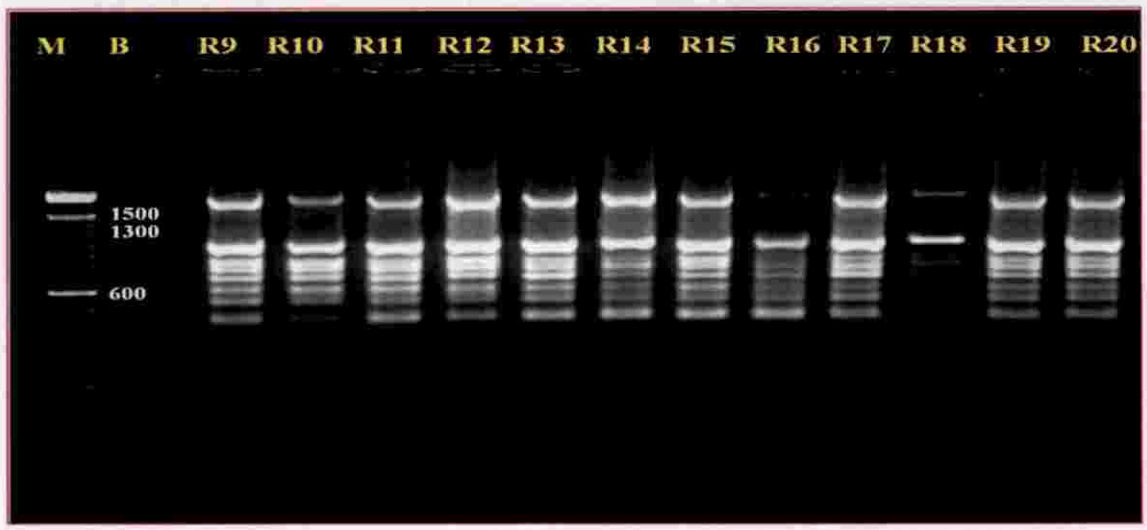
M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.12. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 857 (Genotypes R9 to R 20)



M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2 , **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-**H7, R8**-H8, **S1**-C1, **S2**-C2 , **S3**- C3, **S4**-C4

Fig. 4.13. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 866 (Genotypes R1 to S4)



M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.14. Amplification pattern of twelve cocoa genotypes generated with ISSR primer UBC 866 (Genotypes R9 to R 20)

M-marker 100 bp, **B**- Blank, **R1**-H 1, **R2**-H2, **R3**-H3, **R4**-H4, **R5**-H5, **R6**-H6, **R7**-**H7**, **R8**-H8, **S1**-C1, **S2**-C2, **S3**- C3, **S4**-C4

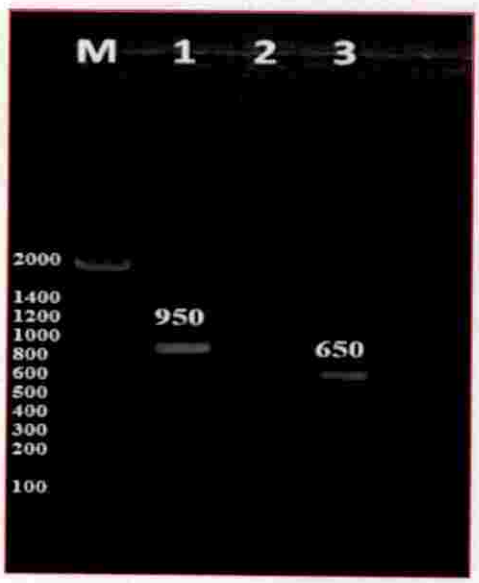
Fig. 4.15. Amplification pattern on twelve cocoa genotypes using the SSR primer mTcCIR 42 (Genotypes R1 to S4)

M-marker100 bp, **B**-Blank, **R9**-H9, **R10**- H10,**R11**-H11, **R 12**-H12, **R13-H13,R14-H 14, R 15-H15, R16**-H16, **R17**-H17, **R18**-H18, **R19**-H19, **R20**-H20

Fig. 4.16. Amplification pattern on twelve cocoa genotypes using the SSR primer mTcCIR 42(Genotypes R9 to R 20)

M-marker-100bp, **B**-Blank,**Lane-1, 2 and 3**-UBC 857 DNA insert.

Fig. 4.17. Colony PCR of recombinant plasmid with DNA insert of marker UBC 857 from resistant lines



M- marker100 bp, **Lane1**- UBC 811, **Lane 3**-UBC 826

Fig.4.18. Reamplification of eluted DNA of polymorphic band generated from marker UBC 811 and UBC 826 in resistant genotypes

Table 4.2 Details of amplification with the 5 primers for ISSR assay in cocoa genotypes

Sl. No.	Primer name	Amplification pattern		
		No. of bands	Type of band	
			Distinct	Faint
1	UBC 811	10	6	4
2	UBC 815	5	3	2
3	UBC 826	10	7	3
4	UBC 857	7	4	3
5	UBC 866	8	6	2

Table 4.3. Details of the amplification pattern of SSR marker mTcCIR 42 generated in 24 genotypes of cocoa

Sl. No.	Primer	Amplification pattern		
		No. of Bands	Type of band	
			Distinct	Faint
1	mTcCIR 42	3	3	0

in twenty four cocoa genotypes had produced an average of eight amplicons including six distinct and two faint band. SSR primer set mTcCIR, on amplification with twenty four genotypes had produced two distinct and one faint band on agarose gel. This SSR primer had shown clear polymorphism.

Table 4.4. Details of microsatellite markers present in twelve genotypes of cacao

Sl. No.		Genotype	Response to VSD	Microsatellite marker					
				ISSR					SSR
				UBC 811	UBC 815	UBC 826	UBC 857	UBC 866	mTcCIR 42
1	H1	M 13.12 X GVI 55	Resistant	+	+	+	+	–	–
2	H2	GII 19.5 X GVI 55	Resistant	+	+	+	+	–	+
3	H3	GVI 4 X GVI 55	Resistant	+	+	+	+	–	+
4	H4	GVI 126 X GIV 18.5	Resistant	+	+	+	+	–	+
5	H5	GVI 126 X GIV 18.5	Resistant	+	+	+	+	–	–
6	H6	GVI 126 X GIV 18.5	Resistant	+	+	+	+	–	–
7	H7	GVI 126 X GIV 18.5	Resistant	+	+	+	+	–	–
8	H8	GVI 126 X GVI 55	Resistant	+	+	+	+	–	–
9	H9	GVI 126 X GVI 55	Resistant	+	+	+	+	–	+
10	H10	GVI 126 X GVI 55	Resistant	+	+	+	+	–	+
11	H11	GVI 126 X GVI 55	Resistant	+	+	+	+	–	+
12	H12	GVI 137 X GVI 55	Resistant	+	+	+	+	–	–

Table 4.5. Details of microsatellite markers present in twelve genotypes of cacao

Sl. No.	Hybrids	Genotype	Response to VSD	Microsatellite marker					
				ISSR					SSR
				UBC 811	UBC 815	UBC 826	UBC 857	UBC 866	mTcCIR 42
13	H13	GVI 137 X GVI 55	Resistant	+	+	+	+	–	–
14	H14	GVI 140 X GVI 55	Resistant	+	+	+	+	–	+
15	H15	GVI 140 X GVI 55	Resistant	+	+	+	+	–	+
16	H16	GVI 140 X GVI 55	Resistant	+	+	+	+	–	+
17	H17	GVI 143 X GVI 55	Resistant	+	+	+	+	–	+
18	H18	GVI 143 X GV I55	Resistant	+	+	+	+	–	+
19	H19	GVI 143 X GVI 55	Resistant	+	+	+	+	–	–
20	H20	GVI 167 X GIV 18.5	Resistant	+	+	+	+	–	+
21	H21	G VI- 50	Susceptible	–	–	–	–	+ +	–
22	H22	G VI- 52	Susceptible	–	–	+	–	+ +	–
23	H23	G VI- 82	Susceptible	–	–	–	–	–	–
24	H24	G VI- 100	Susceptible	–	–	–	–	–	–

4.4. Cloning and characterization of polymorphic band generated from marker UBC 811 in resistant genotypes of cocoa

A) Competency checking of *E-coli* DH 5α cells for cloning of UBC811 marker

The competent cells were prepared as per the procedure given in section 3.3.5.2.1. Competency was checked by transforming with the plasmid vector (pUC18) having ampicillin resistance gene. A large number of blue colonies were developed on LB which indicated a high degree of transformation efficiency. Thus the competent cells prepared were found ideal for cloning (Fig 5.5.a).

B) Ligation

The eluted ISSR product (insert) required for ligation was calculated using the equation described in section 3.3.5.2.3. Insert DNA having concentration 59 ng was used per 50 ng of vector for ligation reaction using pGEMT vector.

C) The transformation of *E. coli* cells with the ligated product of UBC811

The ligated product was used to transform the prepared competent cells using the heat shock method and was incubated at 37°C. Large number of blue and white colonies was developed after overnight incubation (Fig.5.5.b.) The transformed cells were cultured in LB/ampicillin media, overlaid with X gal and IPTG.

D) Screening of the transformed colonies

The transformed colonies grown on LB medium containing ampicillin were screened for recombinant plasmid. Plasmid was isolated from transformed cells and PCR reaction was done with isolated plasmid using SP6 and T7 primer. After screening PCR product was sent for sequencing.

4.5. Cloning and characterization of polymorphic band generated from marker UBC 826 in resistant genotypes of cocoa

The polymorphic ISSR band of size 650 bp generated from marker UBC 826 in resistant lines was eluted. Concentration of eluted DNA was measured using

NanoDrop® ND 1000. Eluted DNA was reamplified to increase the concentration of DNA. Required concentration of DNA was calculated using the equation described in section 3.3.5.2.3. DNA insert was ligated into pGEMT vector and kept at 37° C for 1hr and 24 hr at 4°C for ligation. After complete ligation, it was transformed in competent cells. Cells having ligated product were streaked on LB medium containing ampicillin, X-gal and IPTG. Plates were kept for incubation at 37 ° C for 16 hr. After 16 hr of incubation blue and white colonies were developed on plates (Fig 5.6). The developed white colonies were screened for confirmation of DNA insert. Plasmid was isolated from transformed colonies and PCR was done with isolated plasmid using SP6 and T7 primer. DNA insert was confirmed and PCR product was send for sequencing to Scigenom Cochin.

4.6. Cloning and characterization of polymorphic band generated from marker UBC 857 in resistant genotypes of cocoa

The polymorphic band of size 450 bp generated from marker UBC 857 in resistant lines was eluted using *SIGMA ALDRICH* Gel and PCR clean up kit. The quality and quantity of the eluted DNA was checked. The required concentration of DNA insert was calculated using equation described in section 3.3.5.2.3. The eluted DNA was ligated into pGEMT vector and ligated product was inserted into competent cells. Cells having ligated product were streaked on LB medium containing ampicillin, X-gal and IPTG. Plates were kept for incubation at 37 °C for 16 hr. After 16 hr of incubation blue and white colonies were developed on plates (Fig 5.7). DNA insert was confirmed through colony PCR and PCR product was sent for sequencing.

4.7. Sequence analysis

The nucleotide sequence obtained from marker UBC 811, UBC 826 and UBC 857 after sequencing was annoted in BLASTn programme to check the sequence homology.

Fig. 4.19a. Competency checking of *E-coli* DH 5α cells for cloning of UBC811 marker

Fig.4.19b. The transformation of *E. coli* cells with the DNA eluted from polymorphic band of marker UBC811

Fig.4.20. The transformation of *E. coli* cells with the DNA eluted from polymorphic band of marker UBC826

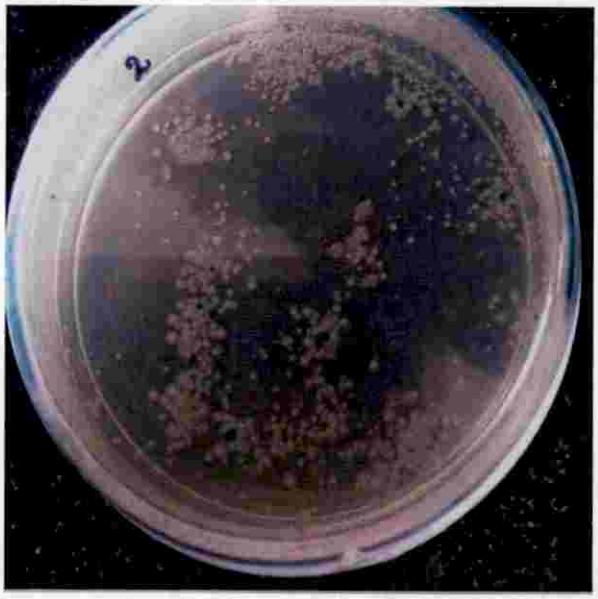
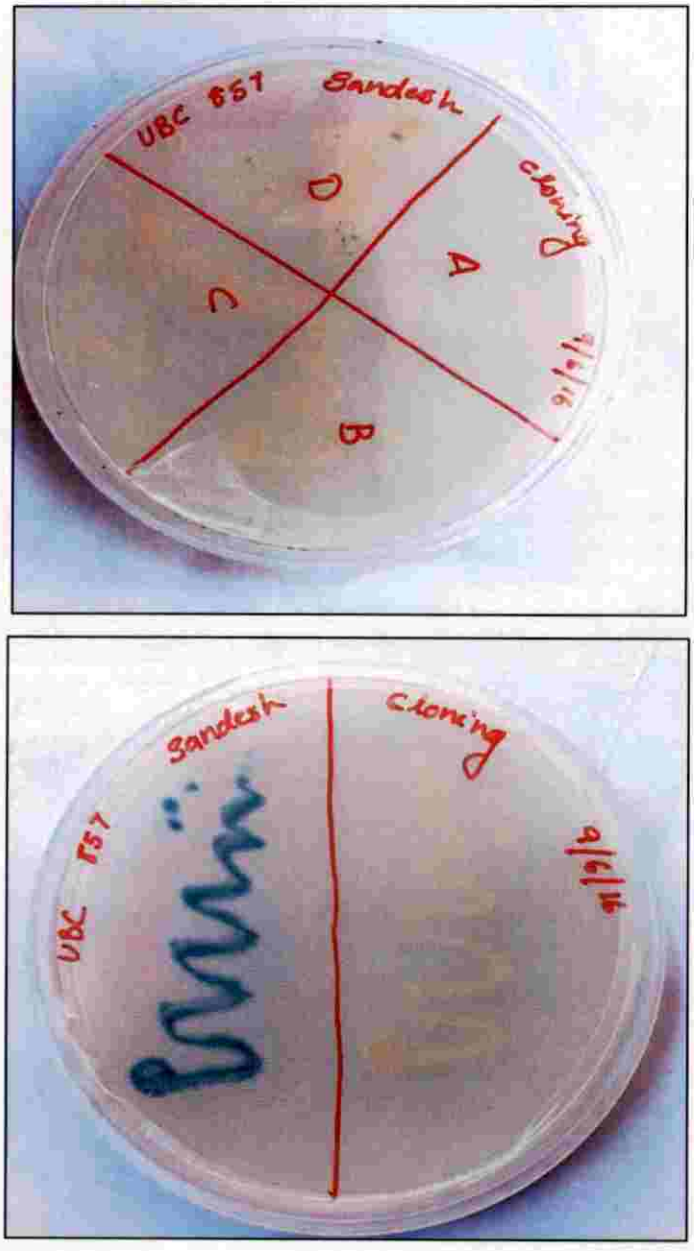


Fig. 4.21.a.The transformation of *E. coli* cells with the DNA eluted from polymorphic band of marker UBC857



4.21.b. Maintenance of recombinant clones on LB medium

4.7.1. BLASTn analysis

4.7.1. 1.BLASTn analysis of sequence obtained from marker UBC 811

A) Sequence 1

>UBC811- A-(927 bp)

5'TGCTCACGTCATAGGCATGAGTTGCATTTGATTGATTTATGTTGATATTT
GCATCCCTATGACTCAATACGTTACTTTTTCTCATGGGGCAGATCGAGCAC
ATGCAACCAAAAGTTAGTAAACATTAATGTGTTAAACAACCCTAAAAACC
CATAAGTTAGGGAAGATGCCACTTCTAATTGACCTTCAAGCATAGTTATC
CCACGCTGCACGTCTTCAGCTTGCCTAGAGTGAACATGTGCCCATGTGAG
CCGCGTGATGCTTTCACTCTTCCGCTCCTCCGAATCACTCTTAGCAGCATC
TGCGACCTCTCTCTCACAGGACTTCCACCATCGCCACAGTCTATATCAAAT
GCCACAATAATGAATGGGGAAGCAAACGTGTGAATCGACGAAATACCTA
AATCCTAGTTACATACGACAATAACATACAAAACGTTTCATTTTCCTGCCA
CTTTCTTATAATTCTAAACTTCAGGATTGCTATGATGTCGAGCATTTAGGA
CTTGCGTGTGAAATTCGGATCTCTACTGTGGCTGCTTACAGTCTATCCTGA
TATCTGATATATACCTCCAAGACTCAAAATCTTTTGATGTTTTTTTCTCTTA
GTCTAAGGGGAGACAAACAAAGATGAGCGAAATAAAAAACCCCGTATGA
GATTTTAACCTATCATGGGACAACGGCATACAGAAAATGCCAAGTAAATT
GGATAATTTTCAAAGAGGAGTGGAAAACCGGCCTGATACACTACAACAT
GATGTGGCGTCGCACGGAGTATTGGTTCAGAACTACTAAGACGAGCAGAC
GTACTTCCAAAATTTCCGATCTCGTTGCAGTATCAAGATATAAGGCACAT
CAGCAAGTGGACGTGCAATAGCTAGCATTATGCATTTTCCACTGACTCAC
CCACACACCATTCTCGTGTGT3'

The sequencing of marker UBC 811 had yielded 927 and 623 nucleotides. The BLASTn analysis of the UBC 811-A (927 bp) nucleotide sequence was found to have 82 per cent identity with *Theobroma cacao* genome assembly chromosome V (Genbank accession number LT 94792.1). BLAST analysis of sequence has shown in (Fig4.22a, b, c).

B) Sequence 2

>UBC 811-D(623 bp)

5'CCAAACCGGCACTCACACACGCGCACACACACACACACTCACACAC
ACTCACAGACACACACACACACACACAATCACTCACACCCACACACACAC
GCACATGCACACACACACACACACACACACACACACACACACACACACA
CACACACACACACACTCACACCCACACAGTTATGGAAGATTCACACACAC
CCACACCTTCACACACAGATATCCAAAGCCGCACGCCCTCAGCTTGCCAA
GAGTCGCAAGTGTCACTCAAACACGCGCGCAGCCTTCACTCTTCAGGTCC
TCAGAATCTTCTTAGCAGCATCTGAAACCTCTCTCTCACAGGCCTTCCACT
GTCACCGGAGTCAATATTAAATGCCAAAATAATGAATGGGACGCAAACG
TGTGTATCCACGTTCTACTCAATCCTAGTAAATTACGACAACAACACACA
AAACGTTTCATTTTCCTGCCACGTTCTTATAAATCTAAACTTAAGGATTGC
TATGATGTCAAGCATTTTGGACTTGCTTGTGATTTCAAATCTCTAGCGGGG
GGGCTGACAGTGTATCCTAACACCTGATATAGACTTCCCAACACTGAACA
TCTGTCTGAGGTGTTTTTTCTCAC3'

The BLASTn analysis of the marker UBC 811-D (623 bp) nucleotide sequence had shown 84 per cent identity to *Theobroma cacao* genome assembly chromosome V (Genbank accession number LT 94792.1).Blast analysis of sequence has shown in (Fig.4.23a, b, c).

4.7.1.2. BLASTn analysis of sequence obtained from marker UBC 826

A) Sequence 1

>UBC 826-A (587 bp)

5'GGTACGTGCACTCATAGCGCCTGGTTCACTTCACTGATTGAACATGCAG
AGTGTCATAATACTTGTTTCAACCCCCCCTTGCCCCCCCTCTTCATATTCC
AGAATCCGGCTTAAAAAAAGGGTTTTTTATGTAATTTTATGTTAATTTTAG
GTCTTGTAAGTTATTTTTAAAAGACTTTCAAAGTCTTTTAAAACATTTTTA
GGTTTTTTTCATGGATTTTTAAGCATTTTTAAGGTCCTTTTAAGGATATTTA
AGGTTTGAATAGGATAGGTGATGAAAATCTTATAAATGTTGTTATGTGTA
GGGGGCGGGTTGGGATGTGTCGAGGTCAACATGGAAAAGGCCCCTAGAG
CGGAGGAAAGCACAAAGAGGATACCGGTTTTCTTTAGGAGAATAGGCAG
AAGAAAAACGAGGGAATTCCAGAGGCTGGAACTGTGGGGGGTTGTGTTG
GGGGGAAGGCTGGAGCAGGCTTTTTAGAGATTTTTCGAGGAATATGAAAT
ATTTTTTTTATGTTTTTAAGTGAATTTTAATTTTTGTTAATGATTTTGTTAC
AATTTTTAAGGATA3'

The sequencing of marker UBC 628 had yielded 587 and 568 nucleotides. In BLASTn analysis of the UBC-826-A (587 bp) nucleotide sequence had shown 99 per cent identity with *Theobroma cacao* genome assembly, chromosome: III (Genbank accession number LT594790.1). Blast analysis of sequence has shown in (Fig.4.24a, b,and c)

B) Sequence 2

>UBC-826-B (568bp)

5'CTTGCCCCCCCTCTTCATATTCCAGAATCCGGCTTAAAAAAAGGGTTTTT
TATGTAATTTTATGTTAATTTTAGGTCTTGTAAGTTATTTTTAAAAGACTTT
CAAAGTCTTTTAAAACATTTTTAAGGTTTTTTTCATGGATTTTTAAGCATTT
TTAAGGTCCTTTTAAGGATATTTAAGCTACGGACCGTTCGTACGCGACGC
ATGGAGGCGCCCTTCACCATTGAACATGTGGTGCCGATACTTGCCACCAA
CTCAACACAGTATCCTTCGTAGGTTCTCTCTAACGCACTTACCATCCGAAC
CCGCGCGACGGCTCTTTAGACATCTTGAACGAGAATATATTGGTGCTTCT
GACTTATGCGACGATATTTGTTTACTTACATGGTAGCAACCGGATCGCTCC
GCGTAGAAAACCACTAGCTGCTGTAGTGAGGTTGTGGCAGAGGCTGATCT
CTACCCGACCTCGCCACCGCGCCACCCTCTGGGTCGCTATCGACTGCAAA
TGGTCTACAGCCTGGATTTTAAAACTACATAATACGAGATGCGGCCAAAT
CTTGTTATTGGAGTTGCTGATTGTGTTTTTG3'

The BLASTn analysis of the sequence UBC-826-B (568 bp) had shown 100 per cent identity with *Theobroma cacao* genome assembly, chromosome: III (Genbank accession number LT594790.1). Blast analysis of sequence has shown in (Fig.4.25a, b,and c)

4.7.1.3. BLASTn analysis of sequence obtained from marker UBC 857

>UBC857-2-(299 bp)

5'CTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTAGTGATTGA
CTGCGTACCAATTCACGGGAACATGTAAGTCGGACCGTAAAAGAAACTT
ATTCCTTCGAAGCAAGGAAACAGCATATATTCTGTTTTCAAAAACTCACA
CACCCTAAGGTTACAACAGACTGTTAGAGCTGAAGCCTGCAATTATAAC
TTCAGAATCGAAGCAAAAGTTTCGCTGTTACTCAGGACTCATCAATCGA
ATTCCCGCGGCCGCCATGGCGGCCGGGGAGCATGCGACGTCGGGCCCAA
TTCGC3'

Sequencing of marker UBC 857 had yielded 299 nucleotides. The 299 bp nucleotide sequence of UBC 857 had shown 93 per cent identity to *Theobroma cacao* genome assembly, chromosome: III with (Genbank accession number LT94790.1). BLAST analysis of sequence had shown in (Fig.4.26a, b, and c)

4.7.2. ORF analysis

The nucleotide sequence obtained from marker UBC 811, UBC 826 and UBC 857, after sequencing was annoted in ORF finder for searching open reading frames (ORFs) present in the DNA sequence. The programme had given the range of each ORF, along with its protein translation.

4.7.2.1. ORF analysis of sequence UBC811-A (927 bp)

A) UBC811- A (927 bp) sequence 1

ORF analysis of sequence had shown ten ORF were present in the sequence UBC811-A (927 bp) (Fig.4.27), among these ten ORFs five ORFs had contributed to uncharacterized protein and five ORF had shown identity with proteins and enzymes deposited on PDB, Swissprot database (Table 4.6). ORF-1, ORF-2, ORF-3, ORF-6 and ORF -9 had 34, 40, 49, 30 and 33 amino acids respectively and, these five ORFs had contributed to uncharacterized protein. ORF-4 had 33 amino acids and it had 60 per cent identity with chain a crystal structure of fold bifunctional protein from *campylobacter Jejuni* (PDB ID 3P2OA). ORF 5 had 32 amino acids and it had shown 50 per cent identity with F box DNA helicase 1, F BOX protein 18 (Swissprot ID F1 ND48.2). ORF-7 had 30 amino acids and it had shown 73 per cent identity with chain X crystal structure archaeal intron encoded homing endonuclease I Tsp 061 I (PDB ID 2DCHX). ORF-8 had 28 amino acids and it had shown 78 per cent identity with chain A crystal structure of putative histidinol phosphate aminotransferase NP 281508.1 from *Campylobacter jejuni* At2.01 A (PDB ID 3GETA). ORF-10 had 33 amino acids and it had shown 47 per-cent identity with chain A crystal structure of carboxy peptidase γ inhibitor complexed with the cognate protease (PDB ID 1WPXA).

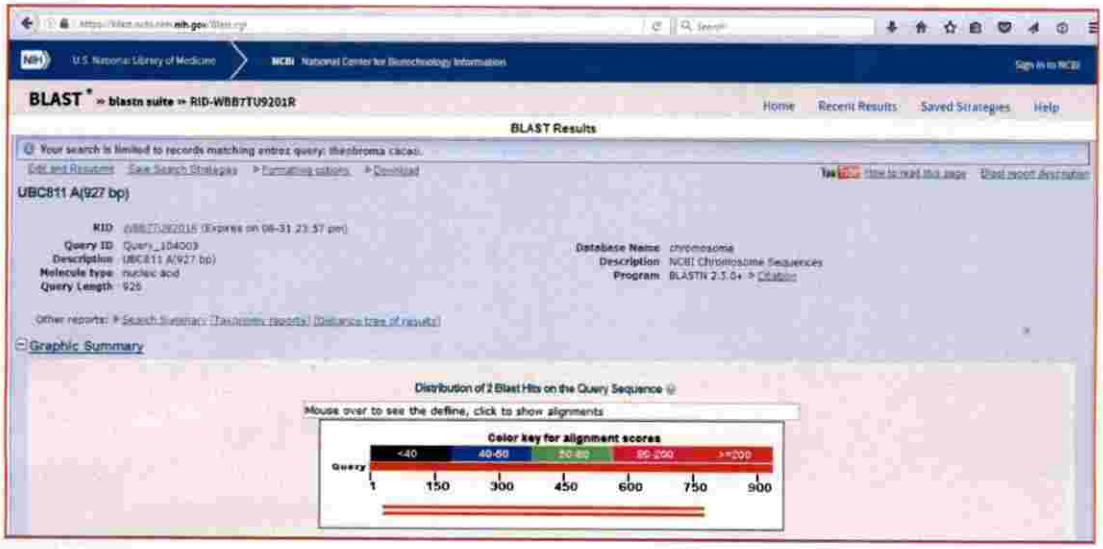


Fig. 4.22.a Graphical summary of nucleotide sequence similarity through BLASTn of the polymorphic band obtained with primer UBC 811.

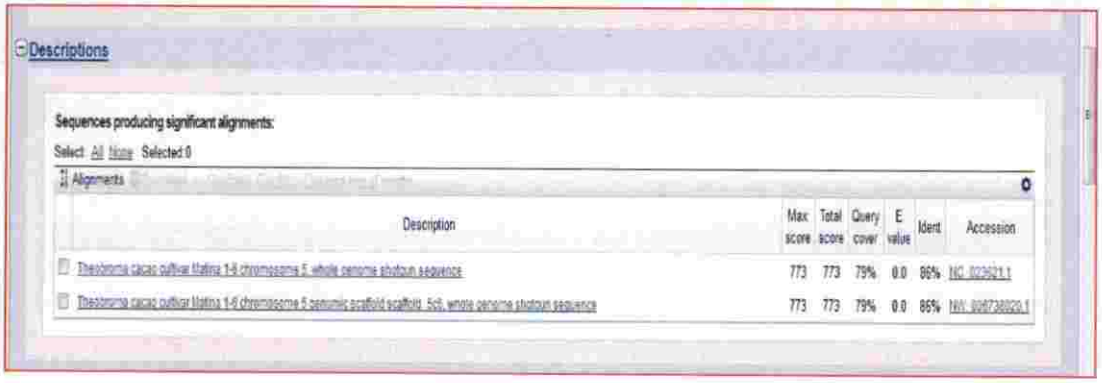


Fig. 4.22.b Results of the similarity search of sequence generated from the polymorphic band obtained with primer UBC 811, using BLASTn

Fig. 4.22.c. Alignment of the 927 bp sequence generated from the polymorphic band obtained with primer UBC 811

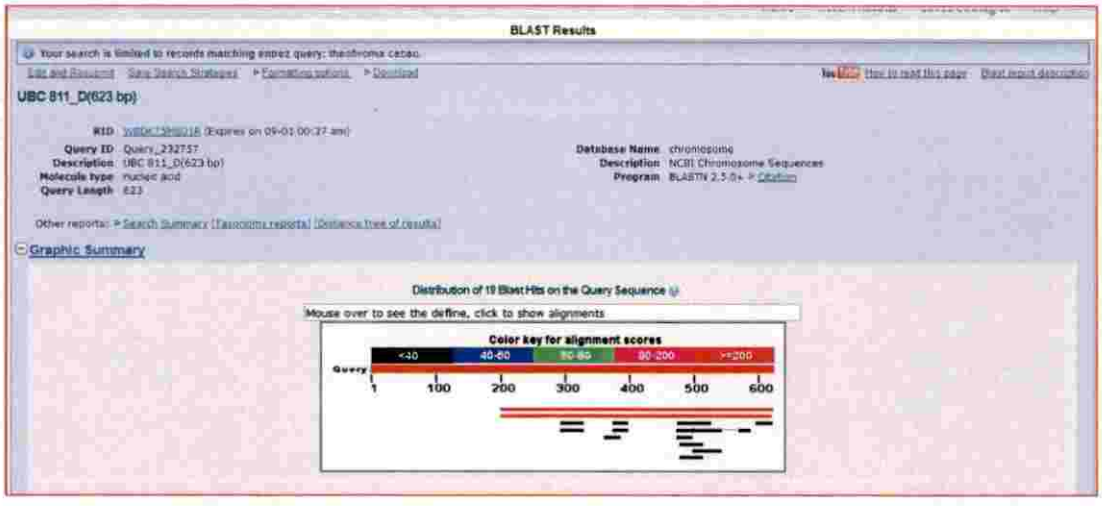


Fig. 4.23a Graphical summary of nucleotide sequence similarity through BLASTn of the polymorphic band obtained with primer UBC 811

Fig. 4.23b Results of the similarity search of sequence generated from the polymorphic band obtained with primer UBC 811, using BLASTn

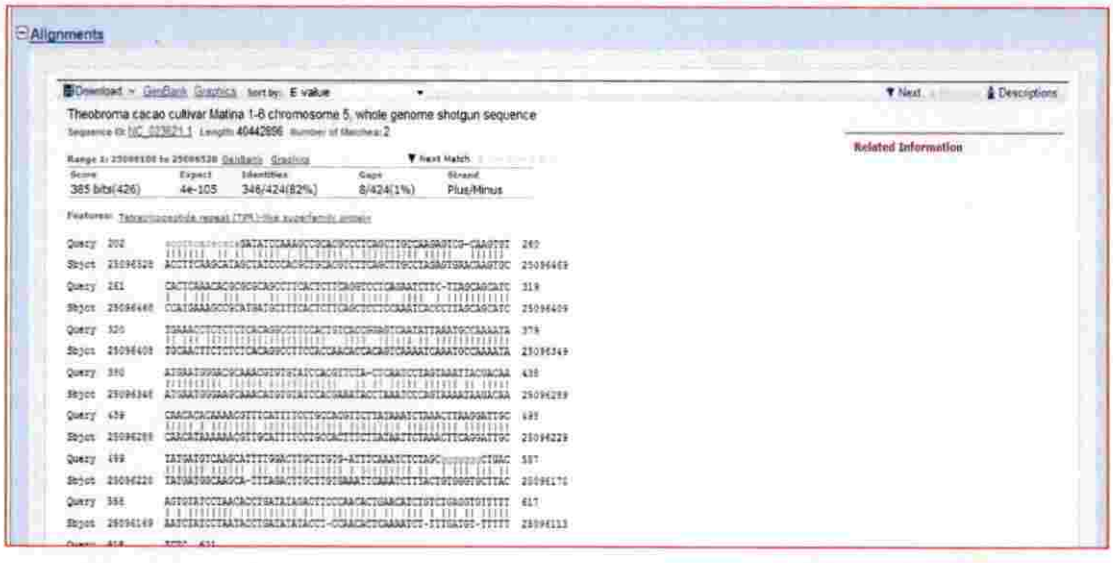


Fig. 4.23c Alignment of the 623 bp sequence generated from the polymorphic band obtained with primer UBC 811

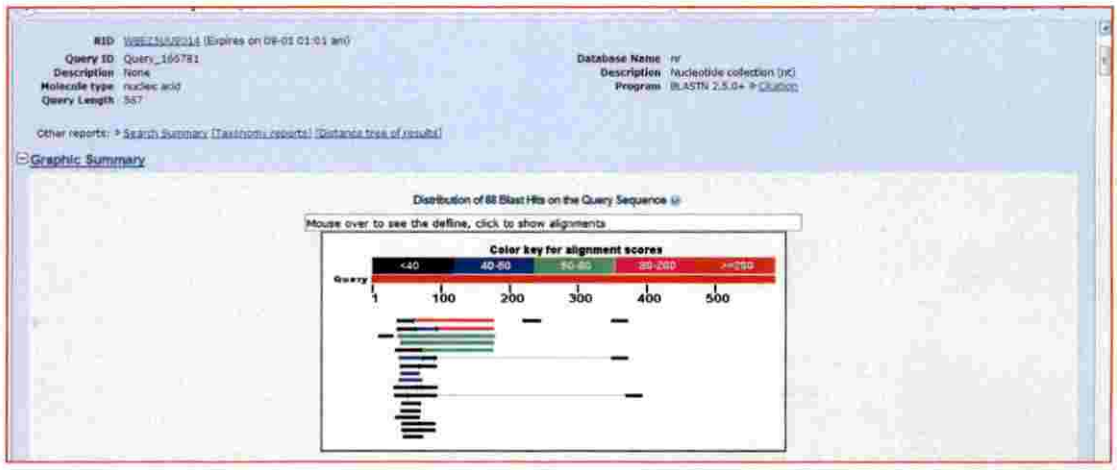


Fig.4.24.a Graphical summary of nucleotide sequence similarity through BLASTn of the polymorphic band obtained with primer UBC 826

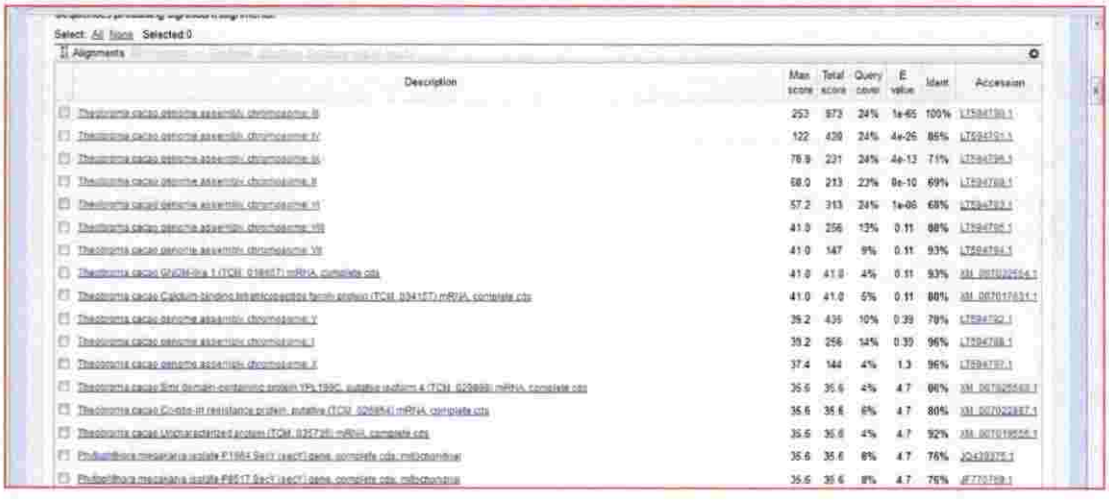


Fig. 4.24b Results of the similarity search of sequence generated from the polymorphic band obtained with primer UBC 826, using BLASTn

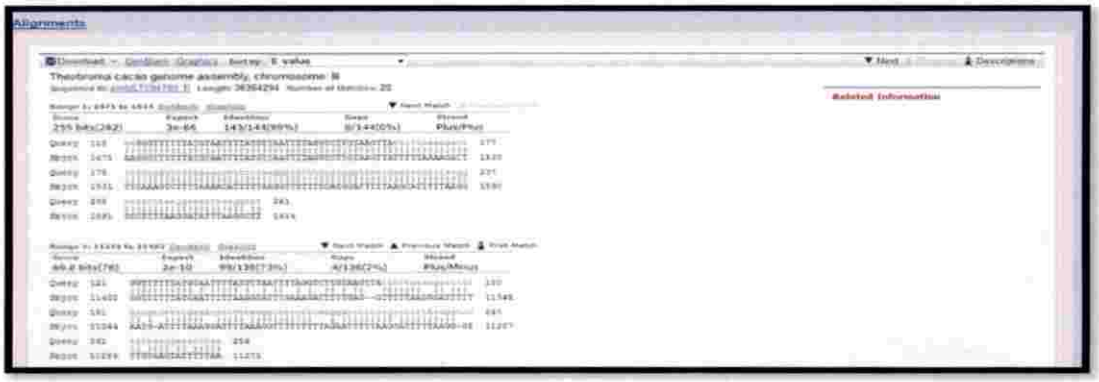


Fig. 4.24.c. Result of nucleotide BLAST of the 587 bp sequence generated from the polymorphic band obtained with primer UBC 826

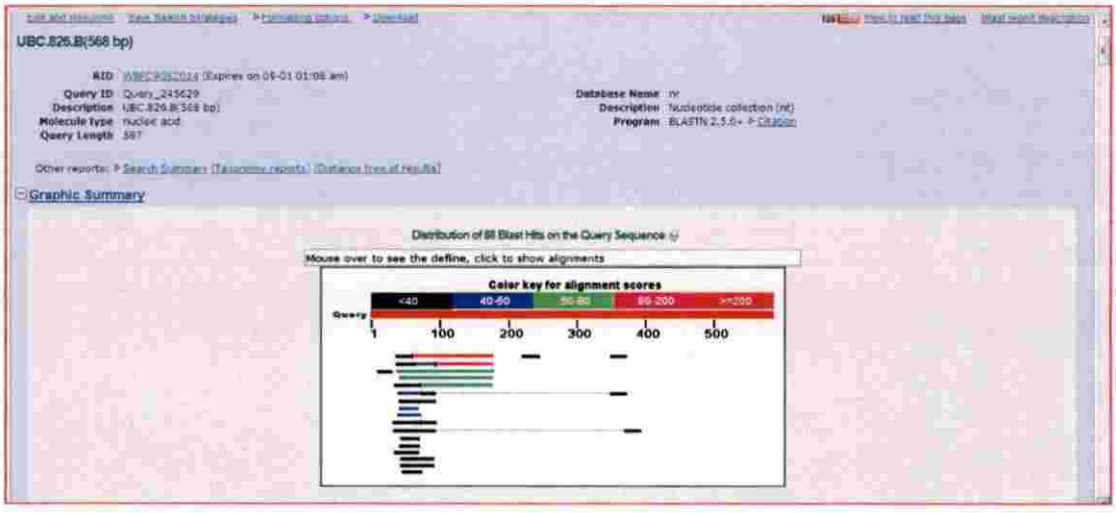


Fig. 4.25a Graphical summary of nucleotide sequence similarity through BLASTn of the polymorphic band obtained with primer UBC 826

Fig. 4.25b Results of the similarity search of sequence generated from the polymorphic band obtained with primer UBC 826, using BLASTn

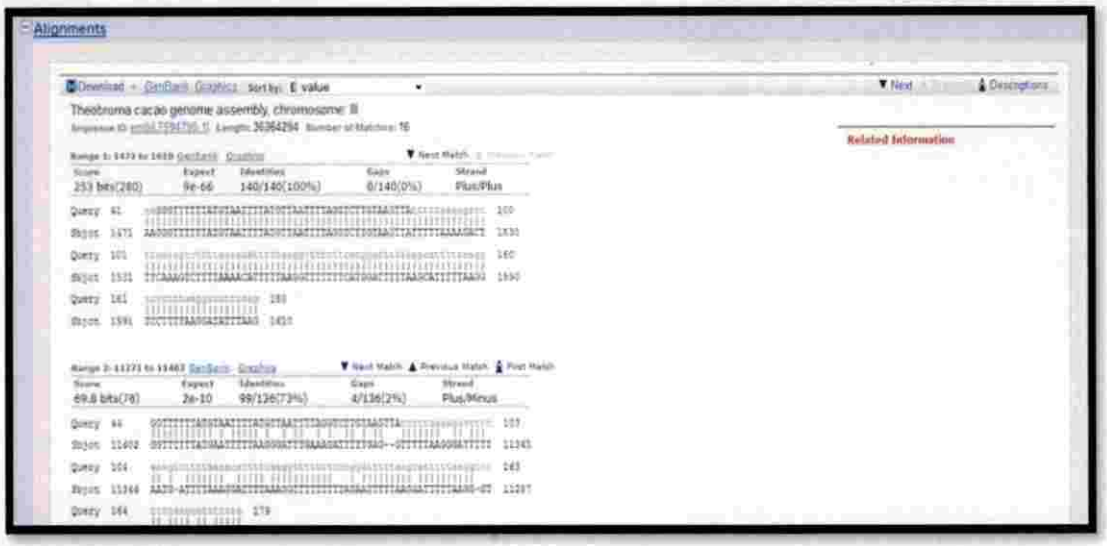


Fig. 4.25c. Result of nucleotide BLAST of the 568 bp sequence generated from the polymorphic band obtained with primer UBC 826

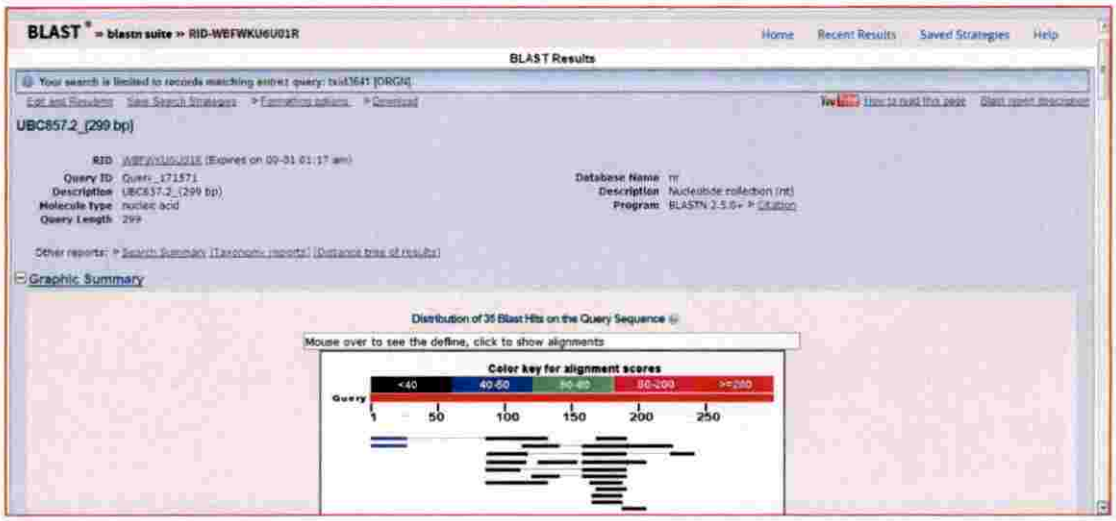


Fig. 4.26.a Graphical summary of nucleotide sequence similarity through BLASTn of the polymorphic band obtained with primer UBC 857

Fig.4.26.b Results of the similarity search of sequence generated from the polymorphic band obtained with primer UBC 857, using BLASTn

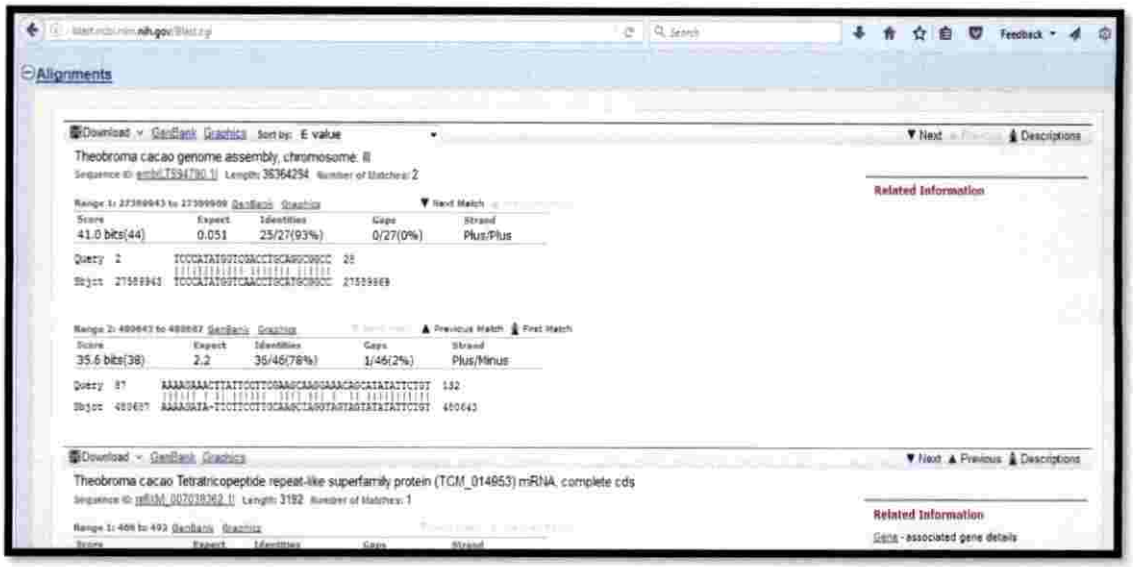


Fig. 4.26c. Result of nucleotide BLAST of the 299 bp sequence generated using the polymorphic band obtained with primer UBC 857

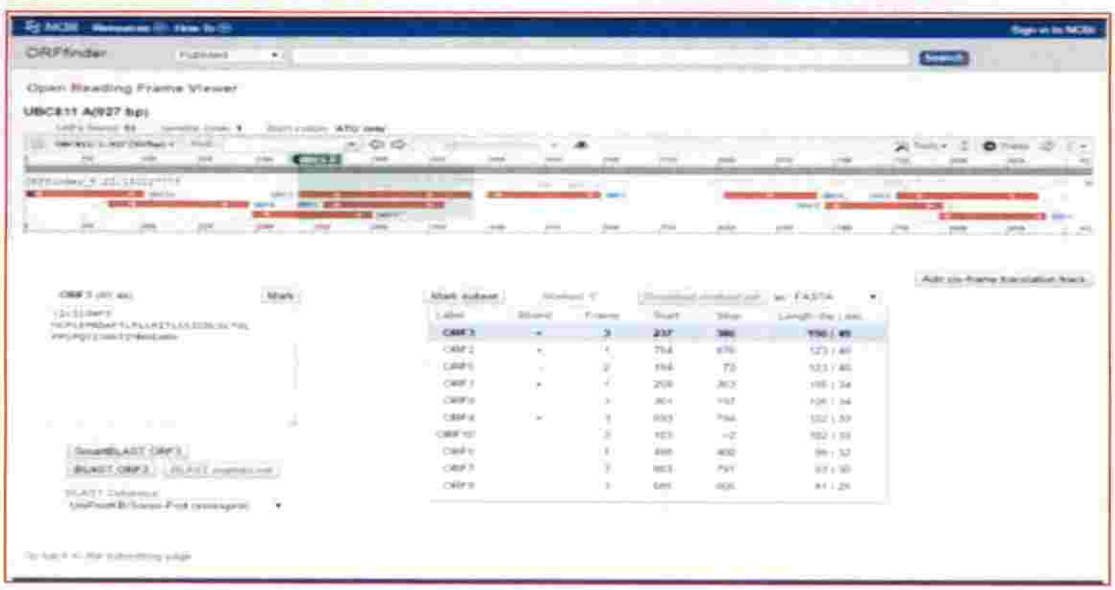


Fig.4.27. Details of ORF present in sequence UBC 811 -A (927 bp)

B) ORF analysis of sequence UBC 811-D (623 bp)

ORF analysis of sequence UBC 811-D (623 bp) had shown five ORFs (Fig.4. 28) and these ORFs were contributing to functional proteins (Table.4.28). ORF- 1 had 69 amino acids and it had 35 per cent identity with aflatoxin biosynthesis regulatory protein (PDB ID P41765.2). ORF -2 had 47 amino acids and it had shown 47 per cent identity with chain A, cryatal structure of activin receptor type iia (acvr2a) kinase domain in complex with dorsomorphin (PDB ID-3Q4TA). ORF-3 having 54 amino acids and it had 32 per cent identity with NAD (P) H dehydrogenase quinone flavoprotein Wrb a, NAD(P)H quinone oxidoreductase NQQ (Swissprot ID Q3J9H5.1). ORF-4 had 41 amino acids and it had shown 47 percent identity with chain a predicted insulinase family proteae from *Yersinia pestis* (PDB ID 3GO9A). ORF-5 had 25 amino acids and it had shown 56 per cent identity with pericentrin, kendrin, pericentrin B (Swissprot ID O95613.4).

4.7.2.2. ORF analysis of sequence UBC 826-A (587 bp)

A) UBC826-A (587 bp)

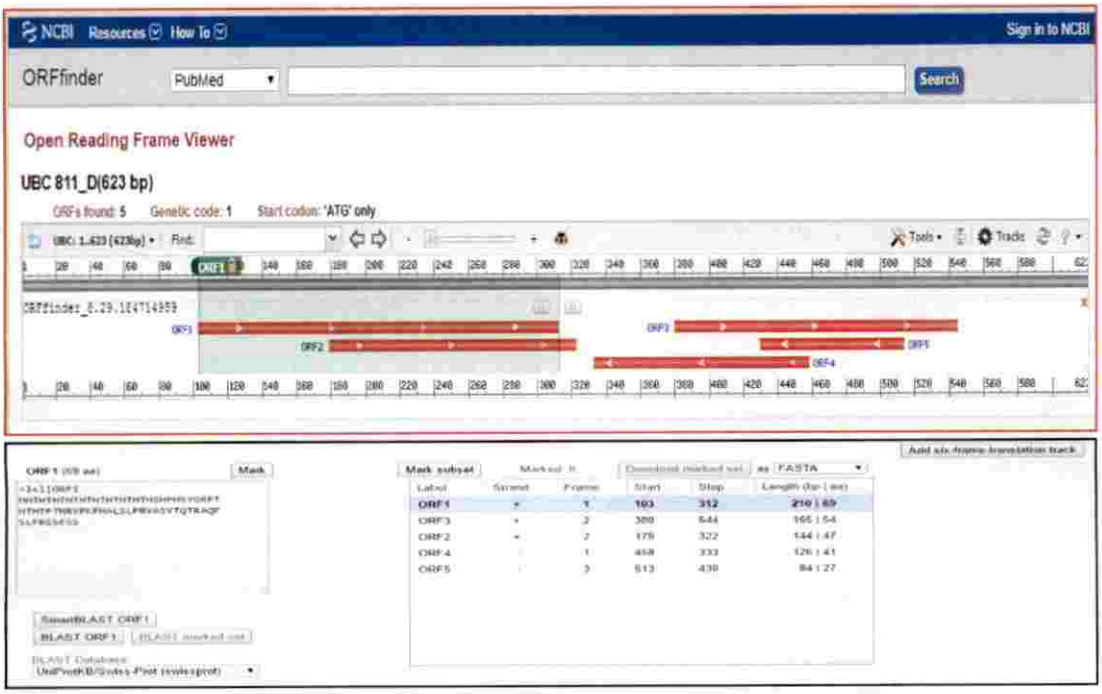
In ORF analysis of sequence UBC 826-A (587 bp) it was observed that, three ORFs were located in the sequence (Fig, 4.29). ORF-1 had 33 amino acids; it contributed for uncharacterized protein (Table 4.8). ORF-2 had 62 amino acids and it had shown 46 per cent identity with light independent protochlorophylide reductase subunit b (Swissprot ID Q6N9K2.1). ORF-3 had 41 amino acids it had shown 45 per cent identity with BAG family molecular chaperone regulator 8, chloroplastic BCl 2 associated athanogene 8 flags precursor (Swissprot ID Q9LIB3.1).

B) UBC 826-B (568 bp) sequence 2

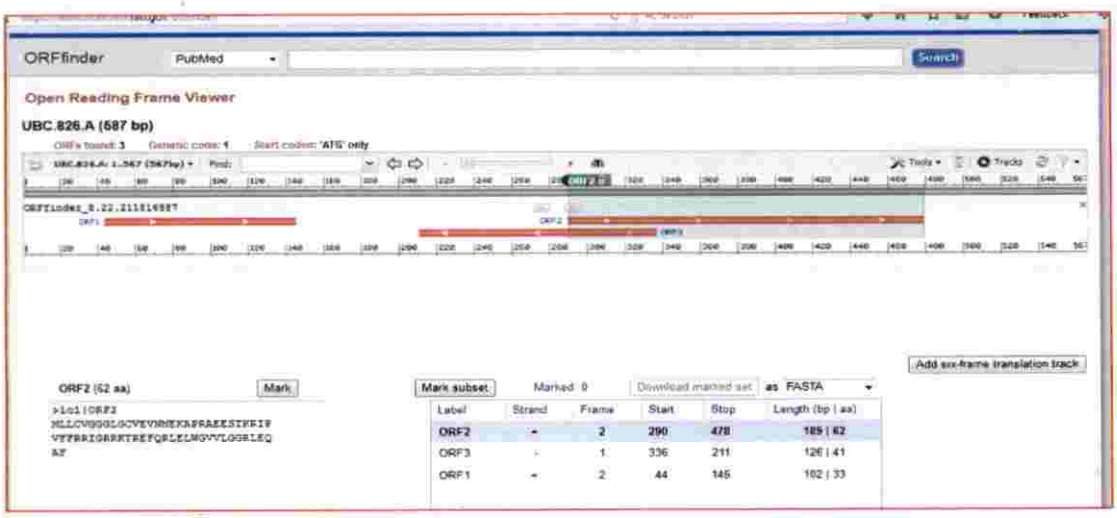
ORF analysis of sequence UBC 826-.B (568 bp) had shown location of two ORFs (Fig.4.30).ORF-1 had 53 amino acids and it shown 38 per cent identity with potassium transporter 27 *Os*HAK 27 (Swissprot ID Q84M84.1) (Table. 4.9). ORF-2 had 54 amino acids and it had shown 50 per cent identity with chain A crystal structure of aquinozolinedione sulfonamide bound to human glur 2 a novel class of competitive ampa receptor antagonists with oral activity (PDB ID 3R7XA).

4.7.2.3. ORF analysis of sequence UBC 857-2 (299 bp)

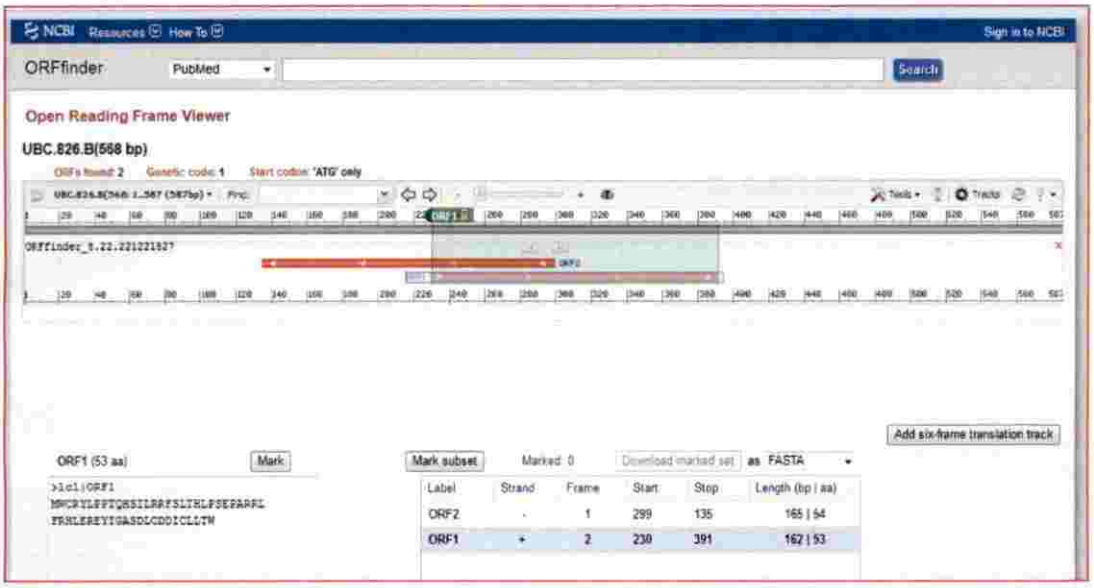
ORF analysis of sequence UBC 857-2 (299 bp) has shown five ORFs (Fig.4.31). ORF-1, 3, and 4 have 41, 36 and 42 amino acids respectively and these three ORFs contributed for uncharacterized protein. ORF-2 had 25 amino acids, it shown 71 per cent identity with chain A mechanosensitive channel of large conductance mscl (PDB ID 2OAR-A). ORF-5 had 42 amino acids it had shown 39 per cent identity with structural polyprotein precursor of VP 2, capsid protein VP2 (Swissprot ID Q96724.1).



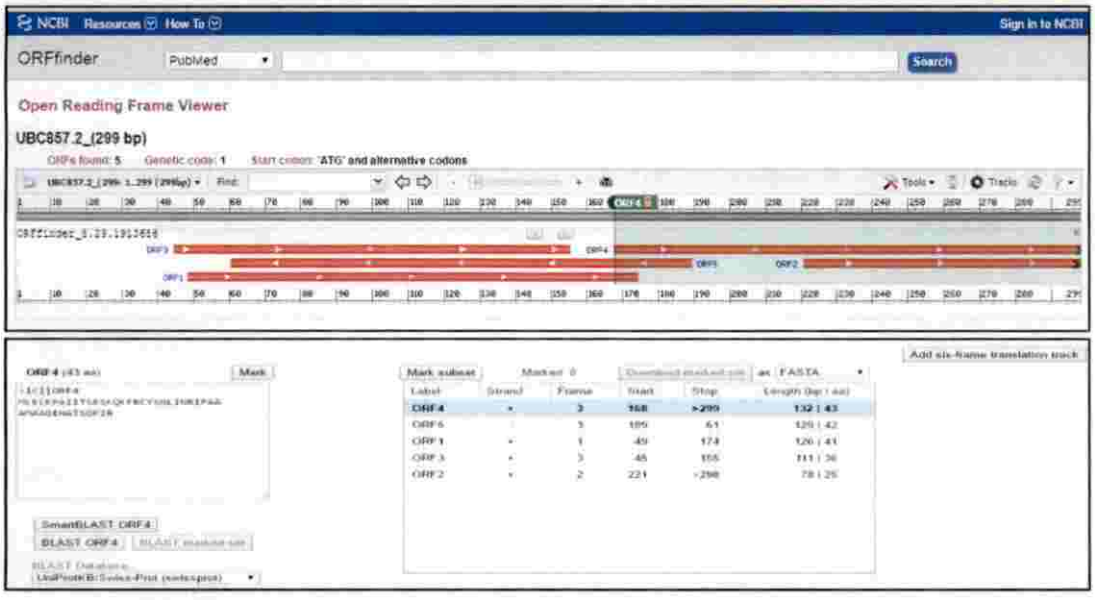
4. 28.Details of ORF present in sequence UBC 811-D (623 bp)



4.29. Details of ORF present in sequence UBC 826-A (587 bp)



4.30. Details of ORF present in sequence UBC 826-B (568 bp)



4.31. Details of the ORF present in sequence UBC 857-2 (299 bp)

Table. 4.6. Details of ORFs present in sequence UBC811 -A (927 bp)

Sl.No.	ORF	No. of amino acids	Sequence identity (%)	Protein	PDB /Swissprot ID
1	ORF1	34	–	–	–
2	ORF2	40	–	–	–
3	ORF3	49	–	–	–
4	ORF4	33	60	Chain A crystal structure of fold bifunctional protein from *Campylobacter Jejuni*	PDB ID 3P2OA
5	ORF 5	32	50	F box DNA helicase 1, FBOX protein 18	Swissprot ID F1 ND48.2
6	ORF6	30	–	–	–
7	ORF7	30	73	Chain X crystal structure archaeal intron encoded homing endonuclease I Tsp 061 I	PDB ID 2DCHX
8	ORF8	26	78	A crystal structure of putative histidinol phosphate aminotransferase NP 281508.1from *campylobacter jejuni* At2.01 A	PDB ID 3GETA
9	ORF9	33	–	–	–
10	ORF 10	33	47	Chain A crystal structure of carboxy peptidase γ inhibitor complexed with the cognate protease	PDB ID 1WPXA

Table.4.7. Details of ORFs present in sequence UBC 811-D (623 bp)

Sl. No.	ORF	No. of amino acids	Sequence identity (%)	Protein	PDB /Swiss prot ID
1	ORF 1	69	35	Aflatoxin biosynthesis regulatory protein	PDB ID P41765.2
2	ORF2	47	47	Chain A, cryatal structure of activin receptor type iia (acvr2a) kinase domain in complex with dorsomorphin	PDB ID 3Q4TA
3	ORF3	54	32	NAD(P)H dehydrogenase quinone flavoprotein Wrb a-NAD(P)H quinine oxidoreductase NQQ	Swiss prot ID Q3J9H5.1
4	ORF4	41	47	Chain A predicted Insulinase family protease	PDB ID 3GO9A
5	ORF5	27	56	Pericentrin, Kendrin, Pericentrin B	Swissprot ID O95613.4

Table.4.8. Details of ORFs present in sequence UBC 826-A (587 bp)

SL. No.	ORF	No. of amino acids	Sequence Identity (%)	Protein	PDB /Swiss prot ID
1	ORF1	33	–	–	–
2	ORF 2	62	46	Light independent protochlorophylide reductase subunit b	Swissprot ID Q6N9K2.1
3	ORF3	41	45	BAG family molecular chaperone regulator 8, chloroplastic BCl 2 associated athanogene 8 flags precursor	Swissprot ID Q9LIB3.1

4.9. Details of ORF present in sequence UBC 826-B (568 bp)

Sl. No.	ORF	No. of amino acids	Sequence Identity (%)	Protein	PDB /Swiss prot ID
1	ORF1	53	38	Potassium transporter27, *Os*HAK 27	Swiss prot ID Q84MS4.1.
2	ORF2	54	50	Chain A crystal structure analysis of aquinozolinedione sulfonamide bound to human Glur2 a novel class of competitive ampa receptor antagonists with oral activity	PDB ID 3R7XA

4.10. Details of the ORF present in sequence UBC 857-2 (299 bp)

Sl. No	ORF	No. of amino acids	Sequence Identity (%)	Protein	PDB /Swissprot ID
1	ORF1	41	–	–	–
2	ORF2	25	71	Chain A mechanosensitive channel of large conductance mscl	PDB ID 2OAR-A)
3	ORF3	36	–	–	–
4	ORF4	43	–	–	–
5	ORF5	42	39	Structural polyprotein precursor of VP 2, capsid protein VP2	Swissprot ID Q96724.1.

4.8. Identification of microsatellites present in ISSR sequence

Microsatellites or SSRs (Simple Sequence Repeats) are sequences in which one or few bases are tandemly repeated for varying numbers of times. The frequency and distribution of SSR motifs, dimers to decamers present in the sequence was determined. Detection and characterization of SSRs motifs between 2 to 10 base pairs was done. In sequence UBC811-A (927 bp) mostly trimer repeats were present followed by tetramer repeats (Table.4.11). Maximum number of dimmers were detected in sequence UBC 811-D (623 bp) (Table.4.12).The sequence UBC 826-A (587 bp) obtained from marker UBC 826 had shown timer repeats followed by pentamers (Table.4.13) and sequence UBC 826-B (568 bp) shown five different trimers repeats and one tetramer repeat (Table 4.14). The identified SSR motif can be used for primer designing.

Table. 4.11. Details of microsatellites present in UBC811-1 marker

Repeats	Sequence	Motif	No. of Repeats	SSR start	SSR end
Dimer repeats	UBC811-1-1	ct	4	309	316
Trimer repeats	UBC811-1-34	aac	2	136	141
	UBC811-1-35	ctc	2	276	281
	UBC811-1-36	agc	2	294	299
	UBC811-1-37	cca	2	327	332
	UBC811-1-38	aat	2	359	364
	UBC811-1-39	cga	2	389	394
	UBC811-1-40	atg	2	485	490
	UBC811-1-41	gct	2	535	540
	UBC811-1-42	gag	2	721	726
	UBC811-1-43	aca	2	749	754
	UBC811-1-44	atg	2	754	759
	UBC811-1-45	act	2	787	792
Tetramer repeats	UBC811-1-46	ttga	2	28	35
	UBC811-1-48	aatg	2	362	369
	UBC811-1-49	acaa	2	620	627
	UBC811-1-50	tagc	2	874	881
Heptamer repeat	UBC811-1-51	ctgatat	2	32	51

Table. 4.12. Details of microsatellites present in UBC 811-2 marker

Repeats	Sequence	Motif	No. of. repeats	SSR start	SSR end
Dimmer repeats	UBC811-2-1	ca	8	24	39
	UBC811-2-2	ca	4	42	49
	UBC811-2-3	ac	9	57	74
	UBC811-2-4	ca	5	88	97
	UBC811-2-5	ca	28	106	161
	UBC811-2-6	ca	4	189	196
	UBC811-2-7	ca	4	207	214
	UBC811-2-8	ct	4	326	333
Trimer repeats	UBC811-2-35	tct	2	304	309
	UBC811-2-36	agc	2	311	316
	UBC811-2-37	aat	2	376	381
	UBC811-2-38	aca	3	435	443
	UBC811-2-39	atg	2	500	505
Tetramer repeats	UBC811-2-43	tcac	2	77	84
	UBC811-2-49	aatg	2	379	386
	UBC811-2-50	cttg	2	520	527
	UBC811-2-51	tctg	2	600	607
Hexamer repeats	UBC811-2-55	cacacc	2	193	204
Decamer repeats	UBC811-2-59	cacacacact	2	32	51

Table. 4.13. Details of microsatellites present in UBC 826-1 marker

Repeats	Sequence	Motifs	No. of. repeats	Start SSR	Stop SSR
Trimer repeats	UBC826-1-20	ata	2	56	61
	UBC826-1-21	tga	2	273	278
	UBC826-1-22	tgt	2	291	296
	UBC826-1-23	gga	2	355	360
	UBC826-1-24	aga	2	401	406
Pentamers repeats	UBC826-1-25	ttcac	2	25	34
	UBC826-1-26	atagg	2	263	272
Hexamer repeat	UBC826-1-27	aatttt	2	524	535

Table. 4.14. Details of microsatellites present in UBC 826-2 marker

Repeats	Sequence	Motifs	No. of. repeats	Start SSR	Stop SSR
Trimer repeats	UBC826-2-22	gtg	2	232	237
	UBC826-2-23	cca	2	248	253
	UBC826-2-24	cga	2	364	369
	UBC826-2-25	gct	2	423	428
	UBC826-2-26	ata	2	535	540
Tetramer repeats	UBC826-2-27	ttac	2	377	384

In BLASTn analysis it was observed that UBC 811 ISSR region is located on chromosome number five and UBC 826 and UBC 857 regions are located on chromosome number three. In ORF analysis it was found that, aflatoxin biosynthesis regulatory protein, NAD(P)H dehydrogenase quinine flavoprotein Wrb, Potassium transporter 27, *Os*HAK 27, structural polyprotein precursor of VP 2, capsid protein VP2. These proteins are involved in resistance to VSD. Microsatellite region present in the sequence were identified and these motif can be used for primer designing.

Discussion

5. DISCUSSION

Cocoa (*Theobroma cacao* L.) is a main tropical crop, known as chocolate tree. Cocoa belonging to family *Malvaceae* and is native of Amazon region of South America. There are 22 species in the genus, but the cocoa tree, *Theobroma cacao* is cultivated widely.

In India commercial cocoa cultivation was started during 1970s. Cocoa is mainly grown in Kerala, Karnataka, Andhra Pradesh and Tamil Nadu. Kerala is the leading state in promoting cocoa cultivation. Brazil is leading producer of cacao, presently 70 per cent of cacao grown worldwide is produced in Africa, including the Ivory Coast, Ghana and Nigeria. In India cacao is cultivated on 78,000 ha and production is 16,050 MT with an average productivity of 475kg/ha (DCCD, 2015).

5.1. Vascular Streak Dieback disease of cocoa

Vascular streak dieback is the major destructive disease of cacao. VSD has been reported from the New Britain Island in Papua New Guinea, Hainan Island, China and Kerala State in India (Keane and Prior 1991). In India for the first time Abraham (1981) and Chandramohan and Kaveriappa (1982) have reported the presence of VSD in Kottayam district of Kerala. The maximum spread of VSD was in Kottayam, followed by Thiruvananthpurm, Kozhikode, Idduki and Pathanamthitta districts of Kerala (Abraham and Ravi, 1991). VSD is caused by the fungus *Ceratobasidium theobromae* (Talbot and Keane, 1971; Keane and Prior, 1991). The fungus cannot sporulate on dead branches, although the fungus readily grows on infected tissue and water agar. *Ceratobasidium theobromae* cannot be cultured in nutrient medium, but the sexual stage has only been induced in tissue culture (Lam *et al.* 1988, Dennis, 1991). Basidiospore of the fungus penetrate into the leaf and colonizes xylem vessels, causing vascular browning in the veins of lamina. Further the fungus moves into the midrib, petiole and at last it reaches to the branch. Chlorosis of leaves with green spots and necrotic blotches, are the symptoms of VSD

(Guest and Keane, 2007). After developing spotted chlorotic symptoms on the leaf, abscission of leaves start within few days. In Malaysia yield losses due to VSD were between 25 to 40 per-cent of total production of cacao. Areas severely affected by VSD yield has reduced considerably in new plantations (Byrne, 1976). In cocoa resistance to VSD is in the form of horizontal resistance. It is polygenic and largely inherited as additive genes. VSD resistance is durable and is inherited quantitatively. In VSD resistance the effect of the gene is additive, each additional gene enhances the expression of the trait by equal increments (Tan and Tan, 1988).

In Kerala Agricultural University between 1995 to 1998 breeding for resistance to vascular streak dieback (VSD) disease was performed. During this period numerous crosses were made using 298 resistant parents, 953 hybrid seedlings shown various levels of resistance (CCRP report, 1998).

Planting with high yielding and resistant cultivars is the only sustainable alternative to counter VSD. Conventional plant breeding relies on the field screening for the confirmation of the transfer of the desired gene in the offspring and this process is influenced by environmental factors. Marker assisted selection is a highly reliable alternative in which the tightly linked markers can be used to confirm the presence of the gene of interest in the selected plants (Salimath *et al.*, 1995).

One SSR marker mTcCIR 42 and 5 ISSR markers namely UBC 811, UBC 815, UBC 826, UBC 857 and UBC 866 which are linked to the gene offering resistance to the VSD in cocoa (Chandrakant, 2014). These six markers were used for validation and characterization study. Twenty VSD resistant hybrids and four susceptible lines (Table 3.1) were selected for the present study on the basis of field screening experiments conducted from 1998-1999 (Minimol and Amma, 2013). The genotypes were identified and maintained at Cocoa Research Centre (CRC), College of Horticulture, Kerala Agricultural University (KAU), Thrissur, Kerala.

5.2. DNA Isolation

Molecular techniques require good quality of genomic DNA; isolation of good quality DNA is the prerequisite thing in molecular analysis. Several protocols are available for extraction of pure and intact DNA from plant tissues (Saghai-Maroof *et al.* 1984; Doyle and Doyle 1990; Haymes, 1996: Drábková *et al.*, 2002; Shepherd *et al.*, 2002; Mogg and Bond, 2003).Crop species belonging to the same family have immense varying in the complexity of the pathways. DNA isolation protocols cannot be reproduced for all plant species (Porebski *et al.*, 1997; Ribeiro and Lovato, 2007).

In the present study, rapid and inexpensive modified Doyle and Doyle method suggested by Chandrakant (2014) was followed for the extraction of DNA. The cocoa leaf contains high amount of polysaccharides, polyphenols and secondary metabolites, these components interfere with DNA isolation procedures and PCR based techniques (Katterman and Shattuck, 1983; John, 1992). Fresh, young and tender leaves were used for DNA extraction. On grinding, the leaves of cocoa turned brown due to phenolic oxidation. The addition of antioxidants like β-mercaptoethanol and PVP helps to inhibit phenolic oxidation and to remove the brown color of DNA pellet. The detergent used in the extraction buffer of CTAB (Cetyl trimethyl ammonim bromide), which helps in the disruption of the cell membrane thereby releasing nucleic acid into the extraction buffer. It prevents co-precipitation of polysaccharide with nucleic acid by acting as a selective precipitant of nucleic acid. Chloroform was used to remove protein and cell debris from the leaf extract. The DNA pellet was washed with the wash buffer (76 per cent ethanol and 10mM ammonium acetate) to remove the stickiness. The EDTA in extraction buffer protects the DNA from endonucleases by cheating the Mg^{2+} ions of DNA. Two times treatment with chloroform: isomyl alcohol followed by centrifugation effectively removes the pigments. The addition of chilled isoproponol precipitates the DNA and washing of the pellet with 70 per-cent alcohol followed by absolute alcohol removes the traces of CTAB (Wettasinghe and Peffley, 1998). DNA extracts contain a large

amount of RNA, proteins, polysaccharides, tannins and pigments. Polysaccharides are more difficult to remove, NaCl together with CTAB is used to remove polysaccharides (Murray and Thompson, 1980; Croy *et al.*, 1993; Paterson *et al.*, 1993). Secondary metabolite present in leaf samples adheres to DNA and causes damage of DNA (Weising *et al.*, 1995). Mixed Alkyl trimethyl ammonium Bromide (MATAB) combined with the use of spheres for maceration, reduce the amount of polyphenols and polysaccharides resulting in a viscous substance (Faleiro *et al.*, 2002). Complex molecules interfere with DNA quality, leading to low yield (Tel-Zur *et al.*, 1999), and inhibiting the action of the *Taq* polymerase (Fang *et al.*, 1992).

5.3. Purification and quantification of DNA

The quality of the isolated DNA was checked in agarose gel electrophoresis and quantification of DNA was done by spectophotometric method using NanoDrop ND 1000. DNA was visualized on 0.8 per cent agarose gel under UV light by staining with ethidium bromide. The extracted DNA has shown RNA contamination, RNase treatment was given to remove RNA. RNase can be used for purification of DNA (Raval *et al.*, 1998; Wettasingf and Peffley, 1998; Gallego and Martinez, 1996). After RNase A treatment, DNA sample on electrophoresis has shown a high molecular weight intact single band.

The absorbance ratio was calculated as OD at 260/280 nm for the every sample using spectrometer NanoDrop ND 1000.The spectrophotometric method is the recent technology for the precise quantification of DNA. The ratio of optical density at 260 and 280 nm was considered to test the quality. DNA samples having O.D. value between 1.8 and 2.0 was considered as good quality.

5.4. Molecular marker analysis

Most of the molecular markers are based on PCR techniques and amplifies unique region on the genomic DNA based on the primer designed for DNA

amplification. Markers have proven as a powerful tool in genetic analysis due to its simplicity, easiness and are independent of the plant growth stage (Milach, 1998). In the present study two marker systems namely, ISSR and SSR was used for validation and characterization of marker in cocoa.

5.4.1. ISSR analysis

In the present study 5 ISSR markers namely UBC 811, UBC 815, UBC 826, UBC 857 and UBC 866 have been employed and these markers are capable to yield the polymorphic band which is linked with resistance to VSD. The ISSR analysis detected a number of bands per primer and a less number of monomorphic markers, indicating high intraspecific genetic variability of the accessions. The same analysis was observed by Chia *et al.* (2011) using ISSR markers in 46 cacao accessions maintained in Tingo Maria -Peru. ISSR marker, in spite of its dominance nature, established clear grouping of Trinitario accessions into a common cluster.

ISSR is DNA fragments of about 100-3000 bp located between adjacent, oppositely oriented microsatellite regions. This technique, reported by Zietkiewicz *et al.* (1994) primers based on microsatellites is used to amplify ISSR region present in DNA sequences. Primers which anchors into the non-repeat adjacent regions (16-18 bp) is used for ISSR assay. From multiple loci 10-60 fragments are generated these fragments are separated by gel electrophoresis and used for scoring. Single Primer Amplification Reaction (SPAR) and Directed Amplification of minisatellite region DNA (DAMD) uses a single primer having a core motif of a microsatellite and minisatellite respectively. Multilocus fingerprinting profiles obtained, through ISSR analysis can be applied in studies involving genetic identity, parentage, clone and strain identification, and taxonomic studies of closely related species. In addition, ISSRs are considered useful in gene mapping studies (Zietkiewicz *et al.*, 1994, Gupta *et al.*, 1994: Godwin *et al.*, 1997).

ISSRs are simple and cost effective and they are able to show a high degree of polymorphism (Salimath *et al.*, 1995; Borba *et al.*, 2005). ISSR does not require

prior sequence information (Reddy *et al.*, 2002). The amplified ISSR products behave like dominant markers (Goulão and Oliveira, 2001).

Amplicons obtained with the primer UBC 811 primer were distinctly polymorphic for the resistance to VSD. The polymorphic band of size 0.950 Kb was present in all resistant lines, but absent in susceptible lines. This marker was able to clearly differentiate resistant and susceptible lines. This marker is tightly linked to gene offering resistance to VSD.

ISSR profile of primer UBC 811 on amplification with eighteen genotypes of *Curcuma longa* shown the polymorphic index value of 87.27 per-cent across all the genotypes. The polymorphism percentage ranged from 78.57 (UBC 812) to 100.00 (UBC 809). The primers 809, 810, 811, and 816 exhibited higher polymorphism percentage among these 809 were exhibited 100.00 polymorphisms and therefore were more informative (Prashanth *et al.*, 2015).

Genetic diversity and phylogenetic relationships was determined using ISSR marker in 90 genotypes of wild and cultivated species of *Oryza* from different geographical regions of the world. The primer UBC-810 and UBC-811 amplified minimum bands (391 and 466, respectively) at only 7 band positions each. The mean polymorphism information content (PIC) ranged from 0.44 to 0.84 and resolving power (Rp) ranged from 8.6 to 23.5. The primers has shown 100 per-cent polymorphisms among all accessions (Haritha *et al.*, 2016)

Using the ISSR primer UBC 815, on an average eight amplicons were obtained in each accession. This primer was found polymorphic for the VSD resistance. The polymorphic band was observed at 750 bp in resistant lines ,but were absent in susceptible lines. This primer was found to be a good candidate with the potential to differentiate between resistant and susceptible lines.

Marker index (MI) and effective multiplex ratio (EMR) was considered for all the primers. The MI values ranged from 0.41 to 3.36. The highest MI (3.36) was observed for the primer UBC-815 and the minimum MI (0.41) was obtained with ISSR primers UBC-822 and UBC-876.The primers that showed higher polymorphism

had higher EMR values. This feature varied from 1.8 to 12 with a mean value of 5.69. MI was positively correlated with PIC (Najaphy *et al.*, 2011)

ISSR analysis of the DNA samples with the primer UBC 826 generated polymorphic amplicons. The polymorphic band of size 650 bp was present in all resistant lines and in one susceptible line (S2) but absent in three susceptible lines, this amplicon is linked to VSD resistance. Since this marker is present in all the resistant lines and are absent in three susceptible lines. This marker has shown the polygenic nature of resistance. Amplicon of size 650 bp was present in one susceptible line (S2) because resistance is polygenic and this may be the minor gene which is present in susceptible line.

Genetic diversity of downy mildew resistant and susceptible cucumber germplasms was analyzed using ISSR. On amplification with ISSR primers eighty four bands were observed, in which 16 bands (19%) were monomorphic and 68 bands (81%) were polymorphic. The percent of polymorphic bands of each primer was ranged from 55.56 (%) for UBC-811 to 100 (%)for UBC-807, UBC-808, UBC-818, UBC-823, UBC-824, UBC-826, UBC-847, UBC-848, and UBC-849 with the mean of 84.52 percent (Innark *et al.*, 2014).

The ISSR primer UBC 857 has shown clear differentiation among cacao genotypes. Amplicon generated from this primer was polymorphic at a band size of 450 bp, this polymorphic band was present in all the resistant lines but it was absent in all susceptible lines. Clear polymorphism between the resistant and susceptible lines was observed with marker UBC 857. This marker was highly repeatable and associated with the VSD resistance in cocoa.

Genetic characterization of pigeon pea was done using RAPD, ISSR and SSR markers, three wilt resistant genotypes (Vaishali, ICPL 87119 and national check ICP 8863) and two susceptible genotypes (T15-15 and national check ICP 2376) and four genotypes grown in Gujarat, India (BDN-2, GT-100, GT-102 and C-11) were screened. ISSR analysis shown that the primer UBC-857 was able to clearly differentiate resistant and susceptible genotypes. PIC was recorded for primer UBC-

873 and UBC-850, the average PIC value was 0.25, ranging from 0 to 0.51. The marker index (MI) for 8 ISSR primers varied from 0 to 40.80, and an average MI was 16.75. Primer UBC-850 was found to be most informative (Swami *et al.,* 2015)

When the genotypes were screened with the ISSR primer UBC 866, polymorphic amplicon had generated. Two polymorphic bands were present at 1.300 kb (marked 'a') and 1.5 kb (marked 'b') in two susceptible lines but absent in all the resistant lines and two susceptible lines. This marker was found to be repeatable and linked with the gene offering resistance to VSD.

Molecular characterization of the melon accessions shown high level of diversity. There was no correlation between accession grouping and collection site. A total of 110 bands were generated on amplification with 13 ISSR primers. Of these bands, 79 were polymorphic and 31 were monomorphic for all accessions. The rate of polymorphism was 71.8 per -cent. Primer UBC-880 had the highest rate (100 %) of polymorphism for all accessions, whereas primer UBC-886 had the lowest rate (42.86 %) of polymorphism. Similarity coefficients among the 83 accessions were calculated using ISSR data and varied between 0.35 and 1.00, with an average coefficient of 0.88 (Mancak *et al.,* 2014).

5.4.2. SSR analysis

Simple sequence Repeats (SSR), are present in the majority of eukaryotic genomes and consist of simple, short tandemly repeated di-to penta-nucleotide sequence motifs (Beckman and Soller, 1996). SSR markers are widely used in many crop species because of its abundance, high level of polymorphism, locus specificity, reproducibility and suitability for multiplexing on automated systems. SSR is co-dominant marker and it requires a low amount of DNA. These properties make SSR as an effective marker for genetic characterization (Bhattramakki *et al.,* 2000). SSR marker provides scope to distinguish between the different forms of genetic change, including chromosome or chromosome arm loss (using several linked SSRs), point mutations leading to the failure of amplification (loss of one SSR allele) through to

slippage mutation (leading to the creation of a new allele). Low-molecular weight SSR allele will tend to amplify with greater efficiency than a high-molecular-weight allele (Clayton *et al.*, 1998).

Microsatellite markers are proven efficient to fingerprint and to resolve identity issues in cacao collections (Figueira, 1998; Risterucci *et al.*,2001; Saunders *et al.*, 2004; Cryer *et al.*, 2006; Zhang *et al.*, 2006)

The development of simple sequence repeat (SSR) markers in cacao (Lanaud *et al.*, 1999) has significantly increased the capacity of molecular characterization of cacao germplasm. More than 300 SSR markers covering the 10 linkage groups have been developed from various genomic DNA and expressed sequence tag (EST) libraries mainly by CIRAD, France, and USDA-ARS, SHRS, USA (Risterucci *et al.*, 2000; Kuhn *et al.*, 2003; Lanaud *et al.*, 2004; Pugh *et al.*, 2004; Borrone *et al.*, 2007).

Efombagn *et al.* (2006) have analyzed the genetic diversity in cocoa using 13 SSR markers. Where 282 alleles were analyzed, farm accessions were differentiated according to the geographical area.

In the present study SSR assay was done with the primer sets mTcCIR 42. The amplicons generated the primer set were polymorphic showing clear differentiation in resistant and susceptible lines. The polymorphic band of size 200 bp is tightly linked with VSD resistance. This polymorphic band was present in 12 resistant lines, but it was absent in four susceptible and four resistant lines. The marker mTcCIR 42 has shown co- dominant nature and this marker has association with VSD resistance. Details of the SSR markers can be viewed at the Cacao Genome Database at www.cacaogenomedb.org (Schnell *et al.*, 2007).

Genetic variability was analyzed between the genitors (Pa 30 and Pa 169) of the segregating population using SSR markers. In bulk analysis it was observed that three genomic SSR primers (mTcCIR 15, mTcCIR 25 and mTcCIR 45) were present in five seedlings of each bulk; bulk were used for the genotyping of the segregating population (200 seedlings). The level of polymorphism between Pa 30 and Pa 169, was compatible in genetic diversity studies of *T. cacao*. The gene is found 38.5 cM

from marker E11 (1650 bp) and is not found in the regions adjacent to loci mTcCIR 15, mTcCIR 25 and mTcCIR 45 in linkage groups 1, 6, and 8, respectively. The SSR data helped for the localization of '*Luteus-Pa*' gene in *T. cacao* (Rehem *et al.*, 2010)

5.5. Sequence analysis

5.5.1 BLASTn analysis

In BLASTn analysis the annoted sequence UBC811-A (927 bp) obtained from marker UBC 811 was found to have 82 percent identity with *Theobroma cacao* genome assembly chromosome V (NCBI accession number LT 94792.1) and sequence UBC 811-D (623 bp) have 84 per cent identity to *Theobroma cacao* genome assembly chromosome V (NCBI accession number LT 94792.1).The BLASTn analysis revealed that the UBC 811 ISSR region is located on chromosome number five. The BLASTn analysis of sequence UBC826-A (587 bp) was found to have 99 per cent identity with *Theobroma cacao* genome assembly, chromosome: III (NCBI accession number LT594790.1) and sequence UBC826-B (568 bp) had shown 100 per cent identity with *Theobroma cacao* genome assembly, chromosome: III (NCBI accession number LT594790.1). The BLASTn results shown that the UBC 826 ISSR region is located on chromosome number three.

Sequence of 299 bp nucleotide obtained from marker UBC 857 had shown 93 percent identity to *Theobroma cacao* genome assembly, chromosome: III (Genbank accession number LT94790.1).In homology search it was found that UBC 857 ISSR result is located on chromosome number three. The UBC 811, UBC 826 and UBC 857 these ISSR regions are present on chromosomes of resistant lines but these three ISSR regions are not present on the chromosome of susceptible plant.

5.5.2.ORF analysis

ORF analysis of sequence UBC811- A (927 bp) had shown ten ORFs. Among these ten ORF five ORFs were contributing to uncharacterized protein and Five ORFs were contributed for proteins which are involved in cell cycle regulation and signal transduction. ORF analysis of sequence UBC 811-D (623 bp) had shown five ORF

and the ORF-1 had 69 amino acids; it has 35 per cent identity with aflatoxin biosynthesis regulatory protein (PDB ID P41765.2). ORF-3 had 54 amino acids and it had shown 32 per cent identity with NAD(P)H dehydrogenase quinone flavoprotein Wrb a -NAD(P)H quinone oxidoreductase NQQ (Swissprot ID Q3J9H5.1).

The aflatoxin biosynthesis regulatory protein and quinine are involved in stress response and defense mechanism. Two fractions from corn seeds were inhibitory to aflatoxin formation. Using a sensitive laboratory assay inhibition of fungal growth and inhibition of aflatoxin biosynthesis was examined. Aqueous extracts from seeds of Tex6, a corn inbred shown to be highly resistant to aflatoxin accumulation in field and laboratory evaluations. Two biologically active fractions were identified. One inhibited growth of *Aspergillus flavus* and, thus, aflatoxin accumulation, and the other inhibited aflatoxin formation with little effect on fungal growth. (Hang *et al.*, 1997)

Plant defense metabolites arise from the main secondary metabolic routes. Quinine and capthotecin are quinoline alkaloids and lysergic acid diethylamide (LSD) is an ergot alkaloid, all these arising from tryptophan these phytochemicals are involved both in resistance against pathogens and in tolerance towards abiotic stresses, such as atmospheric pollution (Iriti and Faoro, 2009)

Sequence UBC826-A (587 bp) had three ORF among these three ORF two ORFs has contributed to characterized protein. In ORF analysis of sequence UBC826-B (568 bp) has shown two ORFs, ORF-1 had 53 amino acids and had 38 per cent identity with potassium transporter 27 OsHAK 27 (Swissprot ID Q84M84.1). sequence UBC 857-2 (299 bp) has shown five ORFs and ORF-5 have 42 amino acids and 39 per cent identity with structural polyprotein precursor of VP 2, capsid protein VP2 (Swissprot ID Q96724.1). All these proteins are identified have definite role in defense related pathways.

High affinity K1 transporters, up-regulates some K1 channels, and activates signaling cascades, some of which are similar to those involved in wounding and

other stress responses. The molecules that signal low K1 status in plants include reactive oxygen species and phytohormones, such as auxin, ethylene and jasmonic acid (Ashley *et al.*, 2006). Systemin, hydroxyproline-rich glycopeptide systemins (Pearce *et al*, 2001) from solanaceous plants and *At*Pep1 peptide from *Arabidopsis* (Huffaker *et al.,* 2006) are involved in Defense mechanism. These peptides are 18 to 23 amino acids in length and are processed from wound- and JA-inducible precursor proteins, and play roles in the activation of local and systemic responses against wounding and pest attack. Systemin is synthesized from prosystemin and stored in the cytoplasm (Narvaez-Vasquez and Ryan, 2004).

5.5.3. Identification of microsatellite within the ISSR sequence

The frequency and distribution of SSR motifs, dimmers to decamers, in these ISSR markers were identified. Detection and characterization of SSRs motifs between 2 to 10 base pairs was done. In sequence UBC811-A (927 bp) mostly trimer repeats were present followed by tetramer repeats (Table.4.11). Maximum number of dimmers were detected in sequence UBC 811-D (623 bp) (Table.4.12).The sequence UBC 826-A (587 bp) obtained from marker UBC 826 had shown timer repeats followed by pentamers (Table. 4.13) and sequence UBC.826.B (568 bp) shown five different trimers repeats and one tetramer repeat (Table 4.14). The identified SSR motif can be used for primer designing.

ISSR marker UBC 811, UBC 815, UBC 826 and UBC 857 were highly reproducible and these four markers have potential to clearly differentiate between resistant and susceptible lines.

Future line of work includes the validated SSR and ISSR markers could be directly employed in breeding programme aimed to develop VSD resistant high yielding cocoa clones. QTL mapping for VSD resistance in cocoa has to be done. Development of SCAR markers for genes involved in resistance.

Summary

6. SUMMARY

The study on "Characterization and validation of microsatellite markers for resistance to vascular streak dieback disease in cocoa (*Theobroma cacao* L.)" was carried out at the Centre for Plant Biotechnology and Molecular Biology, College of Horticulture, Vellanikkara, during the period from 2014 to 2016. The objective of the study was to validate the identified SSR and ISSR markers linked to the gene/s offering resistance to vascular streak dieback disease in cocoa and to characterize the ISSR markers to develop corresponding SSR markers. The resistant hybrids *viz*., M 13.12 X GVI 55, GII 19.5 X GVI 55, GVI 4 X GVI 55, GVI 126 X GIV 18.5, GVI 126 X GIV 18.5, GVI 126 X GIV 18.5, GVI 126 X GIV 18.5, GVI 126 X GVI 55, GVI 126 X GVI 55, GVI 126 X GVI 55, GVI 126 X GVI 55, GVI 137 X GVI 55, GVI 137 X GVI 55, GVI 140 X GVI 55, GVI 140 X GVI 55, GVI 140 X GVI 55, GVI 143 X GVI 55, GVI 143 X GV I55, GVI 143 X GVI 55, GVI 167 X GIV 18.5 and susceptible genotypes G VI- 50, G VI- 52, G VI- 82, G VI- 100 were used in the study. All these accessions were maintained at the Cocoa Research Centre, College of Horticulture, Kerala Agricultural University. Morphological characters and response to disease as scored at Cocoa Research Centre from 1988-1999 onwards were employed in the study.

The outcome of the study is summarized below

I. Modified Doyle and Doyle method suggested by Chandrakant, 2014, using 2X extraction buffer, two time treatment with chloroform: isoamyl alcohol (24:1) and washing with ammonium acetate has yielded good quantity DNA from the tender leaves. RNA contamination was removed by RNase treatment.

II. The quality and quantity of DNA was analyzed using NanoDrop®1000 spectrophotometer. The absorbance ratio ranged from 1.80-1.89 and the quantity of DNA in the isolated samples has ranged between 495-1432 ng/µl.

III. Two molecular marker systems, namely, ISSR and SSR were employed for validation and characterization of markers in the selected genotypes for VSD.

IV. Five ISSR markers namely UBC 811, UBC 815, UBC 826, UBC 857 and UBC 866 were validated in selected genotypes for VSD resistance.

V. DNA samples of twenty four genotypes were amplified with primer UBC 811. Amplicons obtained with the primer UBC 811 primer were distinctly polymorphic for the resistance to VSD. The polymorphic band of size 0.950 Kb was present in all resistant lines, but absent in susceptible lines. This marker was able to clearly differentiate resistant and susceptible lines. This marker is tightly linked to gene offering resistance to VSD.

VI. Using the ISSR primer UBC 815, on an average eight amplicons were obtained in each accession. This primer was found polymorphic for the VSD resistance. The polymorphic band was observed at 750 bp in resistant lines, but were absent in susceptible lines. This primer was found to be a good candidate with the potential to differentiate between resistant and susceptible lines.

VII. ISSR analysis of the DNA samples with the primer UBC 826 generated polymorphic amplicons. The polymorphic band of size 650 bp was present in all resistant lines and in one susceptible line (S2) but absent in three susceptible lines, this amplicon is linked to VSD resistance. Since this marker is present in all the resistant lines and are absent in three susceptible lines. This marker had shown the polygenic nature of resistance. Amplicon of 650 bp was present in one susceptible line (S2) because resistance is polygenic and this may be the minor gene which is present in susceptible line.

VIII. The ISSR primer UBC 857 has shown clear differentiation among cacao genotypes. Amplicon generated from this primer was polymorphic at a band size of 450 bp, this polymorphic band was present in all the resistant lines but it was absent in all susceptible lines. Clear polymorphism between the resistant and susceptible lines was observed with marker UBC 857. This

marker was highly repeatable and associated with the VSD resistance in cocoa.

IX. When the genotypes were screened with the ISSR primer UBC 866, polymorphic amplicon had generated. Two polymorphic bands were present at 1.300 kb and 1.5 kb in two susceptible lines but absent in all the resistant lines and two susceptible lines. This marker was found to be repeatable and linked with the gene offering resistance to VSD.

X. SSR assay was done with the primer sets mTcCIR 42. The amplicons generated the primer set were polymorphic showing clear differentiation in resistant and susceptible lines. The polymorphic band of size 200 bp is tightly linked with VSD resistance. This polymorphic band was present in 13 resistant lines, but it was absent in four susceptible and sevenr resistant lines. The marker mTcCIR 42 has shown co- dominant nature and this marker has association with VSD resistance.

XI. The characteristic and identified marker linked with the VSD resistance was generated using the ISSR primer UBC 811. This marker was eluted and cloned into pGEMT vector.

XII. The BLASTn analysis of sequence UBC811A (927 bp) had shown 82 percent identity with *Theobroma cacao* genome assembly chromosome V (NCBI accession number LT 94792.1) and sequence UBC811D (623 bp) had 84.0 per cent identity to *Theobroma cacao* genome assembly chromosome V (NCBI accession number LT 94792.1). The BLASTn analysis revealed that the UBC 811 ISSR region is located on chromosome number five.

XIII. The BLASTn analysis of sequence from UBC826A (587 bp) had shown 99.0 per cent identity with *Theobroma cacao* genome assembly, chromosome III (NCBI accession number LT594790.1). The sequence of UBC826B (568 bp) had 100 per cent identity with *Theobroma cacao* genome assembly, chromosome III (NCBI accession number LT594790.1). The BLASTn results

had shown that the UBC826 ISSR region is located on chromosome number three.

XIV. The distinct marker generated using the ISSR primer UBC857 was linked with the VSD resistance. The marker was cloned into pGEMT vector and sequenced. Sequence of 299 bp nucleotide obtained from marker UBC857 had shown 93.0 per cent identity to *Theobroma cacao* genome assembly, chromosome III (Genbank accession number LT94790.1).

XV. ORF analysis of sequence UBC811D (623 bp) had shown five ORFs and the ORF-1 had 69 amino acids; it also had 35 per cent identity with aflatoxin biosynthesis regulatory protein (PDB ID P41765.2). ORF-3 had 54 amino acids and shown 32.0 per cent identity with NAD(P)H dehydrogenase quinone (Swissprot ID Q3J9H5.1).

XVI. ORF analysis of sequence UBC826B (568 bp) had shown two ORFs. ORF-1 had 53 amino acids and had 38 per cent identity with potassium transporter 27 *Os*HAK 27 (Swissprot ID Q84M84.1).

XVII. Sequence UBC 857 (299 bp) had shown five ORF and ORF-5 holds 42 amino acids and 39 per cent identity with structural polyprotein precursor of VP2, capsid protein VP2 (Swissprot ID Q96724.1). The identified proteins have definite role in defense related pathways.

XVIII. The frequency and distribution of SSR motifs, dimers to decamers, in these ISSR markers were identified. In sequence UBC811A mostly trimer repeats were present followed by tetramer repeats.

XIX. Maximum number of dimmers was detected in sequence UBC811D. The sequence UBC826A had shown timer repeats followed by pentamers and sequence UBC826B shown five different trimers repeats and one tetramer repeat.

XX. The SSR and ISSR markers validated through this study are linked with various genes involved in resistance to VSD in cocoa and they could be

employed in breeding programme aimed to develop VSD resistance and high yielding cocoa clones.

References

REFERENCES

Abraham, C. S. 1981. Vascular streak dieback of cocoa in India. *Indian Cocoa, Arecanut and Spices J.* **4**: 119-120.

Abraham,K. and Ravi, S. 1991.Vascular streak dieback of cocoa in Kerala occurrence and management. Proc.Kerala Sci. Congr. February-March 1991. p.114-115.

Ahmad,Z. M. A. 1982.Vascular Streak dieback of cocoa in west Malaysia. MSc. (Ag.). Thesis, University of Putra Malaysia, Malaysia. p.97.

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

Amma, S. P. 2013. Cocoa hybrids. In: Minutes of the 35th zonal research and extension advisory committee meeting of central zone of Kerala state, 19 July 2013, Regional Agricultural Research Station, Pattambi, Kerala, India, p.5.

Anderson, R.1989. Avocado an alternate host for *Ceratobasidium theobromae*. *Aust. Plant Pathol.* **53**: 96-97.

Araújo, I. S., Filho, G. A. D., Pereira, M. G., Faleiro, F. G., Queiroz, V. T., Guimarães, C. T., Moreira, M. A., Barros, E. G., Machado, R. C. R., Pires, J. L., Schenell. R. and Lopes, U. V. 2009. Mapping of quantitative trait loci for butter content and hardness in cocoa beans (*Theobroma cacao* L.) *Plant Mol. Biol. Rep.* **27**: 177-183.

Ashley, M. K., Grant, M., and Grabov, A. 2006.Plant responses to potassium deficiencies: a role for potassium transport proteins. *J. of Experimental Botany* **57** (2): 425-436.

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.

Bhattramakki, D., Dong, J., Chhabra. A. K. and Hart, G. E. 2000. An integrated SSR and RFLP linkage map of *Sorghum bicolor* (L.) Moench. *Genome*, **43**(6): 988-1002.

Borba, R. S., Garcia M. S., Kovalleski, A., Oliveira, A. C., Zimmer, P. D., Branco J. S.C. and Malone, G. 2005. Genetic dissimilarity strains of *Trichogramma* Westwood (Hymenoptera:Trichogrammatidae) through molecular markers ISSR. *Neotropical Entomol*. **34**:565-569.

Bowers, J. H., Bailey, B. A., Hebbar, P. K., Sanogo, S. and Lumsden, R. D. 2001.The impact of plant diseases on world chocolate production. *Plant Health Prog*. **5**: 437-441.

Bridgland, L. A., Richardson, J. M. and Edward, I. L.1966. Dieback disease of cacao (Part 1). *South Pacific Plante,r* **1**(6): 13-20.

Byrne, P. N. 1976. Vascular streak dieback in Papua New Guinea and Peninsular Malaysia. *Planter,* **52**: 49-53.

Caetano-Anolles, G., Bassam, B. J. and Gresshoff, P. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Mol. Biotechnol*. **9**: 553-557.

CCRP.1998-99.12 th *Annual report Cadbury- KAU Co-operative Cocoa Research Project*. Kerala Agricultural University, p.26.

Chandrakant, E. S. 2014. Identification of molecular marker linked to the resistance for vascular streak dieback disease in cocoa *(Theobroma cacao* L.). *Thesis,* M.

Sc. (Ag.) in Plant Biotechnology, Kerala Agricultural University, Thrissur, p. 89.

Chandramohan, R. and Kaveriappa, K. M. 1982. Occurrence and distribution of cocoa disease in southern India. *Proc. 8th International Cocoa Conference*, Cartagena, Columbia, 1981.p.445-450.

Charters, Y. 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.

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

Clayton, T. M., Whitaker, J. P., Sparkes, R., and Gill, P. 1998. Analysis and interpretation of mixed forensic stains using DNA STR profiling. *Forensic Sci. Int.* **91**: 55-70.

Clement, D., Lanaud, C., Sabau, X., Fouet, O., Le Cunff, L., Ruiz, E., Risterucci, A. M., Glaszmann, J. C. and Piffanelli, P. 2004. Creation of BAC genomic resources for cocoa (*Theobroma cacao* L.) for physical mapping of RGA containing BAC clones. *Theor. Appl. Genet.* **108**: 1627-1634.

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

Couch, J. A., Zintel, H. A. and Fritz, P. J. 1993.The genome of the tropical tree *Theobroma cacao* L. *Mol. Gen. Genet.* **237**:123-128.

Croy, E. J., Ikemura, T., Shirsat, A. and Croy, R. R. D. 1993. Plant Nucleic Acids. Plant Molecular Biology. Bios Scientific Publishers. Blackwell Scientific Publications. UK.p.682.

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. Crop Evol.* **53**: 1643-1652.

DCCD. 2015.Directorate of Cashew and Cocoa Development. (http://dccd.gov.in/stat2.htm).

Dennis, J. J. C. 1991. Epidemiology and control of vascular-streak dieback of cocoa in Papua New Guinea. Ph.D. thesis, Department of Botany, La Trobe University, Bundoora, Victoria 3086, Australia. p.210.

Dias, L. A. S., Rocha, R. B., and Picoli, E. A. T. 2005. Distinctness of cacao cultivars using yield component data and RAPD markers. *Crop Breed. Applied Biotechnol.* **5**: 47-54.

Dong, Z., Hong, O. S., Li, W. L., Yu, C., Hong, W. Q., Yong, L., Zhong,W.Z., Zhong, X. J.,Yun, Z. D., Yong, W., Xing, C. Y. and Yong, L. Z. 2015. Comparative genetic mapping revealed powdery mildew resistance gene *MlWE4* derived from wild emmer is located in same genomic region of *Pm36* and *Ml3D232* on chromosome 5BL. *J. Integr. Agric.* **14**(4): 603–609.

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

Drábková, L., Kirschner, J., and Vlcek, C. 2002.Comparison of seven DNA extraction and amplification protocols in historical herbarium specimens of Juncaceae. *Plant. Mol. Biol. Rep.* **20**: 161-175.

Efombagn, M. I. B., Sounigo, O., Nyassé, S., Dauleux, M. M., Cilas, C., Eskes, M. A. B and Allen, K. 2006. Genetic diversity in cocoa germplasm of southern cameroon revealed by simple sequences repeat (SSR) markers. *Af. J. Biotechnol.* **5** (16): 1441-1449.

Eujayl, I., Sorrells, M., Baum, M., Wolters, P and Powell, W. 2001. Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. *Euphytica,* **119:** 39-43.

Evans, H. C. 2007. Cacao diseases-the trilogy revisited. *Phytopathol.* **97**: 1640-1643.

Faleiro, F. G., Pires, J. L., Monteiro, W. R., Lopes, U.V.,Yamada, M. V., Piedra, A.G., Adimil, D. M., Gardini, E. A., Marques, J. R. B., Gramacho, K. P., Faleiro, A. S. G. and Santos, M. C. M. 2004.Variability in cacao accessions from the Brazilian, Ecuadorian, and Peruvian Amazons based on molecular markers. *Crop Breed. Appl. Biotechnol.* **4**: 227-233.

Faleiro, F. G., Santos, I. S., Bahia, R. C. S., Santos, R. F., Yamada, M. M., and Anhert, D. 2002. Optimization of extraction and *Theobroma cacao* L. DNA amplification in order to obtain molecular markers. *Agrotrópica,* **14**: 31-34.

Fang, G., Hammer, S. and Grumet, R.1992. A quick and inexpensive method for removing polysaccharides from plant genomic DNA. *Bio Techniques*,**13**:52-55.

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., 1998. Homonymous genotypes and misidentification in germplasm collections of Brazil and Malaysia. *INGENIC Newsletter,* **4**: 4-8.

Figueira, A., Janick, J., and Goldsbrough, P.1992. Genome size and DNA polymorphism in *Theobroma cacao*. *J. Am. Soc. Hort. Sci*.**117**: 673-677.

Gajjar, K. N., Mishra, G. P., Radhakrishnan, T., Dodia, S. M., Rathnakumar, A. L., Kumar, N., Kumar, S., Dobaria, J. R., Kumar, A. 2014. Validation of SSR markers linked to the rust and late leaf spot diseases resistance in diverse peanut genotypes. *Aust. J. Crop Sci*. **8**: 927-936.

Gallego, F. J. and Martinez, I. 1996. Molecular genotyping of rose cultivars using RAPD. *J. Hort. Sci.* **71**(6): 901-908.

Gao, L. F., Jing, R. L., Huo, N. X., Li, Y., Li, P.X., Zhou, R. H., Chang, X. P., Tang, J. F., Ma, Z.Y. and Jia, J. Z. 2004. One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. *Theor. Appl. Genet.* **108**: 1392-1400.

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.

Gonzalez, D., Carling, D.E., Kuninaga, S., Vilgalys, R., and Cubeta, M.A. 2001. Ribosomal DNA systematics of Ceratobasidium and Thanatephorus with *Rhizoctonia anamorphs*. *Mycol.* **93**:1138-1150.

Goulão, L., Oliveira, C. M. 2001. Molecular characterization of cultivars of apple (*Malus* x d*omestica* Borkh.) using microsatellite (SSR and ISSR) markers. *Euphytica,* **122**: 81-89.

Griffiths, L. A. 1958. Phenolic acids and flavonoids of *Theobroma cacao* L. separation and identification by paper chromatography. *Biochem. J.* **70**: 120-125.

Guest, D. I., and Keane, P. J. 2007. Vascular-streak dieback: a new encounter disease of cacao in Papua New Guinea and Southeast Asia caused by the obligate *basidiomycete Ceratobasidium theobromae*. *Phytopathol.* **97**: 1654-1657.

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.

Haritha, G., Sudhakar, T., Chandra, D., Ram, T., B. Divya, B. and Sarla, N. 2016. Informative ISSR markers help identify genetically distinct accessions of *Oryza rufipogon* for use in yield improvement. *Rice Science* **23** (5): 1-17.

Haymes, K. M. 1996. Mini-prep method suitable for a plant breeding program. *Plant. Mol. Biol. Rep.* **14**: 280-284.

Hearne, C. M., Ghosh, S. and Todd, J. A. 1992. Microsatellites for linkage analysis of genetic traits. *Trend. Genet.* **8:** 288-294.

Hodges, T. 1991. *Predicting crop phenology*, CRC Press, Boca Raton, p.33-35

Holmes, R. A., Boston, R. S. Payne, G. A. 2008. Diverse inhibitors of aflatoxin biosynthesis. *Appl. Microbiol. Biotechnol.* **78**:559–572.

Holton, T. A., Christopher, L. J. T, McClure, N. L, Harker, N. and Henry, R. J. 2002. Identification and mapping of polymorphic SSR markers from expressed gene sequences of barley and wheat. *Mol.Breed.* **9**: 63-71.

Hu, J. and Quiros, C. F. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. *Plant. Cell Rep.***10**: 505-511.

Huang, Z., White, D. G. and Payne, G. A. 1997.Corn seed proteins inhibitory to *Aspergillus flavus* and aflatoxin biosynthesis. *Phytopathology* **87**(6): 622-627.

Huffaker, A., Pearce, G., Ryan, C. A. 2006. An endogenous peptide signal in Arabidopsis activates components of the innate immune response. *Proc. Natl. Acad. Sci. USA* **103**:10098-10103.

Innark, P., Ratanachan, T., Khanobdee, C., Samipak, S.and Jantasuriyarat, C. 2014. Downy mildew resistant/susceptible cucumber germplasm (*Cucumis sativus* L.) genetic diversity assessment using ISSR markers. *Crop Prot.* **60**: 56-61.

Iriti, M. and Faoro., F. 2009 Chemical diversity and defence metabolism: How plants cope With pathogens and ozone pollution. *Int. J. Mol. Sci.* **10**: 3371-3399.

Janila, P., Nigam, S. N., Pandey, M. K., Nagesh, P.,Varshney, R. K. 2013. Groundnut improvement: use of genetic and genomic tools. *Plant Sci.* **4**: 1-16.

Jarne, P. and Lagoda, P. J. L. 1996. Microsatellite from molecules to population and back. *Trend. Ecol. Evol.* **11:** 424-429.

Jie, Z. G., Qing,Y.Y.,Ying, M., Feng, Y. X., Yu, C. S., Rui, R.,Gang, W. D., lu, Y. Z., and Jian, Z. H. 2014. Fine mapping and candidate gene analysis of resistance gene *RSC3Q* to soybean mosaic virus in Qihuang. *J. Integr. Agric.***13** (12): 2608-2615.

John, M. E.1992. An efficient method for isolation of RNA and DNA from plants containing polyphenolics. *Nucleic Acids Res.* **20**: 23-81

Jones, E. S., Mahoney, N. L., Hayward, M. D., Xrmstead, I. P., Jones, J.G., and Hurnphreys, M.O. 2002. An enhanced molecular marker based genetic map of perennial ryegrass *(Lolium perenne)* reveals comparative relationships with other Poaceae genomes. *Genome* **45**: 282-295.

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

Katterman, F. R., H. and Shattuck, V. I. 1983. An effective method of DNA isolation from the mature leaves of *Gossypium* species that contain large amounts of phenolic terpenoids and tannins. *Prep. Biochem.* **13**: 347-359.

KAU. 1993. *Sixth Annual Report* (1992-1993) Cadbury-KAU Co-operative cocoa research, College of Horticulture, Kerala Agricultural University. p.140.

KAU. 1995. *Eighth Annual Report* (1994-1995) Cadbury-KAU Co-operative cocoa research, College of Horticulture, Kerala Agricultural University, p.150.

Keane, P. 1992. Diseases and pests of cocoa: an overview. In: Keane P. J., Putter C. A., (eds) Cocoa pest and disease management in Southeast Asia and Australasia. FAO Plant Production and Protection Paper No. 112, FAO, Rome. p.1-11.

Keane, P. J. and Prior, C.1991. Vascular streak dieback of cocoa. *Phytopathol.* **33**: 1-39.

Keane, P. J. 1981. Epidemiology of vascular steak dieback of cocoa. *Ann. Appl. Biol.* **98**: 227- 241.

Keane, P. J. and Prior, C. 1992. Biology of vascular streak dieback of cocoa. FAO, Plant Production and Protection, **112**:p.75-83.

Keane, P. J., Flentje, N. T. and Lamb, K. P. 1972. Investigation of vascular streak dieback disease of cocoa in Papua New Guinea. *Australian J. Biol. Sci.* **25**: 553-564.

Kelly, J. D. 1995.Use of Random Amplified Polymorphic DNA markers in breeding for major gene resistance to plant pathogens. *Hort Science* **30**:461-465.

Kuhn, D. N., Heath M., Wisser, R. J., Meerow, A., Brown, J. S., Lopes, U. and Schnell, R. J. 2003. Resistance gene homologs in *Theobroma cacao* as useful genetic markers. *Theor. Appl. Genet.***107**: 191-202.

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.

Lam, C.H., Varghese, G. and Abidin, Z. M. A. 1988. *In vitro* production of *Ceratobasidium theobromae* basidiospores. *Trans. British Mycol. Soc.* **90**: 505-507.

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.

Lanaud, C., Risterucci, A. M., Pieretti, I., N'Goran, J. A. K. and Fargeas, D. 2004. Characterization and genetic mapping of resistance and defence gene analogs in cocoa (*Theobroma cacao* L.). *Mol. Breed.* **13**: 211-227.

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

Lei, C., Hao, K., Yang, Y.Y., Maa, J., Wanga, S., Wanga, J., Chenga, Z., Zhaoa, S., Zhanga, X.,Guoa, X.,Wangb, C. and Wana, J. 2013. Identification and fine mapping of two blast resistance genes in rice cultivar 93-11. *The crop J.* **1**: 2-14.

Lerceteau, E., Robert, T., Petiard, V., and Crouzillat, D. 1997. Evaluation of the extent of genetic variability among *Theobroma cacao* accessions using RAPD and RFLP markers. *Theor. Appl. Genet.* **95**: 10-19.

Levinson, G. and Gutman, G. A. 1987. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**: 203-221.

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.

Lima, L. S., Gramacho, K. P., Pires, J. L., Clement, D., Lopes, U. V., Carels, N., Gesteira, A. S., Gaiotto, F. A., Cascardo, J. C. M. and Micheli, F. 2010. Development, characterization, validation, and mapping of SSRs derived from *Theobroma cacao* L. -*Moniliophthora perniciosa* interaction ESTs. *Tree Genet. Genomes* **6**: 663-676.

Maddison, W. P., and Maddison, D. R. 2005. MacClade 4. *Analysis of Phylogeny and Character Evolution* 4.08 (ed.) Sinauer Associates, Inc., Sunderland, MA.p.223.

Mancak, I., SARI, N., Solmaz, I., and Ozkan, H. 2014. Determining the relationships between kirkagac and other melon types by using morphological and molecular methods. *Cucurbitaceae* Proc. Bay Harbor, Michigan, USA. p.80-83.

Milach, S. C. H. 1998. Molecular markers in plants. Porto Alegre: UFRGS. p.141.

Minimol, J. S., and Amma, S. P. 2013. Self incompatibility assessment in vascular streak dieback (VSD) disease resistant hybrids of cocoa. *Asian J. Hort.* 8 **(1)**: 114-117.

Minimol, J. S., Suma, B., Ummer, M. and Jayasree, P. A. 2016. Parental contribution analysis in hybrids bred for vascular streak dieback disease resistance in cocoa. *J. Plantation crops*. **44**(1):23-32.

Minimol, J. S., Suma, B., Ummer, M., Chithira, P. G. 2015. Genetic improvement of cocoa by developing superior hybrids. *J. Tropical Agril*. **53**(2):157-165.

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.

Mogg, R. J., and Bond, J. M. 2003. A cheap, reliable and rapid method of extracting high quality DNA from plants. *Mol. Ecol. Notes*.**3**: 666-668.

Mohammadi, S. A. and Prasanna, B. M. 2003. Analysis of genetic diversity in crop plants-salient statistical tools and considerations *Crop Sci.* **43** (4) 1235-1248.

Morgante, M. and Olivieri, A. M. 1993. PCR amplification of microsatellite markers in plant genetics. *Plant J.* **3**: 175-182.

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

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

Nagai, C., Heinig, R., Olano, C., T., Motamayor, J. G. and Schnell, R., J. 2009. Fingerprinting of cacao germplasm in hawaii. *Hawaii Agriculture Research Center* Cacao. *Asian J. Hort.***1**:148-154.

Najaphy, A., Parchin, R. A. and Farshadfar, E. 2011. Evaluation of genetic diversity in wheat cultivars and breeding lines using inter simple sequence repeat markers. *Biotechnol. And Biotechnol. Equipment*. **25**(4): 2634-2638.

Narvaez-Vasquez J, Ryan, C. A. 2004. The cellular localization of prosystemin: a functional role for phloem parenchyma in systemic wound signaling. *Planta* **218**:360-369.

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. *Heridity,* **73**: 589-597.

Nicot, N. V., Chiquet, B. V., Gandon, L. B., Amilhat, F. L., Legeai, P. F., Leroy, M. F., Bernard, M. and Sourdille, P. 2004.Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags, (ESTs). *Theor. Appl. Genet.* **109**: 800-805.

Ning, L., Bin, J. J., Fu, L. J. and Yang, X. X. 2012. Development of molecular Marker Linked to *Cf*-10 gene using SSR and AFLP method in tomato. *J. Northeast Agric. Univ.* **4**:30-36.

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

Park, Y., Hwang, J., Kim, K., Kang, J., Kim , B., Xu, S., and Ahn,Y. 2013. Development of the gene-based SCARs for the Ph-3 locus, which confers late blight resistance in tomato. *Sci. Hort.***164**: 9-16.

Paterson, A. H., Brubaker, C. L., and Wendel, J. F. 1993. A rapid method for extraction of Cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep.* **11**: 122-127.

Pearce, G. Moura, D. S, Stratmann, J., Ryan, C. A. 2001.Production of multiple plant hormones from a single polyprotein precursor. *Nature* **411**:817-820.

Porebski, S. L., Bailey, G., and Baum, B. R. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components.*Plant Mole. Biol. Rep.***5**(1): 8-15.

Prashanth, N., Yugander, A., and Lakshmi Bhavani, N.12015. DNA Isolation and PCR Amplification of turmeric varieties from Telangana state. *Int. J. Curr. Microbiol. Appl. Sci* . **4**(5): 485- 490.

Prior, C. 1978. A method of inoculating young cocoa plants with basidospores of *Ceratobasidium theobromae. Ann. Appl. Biol.* **88**: 357-362.

Prior, C. 1980. Vascular streak dieback. *Cocoa Growers' Bulletin.* **29**: 21-26.

Pugh, T., Fouet, O., Risterucci, A. M., Brottier P., Abouladze M., Deletrez, C., Courtois, B., Clement, D., Larmande, P., Ngoran 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.

Purwantara, A., Iswanto, A., Sukamto, S., McMahon, P., Purung, H., Lambert, S., Guest, D., and Keane, P. 2009. New symptoms of vascular-streak dieback of

cocoa in Southeast Asia possible causes, studies required and control. In: *COPAL International Cocoa Research Conference*. No. 16.p.56.

Ramsay, **L.,** Macaulay, M., McLean, K., Fuller, J., Edwards, K., Turesson, S. 2000. A simple sequence repeat-based linkage map of barley. *Genetics*.**156:** 1997-2005.

Raval, R. R., Lau, K. W., Tran, M. G., Sowter, H. M., Mandriota, S. J., Li, J. L., Pugh, C. W., Maxwell, P. H., Harris, A. L. and Ratcliffe, P. J. 1998. Contrasting properties of Hypoxia-Inducible Factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma. *Mol. Cell Biol.* **25**: 5675-5686.

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

Rehem, B. C., Almeida, A. A. F., Corrêa, R. X., Gesteira, A. S., Yamada, M. M., and Valle, R. R. 2010. Genetic mapping of *Theobroma cacao* (Malvaceae) seedlings of the Parinari series, carriers of the lethal gene *Luteus-Pa. Genet. Mol.Res*. 9(3): 1775-1784.

Ren, Q., Liu, H. J., Zhang Z. Y., Feng, J., Xu, S. C., Pu, Z. J. and Xin, Z. 2012. Characterization and molecular mapping of a stripe rust resistance gene in synthetic wheat CI110. *J. Integr. Agric.* **11**(4): 521-527.

Ribeiro, R. A., and Lovato, M. B. 2007. Comparative analysis of different DNA extraction protocols in fresh and herbarium specimens of the genus *Dalbergia*. *Genetic Mol. Res.* **6**: 173-187.

Rice, R. A. and Greenberg, R. 2003 The Chocolate tree: growing cacao in the forest. *Natural History,* **112**:36-43.

Risterucci, A. M., Eskes, B., Fargeas, D., Motamayor, J.C., and Lanaud, C. 2001.Use of microsatellite markers for germplasm identity analysis in cocoa In: Proceedings of the 3rd International Group for Genetic Improvement of Cocoa INGENIC Accession and plot homogeneity in the ICGT.p.63.

Risterucci, A. M., Grivet, L., N'Goran, J. A. K., Pieretti, I., Flament, M. H., and Lanaud, C. 2000. A high-density linkage map of *Theobroma cacao*. *Theor. Appl. Genet.* **101**: 948-955.

Rivas, L. H., Giustina, L. D., Luz, L. N., Karsburg, I.V., Pereira, T. N. S. and Rossi, A. A. B. 2013. Genetic diversity in natural populations of *Theobroma subincanum* Mart. in the Brazilian Amazon. *Genet. and Mol. Res.* **12** (4): 4998-5006.

Russell, J. R., Hosein, F., Johnson, E., Waugh, R. and Powell, W.1993. Genetic differentiation of cocoa, *Theobroma cacao* L. populations revealed by RAPD analysis. *Mol. Ecol.* **2**: 89-97.

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A. and Allard, R. W. 1984. Ribosomal DNA sepacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci.***81**: 8014-8019.

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

Salimath, S. S., 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.

Sane, A., Chandrasekhar, A., and Manimekalai, R.. 2002. Optimizing DNA amplification parameters for RAPD analysis in cocoa (*Theobroma cacao* L.). *In: Proc National Seminar on Technologies for Enhancing Productivity in cocoa*. November 29-30, 2002, CPCRI,Vittal, Karnataka, p.28-32.

Santos, R.C. D., Pires, J. L., Lopes, U. V., Gramacho, K. P. G., Batistaflores, A., Bahia, R. D. C., Cristineramos, H. C., Corrêa, R. X. and Ahnert, D. 2005. Assessment of genetic diversity on a sample of cocoa accessions resistant to witches' broom disease based on RAPD and pedigree data. *Bragantia campinas* **64** (3): 361-368.

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.

Schnell, R. J., Kuhn, D. N., Brown, J. S., Olano, C. T., Phillips-Mora, W., Amores, F. M. and Motamayor, J. C. 2007. Development of a marker assisted selection program for cacao. *Phytopathol.* **97**: 1664-1669.

Shan, N. J., Yan, J. H., Jun, Y., Qin W . B., Qiang M. A. Z. and Min, S.T. 2010. Development of an STS marker linked to powdery mildew resistance genes *PmLK906* and *Pm4a* by gene chip hybridization. *Agric. Sci. China*. **9** (3): 331-336.

Shaw, D. E.1963. Diseases of cacao in Papua New Guinea. *The Papua and New Guinea Agril. J.* **15**: 79-90.

Shepherd, M., Cross, M., Stokoe, R. L., Scott, L. J. 2002. High throughput DNA extraction from forest trees. *Plant. Mol. Biol. Rep*. **20**: 4-25.

Shepherd, R., Gilbert, J. R. and Cowling, P. G.1977. Aspects of cocoa cultivation under coconut on two estates in Peninsular Malaysia. *Planter* **53** (612): 99-117.

Sujay, V., Gowda, M. V. C., Pandey, M. K., Bhat, R. S., Khedikar, Y. P., Nadaf, H. L.,Gautami, B., Sarvamangala, C., Lingaraju, S., Radhakrishan, T., Knapp, S. J. and Varshney, R.K., 2012. Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). *Mol. Breed.* **30:** 773-788.

Suwastika, I. N., Muslimin, R., Aisyah, N., Rahmansyah, M., Ishizaki,Y., Basri,Z., Shiina, T. 2015. Genotyping based on SSR marker on local cacao (*Theobroma cacao* L.) from central Sulawesi. *Procedia Environ. Sci.* **28:** 88-91.

Swami, R. M., Mahatma, M. K., Parekh, M. J., and Khandelwal.V. 2015. Molecular characterization of wilt resistant and susceptible genotypes of *Cajanus cajan* (L.). *Legume Res.* **38**(6): 726-733

Talbot, P. H. B. and Keane, P. J. 1971. *Ceratobasidium*: a new genus of tulasnelloid fungi. *Aust. J. Bot.* **19**: 203-206.

Tan , G.Y. and Tan, Y . W. K. 1988. Genetic variation in resistance to vascular-streak dieback in cocoa (*Theobroma cacao* L.) *Theor. Appl. Gent.* **75** (5):761-766.

Tel-Zur, N. S., Abbo, S., Myslabodksi, D., and Mizrahi, Y.1999. Modified CTAB procedure for DNA isolation from epiphytic cacti of the genera *Hylocereus and Selenicereus* (Cactaceae). *Plant Mol. Biol. Rep.* **17**: 249-254.

Thondaiman,V., Rajamani, K., Senthil, N., Shoba, N. and Joel, A. J. 2013. Genetic diversity in cocoa (*Theobroma cacao* L.) plus trees in Tamil Nadu by simple sequence repeat (SSR) markers. *Afr. J. Biotechnol.***12** (30): 4747-4753.

Tommerup, I. C., Barton, J. E. and O'brein, P. A. 1998. Reliability of RAPD fingerprinting of three basidiomycetes fungi, *Laccaria*, *Hydnangium* and *Rhizoctonia*. *Mycol. Res.* **99**: 179-186.

Turner, P. D. 1967. Cacao dieback: a review of present knowledge. *FAO Plant. Prot. Bull.* **15**: 81-101.

Van der Vossen, H. 1997. Strategies of variety improvement in cocoa with emphasis on durable disease resistance. Proceedings of the international workshop on the contribution of disease resistance to cocoa variety improvement. Salvador-Bahia, Brazil, INGENIC, United Kingdom, Nov. 24-28, 1996.p.23-32.

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.

Weishing , K., Nybom, H., Wolff, K. and Meyer, W. 1995. DNA isolation and purification. In: DNA fingerprinting in plants and fungi, CRC Press, Boca Raton, Florida.p.44-59.

Welsch, J., Honeycutt, R. J., McClelland, M. and Sobral, B. W. S. 1991. Parentage determination in maize hybrids using the arbitrarily primed polymerase chain reaction (AP-PCR). *Theor. Appl. Genet.* **82**: 473-476.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18**: 7213-7218.

Wettasingf, R. and Peffley, E. B. 1998. A rapid and efficient extraction method for onion DNA. *Plant Breed.* 117: 588-589.

Whitkus, R., Cruz, D. L., M., Mota-Bravo, L. and Gomez-Pompa, A. 1998.Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theor. Appl. Genet.* **96**: 621-627.

Wilde, J., Waugh, R. and Powell, W. 1992. Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. *Theor. Appl. Genet.* **83**: 871-877.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **18**: 6531-6535.

Wood, G. A. R. and Las, B. G. D. 1985. The genetic diversity of cacao and its utilization. (Wallingford, UK: CABI Publishing). p.1-10.

Yang, J. Y., Motilal, L. A., Dempewolf, H., Maharaj, K., and Cronk, Q. C. 2011. Chloroplast microsatellite primers for cacao (*Theobroma cacao*) and other Malvaceae. *Am. J. Bot.* **98**: 372-374.

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. *Mol. Breed.* **1**:375-387.

Yu, J. K., Dake, T. M., Singh, S., Benscher, D., Li, W., Gill, B., Sorrells, M. E. 2004. Development and mapping of EST-derived simple sequence repeat markers for hexaploid wheat. *Genome*, **47**: 805-818.

Zainal, A. M. A., Varghese, G. and Mainstone, B. J. 1984. Aspects of the epidemiology of vascular-streak dieback of cocoa in Malaysia. *Pro. Int.Conference on Cocoa and Coconut*. Malaysia, 15-17 Oct.1984. p.771-779.

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, L. B., Simmons, M. P., Kocyan, A. and Renner, S. S. 2006. Phylogeny of the cucurbitales based on DNA sequences of nine loci from three genomes: implications for morphological and sexual system evolution. *Mol Phylogenet. Evol.* **39** (2): 305-322.

Zhang, N., Yang, W. and Liu, D. 2011. Identification and molecular tagging of leaf rust resistance gene (*Lr24*) in wheat. *Agric. Sci. in China*. **10** (12)**:** 1898-1905.

Zhu, J. G., Qing, Y. Y., Ying, M., Feng, Y. X., Yu, C. S., Rui, R., Zhu, W. H. S., Wu, T. and Zhang, X. 2006. Inheritance analysis and identification of SSR markers linked to late blight resistant gene in tomato. *Agric. Sci.China*. **5** (7): 517-521.

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

Annexures

ANNEXURE I

List of laboratory equipments used for the study

Equipment Name	**Company**
Refrigerated centrifuge	Kubota 6500, Japan
Horizontal electrophoresis system	Wide Mini subset GT, BioRad, USA
Thermal cycler	Life technologies (Proflex) and Agilent Tenologies super cycler
Gel documentation system	BioRad Gel DOC-It™ universal Hood
Nanodrop® ND-1000 spectrophotometer	Nanodrop®Technologies lnc.USA
UV -transilluninator	UV transilluminator (Herolab®)
pH meter	ELUTECH INSTRUMENTS PC 510
Drybath	Labnet International Inc.

ANNEXURE II

Reagents required for DNA isolation

Reagents:

1. 2X CTAB extraction buffer (for 100 ml)

-CTAB (Cetyl trimethyl ammonium bromide): 2g

-Tris base: 1.21 gm

-EDTA: 0.745 gm

-NaCl: 8.18 gm

-PVP: 1.0 gm

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

Ethanol 76 ml and distilled water 24 ml.

Ammonium acetate 0.077 g

Added the 0.077 gm of ammonium acetate in 100 ml 76% ethyl alcohol and mixed thoroughly.

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 (1mM): 0.037g (The solution was prepared, autoclaved and stored at room temperature).

ANNEXURE III

Composition of buffers and yes used for gel electrophoresis

1. TAE Buffer 50X (1000 ml)

Tris base: 242 g

Glacial acetic acid: 57.1 ml

0.5M EDTA (pH 8.0) - 100 ml

2. Loading Dye (6X)

0.25% bromophenol blue

0.25% xylene cyanol

30% 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.

4. Molecular weight marker (λ DNA *Eco*RI / *Hind* III double digest- Invitrogen)

5. RNase -Promega

6. DNA elution kit- SIGMA ALDRICH

7. Cloning Kit- Promega

Abstract

CHARACTERIZATION AND VALIDATION OF MICROSATELLITE MARKERS FOR RESISTANCE TO VASCULAR STREAK DIEBACK DISEASE IN COCOA (*Theobroma cacao* L.)

By

WAGHMARE SANDESH TULSHIRAM

(2014-11-101)

ABSTRACT OF THE THESIS

Submitted in partial fulfillment of the requirement

for the degree of

Master of Science in Agriculture

(Plant Biotechnology)

Faculty of Agriculture

Kerala Agricultural University

CENTRE FOR PLANT BIOTECHNOLOGY AND MOLECULAR BIOLOGY
COLLEGE OF HORTICULTURE
VELLANIKKARA, THRISSUR – 680656
KERALA, INDIA

2016

Abstract

Cocoa is the third important plantation crop next to coffee and tea. The global production and consumption of cocoa is 27.00 lakh MT. Among the fungal diseases, Vascular Streak Dieback (VSD) caused by *Ceratobasidium theobromae* is the main constraint in cocoa growing countries, causing heavy losses in mature trees as well as seedlings. The VSD disease cannot be effectively controlled by chemicals and hence breeding for the development of resistant varieties is the best strategy to tackle the disease.

In order to confirm the transfer of a desired gene into the offspring, conventional breeding methods rely on the field screening which will be highly influenced by the environmental factors. Marker assisted selection is an alternate where the tightly linked molecular markers will be employed to confirm the presence of the gene of interest in the selected plants.

Five ISSR and one SSR markers linked to VSD resistance were identified at Kerala Agricultural University (Chandrakant, 2014). The present study was undertaken with the objective of validating the identified SSR and ISSR markers and to characterize the ISSR markers to identify the corresponding SSR markers.

For validation and characterization, twenty VSD resistant hybrids and four susceptible clones were used.

For molecular analyses, good quality genomic DNA was isolated from twenty four genotypes and ISSR markers UBC 811, UBC 815, UBC 826, UBC 857, UBC 866 and SSR marker mTcCIR 42 were screened. ISSR analysis had shown that all the primers are capable to differentiate resistant and susceptible genotypes. The SSR assay has also differentiated the resistant and susceptible genotypes.

The distinct markers generated in resistant genotypes using UBC 811, UBC 826 and UBC 857 were eluted, cloned to pGEMT vector and sequenced. The

nucleotide sequences were annotated using BLAST, ORF finder and SSR finder. The BLASTn of UBC811A and UBC811D nucleotide sequence have shown that this resistance locus lie in the chromosome V of *Theobroma cacao* genome. BLASTn of UBC826A, UBC826B and UBC857 has positioned these loci in chromosome III.

ORF1 and ORF3 in UBC811D are shown to code for aflatoxin biosynthesis regulatory protein and NAD(P)H dehydrogenase quinine, respectively. ORF1 in UBC826B and ORF5 in UBC857-2 code for potassium transporter 27 (*Os*HAK-27) and structural polyprotein precursor of VP2, capsid protein VP2, respectively. All these proteins are identified to have definite roles in defence pathways. The frequency and distribution of SSR motifs, dimmers to decamers, in these ISSR markers and the corresponding primers were identified.

The reported ISSR and SSR markers were validated and found to be successful in differentiating resistant and susceptible genotypes of cocoa; thereby these markers can be used in marker assisted breeding for VSD resistance.