Development of a genetic linkage map using molecular markers in para rubber tree (*Hevea brasiliensis*)

Thesis submitted to
The University of Kerala
In partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

In

Biotechnology

Under the Faculty of Applied Science

By
Bini. K

THE RUBBER RESEARCH INSTITUTE OF INDIA
RUBBER BOARD
(Ministry of Commerce & Industry, Govt. of India)
Kottayam – 686 009, Kerala, INDIA

JANUARY 2013

DECLARATION

I hereby declare that this Ph.D thesis entitled **"Development of a genetic linkage map using molecular markers in para rubber tree (*Hevea brasiliensis*)"** is an authentic and bonafide record of the research work carried out by me under the supervision of Dr.Thakurdas Saha, Senior Scientist (Genome Analysis Laboratory), Rubber Research Institute of India, Kottayam. The work presented in this thesis has not been submitted earlier for any other degree or diploma elsewhere.

Kottayam

10/01/2013

Bini.K.

भारतीय रबड़ गवेषण संस्थान
THE RUBBER RESEARCH INSTITUTE OF INDIA
(वाणिज्य मन्त्रालय, भारत सरकार)
(Ministry of Commerce, Government of India)

Phone : 91-481-2353311(10 lines)
Fax : 91-481-2353327
Email : saha@rubberboard.org.in
Website : http://www.rubberboard.org.in

रबड़ बोर्ड
RUBBER BOARD
कोट्टयम - 9
KOTTAYAM 686 009

Certificate

This is to certify that this thesis entitled "**Development of a genetic linkage map using molecular markers in para rubber tree (*Hevea brasiliensis*)**" is an original record of the research work carried out by Ms. Bini K. under my supervision at the Genome Analysis Laboratory, Rubber Research Institute of India, Kottayam. The work presented in this thesis is original and has not been submitted anywhere else for the award of any other degree or diploma. The help and sources of information availed during the course of this investigation have been duly acknowledged.

Kottayam
10/01/2013

(Dr. Thakurdas Saha)

DR. T. SAHA
Senior Scientist
Genome Analysis Laboratory
Rubber Research Institute of India
Kottayam - 686 009, Kerala
saha@rubberboard.org.in

Acknowledgement

This thesis has been kept on track and had been seen through to completion with the support and encouragement of numerous people. It is a pleasant task to express my thanks to all those who contributed in many ways.....

Foremost, I extend my deep gratitude to The God Almighty for his shower of blessings throughout my research.

At this moment of accomplishment, I would like to express my deep and sincere gratitude to my research supervisor Dr. Thakurdas Saha, (Senior Scientist, Genome analysis laboratory, RRII). I am truly indebted to him for his peer supervision, endless support and encouragement and without him this thesis would not have been possible. I always remember him as a role model towards dedication and sincerity towards any work in my career.

I wish to extend my special thanks to Smt. Minimol Raveendran and Sri. K. U. Thomas for their immense support, care and loyal assistance provided to me throughout the period of my work in RRII. I can never forget the decisive and energetic support of C. Aishwarya, project fellow at the time of manuscript preparation. I thank Dr. C. Bindu Roy for her timely help at various stages of my work.

I express my sincere thanks to Dr. James Jacob, Director of Research and Dr. N. M. Mathew, (retd. Director of Research), RRII, Kottayam for granting me the permission for conducting this study at the Genome Analysis Laboratory of RRII and also for extending the necessary infrastructure for the successful completion of my PhD programme. It is my pleasure to acknowledge Dr. M.A. Nazeer, Joint Director (Project Monitoring), RRII for his sincere efforts which enabled me to utilize my CSIR fellowship properly and efficiently.

I always remember with gratitude the valuable support and inspiration provided to me by Dr. Jayasree Madhavan, Dr. M. A. Mercy, Dr. C.P. Reghu, Sri. G.P.Rao and all the staff and workers of the Genome and Germplasm division of RRII. The whole hearted support provided to me by Dr. A. Thulasedharan (Deputy Director, Biotechnology), Dr. Jessy M.D (Deputy Director, Agronomy) and Dr. Y. Annamma Varghese, ex- Joint Director (Crop improvement), RRII is kindly acknowledged.

I thank Dr. Kavitha K. Mydin (Joint Director, Crop Improvement) for providing the plant materials for the project, Dr. C. Narayanan, Dr. T. Girish, and Meenakumari, Scientists of Botany Division, RRII, Shri. Ramesh B Nair (Joint Director, Statistics), all the staff of the computer section and library for their valuable suggestions.

I can never forget Supriya. R who stood with me at the time of adversaries as a true friend and for all the support, love and affection. I thank Sri. Satheesh P.R, Smt.Suni and Anuja Dileep, Anichettan for their unconditional support provided to me at the time of manuscript preparation.

I thank all the RRII Scientific staff and those people, who have in some way or other helped me during my tenure.

I can remember only with a deep sense of gratitude the support and true inspiration provided to me by Sri. C. R. Sivaji, Secretary of PCARD Bank Kottayam and Sri. Raveendranathan Nair , Regional Manager, KSCARD Bank Kottayam. I thank Kum. Renju. T soil survey officer and Smt. Minija S. Selvan Assistant Director (Soil Survey), Palakkad for the help rendered to me at the final stage of thesis submission.

I would like to truly admit that my Ph.D endeavor might not be possible without the constant encouragement and support from my family. I thank my husband Sri. S. Suresh, who always stood with effective guidance and kind support for the successful completion of thesis. I specially memorize my mother-in law, for her constant support at every stages of my work and I consider this as a reward to her endless prayers. This achievement may not have possible without the true blessings of my father and mother and the endless support from my brother, sister, uncles and all in-laws. Words fail me to extend my gratitude to Aravind R Nair, my brother in-law for his assistance at the write-up stage of the thesis and also Ananad, (Joy's Computer centre, Puthupally). All these years I always cherished the amusing actions of my sons Aadikrishnan and Aanandkrishnan. I dedicate this thesis to God Almighty.

The financial assistance provided to me by CSIR, New Delhi is acknowledged

Contents

Chapter II

List of figures

List of Tables

AFLP	:	Amplified fragment length polymorphism
APS	:	Ammonium per sulphate
BD	:	Bodjong Datar
BSA	:	Bovine Serum Albumin
CAPS	:	Cleaved Amplified Polymorphic Sequence
cM	:	Centimorgan
CSPD	:	Chloro-5-substituted adamantyl-1,2-dioxetane phosphate ()
CTAB	:	Cetyl Trimethyl Ammonuim Bromide
dNTP	:	Deoxynucleotide triphosphate
EDTA	:	Ethylene diamine tetra acetic acid
EST	:	Expressed Sequence Tag
Gl	:	Glenshiel
GT	:	Godang Tapen
ISSR	:	Inter-Simple Sequence Repeats
LG	:	Linkage Group
MAS	:	Marker Assisted Selection
NTSYS	:	Numerical Taxonomy and Multivariate Analysis System
PB	:	Prang Besar
PCR	:	Polymerase Chain Reaction
PR	:	Proefstation Voor Rubbe
QTL	:	Quantitative trait loci
RAPD	:	Random Amplified Polymorphic DNA
RFLP	:	Restriction Fragment Length polymorphism
RGA	:	Resistance Gene Analogues
RRIC	:	Rubber Research Institute of Ceylon
RRII	:	Rubber Research Institute of India
RRIM	:	Rubber Research Institute of Malaysia
SCAR	:	Sequence-characterized amplified region
SCATC	:	South China Academy of Tropical Crops

SNP	:	Single Nucleotide Polymorphisms
SSR	:	Simple sequence repeat
STMS	:	Sequence Tagged Microsatellite
STS	:	Sequence Tagged Site
TAE	:	Tris acetate EDTA buffer
TBE	:	Tris borate EDTA buffer
TEMED	:	Tetramethyl ethylene diamine
Tjir	:	Tjirandji
UPGMA	:	Unweighted Pair Group Method of Arithmetic averages
VNTR	:	Variable number of tandem repeats

Reagents and solutions

Acrylamide - 40% (100ml)
Acrylamide - 38g
N, N' bisacrylamide - 2g
Dissolve in MilliQ water
Sterilize by syringe filteration.

Ampicillin-100mg/ml
Ampicillin sodium salt - 1 g
Deionized water - 10 ml

Bind silane working solution (1ml)
Ethanol- 950µl
Glacial acetic acid-5 µl
Water-42 µl
Bind silane-3 µl

Blocking Reagent (10X)
Blocking Reagent 10 %(w/v)
Maleic acid buffer

Detection buffer
0.1 M Tris HCl pH 9
0.1 M NaCl
50 mM $MgCl_2$

Developing solution (1litre)
Na_2CO_3 - 30 g
Sterile water - 1000ml
Keep in ice and add 540µl formaldehyde prior to use

Denaturing solution
1.5M NaCl
0.5M NaOH

DNA Extraction Buffer
CTAB - 2% (W/V)
NaCl - 1.4M
EDTA - 20mM
Tris HCl(pH 8.o) - 100mM
β-mercaptoethanol - 0.3% (added just before use)

Ethidium Bromide
10 mg/ml in water

Fixing solution (1litre)
Acetic acid – 10%

Formamide gel loading dye
Formamide - 80%
EDTA-10mM
Xylene cyanole-1mg/ml
Bromophenol blue-1mg/ml

Gel Loading Buffer (6X)
Bromophenol Blue - 0.25%
Xylene Cyanol FF - 0.25%
Glycerol - 50%
MilliQ water - 10ml

IPTG (0.1 M) solution
1.2 g IPTG
MilliQ water to make up final volume of 50 ml

RNase A-(10 mg/ml)
RNase powder - 10mg
Sodium acetate - 0.01 M (pH 5.2)
Heat to 100^0C for 15 min, allow to cool slowly to room temperature.

TE Buffer (10:1)
Tris-HCl(pH 8.0) - 10 mM
EDTA (pH 8.0) - 1 mM

TAE Buffer (50X) per litre
Tris base - 242 g
Glacial acetic acid - 57.1 ml
0.5 M EDTA(pH 8.0) - 100 ml

TBE Buffer (6X) per litre
Tris base - 64.8 g
Boric Acid - 33g
0.5 M EDTA(pH 8.0) - 24 ml

TBE-Urea
6X TBE and 7M urea

Maleic acid buffer
0.1 M maleic acid
0.15 M NaCl.
Adjust pH to 7.5 with NaOH pellets.

Oxidizing solution (1litre)

Concentrated HNO_3 - 15 ml
Sterile water - 985ml

Staining solution (1litre)

$AgNO_3$ - 2g
Sterile water - 1000ml

LB broth (1litre)

Yeast Extract – 5 g
Sodium Chloride – 10 g
Tryptone – 10 g
NaOH (for adjusting pH)
MilliQ water

LB Agar (1litre)

Yeast Extract – 5 g
Sodium Chloride – 10 g
Tryptone – 10 g
NaOH (for adjusting pH)
MilliQ water
Agar – 16 g

X-Gal solution - 50mg/ml

X-Gal - 100 mg
Dimethyl formamide - 2 ml

LB-Amp-Glycerol medium

LB broth - 100 ml
Ampicillin(100mg/ml) - 100 µl
Glycerol - 7.5%

LB top agar

Prepare 1L of LB broth.
Add 0.7 % (w/v) agarose.

Neutralizing solution

1.5 M NaCl
0.5 M Tris pH 8
0.001 M EDTA

SSC - 20X

3M Sodium Chloride
0.3M Sodium Citrate

Abstract

Hevea brasiliensis (para rubber tree) belonging to the family Euphorbiaceae is indigenous to the tropical rain forests of Central and South America and is the major commercial source of natural rubber (*cis*-1,4-polyisoprene). Rubber tree is a perennial tree crop with a long juvenile period. It is heterozygous in nature having a ploidy level of 2n = 2x = 36. Cultivated *H. brasiliensis* clones of Southeast Asian rubber-growing countries were evolved from a small collection of Sir Henry Wickham from Brazil (Wickham clones). Therefore, the genetic base of these clones is assumed to be narrow, which is a constraint for further improvement of this crop.

The aim of the present work is to develop DNA-based molecular markers from *Hevea* genome for their practical application in genetic diversity analysis of cultivated clones and genetic linkage map construction, which is necessary for mapping of traits, marker assisted selection in breeding experiments in rubber.

Different marker systems comprising of dominant markers like RAPD, AFLP, RGA and co-dominant markers: SSR (both di- and tri-nucleotide repeat markers) and SNP were generated successfully from *Hevea* genome and used in polymorphism studies of 28 popular rubber clones. Genetic relationship studies with RAPD, AFLP and SSR (genomic SSR and genic/EST-SSR) markers clearly revealed low levels of genetic diversity (~30%) existing among popular clones under study. The three marker assays differed in the amount of polymorphisms detected among rubber clones. Cluster analysis of combined marker data revealed that the similarity coefficient varied from 0.74-0.9. No specific grouping was observed for the clones based on their origin, except for the Sri Lankan clones. Comparative analysis of the three marker systems showed that SSR markers were the most informative.

Development of genetic maps and subsequent marker assisted selection are the important applications of molecular breeding in any crop. Genetic maps indicate the position and relative distances between markers along the chromosomes. Linkage mapping studies in rubber tree followed a 'pseudo-test cross' strategy involving only the existing pedigrees. In the present study, a linkage map was generated using a segregating progeny population comprising of 61 individuals, obtained from the cross between two popular cultivated *Hevea* clones: RRII 105 and RRII 118.

The informative RAPD, AFLP, SSRs, SNPs and RGA markers, generated through polymorphism studies, were used in the construction of genetic linkage maps. For the development of RAPD markers, 520 'Operon' arbitrary decamer primers were screened, of which 46 primers generating 100 loci were used in the progeny population for segregation analysis. Seventy-eight primer combinations with *Eco*RI and *Mse*I adapter sequences were screened for AFLP marker development and 19 combinations showing more number of scorable polymorphic bands were analyzed in mapping population.

One hundred and four SSR markers based on dinucleotide repeats from the genomic library, developed in our laboratory, were screened and 22 of them were used in the analysis. Thirty-one SSR markers developed from an enriched trinucleotide repeat library were checked for polymorphism between the parents and only two markers could be used for segregation analysis. Markers based on EST derived SSRs can serve the purpose of mapping the gene itself. Hence, 67 EST derived SSR markers from *H. brasiliensis* were tested for polymorphism between the parents. Out of these, 25 were identified as useful markers for segregation analysis.

SNP markers were developed for linkage mapping, since they contribute directly to a phenotype or can be associated with a phenotype as a result of linkage disequilibrium. Primers were preferentially synthesized from the 3' UTR of 12 genes (cDNA) encoding enzymes involved in several important biosynthesis pathways in rubber using the published sequences existing in the GenBank. Detection of SNPs was performed in 16 popular *H. brasiliensis* clones. SNPs were identified only in five out of 12 loci sequenced from all the 16 genotypes. These five loci are geranylgeranyl diphosphate synthase, farnesyl diphosphate synthase, mevalonate kinase, ubiquitin precursor and latex patatin homolog. Genotyping of SNPs in progeny population was performed following the allele-specific PCR amplification (AS-PCR) for geranylgeranyl diphosphate, mevalonate kinase and latex patatin homolog genes, and cleaved amplified polymorphic sequence (CAPS) technique for the gene ubiquitin precursor and sequence tagged site (STS) marker based on a large insertion and deletion in farnesyl diphosphate synthase gene. SNP haplotype variations identified in latex biosynthesis genes will be useful in understanding functional variability of respective genes in both low and high yielding rubber clones for marker-assisted selection (MAS).

A total of 244 markers/loci comprising of 100 RAPD markers, 86 AFLP markers, 52 SSR markers, five SNP based and one RGA markers were finally employed for the construction of a genetic linkage map. The markers generated were analyzed using both

the linkage map construction software MAPMAKER/EXP 3.0 and JoinMap 3.0. The dominant and co-dominant markers following a segregation ratio of 1:1 were analyzed initially using the linkage map construction software MAPMAKER/EXP 3.0, as the other segregation ratios generated from a heterozygous cross could not be handled by MAPMAKER. Marker groups were determined by using a maximum likelihood distance of 40 and a minimum LOD score of 2.0. Linkage groups were detected for each parent separately. In *H. brasiliensis*, the number of linkage groups for a saturated linkage map should be equal to the haploid chromosome number of 18. In this study 23 linkage groups were formed for RRII 105 and also for RRII 118, which could be due to the insufficient number of markers. All the 23 linkage groups in RRII 105 showed a cumulative map distance of 1384.7 cM, whereas RRII 118 had 768.5 cM from 23 linkage groups. Twenty-six markers were unlinked in RRII 105 out of 115 loci analyzed and 33 markers remained unlinked out of 99 markers analyzed in RRII 118.

Segregation analysis of all the markers was repeated using the software JoinMap 3.0 which could handle all types of segregation data. Marker groups were determined using a minimum LOD score of 3.0 and a recombination frequency of 0.4. Segregation distortion was noticed in 4.8% of markers in RRII 105 and 4% of markers in RRII 118, having a high Chi-Square value, and subsequently deleted from the analysis. Twenty-four linkage groups were identified both for RRII 105 and RRII 118. The total genetic distance covered was 762 cM in RRII 105 and 634 cM in RRII 118. Sixty-one markers were unlinked in RRII 105 out of 164 loci analyzed and 62 markers remained unlinked out of 145 markers analyzed in RRII 118. The distance calculated by JoinMap was less compared to that calculated by MAPMAKER due to differences in the computational procedures. Two linkage groups harboring bridge markers common to both the parents were used to merge the respective groups.

The linkage maps created in this study are preliminary but is a significant step towards developing a high-density genetic map incorporating more number of markers and using a population from a wide genetic base resulting from an inter-specific cross combination. This also provides a base for QTL mapping studies in which phenotypic traits of importance can be identified and associated to a specific location of the genome for MAS. The linkage map developed in the present study is the first attempt from the Rubber Research Institute of India.

General introduction

Hevea rubber

Hevea brasiliensis (Willd. ex Adr. de Juss.) Muell. Arg., also known as para rubber tree belongs to the family Euphorbiaceae. It grows wild in the tropical rain forests, South of the Amazon River, from 14^0-15^0 South latitude to 75^0 West longitude covering a wide range of ecological and bioclimatic regions, encompassing the countries of Brazil, Bolivia, Columbia, Peru, Eucador, Venezuela, Surinam and Guyana. It is the principal source of natural rubber (NR) (Fig. 1.1).

Sir Henry Wickham, the father of Natural Rubber, made a collection of rubber seeds from the confluence of river Tapajos and the Amazon (Schultes, 1977). The seeds were distributed to different Southeast Asian countries. The seedling stocks survived at Malaysia had become the base for crop improvement of *Hevea* in 20^{th} century, which is known as the 'Wickham base' (Simmonds, 1989). The source of planting material to the Southeast Asian rubber plantation had its origin from the Wickham collections and from there rubber cultivation spread to all other Asian countries (Schultes, 1977; Simmonds, 1989). In India, rubber cultivation was started in 1878 in Nilambur, Kerala State, as a forest crop, using the planting materials brought from the Royal Botanical Garden, Ceylon (Petch, 1914). However the first commercial rubber plantation of rubber was started in India by European planters in 1902 at Alwaye.

Based on taxonomic description, 10 species *viz., H. benthamiana, H. brasiliensis, H. camargoana, H. guianensis, H. microphylla, H. nitida, H. pauciflora, H. rigidifolia, H. camporum and H. spruceana* have been recognized in the genus (Schultes, 1970; 1977). There are no genetic barriers among the 10 species of *Hevea* and they are inter-crossable (Clement–Demange *et al.*, 2000). There were disputes regarding the ploidy level of the genome. Ong *et al.* (1976) described it as amphidiploids (2n = 36, x = 9) and this contention was disputed at the molecular level by Leitch *et al.* (1998) suggesting a possible allotetraploid origin. However, investigations by Saraswathyamma *et al.* (1984) and the recent molecular studies have confirmed that the chromosome compliment in the somatic cell is 2n = 2x = 36 (Clement-Demange *et al.*, 2007).

Fig. 1.1. Mature plantation of *Hevea brasiliensis*

Natural rubber production and use

The largest producer of natural rubber in the world is Thailand. India being the fourth largest producer of natural rubber in the world contributes 8.2% (0.89 million tones) of the total world production (10.97 million tones) in 2011 (Fig. 1.2) [Source: International Rubber Study Group (IRSG), Singapore].

Rubber producing areas in India can be classified under two major zones: Traditional (Districts of Kerala and Kanyakumari in Tamil Nadu) and non-traditional (coastal regions of Karnataka, Goa, Andhra Pradesh, Orissa, some areas of Maharashtra, Northeastern States mainly Tripura and Andaman & Nicobar Islands from a cultivated land of 7,11,560 ha. Kerala and Tamil Nadu together occupy 86% of the growing area of natural rubber. Kerala alone contributes 90% of India's total production of natural rubber.

The natural rubber produced from latex of the rubber tree accounts for about 41% of the world's total rubber consumption, while 60% is delivered through synthetic process. Para rubber is obtained by tapping the trunks of the trees. Natural rubber is a high molecular weight polymer whose structure is *cis*-1,4-polysoprene, produced in a specialized system (latex vessels) present in all organs of the plant (Mathew, 1992). Cured rubber used for all types of rubber products such as rainwear, diving gear, chemical and medicinal tubing, lining for storage tanks, processing equipment etc. Because of their electrical resistance, soft rubber goods are used as insulation and for protective gloves, shoes and blankets; hard rubber is used for articles such as telephone housings, parts for radio sets, meters, and other electrical instruments. Seeds are source of Para rubber seed oil, recommended for manufacture of soap. In the present scenario the timber from rubber tree is also an additional resource, which is now considered as an alternate source for the depleting natural timber. Hence, research is being conducted for the development of latex timber clones, diverting from the unilateral focus on latex production alone.

***Hevea* breeding**

H. brasiliensis is monoecious with both male and female flowers on the pyramid shaped panicle, ensuring a high degree of cross pollination. Sticky pollen and stigmatic surfaces

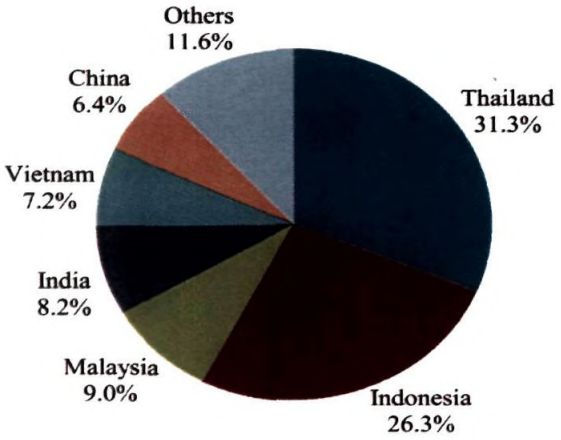


Fig. 1.2. Diagram showing the share of natural production by major rubber growing countries

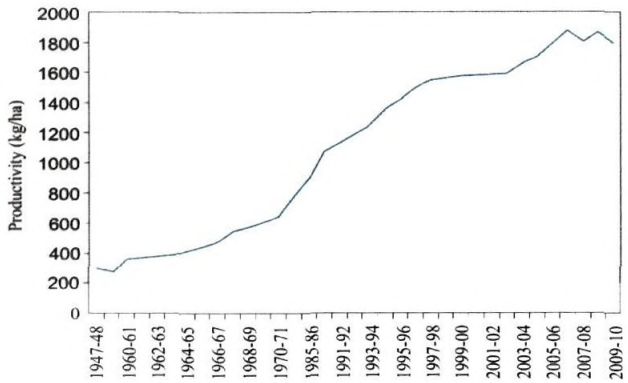


Fig. 1.3. Graph representing the increase in natural rubber productivity over the years

indicate typical entomophilous nature of the flower. Rubber is reported to tolerate annual precipitation of 1500 to 4000 mm, annual temperature of 23.1 to 27.5°C and pH of 4.3 to 8.0. The economic life span of the tree is around 32 years and the plantation starts its yield from the sixth year onwards.

The ultimate objective of breeding in *Hevea* is to develop high yielding clones combined with desirable secondary attributes. However the specific objectives vary depending on the agroclimatic and socio-economic requirements. Being a tree crop, there are lots of breeding constraints in *Hevea brasiliensis.* The common procedures employed for crop improvement involves introduction, plus tree/ mother tree/ ortet selection, hybridization followed by selection, special breeding techniques involving mutation breeding and polyploidy breeding can also be adopted. It is evident that the breeding efforts were fruitful with the development of high yielding clones from the gene pool introduced to Southeast Asia by Sir Henry Wickham in 1876 with an average yield of 200 to 300 Kg/ha/year (Panikkar *et al.*, 1980) (Fig. 1.3). India stands first among the world major producers in productivity with an average yield of 1842 kg/ha/yr in the year 2011-2012.

Molecular markers

Modern plant breeding methods are based on scientific principles of genetics, which began only with the rediscovery of Mendel's paper that was originally published in 1866 (Mendel, 1866). Science of genetics has progressed at a rapid pace and has laid the foundation for molecular plant breeding with the integration of new advances in biotechnology, genomics and molecular marker applications. The development of marker systems was initiated with the mutations at the loci controlling plant morphology (Stadler, 1929). But the morphological markers provided little information on many complex genetic factors and led to the development of molecular markers. The advent of molecular marker technology has revolutionized the entire scenario of biological sciences and widened the existing knowledge on polymorphism between individuals – 'from morphological to molecular level'. The hype on the molecular marker technology was so prominent that this resulted in the development of new types of molecular markers and various technological simplifications to resolve the problems associated with molecular marker technology. The recent PCR based approach, gel free visualization of PCR

products and automation at various steps are the boons to the molecular marker approaches adopted for genome mapping and genetic diversity analysis in any organism.

Molecular markers are based on naturally occurring polymorphisms in the DNA sequence, *ie.*, base pair deletions, substitutions or additions (Gupta *et al.*, 1996). They are considered as landmarks on chromosomes, which are essential in finding out where the genes are placed in a genetic map. Molecular markers can identify a particular aspect of phenotype or genotype or both, whose inheritance can be easily followed from generation to generation. Isozymes were the first developed markers (Markert and Moller, 1959) before the advent of DNA based molecular markers. The concept of using variations at DNA level as genetic markers started with Restriction Fragment Length Polymorphism (RFLP) and its first documentation came from viruses (Grodzicker *et al.*, 1974) followed by a subsequent elegant demonstration made in the human α-globin gene cluster (Jeffereys, 1979).

Desirable properties of an ideal marker

Useful genetic markers should possess the following attributes:

- Highly polymorphic in nature
- Co-dominance (Ability to distinguish homozygous and heterozygous states of diploid organisms).
- Genomic abundance
- Neutral behaviour to environmental or management practices
- Locus specificity
- Small quantities of template DNA required
- High reproducibility
- Demanding less technical skills
- Easy access
- Easy and fast assay
- Amenability to automation
- Easy exchange of data between laboratories
- Less operational and developmental costs

Depending on the type of the study to be undertaken, a marker system can be identified that would fulfill some of the above mentioned characteristics.

Classification of molecular markers

Three main types of molecular markers are a) biochemical, b) hybridization based, and c) polymerase chain (PCR) based (Kumar *et al.*, 2009). Based on chronological evolution, a series of molecular marker system, which became available during the last two decades, can be broadly classified into three classes (a) the first generation molecular markers including Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNAs (RAPDs) and their modifications, (b) the second generation molecular markers including Simple Sequence Repeats (SSRs), Amplified Fragment Length Polymorphisms (AFLPs) and their modification forms, and (c) the third generation molecular markers including Expressed Sequence Tags (ESTs) and Single Nucleotide Polymorphisms (SNPs).

Biochemical markers (Isozyme)

Isozyme markers (Markert and Moller, 1959) have been used for over 60 years for various research purposes in biology. They are electrophoretically separable variants of an enzyme with qualitatively the same catalytic function (Bergmann *et al.*, 1989). Isozymes have been proven to be reliable genetic markers in breeding and genetic studies of plant species and have been successfully used in several crop improvement programmes (Vallejos, 1983; Glaszmann *et al.*, 1989, Baes and Custsem, 1993). The main drawback of isozyme is their relatively low abundance and low level of polymorphism (Schmidt, 1995).

Hybridization based molecular markers

Restriction fragment length polymorphisms (RFLPs) (Botstein *et al.*, 1980) are the most widely used hybridization based marker. The first DNA marker used in the construction of genetic maps was reported by Botstein *et al.* (1980) in human and thereafter used in plant research (Beckman and Soller 1986; Tanksley *et al.*, 1989; Weber and Helentjaris, 1989). These markers are co- dominant in nature and provide extremely reliable method for DNA typing, determining genetic relationships and genetic linkage mapping studies. But the disadvantages are, this technique requires large quantity of DNA (50-200 µg), less polymorphism exhibited and the lengthy procedure associated with it. Minisatellites/variable number of tandem repeats (VNTR) are similar to RFLPs, introduced by Jefferys *et al.* (1985).Due to the high mutation rate of minisatellites, the

PCR-based markers

The PCR based markers are mainly of two types, arbitrarily-primed markers and sequence tagged site (STS) markers of known sequence. Arbitrarily-primed markers include randomly amplified polymorphic DNAs (RAPDs) (Welsh and McLelland, 1990; Williams *et al.*, 1990) amplified fragment length polymorphisms (AFLPs) (Vos *et al.*, 1995) and inter-simple sequence repeats (ISSRs) (Zietkiewicz *et al.*, 1994). These marker technologies could generate large number of markers in a single reaction. But these markers suffer from the disadvantages of lack of reproducibility, especially in the case of RAPDs (Jones *et al.*, 1997) where the nucleotide sequence of the markers is unknown, and difficult to transfer between studies without converting them to sequence characterized markers (SCARs) (Paran and Michaelmore., 1993). AFLP technique is an intermediate between RFLPs and PCR and is based on the detection of genomic restriction fragment by PCR amplification and can be used for DNA of any origin or complexity (Martin *et al.*, 1991). These fragments are viewed on denaturing polyacrylamide gels either through autoradiography or flourescence (Vos *et al.*, 1995), which are detected as dominant markers (presence or absence of band). Selectively amplified microsatellite polymorphic locus (SAMPL) Witsenboer *et al.* (1997) is a modified AFLP technique, where the SAMPL analysis uses one AFLP primer in combination with a primer complementary to microsatellite sequences. Other modifications of AFLP technique include Methylation Sensitive Amplification Polymorphism (M-SAP) and Three Endonuclease AFLP (TE-AFLP) (Van der Wurff *et al.*, 2000).

Sequence tagged sites (STS) (Olsen *et al.*, 1989) are single locus markers with known sequences. They include microsatellite (SSR) markers which are technically difficult to produce but very reliable. They can be easily integrated into plant breeding programme for marker assisted selection of the trait of interest. The term microsatellite was coined by Litt and Lutty (1989). DNA consisting of tandemly repeating mono, di, tri, tetra or penta nucleotide units are distributed throughout the genomes of most eukaryotic species (Powell *et al.*, 1996). Microsatellites are co-dominant, highly abundant in genomic region of eukaryotes and randomly distributed throughout the genome with preferential association in low copy regions (Morgante *et al.*, 2002).

Expresses Sequence Tag (EST) markers and Single Nucleotide Polymorphic markers belongs to the new generation markers. ESTs are short DNA molecular (300-500

bp) reverse transcribed from a cellular RNA population (MacIntosh *et al.*, 2001) and the term was introduced by Adams *et al.* (1991). Single Nucleotide Polymorphisms (SNP) are single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in a population. SNP and insertions /deletions (Indels) are highly abundant and distributed throughout the genome in plants (Garg *et al.,* 1999). Automated methods for scoring SNPs in segregating population make analysis very effective in mapping approach (Landegren *et al.,* 1998). In plants, SNP seems to be more abundant than those in human genome as evidenced by the existence of one SNP per 20 bp in wheat (Picoult *et al.*, 1999) and one SNP per 70 bp in maize genome (Bhattramakki *et al.*, 2000). Studies of SNPs provide a frame work for examining how population history, breeding system and selection affect genetic loci and delineate the mechanisms that lead to evolutionary diversification of genomes (Kota *et al.*, 2003).

Another group of markers are the Resistance Gene Analogues (RGAs) which have been cloned from numerous plant species, which confer resistance to a wide range of plant pathogens, and encode proteins with a high degree of structural and amino-acid sequence conservation. Most of these proteins have domains involved in signal transduction or protein-protein interactions (Hammond-Kosack and Parker, 2003).

Applications of molecular markers

The molecular markers produced a great impact on various aspects of crop improvement from diversity studies to genome mapping, gene tagging, map based cloning and physical mapping of genes.

Diversity analysis of exotic germplasm and genotyping of cultivars

STMS, ISSR, AFLP markers etc. can provide useful criteria for evaluating the gene pool of crop plants. Diversity analyses can classify the existing biodiversity among plants, which can be further useful in wild gene introgression programmes (Brown and Kresovich, 1996). One of the simplest and most pervasive applications of DNA markers in plants is DNA finger printing, a technique that was forecasted by Soller and Beckmann (1983) for the detection of genetic variation, cultivar identification and genotyping. Microsatellite markers are the widely accepted markers of choice for this purpose. The multi-allelism of these markers facilitates comparative allelic variability detection reliably across a wide range of germplasm and allows individuals to be ubiquitously genotyped, so that gene flow and paternity can be established.

Phylogeny and evolution

The technique of molecular biology provides path-breaking information regarding time scale on which closely related species have diverged and types of genetic variations are associated with species formation (Ratnaparkhe *et al.*, 1998). RFLPs, microsatellites, minisatellites, transposons, RAPD, ISSR and specific Sequence Tagged Microsatellite (STMS) or STS markers have proved to be extremely valuable in the analysis of gene pool variation of crops during the process of cultivar development and classification of germplasm.

Construction of genetic linkage maps

Development of genetic maps is one of the important applications of molecular markers in any species. Genetic maps indicate the position and relative distances between markers along chromosomes (Paterson, 1996) and it summarizes much of the genetic information available for that species.

Gene tagging

Monogenic and polygenic loci in many plants have been identified using molecular markers and once mapped these markers can be used for tagging of individual traits, which are important for breeding programme. The first report of gene tagging was from tomato where markers linked to agronomic and resistance characters were discussed (Weller *et al.*, 1988; Williamson *et al.*, 1994).

Mapping genes and map based cloning

Molecular markers offer a tool for locating genes governing agronomically important characters through a genetic linkage map. Major genes can be mapped establishing association between molecular markers and simply inherited traits. The advent of molecular marker technology provides new tools for identifying traits, which are controlled by polygenes and thereafter mapping of these QTLs through the construction of a saturated genetic linkage map. Map-based cloning follows physical mapping of gene of interest, which is identified by adding overlapping DNA pieces from the position of molecular markers.

Marker Assisted Selection (MAS)

The term 'markers-assisted selection' (MAS) was used in literature by Beckmann and Soller (1986). MAS involve selection of plants carrying genomic regions that are involved in the expression of traits of interests through molecular markers. For successful adoption of a marker aided selection, a tight linkage between marker and the gene of interest is essential. A population which is highly polymorphic for the marker and the gene of interest showing extreme linkage disequilibrium is required. In the context of MAS, DNA markers can trace favorable alleles, which are both dominant and recessive across generations to identify most suitable individual among the segregating progeny based on allelic composition. Molecular markers are employed for marker assisted foreground selection (Melchinger, 1990) and background selection for marker assisted breeding in a backcross programme (Hospital and Charosset, 1997). The first substantive article on the application of MAS in plant breeding using DNA markers was for soybean cyst nematode (*Heterodera glycines*) resistance (Concibido *et al.*, 1996).

Application of molecular tools in rubber tree improvement was lagging behind because of limited knowledge of the genome and lack of genomic resources. Therefore, an effort was made through the present study to generate different types of molecular markers and their application in genetic linkage map construction. The content of the thesis was partitioned in to two chapters. The first chapter focuses on the development of molecular markers and their application in genetic diversity analysis of the cultivated clones from Southeast Asian rubber growing countries including India. The second chapter aimed at developing genetic linkage map employing all the developed markers for future crop-improvement through molecular breeding.

Objective of the present study

In *Hevea brasiliensis*, breeding process is laborious and time consuming. Long juvenile period of rubber tree makes the generation based improvement and evaluation difficult. Moreover, the commercially cultivated clones of *H. brasiliensis* represent a very narrow genetic base. In order to overcome these hurdles of conventional rubber tree breeding, molecular breeding approaches are applied to the crop. Application of DNA based molecular markers makes breeding schemes more efficient through marker assisted selection (MAS) of traits for early selection of young seedlings. In *Hevea,* the existing diversity can be made use of in future breeding programmes by determining the extent of diversity among the cultivated 'Wickham clones' using molecular markers. Further, these markers are used for the construction of genetic linkage maps, which is essential to assign genes on linkage groups corresponding to chromosomes for cloning and fine mapping of genes, especially for the latex biosynthesis and also to identify the markers flanking major QTLs related to biotic and abiotic stress tolerance and plant growth. Therefore, major objectives of the present work are as follows:

1) Development of DNA based molecular markers for rapid and precise characterization of popular clones for their use in breeding programme
2) Diversity analysis to know the extent of genetic variation existing among the cultivated clones
3) Construction of genetic linkage maps for establishing marker trait relationship, essential for map based cloning and genetic dissection of quantitative traits (QTLs)

Chapter I

Development of molecular markers and phylogenetic analysis

Chapter I

1.1 Introduction

Hevea brasiliensis is a perennial cross pollinated tree crop belongs to the family Euphorbiaceae. *H. brasiliensis* is the only cultivated species and contributes to 95% of the world natural rubber production. The breeding of rubber tree faces many constraints due to its perennial nature, long breeding cycles, extended juvenile period and the low genetic diversity exhibited by the cultivated clones due to the unidirectional selection for yield characters. In India, *Hevea* breeding was initiated in the year 1954 in Rubber Research Institute of India (RRII), Kottayam evolving the first set of hybrids including the most popular cultivated clone RRII 105. But the conventional method of breeding with morphological markers has its own limitations, since they are very less in number and also highly influenced by environment.

In order to increase the efficiency and precision of crop improvement programmes in rubber, the molecular marker techniques are now widely used. Molecular markers play a significant role in plant breeding efforts from the detection of polymorphism, thereby assessing the genetic diversity and to the construction of genetic linkage maps, which are essential for marker assisted selection, QTL analysis and finally map-based cloning of agronomically important traits.

Diversity is a pre-requisite for any crop improvement programme since they could be incorporated in breeding process to evolve superior genotypes with wide adaptability. In *Hevea*, molecular markers like RAPD, AFLP (Lespinasse *et al.*, 2000), SSR (Feng *et al.*, 2009; 2010; Roy et al., 2004; 2012; Saha *et al.*, 2005; 2012), SNP (Pootakham *et al.*, 2011) were developed and used in various genetic analysis. In the present study, an attempt was made to develop RAPD, AFLP, SSR, SNP and STS markers for getting an insight into the extent of diversity of the popular clones of *Hevea brasiliensis* and for further crop improvement programmes.

1.2 Review of literature

Genetic markers are essential tools for diversity analysis, fingerprinting, construction of genetic linkage maps and marker assisted selection. These markers are advantageous over the morphological markers as they can accurately quantify the genetic relatedness among organism belonging to a species. A large volume of data is available regarding marker analysis, diversity assay and genetic linkage map construction. However the present review of literature is mainly restricted to rubber and other tree crops; mainly apple, Eucalyptus, grapes, pear, almond. The different types of markers used in the present study include Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), microsatellites/simple sequence repeats (SSRs), Single Nucleotide Polymorphisms (SNPs), Sequence Tag Site (STS) and Resistance Gene Analogue (RGA) markers.

1.2.1 Random Amplified Polymorphic DNA (RAPD) markers

RAPD markers first described by (Wiliams *et al.*, 1990) offered highest potential for generating large number of markers with ease (Russell *et al.*, 1997). RAPD markers were shown to be very useful for DNA fingerprinting, detection of genetic polymorphism, varietal identification, evaluation of gene flow between species (Arnold *et al.*, 1991), detection of gene introgression in various plant species (Orozco Castillo *et al.*, 1994) and mapping (Grattapaglia and Sederoff, 1994).

In *Hevea brasiliensis* RAPD markers were initially developed for varietal identification and genetic diversity analysis. The applicability of RAPD markers for genetic analysis in *H. brasiliensis* was evaluated in a set of 24 clones in breeding pool of RRII (Varghese *et al.*, 1997). Maximum genetic distance was displayed by the clone RRIC 100 and it was used as parent in hybridization programmes resulted in highly heterotic hybrids (Licy, 1997). The genetic relationships among 37 cultivated clones of *H. brasiliensis* was established by Venkatachalam *et al.* (2002) using this marker system. Mathew *et al.* (2005) conducted phylogenetic relationship studies in three species of *Hevea* (*H. brasiliensis, H. benthamiana* and *H. spruceana*) using RAPD markers along with other marker systems such as chloroplast DNA PCR-RFLP and heterologous chloroplast microsatellites. A dwarf genome specific RAPD marker was identified, cloned and sequenced from rubber by Venkatachalam *et al.* (2004). RAPD marker was

developed which showed partial homology to proline-specific permease gene (Venkatachalam *et al.*, 2006). A differentially expressed thymidine kinase gene related to tapping panel dryness syndrome in the rubber tree was also identified by RAPD markers (Venkatachalam *et al.*, 2010). In rubber, a number of RAPD markers have been used to identify clones (Nurhaimi-Haris *et al.*, 1998) and to identify markers related to diseases (Toruan-Mathius *et al.*, 2002). The genetic diversity of early introduced clones of *H. brasiliensis* in Southern Thailand was studied by Nakkanong *et al.* (2008) using RAPD markers. RAPD technique with eight arbitrarily pre-selected primers was used to evaluate genetic variability of *Hevea* population, which were controlled-pollinated hybrids of maternal clone PB 260 with five Amazonian genotypes as male parents. The structure and genetic diversity of the IRRDB'81 germplasm collection was studied using six selected RAPD primers by Lam *et al.* (2009) and the UPGMA clustering conformed to the geographical regions of the collections. The phylogenetic relationships among 45 rubber clones was analyzed using 12 RAPD primers and based on the genetic distance obtained an attempt was made to select the parent trees for further crop improvement programmes (Oktavia *et al.*, 2011). The genetic similarity among various rubber clones from different Brazilian regions were identified (Bicalho *et al.*, 2008) with 19 random primers. Other than the genetic diversity analysis RAPD markers were used to construct a genetic linkage map in *Hevea* in an F2 population along with other markers such as RFLP and isozyme loci. Shoucai *et al.* (2004) identified mildew resistance genes by RAPD techniques.

RAPD analysis has been widely employed for assessing genetic diversity and relationships in tree crops like *Prunus* (Lu *et al.*, 1996), almond (Joobeur *et al.*, 2000), apricot (Hurtado *et al.*, 2002), *Eucalyptus* (Rebecca et al., 2002), date palm (Kunert *et al.*, 2003), pomegranate (Talebi Bedaf *et al.*, 2003; Sarkhosh *et al.*, 2006; Dorgac *et al.*, 2008) and mulberry (Shabir *et al.*, 2010). The differences in morphology and geographical origin of the genotypes was reflected in the RAPD patterns (Orozco-Castillo *et al.*, 1994). RAPD markers were also identified to be associated with resistance to certain diseases in coffee (Agwanda *et al.*, 1997) and tea (Sriyadi *et al.*, 2002). The technique was effectively used to develop strain-specific sequence-characterized amplified region (SCAR) marker (Hermosa *et al.*, 2001). Variants of the RAPD technique include Arbitrarily Primed Polymerase Chain Reaction (AP-PCR), which uses longer arbitrary primers than RAPDs, and DNA Amplification Fingerprinting (DAF) that uses shorter, 5–8 bp primers to generate a larger number of fragments. Multiple Arbitrary

Amplicon Profiling (MAAP) is the collective term for techniques using single arbitrary primers (Kumar *et al.*, 2009).

1.2.2 Amplified Fragment Length Polymorphism (AFLP) markers

This multi-locus marker technology was developed by Vos *et al.* (1995). This was based on the detection of genomic restriction fragments by PCR amplification. AFLP is considered to be information rich due to its ability to analyze a large number of polymorphic loci simultaneously with a single primer combination on a single gel as compared to RAPDs and RFLPs (Powell *et al.*, 1996). They are more reliable and allow the establishment of a saturated genetic linkage map. But AFLPs are non-locus specific, dominant markers that can exhibit only two states in a species (presence versus absence of a band) (Seguin *et al.*, 2001).

In *H. brasiliensis*, there are only a few reports regarding the assay of AFLP markers. The genetic diversity was assessed by AFLP markers in *Hevea* by An-Ding *et al.* (2001). AFLP markers were utilized by Lespinasse *et al.* (2000) for the construction of a genetic linkage map in *Hevea* in the progeny derived from an interspecific cross. Genetic diversity analysis was performed in Apricot cultivars with five different primer combinations of AFLP generating 203 polymorphic fragments (Krichen *et al.*, 2008). AFLP markers were chosen for their distribution through out the whole genome in apricot (Lambert *et al.*, 2004), phylogeographic analysis (Despres *et al.*, 2003). AFLP was demonstrated to be very powerful for assessing species diversity in peach (Aranzana *et al.*, 2003), cherry (Tavaud *et al.*, 2004) and pummelo (Nartvaranant and Nartvaranant, 2011). AFLP markers were also used to measure the genetic diversity and relationships in fruit species; such as cherry (Gerlach and Stosser, 1997), mango, peach and pear (Monte-Corvo *et al.*, 2000) and Chinese pomegranate (Yuan *et al.*, 2007).

1.2.3 Microsatellite/SSR markers

Microsatellite DNA or simple sequence repeats, a relatively new class of DNA markers, are highly informative genetic markers. The existence of microsatellite sequences in plants and algae was first reported in 1986 (Tautz *et al.* 1986). The term "microsatellites" was coined by Litt and Luty (1989). These are abundant, dispersed throughout the genome and show higher levels of polymorphism than other genetic markers (Schlotterer and Tautz, 1992). These features coupled with their ease of detection, have made them

useful molecular markers. Their potential for automation and their inheritance in a co-dominant manner are additional advantages (Morgante and Olivieri, 1993; Thomas and Scott, 1993). Simple sequence length polymorphisms are based on the difference in the number of the DNA repeat units (Cho *et al.*, 2000). One of the first reports, describing the presence of microsatellites in plant genomes was in the forest trees (Conditt and Hubbell, 1991). The SSR markers are of interest to geneticists and breeders and have been successfully used to infer about genetics, pedigree, phylogeny, and/or identity of various traits and/or germplasm accessions.

In *H. brasiliensis* recently much importance is given for the development of microsatellite markers. The first report of DNA fingerprints in *H. brasiliensis* using heterologous minisatellite probes from humans came from Besse *et al.* (1993). Low *et al.* (1996) for the first time, detected microsatellites in the *Hevea* genome through the database search of some *Hevea* gene sequences. A systematic effort towards developing microsatellite markers was made by Roy *et al.* (2004) through screening of *Hevea* genomic library. Microsatellites markers were used to identify 27 *Hevea brasiliensis* clones in our laboratory (Saha *et al.*, 2005). The polymorphisms observed could be used for developing markers for screening various traits in *Hevea* breeding programmes. Along with RAPD markers, four microsatellite pairs (*hmac*4, *hmac*5, *hmct*1, *hmct*5) reported by Saha *et al.* (2005) were used for genetic diversity analysis in 53 early introduced clones of *Hevea* collected from different areas of Southern Thailand by Nakkanong *et al.* (2008). Dinucleotide $(CT)_n$ repeats detected in Mn - SOD had been used as SSR markers for genetic relationship studies by Lespinasse *et al.* (2000) and Lekawipat *et al.* (2003). Lespinasse *et al.* (2000) described the construction of a genetic linkage map of rubber tree using SSR markers along with different molecular markers. The polymorphic microsatellite loci isolated and characterized from an enriched genomic library of *H. brasiliensis* was highly useful in understanding genetic diversity and gene flow among *Hevea* species (Souza *et al.*, 2009). Two hundred and ninety six new polymorphic microsatellite markers were introduced by Le Guen *et al.* (2010) for genomic studies in *H. brasiliensis* through screening of an enriched genomic library. Dinucleotide repeats $(AG)_n$ was detected at the 3'-UTR of mRNA of HMG-CoA reductase encoded by the gene *HMGR*, which is a key enzyme involved in latex biosynthesis in rubber (Saha *et al.*, 2005). SSR polymorphism at this locus was successfully used for studying the allelic diversity in wild accessions of rubber (Saha *et al.*, 2007). Cross-species amplification of the markers, developed for *H. brasiliensis* was

also found successful in the wild *Hevea* species *H. guianensis, H. rigidifolia, H. nitida, H. pauciflora, H. benthamiana and H. camargoana* (Saha *et al.*, 2005; Souza *et al.*, 2009). They reported that a high degree of sequence homology existed in the microsatellite flanking regions of these species. Garcia *et al.* (2011) identified microsatellite markers from geneBank sequences and utilized 15 markers for the genetic diversity analysis of *Hevea* clones. Microsatellite markers were used in diversity analysis and their cross species amplification was tested in six *H. brasiliensis* species by Mantello *et al.*, (2012)

Molecular approach is one of the methods for the development of SSR markers. Here genomic libraries are constructed (with or without enrichment for SSRs), screened, candidate clones are sequenced and SSR motifs are identified either manually or using computer programs. Microsatellite enrichment has been developed to increase the proportion of clones in a given library containing the microsatellite motif of interest. Several strategies for microsatellite enrichment have been reported (Kijas *et al.*, 1994; Edwards *et al.*, 1996; Fischer and Bachman, 1998; Koblizkova *et al.*, 1998; Hamilton *et al.*, 1999; Paetkau, 1999; Phan *et al.*, 2000; Zane, 2002). SSRs, once developed, are extremely valuable, though their development is time consuming, laborious and expensive. The construction of a microsatellite-enriched library in *Hevea brasiliensis* was reported by Atan *et al.* (1996). Microsatellite-enriched library were produced by selective hybridization of genomic DNA of *Populus nigra* restricted with *Alu* I by Van de Schoot *et al.* (2000). Around 10 microsatellite markers have been developed in *Frageria* by James *et al.* (2003), which were useful in its synteny analysis.

The computational or bioinformatics approaches take advantage of the available sequences such as those in the public databases and by scanning through them for identifying the ones that contain SSRs. They supplement the molecular approaches by identifying SSR repeats in candidate sequences derived from the libraries. Data mining of microsatellites from ESTs make use of this approach and has proven effective for generating markers for fingerprinting, genetic mapping and comparative mapping among species (Varshney *et al.*, 2005). Developing these markers is less costly and time effective, and may provide abundant information. Sequences from many genomes are continuously made freely available in the public databases and mining of these sources using computational approaches permits rapid and economical marker development.

Expressed sequence tags (ESTs) are ideal candidates for mining SSRs not only because of their availability in large numbers but also due to the fact that they represent expressed genes. EST - SSR markers were developed first in rubber trees by Feng *et al.* (2009). Multivariate techniques and microsatellite markers were used for genetic divergence estimation in rubber by Gouvea *et al.* (2010). Genetic diversity analysis of wild and cultivated clones of *Hevea* was estimated by Perseguini *et al.* (2012) using EST derived SSR markers and the cross species amplification of these markers were also estimated successfully. It was found that since rubber tree is highly heterozygous due to its out crossing nature, cross pollinated, transcriptome sequencing is an attractive alternative to whole-genome sequencing which focuses on the coding regions of the genome which can be used for SSR mining. Earlier transcriptome studies of *H. brasiliensis* focused mainly on latex in order to gain insight into the rubber biosynthesis pathways (Han *et al.*, 2000; Ko *et al.*,2003; and Chow *et al.*, 2007). The transcriptome from rubber tree bark was sequenced with Illumina paired-end sequencing and 106 EST-SSR markers were developed by Li *et al.* (2012). In the studies of Triwitayakorn *et al.* (2011), the transcriptome from the vegetative shoot apex was sequenced and synthesised 323 EST derived SSR primer pairs which were amplifiable in *H. brasiliensis* clones. From these a selected number of 47 SSR markers were utilized for the genetic similarity analysis of 20 *H. brasiliensis* clones.

The occurrence and characteristics of nuclear microsatellites has been reported in several forest-tree species such as *Pinus radiata* (Smith and Devey, 1994), *Pinus sylvestris* (Kostia *et al.*, 1995), *Eucalyptus nitens* (Byrne *et al.*, 1995), *Pithecellobium elegans* (Chase *et al.*, 1996), *Pinus strobus* (Echt *et al.*, 1996), *Picea stechensis* (Van de Ven and McNicol, 1996), *Swietenia humilis* (White and Powell, 1997), *Quercus macrocarpa* (Dow and Ashley, 1998) and *Quercus suber* (Aljorna *et al.*, 2007).

Achievements in the use of microsatellites have been made in areas like selection and diagnostics in segregating populations, genome selection during gene introgression (in back cross programmes), genome mapping, gene tagging, cultivar identification, germplasm characterization and in estimation of genetic relatedness (Gupta *et al.*, 1996). DNA fingerprinting assays were conducted in peach by Testolin *et al.* (2000) using SSR markers isolated from enriched genomic library for dinucleotide repeats. Eight functional Mendelian microsatellite markers were developed from a non-enriched genomic library of *Eucalyptus sieberi* (Glaubitz *et al.*, 2001). Development of

microsatellites in apple had made it possible to efficiently select and breed apple varieties of high quality with multiple disease resistances (Liebhard *et al.*, 2002). Pedigree reconstruction of ancient alpine grape cultivars was performed with microsatellite analysis by Vouillamoz *et al.* (2003). Microsatellite markers for avocado (Lavi *et al.*, 1994; Sharon *et al.*, 1997; Ashworth *et al.*, 2004) is limited and some of it do not consistently amplify in all varieties (Ashworth and Clegg 2003; Ashworth *et al.*, 2004). Only 14 of 35 microsatellites were suitable to fingerprint diverse collections of avocado (Schnell *et al.*, 2003).

Microsatellites markers are found throughout both the transcribed and non-transcribed regions of a genome (Varshney *et al.*, 2005). Their role in gene regulation and genome evolution has also been discussed widely (Aishwarya and Sharma, 2007). They are considered as valuable tools for many purposes such as phylogenetic, fingerprinting and molecular breeding studies (Sharma *et al.*, 2008). In avocado trees, maternal effects may interfere with identifying truly root rot disease tolerant plants (caused by *Phytophthora cinnamomi*). So isolation of elite parent trees is to be maintained. Here microsatellite markers come handy in parentage analysis where a high proportion of the putative parents are closely related (Violi *et al.*, 2009). An SSR-analysis of rootstock of grapevine cultivar clones revealed the allelic characteristics of microsatellite loci which could be used to identify and certify grapevine clones. Microsatellite analysis also found useful in better germplasm management and for devising strategies for identifying core selection (Upadhyay *et al.*, 2010).

1.2.4 Single nucleotide polymorphism (SNP)

SNPs represent the most common variations across a genome (Gupta *et al.*, 2001; Kwok *et al.*, 2001). SNPs are stable and the relative fidelity of their inheritance is higher than that of the other marker systems like SSRs and AFLPs (Semagn *et al.*, 2006). SNPs occur at a frequency of about one SNP in 1000 nucleotides in human genome (Wang *et al.*, 1998). SNPs are consequences of either transition or transversion events. SNP fall into several classes depending on (1) their precise location in the genome and (2) impact of their location within coding or regulatory regions onto the encoded proteins or phenotype. Given that majority of SNPs are located in noncoding DNA, they are called noncoding SNPs. SNP that reside on exons and corresponding cDNA are called coding SNPs or exonic SNPs. Exonic SNPs that do not change the composition of the encoded

domains or proteins are called synonymous SNPs (Weising *et al.*, 2005). SNPs, when synonymous, can change the structure and stability of the messenger RNA, whereas non-synonymous can change the structure and function of the protein and consequently affect the amount of protein produced, as it may be a deleterious mutation (Guimarães and Costa, 2002). SNPs contribute directly to a phenotype or can be associated with a phenotype as a result of linkage disequilibrium (Risch and Merikangas, 1996). Direct analysis of genetic variation at the DNA level has made SNPs attractive as genetic markers (Bhattramakki *et al.*, 2002; Rafalski, 2002). In plants, SNP seems to be more abundant than those in human genome as evidenced by the existence of one SNP per 20 bp in wheat (Picoult *et al.*, 1999) and one SNP per 70 bp in maize genome (Bhattramakki *et al.*, 2000).

Rapid advances in the high-throughput next-generation sequencing (NGS) technologies have facilitated extensive SNP discovery projects in several plant species (Barbazuk *et al.*, 2007). Once discovered, SNPs can be converted into genetic markers that can be assayed in a high-throughput manner (Gut, 2001; Kwok, 2001). SNP markers are very useful both for marker-assisted selection and for gene isolation when found in proximity of the coding sequences. Since SNPs are abundant, they can be used as genetic markers in many applications such as cultivar identification, construction of genetic maps, assessment of genetic diversity or marker-assisted breeding (Flint-Garcia *et al.*, 2005; Chágne *et al.*, 2008; Wu *et al.*, 2008).

1.2.4.1 SNP identification methods

The broad approaches for SNP discovery are resequencing approach, *in-silico* SNP mining approach. There are other techniques also like denaturing/temperature gradient gel electrophoresis (D/TGGE) (Myers *et al.,* 1988), Single stranded conformational polymorphism (SSCP) (Orita *et al.,* 1989) which has been employed to detect polymorphisms in forest trees (Plomion *et al.*, 1999), horticultural trees (Etienne *et al.,* 2002).

EST sequence data is a powerful source of biologically useful SNPs due to the relatively high redundancy of gene sequence, and the fact that each SNP would be associated with an expressed gene (Picoult-Newberg *et al.,* 1999). Direct amplicon sequencing first reported by Wang *et al. (1998)* was one of the simplest forms of SNP discovery. In cloned amplicon sequencing, heterozygous SNPs may be unambiguously

determined in primary analysis and there by determining the haplotype structure. In *Hevea brasiliensis* there are only limited reports regarding the identification and utilization of SNP markers. The only reports of SNP identification and marker validation of SNP markers was from Pootakham *et al.* (2011) They identified 5883 biallelic SNPs by transcriptome sequencing and 50 of them were validated in *Hevea* clones. In other tree crop, scots pine populations were examined for SNPs in gene encoding several structural proteins or enzymes like phenylalanine lyase (Dvornik *et al.*, 2000). Single nucleotide polymorphisms including insertion/deletions (indels) can provide a rich source of useful molecular markers in genetic analysis for many crop species (Ronaghi *et al.*, 1996).

1.2.4.2 SNP genotyping

High-throughput, high density SNP genotyping has become an essential tool for QTL mapping, association genetics, gene discovery etc. and applied in several crops (Zhu *et al.*, 2001; Rafalski, 2002). In *Eucalyptus*, the Golden Gate Genotyping Technology (GGGT) developed by Illumina (Ganal *et al.*, 2009) was used for the detection of SNPs (Grattapaglia *et al.*, 2011).

Large scale genotyping in minimally equipped laboratories require a low cost technology. The restriction based techniques was the earliest method used for the detection of single nucleotide polymorphism (Botstein *et al.,* 1980). The cleaved amplification polymorphic site or polymerase chain reaction-restriction fragment length polymorphism (CAPS or PCR-RFLP) and the derived CAPS (dCAPS) (Michaels and Amisino, 1998; Neff *et al.*, 1998) are widely applied (Iwaki *et al.* 2002; Yanagisawa *et al.*2003; Yamanaka *et al.* 2004). Another alternative is the allele-specific PCR amplification (AS-PCR) also called PCR allele-specific amplification (PASA) (Sommer *et al.*, 1992) or amplification refractory mutation system (ARMS) (Lo, 1998). The allele specific PCR amplification follows a simple procedure with a common reverse primer and two allele specific forward primers in normal conditions and allow the discrimination of alleles by gel electrophoresis (Dutton and Sommer, 1991). Detection of amplicons generated through PASA can be effectively done through other techniques also such as fluorescence detection, real-time fluorescence detection, and sequencer (Germer and Higuchi, 1999; Ishiguro *et al.*, 2005; Hansson and Kawabe, 2005; Wu *et al.*, 2005; Hinten *et al.*, 2007). Kwok *et al.* (1990) found that the insertion of a destabilizing mismatch within five bases at the 3' end will improve the specificity of the PCR amplifications. A

variant of the PASA method is the bidirectional-PASA (Bi-PASA) was reported by Liu *et al.* (1997). A modified technique was developed by Soleimani *et al.* (2003) where, one of the alleles is amplified by a PASA reaction in one direction, and the second allele was amplified in the opposite direction for validation of SNP markers in barley. Although no single genotyping method is ideally suited for all applications, a number of good genotyping methods are available to meet the needs of many study designs. The challenges for SNP genotyping in the near future include increasing the speed of assay development, reducing the cost of the assays, and performing multiple assays in parallel (Kwok, 2001).

STS markers could be generated using PCR amplifications and detected through agarose gel electrophoresis and ethidium bromide staining. In the tree species sugi, Ujino-Ihara *et al.* (2010) used STS markers for genotyping. The variation among the different cultivars of Norway spruce was detected through STS marker analysis (Perry *et al.*, 1999). STS markers generated based on varying length of indels was used to genotype different black spruce species (Perry and Bousquet, 1998).

Table 1.1. List of popular clones of *Hevea brasiliensis* originating from different rubber growing countries

Sl. No.	Clone	Parentage	Country of origin
1	RRII 105	Tjir 1 x GL 1	India
2	GL 1	primary clone*	Malaysia
3	PB5/51	PB 56 x PB 24	Indonesia
4	PB 86	Primary clone*	Indonesia
5	PB 217	PB 5/51 x PB 6/9	Indonesia
6	PB 235	PB 5/51 x PB S/78	Indonesia
7	PB-314	RRIM 600 x PB 235	Indonesia
8	PB 28/59	Primary clone*	Indonesia
9	PR 107	Primary clone*	Indonesia
10	RRIM 600	Tjir 1 x PB 86	Malaysia
11	RRIM 703	RRIM 600 x RRIM 500	Malaysia
12	RRIM 701	44/553 x RRIM 501	Malaysia
13	RRII 118	Mil3/2 x Hil 28	India
14	RRII 33	Primary clone*	India
15	RRIC 52	Primary clone*	Sri Lanka
16	RRIC 104	RRIC 52 x Tjir 1	Sri Lanka
17	RRIC 100	RRIC 52 x PB 83	Sri Lanka
18	Tjir 16	Primary clone*	Indonesia
19	BD 10	Primary clone*	Indonesia
20	GT 1	Primary clone*	Indonesia
21	KRS 25	Primary clone*	Thailand
22	RRII 5	Primary clone*	India
23	PR 255	Tjir 1 x PR 107	Indonesia
24	PR 261	Tjir 1 x PR 107	Indonesia
25	SCATC 88/13	RRIM 600 x Pil B 84	China
26	SCATC-93/114	TR 31-45 x HK 3-11	China
27	Haiken 1	Primary clone*	China
28	Tjir 1	Primary clone*	Indonesia

* Elite selected clones with unknown parental origin

1.3 Materials and methods

1.3.1 Plant material

The plant materials used in the present experiments were 28 *Hevea brasiliensis* clones originating from Southeast Asian rubber growing countries including India. Out of these 13 were primary Wickham clones (Wycherly, 1968; Dean, 1987) and rest of the clones were derived from hybridization of primary clones in Southeast Asian rubber growing countries (Table 1.1).

1.3.2 Genomic DNA extraction and purification

1.3.2.1 Collection of leaf samples

Tender healthy green leaves of the *Hevea* clones were collected from germplasm division of Rubber Research Institute of India (RRII) campus and immediately kept in ice and brought to the laboratory for DNA isolation. Leaves were washed with sterile distilled water supplemented with a few drops of Tween-20 for 2- 3 min. Then washed in sterile water for few min and blotted dry on Whatman filter paper. The margins of the leaves were removed and one gram samples were weighed, excluding the midrib and disease spots on the lamina.

Extraction and purification of the total genomic DNA were carried out following a modified CTAB method of Doyle and Doyle (1990) with minor modifications (Saha *et al.*, 2002). Leaf samples were ground to fine powder in the presence of liquid nitrogen using a pre-chilled mortar and pestle. The powder was homogenized in 10 ml of 2x CTAB buffer preheated to 60 ^{0}C [2% (w/v) CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris–HCl (pH 8)] β-mercaptoethanol [0.3% (v/v)] was added just before use. The entire content was transferred to polypropylene tubes and was incubated in a water bath at 65°C for 45 min with slow intermittent mixing by swirling at an interval of 10 min. After cooling, equal volume of 24:1 chloroform:isoamyl alcohol was added. Mixed well by gently inverting the tubes to and fro for 10 min. Centrifuged at 10,000 rpm for 10 min at room temperature and supernatant was transferred to fresh tubes without disturbing the interphase. It was treated with 25 µl RNase (10 mg/ml) and incubated at 37°C for one hour. An equal volume 24:1 chloroform: isoamyl alcohol was added and mixed well by slow inversion for 10 min. Spinned at 10,000 rpm for 10 min, at room temperature. The aqueous phase was transferred to a fresh tube and extracted with 0.6 volume of ice cold isopropyl alcohol and

mixed slowly. The precipitated DNA was spooled and transferred to a new 1.5 ml eppendorf tube containing 1 ml 75% ethanol with a pipette tip. centifuged for 10 min at 4 °C at 10,000 rpm. Supernatant was removed and added 1 ml 75% ethanol again. Spinned as above. The supernatant was discarded and the pellet was kept for air drying. The pellet was dissolved in 500 µl sterile water.

1.3.2.2 Quantifying DNA concentration and purity

The DNA concentration and the purity were determined by NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies) and stored at 4 °C.

1.3.2.3 Gel electrophoresis and photography

The DNA was electrophoresed on a 0.7 % (w/v) agarose gel at 60V for about 1-2 hours. Lambda DNA/*Eco*RI+*Hin*d III double digest was used as molecular weight marker. The gels were visualized on UV transilluminator and photographed with Stratagene Eagle Eye II digital gel documentation system.

1.3.3 Development of molecular markers

1.3.3.1 Random Amplified Polymorphic DNA (RAPD) markers

1.3.3.1.1 PCR procedure

PCR amplifications were carried out according to Williams *et al.* (1990). Five hundred and twenty arbitrary decamer primers belonging to 39 'Operon' series (OPA–OPJ, OPO, OPR, OPX and OPAA-OPAZ series (Operon Technologies Inc.) were evaluated with four *H. brasiliensis* clones to identify the polymorphic primers with strong amplification profiles.

PCR amplifications were performed in a total volume of 25 µl by mixing 50 ng of template DNA with 10 picomoles of single primer, 0.2 mM of each dNTP, 0.7 unit of Taq DNA polymerase (Promega, USA), 2.5 µl of DNA Polymerase buffer 10x [100 mM Tris-HCl (pH 9), 500 mM KCl, 20 mM $MgCl_2$]. Amplifications were performed in a thermal cycler (GeneAmp PCR System 9600, Perkin Elmer Cetus, USA) with an initial denaturation step at 94°C for 3 min, followed by 40 cycles of 30 sec at 94°C, 1 min at 36°C and 2 min at 72°C with a final extension at 72°C for 7 min. Amplified products were analyzed along with Lamda DNA marker, digested with *Eco* RI and *Hin*d III as molecular size reference by electrophoresis on a 1.5% agarose gel in 1X TAE buffer at 70 V for 3 hours and then stained with ethidium bromide (0.5 µg/ml) and visualized on a UV

transilluminator. Photograph was taken using gel documentation system. The scoring of the bands was done manually as presence and absence of bands.

1.3.3.2 Amplified Fragment Length Polymorphism (AFLP) markers

AFLP analysis was performed according to Vos *et al.* (1995) with minor modifications. Both conventional technique and also the AFLP® Starter Primer Kit (Invitrogen) were used.

1.3.3.2.1 Conventional method of AFLP analysis

1.3.3.2.1.1 Restriction endonuclease digestion of the DNA

In the conventional method, the genomic DNA (500ng) was digested with 10 U of the restriction enzyme *Eco*RI [(G/AATTC), New England Biolabs] and then with 10 U of *Mse* 1 [(T/TAA), Roche Applied Sciences]. The restriction digestion was performed in 1x OnePhorAll buffer (Amersham biosciences) in a final volume of 50 µl with 0.5 µl BSA (10 mg/ml) and incubated at 37 °C for 3 hours. The reaction was heat inactivated at 70°C for 15 min. Digested samples were precipitated through ethanol precipitation and dissolved in 10 µl water.

1.3.3.2.1.2 Adapter preparation and ligation

The *Eco*RI and *Mse*I oligos were synthesized at Sigma Aldrich, Bangalore for adapter preparation. The sequence details of adapters are given in Table 1.2.

The two oligonucleotides were mixed together at 1:2 ratios. For each reaction, 5 pmoles of oligo-1 and 10 pmoles of oligo-2 were mixed with STE buffer [0.1 M Nacl, 10 mM Tris HCl (pH 8, 1 mM EDTA]. Each adapter was prepared in a total volume of 50 µl and the reaction contents were 12.5 µl of oligo-1, 25 µl of oligo-2, 5 µl of 10x STE buffer and 7.5 µl of water. The reagents were mixed and the tubes were placed at 70°C for 10 min and was allowed to cool to room temperature slowly.

Both *Eco*RI (5 pmol) and *Mse* I (50 pmol) were ligated onto the restricted DNA fragments at 16°C using 2U of T4 DNA ligase (Amersham Biosciences) in a final volume of 20 µl. Ligations were diluted to 10 fold in TE [10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA] and used as templates for the pre-amplification PCR reactions.

1.3.3.2.1.3 Pre-amplification of ligated DNA

The pre-amplification reactions were performed in a total volume of 25 µl and the components being 0.2 µM of each primer that was complementary to the *Eco*RI and *Mse*I

Table 1.2. Sequence details of adapters and primers for pre-amplification reactions used in the AFLP analysis

Adapters	Sequence (5'-3')
Eco R I adapter Oligo-1	CTCGTAGACTGCGTACC
Oligo-2	AATTGGTACGCAGTC
Mse I adapter Oligo-1	GACGATGAGTCCTGAG
Oligo-2	TACTCAGGACTACT
Pre-selective primers	
*Eco*R I –A (E-A)	GACTGCGTACCAATTCA
Mse I- A (M-A)	GATGAGTCCTGAGTAAA
Selective primers	
E-AGA	GACTGCGTACCAATTCAGA
E-AGC	GACTGCGTACCAATTCAGC
E-AGG	GACTGCGTACCAATTCAGG
E-AGT	GACTGCGTACCAATTCAGT
E-ACA	GACTGCGTACCAATTCACA
M-AAC	GATGAGTCCTGAGTAAAAC
M-ACA	GATGAGTCCTGAGTAAACA
M-AGA	GATGAGTCCTGAGTAAAGA
M-ACT	GATGAGTCCTGAGTAAACT
MA-AGT	GATGAGTCCTGAGTAAAGT

'E' and 'M' denotes the adapter sequences of *Eco*R I and *Mse* I respectively.

adapter sequences synthesized at Sigma Aldrich, Bangalore (Table 2) and 1U of Taq DNA polymerase (Roche Applied Sciences). The PCR reactions were carried out for 20 cycles at the temperature cycle profile of an initial denaturation at 94 °C for 30 sec, annealing at 56°C for 30 sec and an extension at 72 °C for one min using the Perkin Elmer GeneAmp PCR system 9600 (Perkin Elmer). The pre-amplified DNA was diluted 10 fold with sterile water and used as template for selective amplification.

1.3.3.2.1.4 Selective PCR amplification and detection of bands

The PCR reactions were carried out using *Eco*RI and *Mse*I primers having three additional nucleotides at their 3' ends (*Eco*RI-ANN and *Mse*I-ANN where N represents A, C, G or T). The reactions were performed in a total volume of 25 µl containing 5 µl of 10 fold diluted pre-amplified template DNA, 2 mM of $MgCl_2$, 200 mM of each dNTPs, 1.2 µM of *Mse* I, 0.2 µM of *Eco*R I primer and 1U of Taq DNA polymerase (Promega) in the same thermal cycler. The PCR profile includes: one cycle of initial denaturation at 94°C for 30 sec, 65°C for 30 sec and 72°C for one min. The annealing temperature was lowered down by 0.7°C at each cycle for 12 cycles with the touch down PCR programme. This was followed by a normal PCR of 23 cycles of 94°C for 30 sec, 56°C for one min and 72°C for 1 min and an extension at 72°C for 10 min. The amplified products were added with equal quantity of the loading buffer [0.1% each of bromophenol blue and xylene cyanol FF, 10 mM EDTA (pH 7.5) and 98% deionized formamide] was added immediately to stop the reaction and stored at -20°C. Initially the amplified products were run on a mini-PAGE (Bio-Rad) and stained following silver staining technique. Twenty five primer combinations (5 each of *Eco* RI and *Mse*I primers with three selective nucleotides) (Table 1.2) were screened in two clones RRII 105 and RRII 118 to identify the informative primer combinations (Fig. 1. 4).

1.3.3.2.1.5 Silver staining

From 25 primer combinations, nine primer combinations (E-AGA/M-AAC, E-AGC/M-AAC, E-AGT/M-ACA, E-AGT/M-AGA, E-AGT/M-AAC, E-AGG/M-ACA, E-AGA/M-ACT, E-AGC/M-AGA and E-AGT/M-ACA) based on the amplification profiles generated on the mini-PAGE were selected for genetic diversity analysis of the 28 popular *Hevea* clones and they were later run on 6% denaturing PAGE (sequencing gel, BioRad) and electrophoresed using 1x TBE buffer (Sambrook *et al.*, 1989) and silver stained to visualize the AFLP profiles using the protocol of Creste *et al.* (2001) with minor

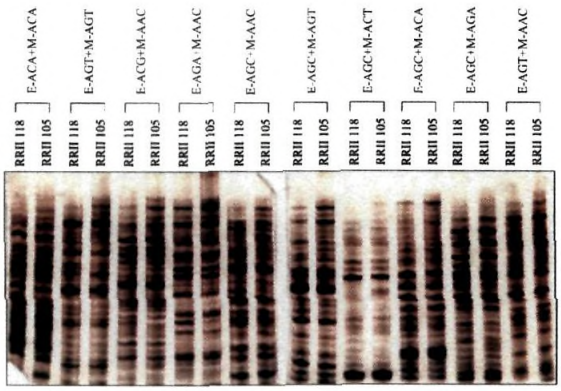


Fig. 1.4. Photograph showing the silver stained AFLP profiles of RRII 105 and RRII 118 on denaturing mini PAGE for screening of primer combinations

modifications (Roy *et al.*, 2012). The glass plate containing the gel was removed from the electrophoretic apparatus and was passed on to one liter each of the following solutions: Fixing solution (acetic acid-10%) for 30 min, Oxidizing solution (nitric acid - 1.5%) for 3 min, followed by staining solution (silver nitrate – 2 g/l) for 30 min and then to freshly prepared cold developing solution (sodium carbonate – 30 g/l, formaldehyde - 540 µl/l) for 30 min. Sterile water wash was performed between each step and finally transferred into the fixing solution for a few min, then rinsed with distilled water, dried and was scanned using GS 800 Calibrated Densitometer (BioRad).

1.3.3.2.1.6 Isotopic detection of bands

For the isotopic detection of bands, *Eco*R I primers (with 3 additional bases at their 3' end) were end labeled with γ^{33}P-ATP prior to selective amplification.

Primer labeling

The *Eco*RI primer labeling with γ^{33}P-ATP was performed in a volume of 1 µl for one reaction and the components being: 0.2 µl *Eco*R I primer, 0.1 µl 10x kinase buffer (Roche Applied Sciences), 0.10 µl γ^{33}P-ATP and 0.5 µl T4 Polynucleotide kinase (Roche Applied Sciences). The labeling was done at 37°C for 60 min and 70°C for 10 min. The contents were collected at the bottom and stored at -20°C.

Selective amplifications were carried out with the labeled *EcoR* I primer and the unlabeled *Mse* I primers complementary to the adapter sequences extended with three selective nucleotides at their 3' end in a total volume of 10 µl containing 2 µl of 10 fold diluted pre-amplified DNA. Temperature cycling remained the same as described earlier and the amplicons were run on 6% denaturing PAGE (Bio-Rad) using 0.6x TBE buffer at a constant power of 55 W. The gels were then dried and autoradiographed on X-ray film using standard procedures.

The conventional method of AFLP analysis is laborious and time consuming. Hence for the generation of large number of markers, different primer combinations were screened in parents of the progeny population using the AFLP analysis system AFLP® Starter Primer Kit (Invitrogen) as per the instructions provided in the manual. The detection of amplified products was performed using the silver staining technique.

1.3.3.3 Microsatellite markers

Microsatellite markers were either generated through the screening of *Hevea* genomic library constructed in the laboratory or through the mining of the ESTs as well as the genomic sequences available in the NCBI GenBank, USA.

1.3.3.3.1 Dinucleotide microsatellite markers

Dinucleotide microsatellite markers were developed from 104 synthesized primers based on the flanking sequences of positive clones from a small insert genomic library of *Hevea brasiliensis* clone GT1, developed in our laboratory (Roy *et al.*, 2004; 2012). Initial screening was performed with parents of the progeny population and 17 markers were selected for the diversity analysis among 28 selected popular clones (Table 1.3). The PCR reactions and the temperature cycling conditions were as follows: The PCR reaction was carried out in a 20 µl final volume containing 40 ng of genomic DNA, 0.4 µM of both forward and reverse primers, 200µM dNTPs, 0.7 units of Taq DNA polymerase (GE Healthcare, UK) along with the buffer supplied. The temperature cycle profile involved an initial denaturation step of 5 min at 94°C followed by a touch down PCR program. Temperature profiles of the touch down PCR for 7 cycles were as follows: 94°C for 30 sec, 63°C for 1 min, $\Delta\downarrow$ 1°C for 7 cycles, 72°C for 1 min. This was followed by a normal cycling of 94°C for 30 sec, 56°C for 1 min, 72°C for 1 min for 23 cycles and a final extension at 72°C for 7 min.

Once the PCR was completed, the reactions were stopped immediately by the addition of 10 µl formamide loading buffer [(0.1% each of bromophenol blue and xylene cyanol FF, 10 mM EDTA (pH 7.5) and 98% deionized formamide)] and stored at -20°C. Amplification products were run on a 6% denaturing polyacrylamide gel containing 7 M urea using 0.6x TBE buffer at a constant power of 55 W and the amplifications were detected through silver staining and also through autoradiography using standard procedures. Gel scoring for the identification of alleles was performed directly from the gel or from the scanned photographs.

1.3.3.3.2 EST-derived microsatellites/SSRs

Sixty-seven EST derived SSR markers reported by Feng *et al.* (2009) were screened in a small progeny population derived from a biparental cross between RRII 105 and RRII 118 to identify the polymorphic and robust markers. Out of these, 14 markers were utilized for

Table 1.3. Details of selected dinucleotide SSR markers used in diversity analysis

Sl. No.	Clone	GenBank accession No.	SSR motif	Expected allele size	Forward primer (5'-3')	T_m	Reverse primer (5'-3')	T_m
1	Hmct-1	AY135651	$(CT)_{12}AG(GT)_8$	195 bp	AACCAGAAGGGTGTCATGCT	63.5	GGAATCCCATGACAATCCAC	64.0
2	Hmct-16	AY439296	$(AG)_{16}$	178 bp	CATGCAAATAACGAACCAGA	53.2	TTCCGTTCACTATATCGCTCA	55.9
3	Hmct-17	AY439298	$(AG)_{10}$	229 bp	GGAAAATGAAGCTTTAACGG	53.2	AATGAGGACGTGGATGGAAT	55.3
4	Hmct-19	AY439301	$(CA)_5$ $(TG)_2$ $AAA(AG)_{13}$	202 bp	CAGATGGGCATTGGTCTTTT	55.3	TCAGAAAATGGATATGGGGC	55.3
5	Hmct-20	AY439302	$(TTTCC)_1TCCG$ $(TTTCC)_1$	200 bp	AGTGGAAACTTGGAAGGGCT	57.3	AGACAGGCATATGGGCTTTG	57.3
6	Hmct-21	AY439303	$(CT)_{14}$ $(AT)_7$	239 bp	TTTGGGCATAAGGAATTTCG	53.2	TGCATGCAACAGAGAATGGA	55.2
7	Hmct-27	AY439314	$(AG)_3$ (AC) $(AG)_8$	249 bp	TGAGCAACGGAGGAGAGAAC	59.4	AAACACCCAAACCCAATTCA	53.2
8	Hmct-45	AY962213	$(CTT)_9$	215 bp	CGCCTACCTGCAGCTCTCTT	65.8	GAGACCCAGCTGCCAAAAAC	65.8
9	Hmct-53	AY962223	$(CT)_{13}$	163 bp	GCCCCTCTCCCTCTTCAGAT	65.8	CATTCGGGTGGGAAAAAGAA	65.7
10	Hmct-58	AY962229	$(AG)_{10}$	141 bp	GGAGGCAGTGAGCACTGAAA	65.6	TTCTTTGTTGACGTCCCCTTC	65.1
11	Hmct-61	AY962233	$(GA)_4$ $(GG)(GA)_4$	194 bp	ACGATTGCACACCCCTAGA	66.6	CAACAGACTCTTCGCCCAAA	65.2
12	Hmac-4	AY135656	$(CT)_{17}$ $(CA)_3$ $(CT)_2$ $(GT)_{16}$	216 bp	GTTTTCCTCCGCAGACTCAG	63.9	ATCCACCAAA TAAGGCATGA	61.7
13	Hmac-5	AY135657	$(CA)_{18}$	270 bp	TCGGTTGGTTTACCATGACA	63.8	ACATCACATGAGTGTATCTGATCTC	60.9
14	Hmac-13	AY962198	$(AC)_{11}$	246 bp	CATCCCTGCATTCCCCTAAT	64.7	ATGGGATGGGACCGTAACAA	66.1
15	Hmac-14	AY962202	$(CA)_{13}$	228 bp	CTCCACTTGTGCAAGCGTTC	65.9	GCAATCAAACGTGCCAAGAA	65.9
16	Hmac-17	AY962212	$(CT)_6$ $(CA)_{14}(TA)_4$	217 bp	CGCTGTGCCTTGTTGATCTT	65.2	CACTGGCACGTGAATCCCTA	66.0
17	*Hev*-glu	JQ 650524	$(CT)_{18}$	146 bp	TAGTTCCCTGCCCTCAAGAA	57.3	GAGCATCGTCAACTTTAACATGAG	59.3

Table 1.4. Details of primer pairs for EST-SSR markers used in the genetic diversity analysis

Sl. No.	Clone	GenBank accession No.	SSR Motif	Expected allele size	Forward primer (5'-3')	T_m	Reverse primer (5'-3')	T_m
1	HBE 002	EC 609891	$(TCT)_8$	153	AACTTCTTGTTTGCTTCCCTT	53.96	CGGGCGATACAAGCAGTC	58.24
2	HBE 003	EC 609774	$(TTC)_6$	231	AGCCAGTCCCTTGTCTTTTC	59.34	TGTTTGGCTTTCTTTAGTTTTTAC	54.17
3	HBE 010	EC 609548	$(AG)_{14}$	265	GGTTGAAACTTCGCACGC	55.96	TTTTGCTGCCTGCTGTCG	53.68
4	HBE 017	EC 608908	$(ATG)_6$	199	AGCTGGTGCAAACATCCTG	56.66	CGTTCCTCCAAAAGTTATTCAA	54.66
5	HBE 033	EC 608405	$(TT)_9$	171	TTCATCAATCATCTTGCCTTTT	52.80	CCCAAAACTGAGCTTGAAAG	55.91
6	HBE 034	-	$(TTC)_6$	222	AGCTTCTGCATGGGTTTGG	56.66	CTAAATCCAGGAGGCAGTCTAA	58.39
7	HBE 044	-	$(TTC)_8$	157	TTAAAGAGCTAGAATTGGGG	55.91	GCACACCGTTCTTGACAAAAT	55.91
8	HBE 068	EC 607281	$(GAA)_8$	213	TCGGGTCACCAACAGCAA	55.96	CTCCCAATCGCCGTATCG	58.24
9	HBE 092	EC 606292	$(ACC)_6$	172	TTCTTGTGAGCGAGTTGCC	56.66	TGCCGTTACCAGGGACAA	55.96
10	HBE 101	EC 606085	$(TTA)_7$	176	CCAGAAACTGACGATGAAGAA	55.91	AACAAGAGAGGACATTGACCC	57.87
11	HBE 122	EC 605312	$(ATA)_8$	212	GATGCCCGATGGACCTCT	58.24	CTACAAAACCCGACCCAATA	55.24
12	HBE 126	EC 605124	$(AAG)_{10}$	242	ACTTCTGGGTTTGATTTTGATG	54.66	CCTCTTTCTTCGTCGTTGC	56.66
13	HBE 170	EC 602995	$(AAG)_{10}$	268	ATTTGGGTGACCTGTTTGG	54.51	CATGACCTCTTGTACTCGTTGC	60.25
14	HBE 190	EC 601511	$(CTT)_{11}$	278	CCCTTCCTTCTGTCCCTCTC	61.40	GGGTAGATTCTGGAGGTCGG	61.40

Table 1.5. Details of the M-series of primers screened for polymorphisms in parents

Sl. No.	Clone	GenBank accessionNo.	Forward primer (5'-3')	T_m	Reverse primer (5'-3')	T_m	Linkage group (LG) assignment in French map	Profiles
1	M-72	AF 221696	ATAGGCCAAGAAATCTCACCTTC	58.9	AGATTCACTCCCTCTGGTTTTTC	58.9	LG-8	complex
2	M-124	AF 221697	GTTCACCCGTGTTATTTCACTTC	58.9	CGAGTACTTGGATGTGTCAATGG	60.6	LG-14	simple
3	M-127	AF 221698	ATAATAGATGGGAAGCCCCTT	58.4	GGGGAGACAATTAGACGTAAAGT	58.9	LG-7	complex
4	M-197	AF 221711	TGGAGACCCTCTTAACCAAAGTT	58.9	GGAATGCAATTTCCCCTTGTATC	58.9	Unlinked	complex
5	M-249	AF 221699	GCGACTGTTTTCGAGAGAGTGAT	60.6	TTCAATCGCAAGAGAAGACAGAG	58.9	Unlinked	complex
6	M-291	AF 221700	TGCTTGAAGTAGAAATGCCCCTA	58.9	TGCCAGAAGATATTCCTACATGC	58.9	LG-1	complex
7	M-379	AF 221703	TCAAAGAGAAGATGCCAAGAGAG	58.9	TGAGGTTCAGACTTTTACTCCAG	58.9	LG-11	complex
8	M-508	AF 221705	GCAGAGCTAACCCTCTCTTCATT	60.6	TCTCAGAAGACCGGAAACAAATC	58.9	LG-1	complex
9	M-574	AF 221706	GCCTTGTGTCCTCTACTTGTC	60.6	GCTGGTTCACCGATCTACTTTT	58.4	Unlinked	simple
10	M-613	AF 221707	GCATAGGATGCAAGTAAGTAGGC	60.6	GTAATGATGTAGCTGCACAACGA	58.9	LG-11	complex
11	M-622	AF 221709	CCCTAAGAAACACCTGACATCAA	58.9	GAGTTCAGGACTTCAGGTTTCAA	58.9	LG-2	complex
12	M-630	AF 221710	GCAGTTGCAAACTCCAATACAAC	58.9	TCTATCTGTATCGGTGGTCATTG	58.9	LG-8	complex

the genetic diversity analysis of the selected 28 popular *Hevea* clones (Table 1.4). The reaction procedure and PCR conditions remained the same as described earlier for other SSR markers.

1.3.3.3.3 SSR markers from the genetic linkage map developed by French group

Twelve SSR markers (M-series) reported by Lespinasse *et al.* (2000) were screened in RRII 105 and RRII 118 as described above to identify the polymorphic markers (Table 1.5). The same PCR reaction protocol with touch down PCR programme was followed.

1.3.3.3.4 Development of an enriched trinucleotide repeat library in *Hevea*

1.3.3.3.4.1 Selection of restriction enzyme and detection of trinucleotide repeats in the genomic DNA of *Hevea*

Genomic DNA of the clone RRII 105 was digested with a set of restriction enzymes for the selection of the appropriate enzyme based on size of the restricted fragments (200-800 bp). The complete restriction of 20 µg of genomic DNA was performed using the restriction enzymes *Eco* RI, *Sau* 3A, *Mse* I and *Tsp*509I (Fig. 1.5) in a total volume of 50 µl reaction mix using 40 U of each restriction enzyme with its buffer and supplemented with BSA. *Eco*RI and *Sau* 3A were incubated at 37°C overnight while *Mse*I and *Tsp*509 I were incubated at 65°C overnight. *Eco*RI and *Sau* 3A were heat inactivated at 65°C for 20 min after incubation. Complete digestion was ensured by agarose gel electrophoresis. Fifteen microliters each of restricted products of genomic DNA was electrophoresed on a 1% agarose gel and then blotted onto a nylon membrane. The blot was then hybridized with the radiolabelled probe $(AAG)_n$ (Fig. 1.6). The repeat was end labeled with the radioisotope γ^{32}P-ATP. The total reaction was performed in 20 µl and the contents being: 2 µl trinucleotide repeat (20 pmols), 1 x kinase buffer, 5 µl γ^{32}P-ATP 5 µl and T4 PNK enzyme 15 U/reaction. The mixture was incubated at 37°C for 30 min, inactivated with 1 µl of 0.5 M EDTA after incubation and added 90 µl 1x TEN buffer. The labeled probe was added to the hybridization chamber and hybridized at 37°C overnight after pre hybridization at 42°C for 45 min. The hybridized blots were then dried on a filter paper and exposed to X-ray film and the autoradiography was performed.

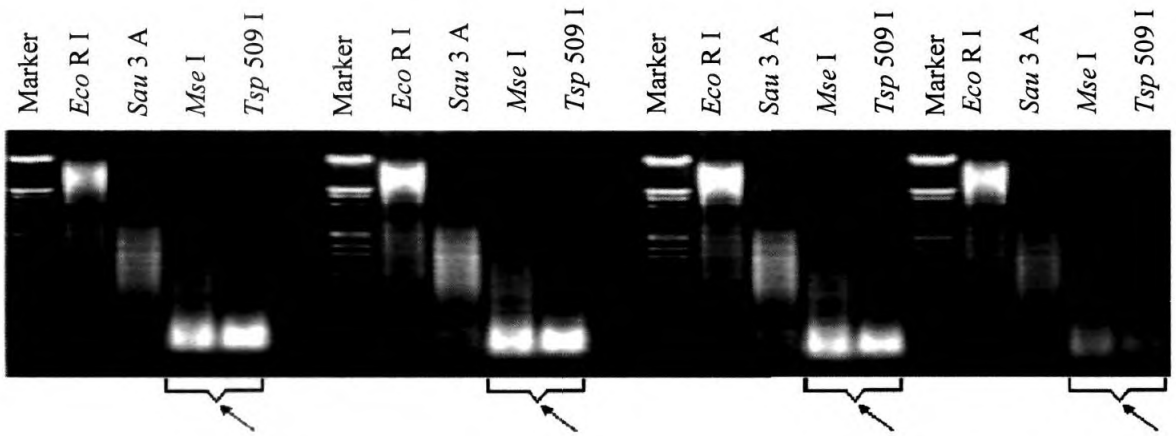


Fig. 1.5. EtBr stained agarose gel photograph showing screening of restriction enzymes for digestion of genomic DNA for the development of enriched trinucleotide repeat library. Four sets of digested products were electrophoresed for Southern hybridization of the blot with oligonucleotide repeat probe. Arrowheads show the restriction enzymes with the appropriate size range after digestion. λ DNA/*Eco* R I+*Hin*d III digest was used as molecular weight marker

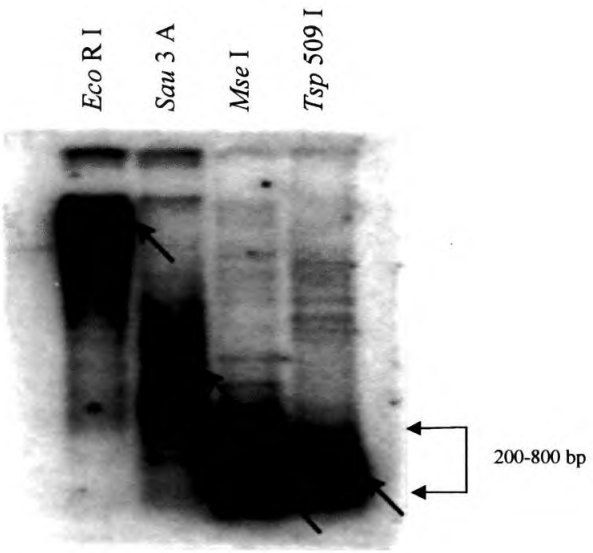


Fig. 1.6. Autoradiogram of the hybridized blot with trinucleotide repeat probe showing the abundance of (AAG) repeat in *Hevea* genome. Arrowheads show the intensity of hybridization signals. *Tsp* 509 I digestion was selected for desired fragment size range 200-800 bp

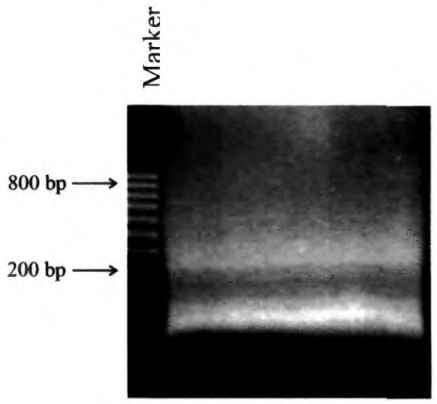


Fig. 1.7. EtBr stained agarose gel photograph showing digested genomic DNA of RRII 105 with the restriction enzyme *Tsp* 509I. Restricted fragments of size range 200-800 bp ware selected for the library construction. ΦX 174 DNA/ *Hae* III digest was used as molecular weight marker

1.3.3.3.4.2 Enrichment procedure of the genomic DNA with trinucleotide repeats

Restriction digestion of genomic DNA

The total genomic DNA of 60 µg was digested in a total volume of 100 µl with 10 U of *Tsp*509 I restriction enzyme at 65°C for 8 hours in 1x reaction buffer and 1 µl BSA (10 mg/ml). Fragments size of 200 to 800 bp were selected (Fig. 1.7). Fragments were purified from TAE agarose using GFX Microspin column (GE Healthcare).

Preparation of oligonucleotide adapters

Adapters were prepared to flank the ends of the digested DNA fragments (size selected). Tsp-long and Tsp-short oligos (Sigma Genosys) with an internal *Eco*RI site were used as oligonucleotide adapters.

Tsp-short: 5'-CGGAATTCTGGACTCAGTGCC-3', Tm: 68°C, (GC%: 57.14)

Tsp-long: 5'-AATTGGCACTGAGTCCAGAATTCCG-3', Tm: 71.5°C, (GC%: 48)

The annealing mixture contain 20 µl (100 µM) of each adapter and 1x annealing buffer (STE) in a total volume of 100 µl. The mix was heated to 70°C and incubated for 10 min before being allowed to cool slowly to room temperature.

Ligation of the adapters to the Tsp 509 I restricted DNA

Synthetic adapter was ligated to restricted fragments to provide annealing sites for the enrichment of ligated fragments by PCR amplification. Ten microlitre of the above combined adapter mix, 5 µl of 1x ligase buffer and 10 µl of T4 DNA ligase (10 U) were mixed to 10 µl of *Tsp*509I DNA digested in a total volume of 50 µl. The incubation was performed at 16°C overnight.

3' end labeling of synthetic trinucleotide repeat oligos

For the enrichment of the genomic DNA with the microsatellite containing fragments, the following synthetic oligonucleotides: $(AAG)_5$, $(AAT)_5$, and $(GTG)_5$ (Sigma Genosys) were used for end labelling. The melting temperatures of the oligos were found to be 38.9°C, 27.2°C and 57.2°C respectively. Oligo $(AAT)_5$ was given heat treatment at 90 °C for 3 to 4 min and then immediately kept on ice to prevent self annealing. The labeling reaction was performed with 10 µl of each oligonucleotide (100 pmoles each), 12 µl of cobalt chloride, 3 µl of biotin ddUTP (Biotin-16-2',3'dideoxy uridine-5'-triphosphate), 2 µl terminal transferase (40 units/µl) in 12 µl of 5 x buffer in a total volume of 60 µl. The mixture was incubated at 37°C, for 20 min. The labeled DNA oligos were precipitated, with 100%

ethanol and 3 M sodium acetate (6 µl) and centrifuged at 10,000 rpm for 30 min, rinsed twice with 80% ethanol at 10,000 rpm for 5 to 10 min. Pellet was dried, resuspended in 100 µl sterile water and stored at 4°C.

Ligation of oligonucleotides to the beads and hybridization

Streptavidin coated magnetic beads, which are uniform paramagnetic beads, covalently coated with purified streptavidin, were used in this experiment. Beads were thoroughly mixed and 100 µl of it was washed thrice with 0.1 % bovine serum albumin (BSA) in phosphate buffered saline (pH 7.4). Washing was done with 1x binding washing buffer (100 µl). Resuspended in 100 µl of 2x binding washing buffer and added 100 µl oligos to it. Incubated at room temperature for one hour and stored at 4°C. The beads were hybridized at 55°C overnight and the enriched DNA was suspended in 100 µl of 10x PCR buffer.

PCR amplification and concentrating by Microcon YM 100

PCR reactions were performed in a total volume of 50 µl and each reaction mix consisted of: 0.5 mM magnesium chloride (2µl), 10 mM dNTPs (1 µl), *Tsp* short primer (2.5 µl), sterile water (39 µl), *Taq* DNA polymerase 5 units/µl, (0.4 µl) and the DNA (5 µl). Amplifications were performed in an Eppendorf Mastercycler, using following PCR conditions: initial denaturation at 94°C for 3 min followed by 22 cycles of denaturation at 94°C for 30 sec, annealing at 64°C for 45 sec and elongation at 72°C for 1 min with a final extension step of 7 min at 72°C. Amplified product was concentrated using Microcon Centrifugal Filter Devices (Amicon Bioseparations YM 100) and stored at 4 °C. Enriched library was constructed in two vectors: a) lambda phage and 2) plasmid vector system.

1.3.3.3.4.3 Construction of lambda phage library

Digestion of enriched DNA samples with EcoRI

Enriched, PCR amplified DNA was digested with restriction enzyme *Eco*RI The reaction was carried out in a total volume of 30 µl and the components of the reaction were 5 µl of PCR amplified DNA (434 ng/µl), 0.3 µl of BSA (50 mg/ml), 3 µl of Buffer (10x) and 1 µl *Eco*RI (40 U/ µl) enzyme. The reaction mixture was incubated overnight at 37°C for complete digestion. Heat inactivation of the enzyme was performed at 65°C for 10 min and the mixture was spinned down and was loaded onto a 1% gel for size selection. Three hundred to 900 bp size range was selected and the fragments were eluted.

Ligation of digested DNA fragments to lambda vector (Lambda Zap II predigested vector)

*Eco*R I digested, size selected fragments were ligated to lambda ZAP II vector (Fig. 1.8) for phage packaging and incubated at 4°C. Five microlitre of the ligation mix included: 1 µl of Lambda ZAP II predigested vector, 2.5 µl of sample insert, 0.5 µl of 10x ligase buffer, 0.5 µl of 10 mM ATP, pH 7.5 and 0.5 µl T4 DNA ligase.

Phage packaging and enriched microsatellite library development

Ligated DNA samples were added to the packaging extract. Stirred the tube gently with a pipette tip to mix well. Spinned and incubated at 22°C for 2 hours. Added 500 µl SM buffer, 20 µl chloroform and spinned to sediment the bacterial debris. The supernatant containing the phage is stored at 4°C. For amplifying the library, the host bacteria, XL-1 Blue cells were streaked on an LB-Tetracycline medium and incubated overnight for obtaining single colonies. A single colony was inoculated onto LB medium supplemented with 10 mM $MgSO_4$ and 0.2% (w/v) maltose. Incubated overnight at 30°C, 200 rpm. Spinned the cells at 500 xg for 10 min and discarded the supernatant. Resuspended in sterile 10 mM $MgSO_4$. The host cells, (XL-1 Blue, 0.5 OD; 200 µl) mixed with 1 µl of 1:10 diluted packaged trinucleotide repeats were incubated at 37°C for 15 to 30 min. After incubation 3.5 ml of LB top agar was mixed well and pour plated onto LB plates. The plates were incubated at 37°C for 6 – 8 hours. Plaques were visible after 6 hours. Efficiency of plaque formation was calculated according to the formula:

$$\frac{\text{Number of plaques x dilution factor x total packaging volume}}{\text{Total quantity of DNA (µg) x number of microlitres (µl) plated}}$$

1.3.3.3.4.4 Screening of the lambda phage library

1.3.3.3.4.4.1 Primary screening

Plaque lifts onto nylon membrane

The phage plaques were transferred to positively charged nylon membrane, denatured by submerging the membrane in denaturing solution (1.5 M NaCl and 0.5 M NaOH), neutralized in neutralizing solution [1.5 M NaCl and 0.5 M Tris-HCl (pH 8)] and finally the blot was washed in 2x SSC twice, dried and crosslinked with the UV crosslinker (Hoefer UVC-500).

Probe preparation

Forty picomoles of trinucleotide repeats [(AAG), (AAT) and (GTG)] were endlabelled with ^{32}P- ATP in a total volume of 30 µl using 30 U of the T4 Polynucleotide Kinase enzyme at 37°C for 45 min. The reaction was stopped by adding 3 µl of 0.5M EDTA (pH 8) and finally made up to 50 µl with TEN buffer.

Hybridization of blots with labeled probe and selection of positive plaques

Pre-hybridization was performed in the hybridization chamber at 45°C for 2-3 hours and hybridization was performed at 37°C overnight. Washing of the unbound probes were done with 2x SSC for 15 min followed by 0.5x SSC for 15 min. The blots were then dried on a filter paper and exposed to X-ray film and the autoradiography was performed. Those plaques showing positive signals detected as black spots on film were transfered from the LB agar plate with the pipette tip and stored in a sterile microcentrifuge containing 500 µl SM buffer and 20 µl of chloroform.

1.3.3.3.4.4.2 Secondary screening of positive plaques and *in vivo* excision

The positive plaques after primary screening were again subjected to second round of screening through blot hybridization with labeled probes as described above. For *in vivo* excision, XL-1 blue and SOLR cells were grown overnight in 20 ml LB broth with supplements (0.2 % maltose and 10 mM $MgSO_4$, 200 rpm at 30°C). The SOLR cells were pelleted (1000 xg, 10 min). The pellet was resuspended in 10 ml of 10 mM $MgSO_4$ and the concentration of cells was adjusted to an OD_{600} of 1. The following components were mixed in a polypropylene tube: 100 µl of XL-1 Blue cells, 125 µl of phage stock and 0.5 µl of the Ex-assist helper phage, incubated at 37°C for 15 min to allow phage to attach to the cells. LB broth with supplements (3 ml) was added to the mixture and incubated at 37°C overnight with shaking. The following day tube was heated at 65- 70°C for 20 min to lyse the lambda phage particles and the cells were spinned at 1000 xg for 15 min to pellet the cell debris. The supernatant was decanted into a sterile tube containing the excised pBluescript phagemid packaged as filamentous particles. 10 µl of this phage supernatant was added to 200 µl of SOLR cells and incubated at 37°C for 15 min. One hundred microlitre of the cell mixture was plated on LB-ampicillin agar plates ampicillin (100 µg/ml) and incubated overnight at 37°C. Single colonies from the plates were streaked on a

new LB-ampicillin agar plate and phagemids were isolated using the Plasmid Prep Mini Spin kit (GE Healthcare) isolation kit. The plasmid concentration was checked spectrophotometrically.

An alternative method was developed for an efficient screening of large number of plaques at a time, where the LB plates were incubated with XL-1 Blue cells first and the plate was divided into different grids by marking on the bottom side of the plate. One microliter of the diluted phage (1:10 dilution) was added on each grid with micropipette and incubated at 37°C for 5-7 hours till the plaques attained the size of a pinhead. Blotting and hybridization of blots were performed following standard procedures as mentioned elsewhere. Autoradiography was performed for the detection of positive plaques. Large number of primary plaques could be screened simultaneously following this procedure.

The recombinant lambda vector from the positive plaques was converted to pBluescript phagemids by *in-vivo* excision that allowed insert characterization in the plasmid system. Plasmids were isolated using GenElute HP Plasmid Miniprep Kit (Sigma) and sequenced at Macrogen Inc., Republic of Korea.

1.3.3.3.4.5 Construction of plasmid (pGEM-T vector) library

The vector used for the ligation reaction was pGEM-T (Fig. 1.9). The reaction was performed in a total volume of 10 µl and the components were: 1 µl of vector (50 ng/µl), 5 µl of 2 x ligation buffer, 1 µl of PCR product (434 ng/µl) and of 1 µl of T4 DNA ligase (3U/ µl. The mixture was vortexed briefly and centrifuged for 3 to 5 sec. Incubated the ligation mix at 4°C overnight. Competent DH5α cells were prepared and transformation was performed. The transformed cells were plated onto LB plates (Sambrook *et al.*, 1989) containing antibiotic ampicillin (50 mg/ml), spread plate with 40 µl [100 mM IPTG (isopropyl β-D thiogalacto-pyranoside] and 20 µl (50 mg/ml) X-Gal. Single colonies of the recombinant colonies were streaked onto LB plates supplemented with antibiotic ampicillin (50 mg/ml). Colony PCR was performed to confirm the presence of the insert. Ten microlitre reaction mixes were made with 0.2 µl of 10 mM dNTPs, 1 µl of 10x PCR buffer, 0.8 µl Taq DNA polymerase (5U/µl, Promega), 8.32 µl of sterile water, 0.2 µl of T7 forward primer (5'-TAATACGACTCACTATAGGG-3') and 0.2 µl of SP6 reverse primer (5'-ATTTAGGTGACACTATAG-3') synthesized at Sigma Genosys. Inoculation of single colony was made with toothpick to each 10 µl PCR mix. The reaction was carried out in Eppendorf Mastercycler. The PCR programme was an initial denaturation at 95 °C for 3

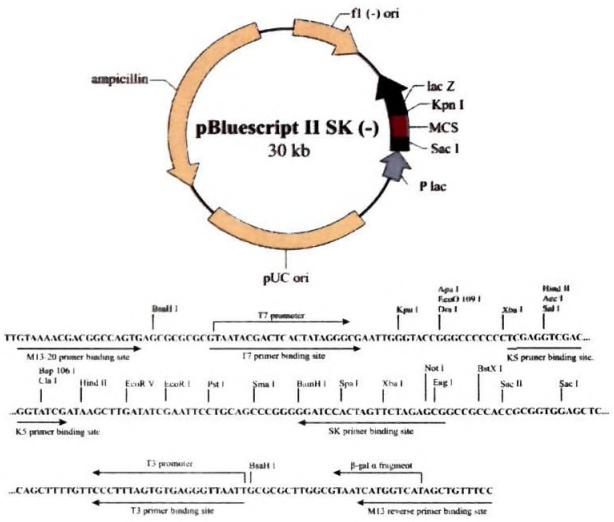


Fig. 1.8. Restriction map, multiple cloning site and sequence of vector pBluescript II SK(-)

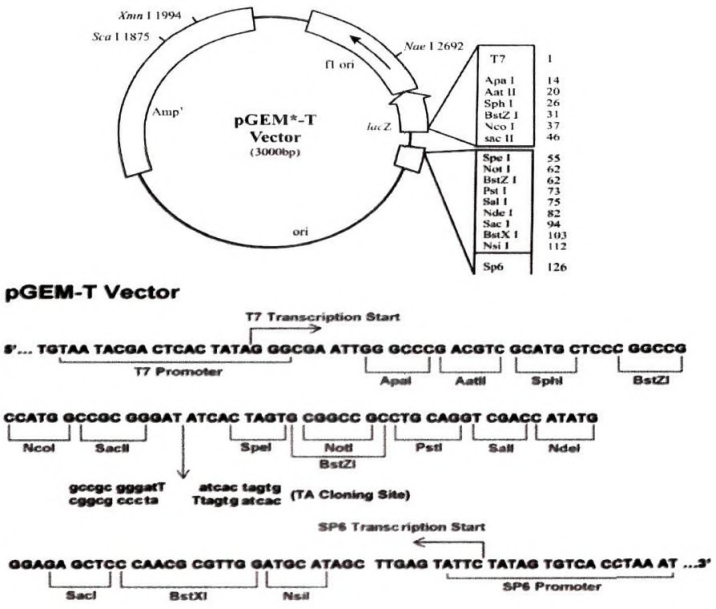


Fig. 1.9. Restriction map, multiple cloning site and sequence of pGEM-T vector

min followed by denaturation at 94°C for 30 sec, annealing at 50°C for 45 sec, extension at 72°C for 1.5 min for 35 cycles and a final elongation at 72°C for 7 min.

1.3.3.3.4.5.1 Southern blotting, hybridization of PCR amplified genomic inserts for detection of positive clones

Blotting of colonies directly for screening is not always reliable because there are chances of false positives. We observed that even after thorough washing of blots the problem of adherence of debris resulting in false signals with radioisotope still remained. Therefore another technique was followed where the recombinant colonies were streaked individually and single colonies were PCR amplified using vector directed primers (T7 and SP6). Southern blotting of the PCR amplified products of recombinant colonies were performed and hybridized with DIG-labeled trinucleotide repeat probes. The labeling reaction was carried out in a total volume of 30 µl and the components being: 14.5 µl of trinucleotide repeat oligo mix (50 pmoles each), 6 µl of cobalt chloride, 6 µl of 5x buffer, 1.5 µl of DIG-ddUTP and 2 µl of terminal transferase enzyme. Incubation was given at 37°C for 20 min and inactivation of the enzyme was performed by adding 0.2 M EDTA (pH 8). The positive clones were detected by both colorimetric detection assay using Anti-DIG AP-conjugate and chemiluminiscence substrate CSPD. Hybridization was performed in the hybridization chamber at 37°C overnight.

1.3.3.3.4.5.2 Detection procedure

Both colourimetric and chemiluminiscent detection methods were adopted.

Colourimetric detection

Detection procedure was started with maleic acid buffer wash (15 min). Incubated in 100 ml 1x blocking solution for 30 min. Diluted Anti-DIG-AP conjugate in (1x) blocking solution (2 µl in 20 ml blocking solution). Incubated the membrane, in it for 30 min. Washed twice in 100 ml maleic acid buffer (15 min each) and equilibrated with 20 ml detection buffer for 2-5 min. Added colour development substrate solution (45 µl NBT and 35 µl X-phosphate in 10 ml detection buffer) onto the membrane. Kept in dark till colour development started. Washed in sterile water, dried and photographed.

Chemiluminiscent detection

After equilibriating with detection buffer (procedures same as colourimetric reaction), 2ml of diluted CSPD solution (1:100 in detection buffer) was applied to the blotted membrane. The membrane was immediately covered with sheet of folder to spread the substrate evenly and without air bubbles over the membrane. Incubated for 5 min at room temperature and after squeezing out the excess fluid, incubate the damp membrane for 5-15 min at 37°C to enhance the luminescent reaction. The membrane was exposed to X-ray film for 10 min and developed.

1.3.3.3.4.5.3 Sequencing and primer designing

Plasmid DNA was isolated from positive clones and sequenced at Macrogen Inc., South Korea. Thirty-one primer pairs were designed based on the flanking sequences of the repeat regions from both plasmid and lambda phage library (Table 1.6). Primer designing was performed using the computer software DNASIS MAX. (Version 2.5), Mirai Bio, Hitachi Software Engineering Co. Ltd. and was synthesized by Ocimum Biosolutions. The length of the primers were about 18 to 26 bases and an average amplicon size was limited to 200 to 300 bp and optimium T_m was 55 °C. GC content was ranged between within 40 % to 60 %. All these primers were initially screened in a limited number of popular *Hevea* clones. The PCR reactions and thermal cycles were the same as mentioned elsewhere.

1.3.4 Computation of genetic diversity in cultivated popular clones of *Hevea brasiliensis*

1.3.4.1 Data scoring, cluster and principal coordinates analysis

The data obtained with the three molecular marker techniques (RAPD, AFLP and SSR) were scored as binary unit character as presence and absence (1 and 0) of bands and each band was treated as an independent locus with two alleles for the dominant markers. The data was compiled in separate data sheets for each marker based on unique and shared fragments. The efficiency of each marker determined through the polymorphic information content (PIC), was computed for dominant markers using the formula $PIC = 1- [f^2 + (1-f)^2$ (De Riek *et al.*, 2001), where PIC is the polymorphic information content and *f* is the frequency of the amplified allele. For the codominant marker SSRs, the PIC was determined as $PIC = 1- \sum Pi^2$ (Simpson, 1949), where Pi is the frequency of the i^{th} allele.

Table 1.6. Details of primer pairs generated for trinucleotide SSR marker development

Sl. No.	Clone	GenBank accession No.	SSR motifs	Expected allele size	Forward primer (5'-3')	T_m	Reverse primer (5'-3')	T_m
1	Hbtnr-1	FJ 160552	$(AAG)_5AAA\ (AAG)_5$	254 bp	CGAACAGAGCAGCAGCAGGA	61.4	CCTCAACCTCGACTTCACGACA	62.1
2	Hbtnr- 2	FJ 160553	$(CTT)_{18}$	265 bp	CGCCCTTTCGGGTTTTCATT	57.3	GCGTGGGGACCTGATATTGG	61.4
3	Hbtnr- 4	FJ 160554	$(CCA)_5N_6(CCA)_3$	250 bp	TCCACCATACCTTCAAGCACTCC	62.4	AGAGGTGTGGGCGGAATGAG	61.4
4	Hbtnr-9	FJ 160555	$(CTT)_{14}\ (CTT)_5\ (CTT)_4$	257 bp	GCTGCCTCCATTACCATTTCCA	60.3	GCACTGAAGCTGGAGAAACCA	59.8
5	Hbtnr-21	FJ 160557	$(TTC)_9$	237 bp	GCCTTCAAAACAAAGACCAAAGACC	61.3	TGAAGGGAAGAGGCTGGGTGT	61.8
6	Hbtnr-23	FJ 160558	$(ACC)_4$	261 bp	CAGGAGGGAACCTGCTCCAC	63.5	CAGTGCCAATTCTGCTCTGGA	59.8
7	Hbtnr-26	FJ 160559	i) $(CTTT)_3(CTT)_5(CT)_2CTTT(CTT)_2GCTT$ ii)($CTT)_3ATT(ATTT)_2(CTTT)_5(CTT)_4(CT)_2CTTT$ $(CTT)_2GCTT$ iii) $(CTT)_2(CT)_2CTT$	306 bp	GCCTTACAGAGGGGTAAGGTGTGA	64.4	CCAGTGAGGTGCAACCAGATCA	62.1
8	Hbtnr-31	FJ 160560	$(AAG)_{10}(AG)_8$	271 bp	GCAGCCCAGCAGGAAAAGAA	59.4	GCAGCCACCACCAGCAGATA	61 4
9	Hbtnr-32	FJ 160561	$(ACC)_6$	247 bp	TCAGAGAGCCGTTGAGGAGGA	61.8	GAAATCCACTGCCCCAAACTTTC	60.6
10	Hbtnr-35	FJ 160563	$(TGG)_2\ T(TGG)GAAGG(TGG)_3(TTC)_4$	232 bp	TGAATGGGAAGGGTGCTTCA	57.3	GATTGGAAGAGACACGGCAAA	57.9
11	Hbtnr-37	-	$(TGG)_3\ CTA(TGG)TTG(TGG)$	252 bp	CGATATGGCCACCCACCATT	59.4	CAGCCACAGCCATAGCCACA	61.4
12	Hbtnr-38	FJ 160564	$(TGG)_3AA(TGG)(GAA)_9\ (GA)_{16}$	312 bp	AAATCGCAGGCCAAGTCGTC	59.4	CACGAAGACGACGACAACAGG	61.8
13	Hbtnr-46	FJ 160565	$(CCA)_3(CAACCA)_4$	235 bp	GGCAAATGATGGGGGCAAT	56.7	AGGGACACAGCACCCTCCAG	63.5
14	Hbtnr-50	FJ 160567	$(CCG)_3CAG(CCG)(CCA)_5$	256 bp	CCAGCGTCTTCTTTTCAAGAATCA	59.3	GGACTCAGTGCCCCAGCTTC	63.5
15	Hbtnr-54	FJ 160568	$(CT)_6(CTT)_{11}$	256 bp	CAGCCATTCCACCAGCAGAG	61.4	AGCAGCGCAGCAGGAAAGAA	59.4
16	Hbtnr-55	FJ 160569	$(TGG)_7$	251 bp	CCGAAAAAGTCCACTGCTCCA	59.8	CCCACATCTCAGACGGCAGTT	61.8
17	Hbtnr-57	FJ 160570	$(CTT)_{17}CTA(CT)_{23}$	200 bp	ATGGAGAGGAGGGGCTTTTG	59.4	TGAGTTTGAGGGGGAGGACA	59.4
18	Hbtnr- 58	FJ 160571		247 bp	TTCTGGGGAATGGGTGTCTG	59.4	CATTTCTTCCGGCATTCTCG	57.3
19	Hbtnr- 71	FJ 160574	$(CTT)_{11}$	255 bp	GAAAAGCCCAATGCCTATGCT	57.9	CGGAGACCTGTCCCTGTAGGT	63.7
20	Hbtnr-75	FJ 160575	(GGT)GGG(GGT	244 bp	GTGCCAATTTTCGACTGCAA	55.3	AGCAGCTGAGACTGGGGGTA	61.4
21	Hbtnr-76	FJ 160576	$(CCA)CCCAGCACC(CCA)_4$	256 bp	TGGCGAGAACCAAGATGTGG	59.4	CGGGTAAGGGGTTTCATTTTG	57.9
22	Hbtnr-79	FJ 160577	$(TGG)TG(TGG)_4(TGG)_4$	246 bp	TGGCTCTAGTGCTAGTGGGAGA	62.1	CACCACCACTTGACCACTGC	61.4
23	Hbtnr-81	FJ 160578	$(CT)_{12}(CTT)TTAT(CTT)_6(CT)_5CACGA(CTT)_5$	255 bp	CTGATGCAGCCACATCACTT	57.3	CGTGAGAGAGAGATCGAACAGA	60.3
24	Hbtnr-83	FJ 160579	$(GGT)_5$	243 bp	TCGCTCAACCTCGACTTCAA	57.3	GGGACGTGAGAGCGATGAAC	61.4
25	Hbtnr-90	FJ 160580	$(ACC)_2ACA(ACC)_8$	272 bp	ACTGGGCCACTATTGCTATTACC	60.6	GGTAGTGGTGGCGACAGTGA	61.4
26	Hbtnr-91	FJ 160581	$(TGG)_6TGC(TGG)$	219 bp	CCTTTCCAGGTGTGATTTCTG	60.3	CAGAGAGCCGTTGAGGAGGA	61.4
27	Hbtnr-94		$(CTT)_5\ CCT(CTT)_2(CTT)_{28}$	262 bp	TGTTTCTTTCCCTTCTCCCTC	58.4	GGTGAAACTAGTAACGCAGGCAC	62.7
28	Hbtnr-96	FJ 160582	$(GAA)_5(GAA)_5ACA(GAA)_2(GAA)_4$ $(GAT)_4(CCTGCT)_2GCTCCT$	263 bp	TCCAATCAACATCACTCCCCT	58.4	CCATCTTTGCTGCCTCCATT	57.3
29	Hbtnr-99	FJ 160583	$(GGT)TGT(GGT)_3(TGG)_7$	240 bp	GGTGGAGGTTATGGCTGTGG	61.4	CCACCGTTATTCGACCATCA	57.3
30	Hbtnr-101	FJ 160584	$(CCA)_4CAATCA(CCA)_2AGCCCG(CCA)_3CCTTCA(CCA)_2$	244 bp	TGCTAGCACAAACCCTACCC	59.4	GAGGCTTTCTTGGTTCTACCT	60.3
31	Hbtnr-102	FJ 160585	$(GAA)_4ATAGAAAAG(GAA)_5ACA(GAA)_4(GAA)_4$	229 bp	GAGAACGAAAAGGGGCATCC	59.4	GAGAACGAAAAGGGGCATCC	57.9

1.3.4.2 Cluster analysis

Pairwise comparisons, were computed using simple matching coefficient to produce a similarity matrix and was analyzed using the similarity of qualitative data (SIM-QUAL) programme of Numerical Taxonomy Multivariate Analysis System (NTSYSpc) version 2.2i software package (Rohlf, 1998). Cluster analyses was performed using Unweighted Pair Group Method of Arithmetic averages (UPGMA) clustering (Sokal and Michener, 1958) to construct a dendrogram using the SAHN programme of NTSYSpc. Statistical stability of the branches in the clusters was estimated by bootstrap analysis with 1000 replicates using WINBOOT software program (Yap and Nelson, 1996).

1.3.4.3 Mantel's test

To test the statistical significance of groups, determined by cluster analysis using different marker systems, Mantel test was performed (Mantel, 1967). Similarity matrices obtained by genetic data were compared pairwise (SSR–RAPD, SSR–AFLP, genic SSRs – genomic SSRs and RAPD – AFLP) with the MXCOMP procedure (NTSYS) which produces the Pearson product-moment correlation, r, and the Mantel test statistic, Z. Mantel statistics calculate a linear correlation between the generated matrix of different marker systems and test its significance. If the two matrices show similar relationships, Z is large in comparison to chance expectation. The estimated Z was compared with its permutational distribution obtained from random samples of all possible permutations of the matrices.

1.3.4.4 Principal Coordinates Analysis (PCoA)

NTSYSpc was used to perform to show multiple dimensions of the distribution of the genotypes in a scatter-plot. DCENTER and EIGENVECTOR programmes were run for each matrix. For graphical representation of principal co-ordinate analysis, three-dimensional plots for populations were produced by NTSYS. This multivariate approach was used to complement the information obtained from cluster analysis because it is more informative regarding distances among major groups (Taran *et al.*, 2005).

1.3.3.5 Single Nucleotide Polymorphisms (SNPs)

1.3.3.5.1 Plant material

Sixteen popular *Hevea* clones (Table 1.7) were used in this study for the identification of single nucleotide polymorphisms in *Hevea.* The criteria of selection of the clones were a) mostly primary clones b) based on the diversity analysis with AFLP markers showing

Table 1.7. List of popular clones used for the identification of single nucleotide polymorphism (SNP)

Sl. No.	Clone	Parentage	Country of origin
1	RRII 33	Primary clone*	India
2	RRII 5	Primary clone*	India
3	RRIC 52	Primary clone*	Sri Lanka
4	BD 10	Primary clone*	Indonesia
5	PB 86	Primary clone*	Malaysia
6	PB 28/59	Primary clone*	Malaysia
7	PR 107	Primary clone*	Indonesia
8	Tjir 1	Primary clone*	Indonesia
9	Tjir 16	Primary clone*	Indonesia
10	GT-1	Primary clone*	Indonesia
11	GL-1	Primary clone*	Malaysia
12	Haiken 1	Primary clone*	China
13	KRS 25	Primary clone*	Thailand
14	RRII 105	Tjir 1 x GL-1	India
15	RRII 118	Mil3/2 x Hil 28	India
16	PB 235	PB 5/51 x PB S/78	Malaysia

* Elite selected clones with unknown parental origin

maximum diversity and c) parents of the mapping population for further segregation and linkage analysis.

1.3.3.5.2 Identification of SNPs

1.3.3.5.2.1 Designing of PCR primers

PCR primers were designed to amplify fragments of around 500 bp mainly in the 3' UTR of 16 genes (cDNA) encoding important enzymes catalyzing several biochemical reactions (Table 1.8) in all the rubber genotypes/clones using the published sequences in the GenBank.

1.3.3.5.2.2 PCR amplification and processing of amplified products for sequencing

The PCR reaction was carried out in a 25 µl final volume containing 50 ng of genomic DNA, 0.2 µM each of the forward and reverse primers, 200 µM dNTPs and 1 unit of *Taq* DNA polymerase (GE Healthcare, UK) along with the buffer supplied. The temperature cycle profile involved an initial denaturation step of 5 min at 95^0C followed by a touchdown PCR program. Temperature profiles of the touchdown PCR for 12 cycles were as follows: 94^0C for 30 sec, 63^0C for 1 min, $\Delta\downarrow$ 0.6^0C for 12 cycles, 72^0C for 1 min. This was followed by a normal cycling of 94^0C for 30 sec, 56^0C for 1 min, 72^0C for 1 min for 35 cycles followed by a final extension at 72^0C for 10 min. The PCR products were run on a 1% agarose gel stained with ethidium bromide and viewed using a gel documentation system (Eagle Eye II, Stratagene). Successfully amplified 12 loci out of 16 loci were processed for further sequencing reactions. Multiple PCRs were set for each gene to get sufficient products for sequencing. PCR products were purified using purification kit (GFX PCR DNA and Gel Band purification Kit, Amersham Biosciences). Concentration of PCR products was adjusted to 50 ng/µl using MilliQ water for sequencing. Each fragment was sequenced in both directions in each individual (12 loci x 16 genotypes) to verify sequence variants.

1.3.3.5.2.3 Sequence analyses and identification of SNPs

All sequencing reactions were performed using ABI Prism Big Dye Chemistry and were read on ABI 3700 automated sequencer (Applied Biosystems) at the sequencing facility of Macrogen Inc., Republic of Korea. Contigs were constructed and aligned using Phred/Phrap suite (Ewing *et al.*, 1998), and viewed through Consed (Gordon *et al.*, 1998). Heterozygous condition of the bases/nucleotides is identified by viewing the overlapping

Table 1 .8. Details of primers for amplification of genes selected for SNP analysis

GenBank Accn. No.	Gene designation	Expected amplicon size	Forward primer sequence (locus specific)	T_m	Reverse primer sequence(locus specific)	T_m
AB055496	Geranylgeranyl diphosphate synthase	500 bp	AGTTGTTGGAGGGGGCTGT	65.8	AATTGCGCTGGTCAAAGAG	64.9
AF 429384	Mevalonate kinase	499 bp	TGATGTGTCCATAACTGAGAAGG	62.8	CCCCAGAAACTTCAAAGAAG	61.5
AY 135188	Farnesyl diphosphate synthase	529 bp	CGGGTGAGAATCTGGACAAT	63.8	TGTGCTACAAGCAGAATAAGCAA	63.5
AF 193438	Ubiquitin precursor	494 bp	GAAGAAGAAGAAGAAAGGAGACGA	62.7	CCAGAATGTTTACATAGGCGAGT	62.9
U 80598	Latex patatin homolog	500 bp	CCCATCTCTCTGCCCTTTTT	64.2	ACTCGCATAAATTTGGAAGC	60.6

Table 1.9. Details of allele specific primers used for the SNP genotyping of mapping population

Marker	Primer sequence (5'-3')	Tm
GGDPS_88G	TTGAGGAAATATGCTAGGGG	55.3
LPH5R_313G	ATAGCTGCTTCATTAGTAGG	53.2
MK3F_197C	CACTGTTACCAACCCGTATCC	59.8
STS marker (farnesyl diphosphate synthase)	F- CGTATACACATGTTTGTG R-TGCCAAGAAGTTAAAGGA	58.9 56.4

peaks on the chromatogram. Poor quality sequences from the ends of the chromatogram were not considered for SNP detection.

1.3.3.5.2.3.4 Cloning and recombinant plasmid isolation for haplotype detection

The gel-purified PCR product was ligated to pGEM–T Vector (Promega, USA). Ligation mix was prepared by taking 2x rapid ligation buffer, pGEM-T Vector (50 ng/µl), T4 DNA ligase (3U/µl) and purified DNA. All the reagents ware mixed gently and kept at 4^0C overnight.

E. coli DH 5α cells were transformed with ligated products and plated onto LB medium supplemented with ampicillin (100 µg/ml) as well as IPTG and X Gal for blue/white selection of transformed colonies. Plates were incubated overnight at 37^0C. Colony PCR was performed using vector directed primers (T_7 and SP_6) in a 10 µl final volume containing 0.2 µM each of T_7 and SP_6 primers, 200 µM dNTPs and 0.8 U of *Taq* DNA polymerase (GE Healthcare), prescribed PCR buffer and a single white colony.

For plasmid isolation, single positive colony was streaked on LB-ampicillin plate. A single colony from the plate was grown overnight in LB broth supplemented with ampicillin at 37^0C. Plasmid was isolated using the kit (Illustra Plasmidprep Mini Spin kit, GE Healthcare). The concentration of isolated plasmid DNA was checked and adjusted to 50 ng/µl for sequencing.

1.3.3.5.3 SNP genotyping

SNP genotyping was performed following the techniques: allele-specific PCR, Cleaved Amplified Polymorphic Sequence (CAPS) and Sequence Tagged Site (STS). The genotyping techniques were initially standardized with two parental genotypes (RRII 105 and RRII 118) whose sequence informations were known.

1.3.3.5.3.1 Allele specific-PCR

The allele-specific primers were designed manually at the upstream of the substituted nucleotides with the last 3' base as the polymorphic base/allele (Table 1.9). Simple presence/absence PCR amplification was performed directly on total genomic DNA and the fragment was amplified only for one of the alleles. For the analysis of single nucleotide polymorphism, the PCR amplification was performed with three primers; one allele-specific primer and the gene/locus specific forward and reverse primers which were amplified as a positive control. Allele-specific PCR amplifications were carried out

for the three genes 1) geranylgeranyl diphosphate synthase, 2) mevalonate kinase and 3) latex patatin homolog. Initial screening was performed with the parental genotypes for identifying the allele specificity, whose sequence information was available. For achieving the allele specificity, primers were tested initially with varying annealing temperatures. The PCR reactions were performed in a total volume of 20 µl with 0.2 µM of the allele- specific primer, 0.4 µM of each gene specific primer, 200 µM of dNTPs, 0.1 µl of Advantage ® 2 polymerase mix (Clonetech) along with its buffer (10x advantage 2 SA) and 5 ng of genomic DNA. The temperature cycling parameters followed were an initial denaturation at 95^0C for one min followed by a touchdown PCR programme of 12 cycles : 95^0C for 30 sec, 65^0C for 45 sec, 68^0C for 1 min with, $\Delta\downarrow$ 0.7^0C for 12 cycles. This was followed by a normal cycling of 95^0C for 30 sec, 56^0C for 30 sec, 68^0 C for 1 min for 35 cycles followed by a final extension at 68^0C for 5 min. The PCR products were electrophoresed on a 1.4% agarose gel for 3 hours in 1x TAE buffer at 70 V, stained with ethidium bromide (0.5 µg/ml) and visualized on a UV transilluminator.

1.3.3.5.3.2 Cleaved Amplified Polymorphic Sequence (CAPS)

The SNP site in the locus ubiquitin precursor was utilized for developing CAPS marker. Aligned sequences of the loci under study derived from RRII 105 and RRII 118 were subjected to the computer program SNP2CAPS to generate CAPS marker. Here the restriction sites of all enzymes within the sequence were shown and the differences in the restriction site due to SNPs were identified. *Xho* I restriction enzyme was tested with the PCR product of the locus ubiquitin precursor. Here the recognition sequence of the enzyme was C/TCGAG. SNP was present in the parental genotype RRII 105 where the recognition site was functional, whereas in the case of RRII 118, the sequence was C/TCGAA where the site was non-functional. The expected fragment size of digested products was 200 and 290 bp in the case of RRII 105, and for RRII 118 the fragments remained undigested. The restriction digestion was carried out in a total volume of 30 µl, with 5 U of the restriction enzyme *Xho* I/ at 37°C overnight. The digested samples were run on a 1.5% agarose gel and electrophoresed for 3 hours in 1x TAE buffer at 70 V and then stained with ethidium bromide (0.5 µg/ml) and visualized on a UV transilluminator.

Restriction digestion was performed with *Xho* I in the progeny populations following the same reaction conditions and agarose gel electrophoresis was performed for

the detection of digested products. Segregation pattern obtained was scored for each individual following the instructions given in JoinMap 3 software.

1.3.3.5.3.3 Sequence Tagged Site (STS) marker

A 22 bp insertion-deletion polymorphism was identified from different clones at the 5' end of the *FDPS* gene sequence which was converted to a STS marker. The details of primer pairs utilized are given in Table 1.9. The initial screening was done on a set of five selected clones based on their diversity. Three distinct fragments were obtained resulting from 22 bp indels, which was utilized for genotyping the entire progeny for mapping. The PCR reactions were performed in a total volume of 20 µl with 0.2 µM each of the allele specific primer, 200 µM of dNTPs, 0.16 µl Taq DNA polymerase (5U/ µl) (Amersham) along with its buffer and 40 ng of genomic DNA. The PCR temperature profile followed were an initial denaturation at 94^0C for 1.30 min, annealing at 54^0C for 30 sec and 72^0C for one min followed by a touch down PCR programme of 10 cycles : 94^0C for 30 sec, 64^0C for 30 sec, 72^0C for 1 min with $\Delta\downarrow$ 1^0C for 10 cycles. This was followed by a normal cycling of 94^0C for 30 sec, 54^0C for 30sec, 72^0 C for 1 min for 35 cycles followed by a final extension at 72^0C for 5 min.

1.3.3.6 Resistance Gene Analogues (RGA)

PCR-based approach was adopted to amplify RGAs from rubber. Degenerate forward primers based on conserved motif of NBS (GVGKTT) and reverse primers based on membrane-spanning motif (GLPLA) were used for PCR amplification. The NBS regions were amplified from the genomic DNA of RRII 105 and RRII 118 using RGF3+RGR15 combination of degenerate primers:

RGF3: AAGAATTCGGNGTNGGNAAGACAAC (*Eco*R I)

RGR15: AACTCGAGAGNGCNAGNGGNAGTCC (*Xho* I). The amplicons were initially checked on agarose gel and later they were run on denaturing PAGE to detect polymorphic bands.

1.4 Results

1.4.1 Genomic DNA

The genomic DNA was isolated from 28 cultivated clones of *H. brasiliensis*. The yield of genomic DNA ranged from 164-954 µg/g of leaf tissue and the purity of DNA at A_{260}/A_{280} was in the range of 1.8-1.86. The same DNA was used for all the reactions with different marker systems under study.

1.4.2 Polymorphism studies in rubber clones

1.4.2.1 Random Amplified Polymorphic DNA (RAPD) markers

One hundred and thirty eight polymorphic primers with strong and reproducible amplification profiles were obtained from the screening of the 520 Operon arbitrary decamer primers. From these, a set of 16 informative primers were utilized for characterization of 28 selected popular *Hevea* clones (Fig. 1.10). One hundred and eight polymorphic bands were identified from a total of 138 amplified fragments (78%) (Table 1.10). The maximum numbers of 14 bands/loci were amplified with the primer OPH-03 and minimum number of 4 bands were amplified with the primers OPG03 and OPF10. Polymorphic bands were ranged between two to 13 with an average number of seven polymorphic fragments per primer. The Polymorphism Information Content (PIC) values ranged from 0.15 for the marker OPF-10 to 0.30 for the markers OPD-08, OPG-17 and OPI-06 and the average PIC was found to be 0.22 (Table 1.11).

1.4.2.2 Amplified Fragment Length Polymorphism (AFLP) markers

AFLP markers were used to characterize the same set of *Hevea* clones with nine selected primer combinations of *Eco*R I-ANN and *Mse* I-ANN (Fig. 1.11). In total 434 loci were amplified with an average of 48.3 loci per primer combination. The primer combination E-AGA/M-AAC showed the maximum number of amplified loci (83), whereas the least number of loci were amplified with the primer combination E-AGT/M-ACA (17). The polymorphic loci generated ranged between nine (E-AGT/M-ACA) to 49 (E-AGA/M-AAC) with an average of 29.11 fragments per primer combination. Out of 434 loci, 262 were polymorphic (60.4%). The relative number of polymorphic fragments to the total number of amplified fragments ranged between 52.29 to 71.43 with an average of 66.37% for all the primers expressed as percentage. The PIC values ranged from 0.1 to 0.23 for

Table 1.10. List of 'Operon' primer series selected for RAPD analysis

Sl. No.	Operon primers (1-20)	No. of polymorphic primers	Sl. No.	Operon Primers (1-20)	No. of polymorphic primers
1	OPA series	3	21	OPAH series	3
2	OPB series	3	22	OPAI series	3
3	OPC series	2	23	OPAJ series	3
4	OPD series	2	24	OPAK series	6
5	OPE series	2	25	OPAL series	5
6	OPF series	2	26	OPAM series	5
7	OPG series	2	27	OPAN series	5
8	OPH series	2	28	OPAO series	6
9	OPI series	2	29	OPAP series	6
10	OPJ series	2	30	OPAQ series	4
11	OPO series	3	31	OPAR series	3
12	OPR series	4	32	OPAS series	5
13	OPX series	6	33	OPAT series	2
14	OPAA series	6	34	OPAU series	3
15	OPAB series	5	35	OPAV series	4
16	OPAC series	4	36	OPAW series	2
17	OPAD series	5	37	OPAX series	3
18	OPAE series	4	38	OPAY series	3
19	OPAF series	4	39	OPAZ series	3
20	OPAG series	2	Total		138

Table 1.11. The percent polymorphism and polymorphism information content (PIC) of RAPD markers used in genetic diversity analysis in *Hevea*

Sl. No.	Primer	Total bands	Polymorphic bands	Percent polymorphism (%)	Polymorphism Information Content (PIC)
1	OPA-01	7	6	75	0.25
2	OPA-10	13	9	69.2	0.16
3	OPB-15	8	7	87.5	0.20
4	OPC-05	9	8	88.8	0.24
5	OPD-08	8	8	100	0.30
6	OPE-03	8	6	75	0.23
7	OPG-03	4	2	50	0.06
8	OPG-10	7	5	71.4	0.27
9	OPG-17	10	9	90	0.30
10	OPH-18	3	2	66.6	0.16
11	OPH-03	14	12	80	0.28
12	OPI-06	12	7	70	0.30
13	OPJ-19	11	8	72.7	0.25
14	OPJ-20	12	11	91.6	0.25
15	0PI-12	8	6	75.0	0.21
16	OPF10	4	2	50.0	0.15
Total		138	108		0.22

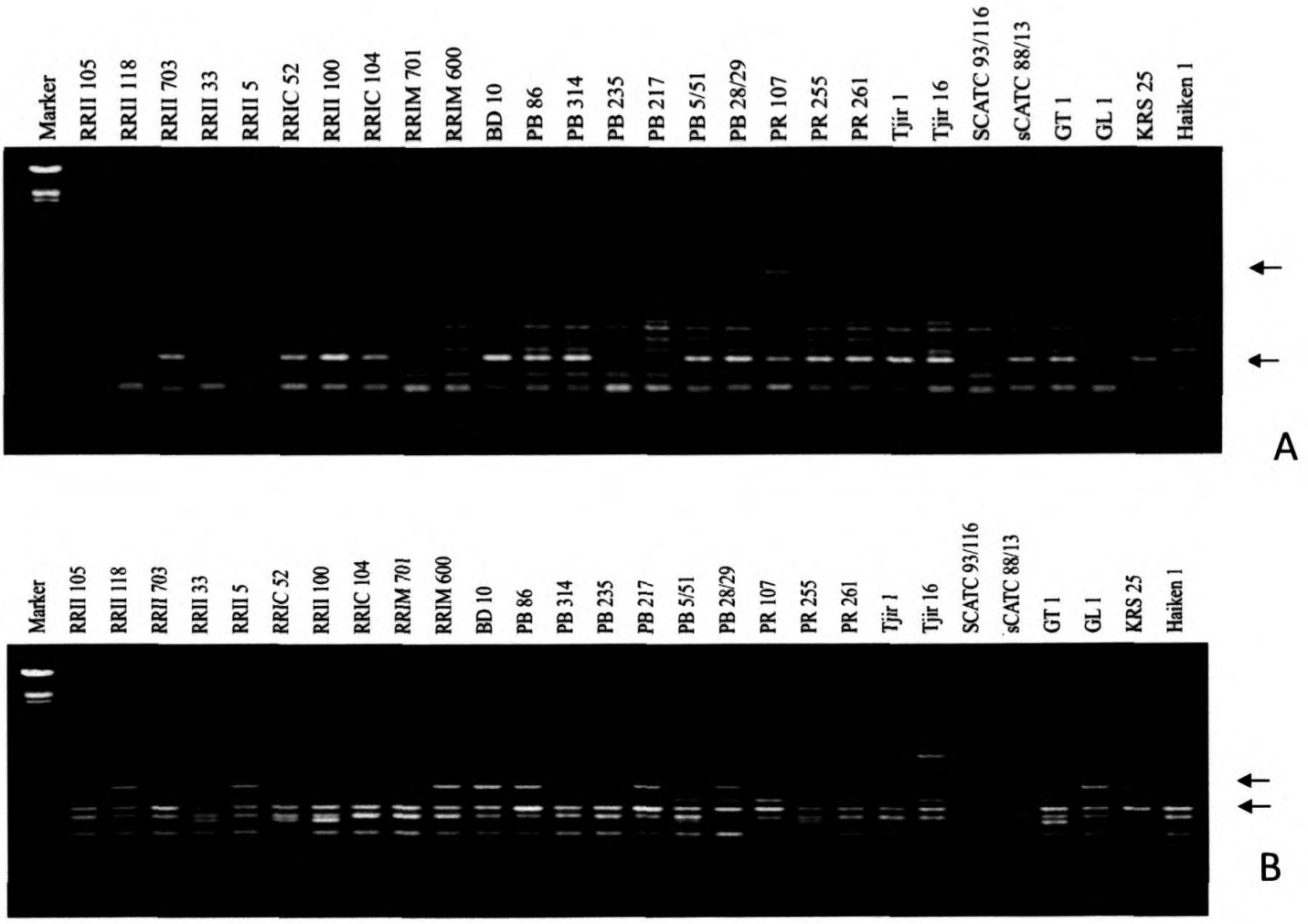


Fig. 1.10. Representative EtBr stained agarose gel photograph showing the RAPD profiles of 28 popular *Hevea brasiliensis* clones with the arbitrary primers OPH-03 (A) and OPA-03 (B) for assessment of genetic diversity. The arrowheads indicate polymorphic bands

Table 1. 12. The details of amplified/ polymorphic loci generated with nine AFLP primer combinations used in genetic diversity analysis of *H. brasiliensis* clones

Primer combinations	No. of bands amplified	No. of polymor-phic bands	Percentage of polymorphism	Polymorphism information content (PIC)
E-AGA/M-AAC	83	49	59.04	0.169
E-AGC/M-AAC	27	19	70.37	0.235
E-AGT/M-ACA	17	9	52.94	0.148
E-AGT/M-AGA	34	18	52.29	0.102
E-AGT/M-AAC	40	22	55.00	0.159
E-AGG/M-ACA	49	35	71.43	0.223
E-AGA/M-ACT	57	32	56.14	0.177
E-AGC/M-AGA	64	43	67.19	0.192
E-AGT/M-ACA	63	35	55.55	0.142
Total	434	262	60.37	1.547
Average	48.22	29.11	66.7	0.17

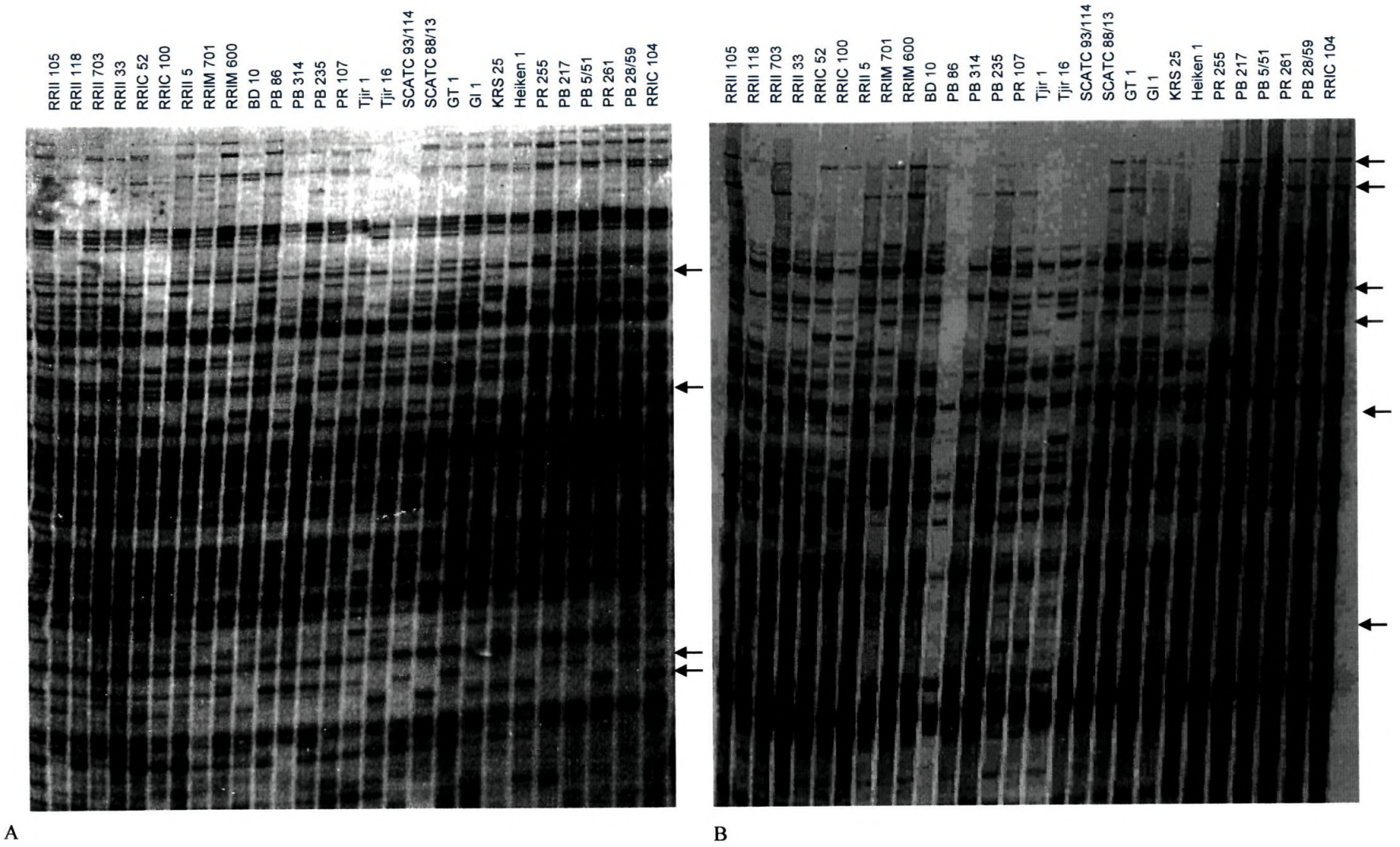


Fig. 1.11. Representative autoradiogram showing AFLP profiles of 28 popular *Hevea* clones for assessment of genetic diversity. A) primer combination used was E-AGC+M-AGA, B) primer combination used was E-AGC+ M-AAC. Arrowheads indicate the polymorphic bands

Table 1.13. Details of the alleles generated by genomic microsatellite markers used in genetic diversity analysis with their polymorphism information content

Sl. No.	Marker	No.of alleles identified	Polymorphism Information Content (PIC)
1	Hmct-1	2	0.45
2	Hmct-16	5	0.42
3	Hmct-17	3	0.61
4	Hmct-19	4	0.68
5	Hmct-20	8	0.80
6	Hmct-21	6	0.51
7	Hmct-27	3	0.56
8	Hmct-45	3	0.52
9	Hmct-53	3	0.57
10	Hmct-58	4	0.53
11	Hmct-61	2	0.59
12	Hmac-4	6	0.58
13	Hmac-5	5	0.53
14	Hmac-13	3	0.53
15	Hmac-14	4	0.71
16	Hmac-17	6	0.68
17	*Hev*-glu	7	0.79
Total		74	10.06
Average		4.35	0.59

the polymorphic data set and the average PIC was found to be 0.17 for the AFLP markers (Table 1.12).

1.4.2.3 Microsatellite/SSR markers

1.4.2.3.1 Dinucleotide SSR markers

A set of 17 markers were selected for genetic analysis of popular *Hevea* clones from 67 dinucleotide derived microsatellite markers generated as described earlier. Seventeen genomic SSR markers used for genetic diversity analysis generated 74 alleles with an average of 4.34 alleles per marker. The PIC value was calculated to estimate the informativeness of each primer and it varied from 0.42 to 0.80 with an average of 0.59. Maximum number of 7 alleles were amplified by the marker *Hev*-glu and minimum of two alleles were amplified by the markers Hmct-1 and Hmct-61 (Table 1.13). Representative SSR profiles of the *Hevea* clones are given in Fig. 1.12.

1.4.2.3.2 EST- derived/ genic microsatellite markers

From the SSR-ESTs available for *Hevea,* in the public domain, 14 markers were analyzed to estimate the genetic characterization (Fig. 1.13) which had generated a total of 51 alleles with an average of 3.64 alleles per marker. A maximum of 5 alleles were detected by the primers (HBE-010, HBE-190 and HBE-126). The PIC value of each marker was detected and the values ranged from 0.2 to 0.76 with an average of 0.56 (Table 1.14). Twenty five primer pairs which were polymorphic and showing simple amplification patterns within the parents of the progeny population were finally utilized in the segregating population. It was found that the trinucleotide derived repeats generated more polymorphic markers (Table 15).

1.4.2.3.3 SSR markers from other sources

The 'M' series of primers were screened in selected popular clones of *Hevea* and two markers M-127 and M- 574 were identified as polymorphic with simple amplification profiles and used in the segregation analysis of the progeny population.

Table 1.14. Details of the alleles generated with genic microsatellite/EST-SSR markers in genetic diversity analysis of popular clones

Sl. No.	Marker	No.of alleles identified	Polymorphism Information Content (PIC)
1	HBE 002	3	0.61
2	HBE 003	2	0.22
3	HBE 010	5	0.65
4	HBE 017	3	0.54
5	HBE 033	4	0.70
6	HBE 034	3	0.49
8	HBE 068	3	0.20
9	HBE 092	3	0.53
10	HBE 101	3	0.63
11	HBE 122	6	0.74
12	HBE 126	5	0.70
13	HBE 170	2	0.32
14	HBE 190	5	0.76
Total		51	7.79
Average		3.64	0.56

Table 1.15. Comparison of RAPD, AFLP and SSR markers used in genetic diversity analysis of popular clones

Sl.No.	Marker	PIC
1	RAPD	0.22
2	AFLP	0.17
3	Genomic SSR (gSSR)	0.59
4	Genic/ EST-SSR	0.56

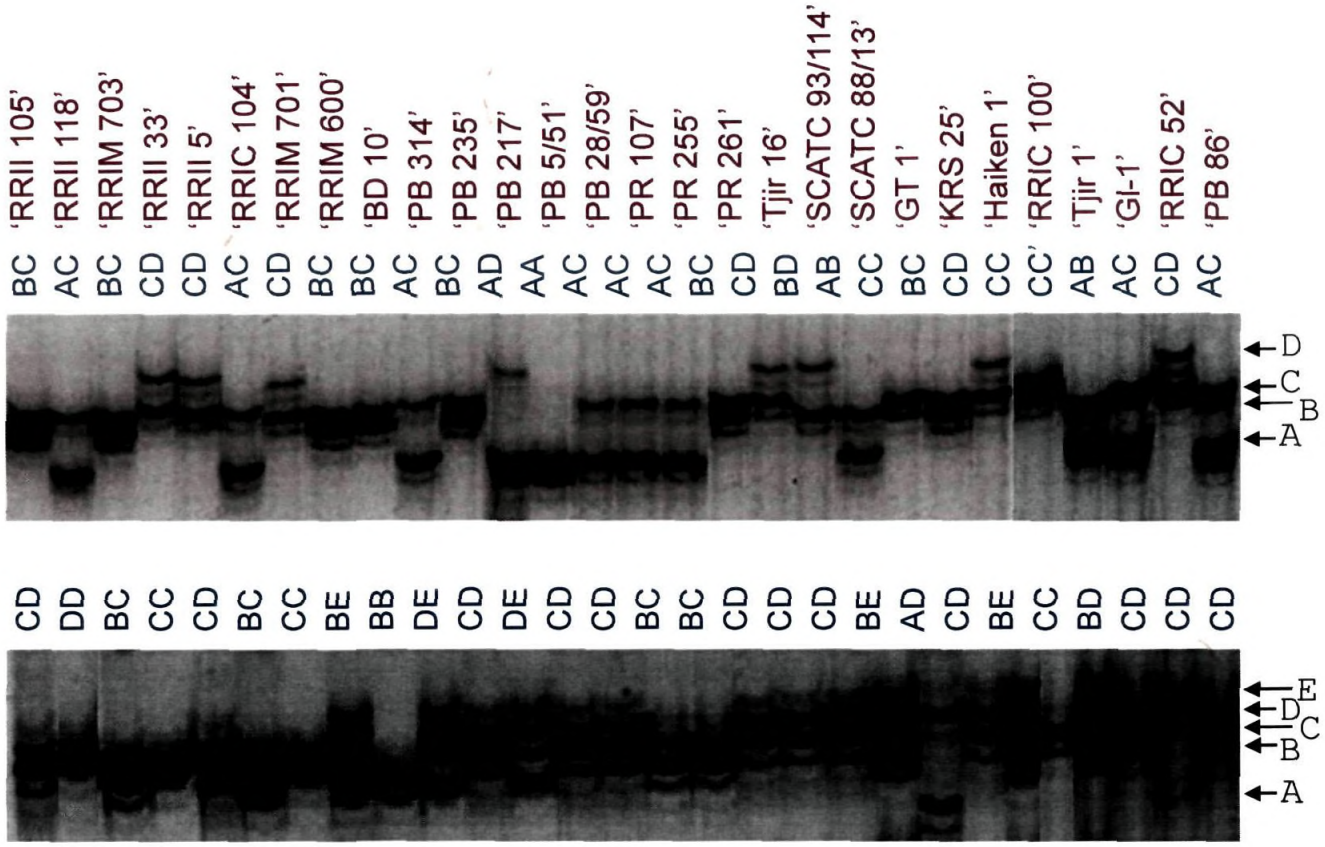


Fig. 1.12. Representative autoradiogram showing alleleic profiles of 28 *H. brasiliensis* clones generated with hmac-14 (upper) and hmct-16 (lower) primer pairs. Four microsatellite alleles, designated as A ,B,C and D (upper) were identified among *H. brasiliensis* clones at hmac-14 locus. Similarly, five alleles (A-E) were detected at the locus hmct -16 among these clones. Genotypes are indicated on each lane

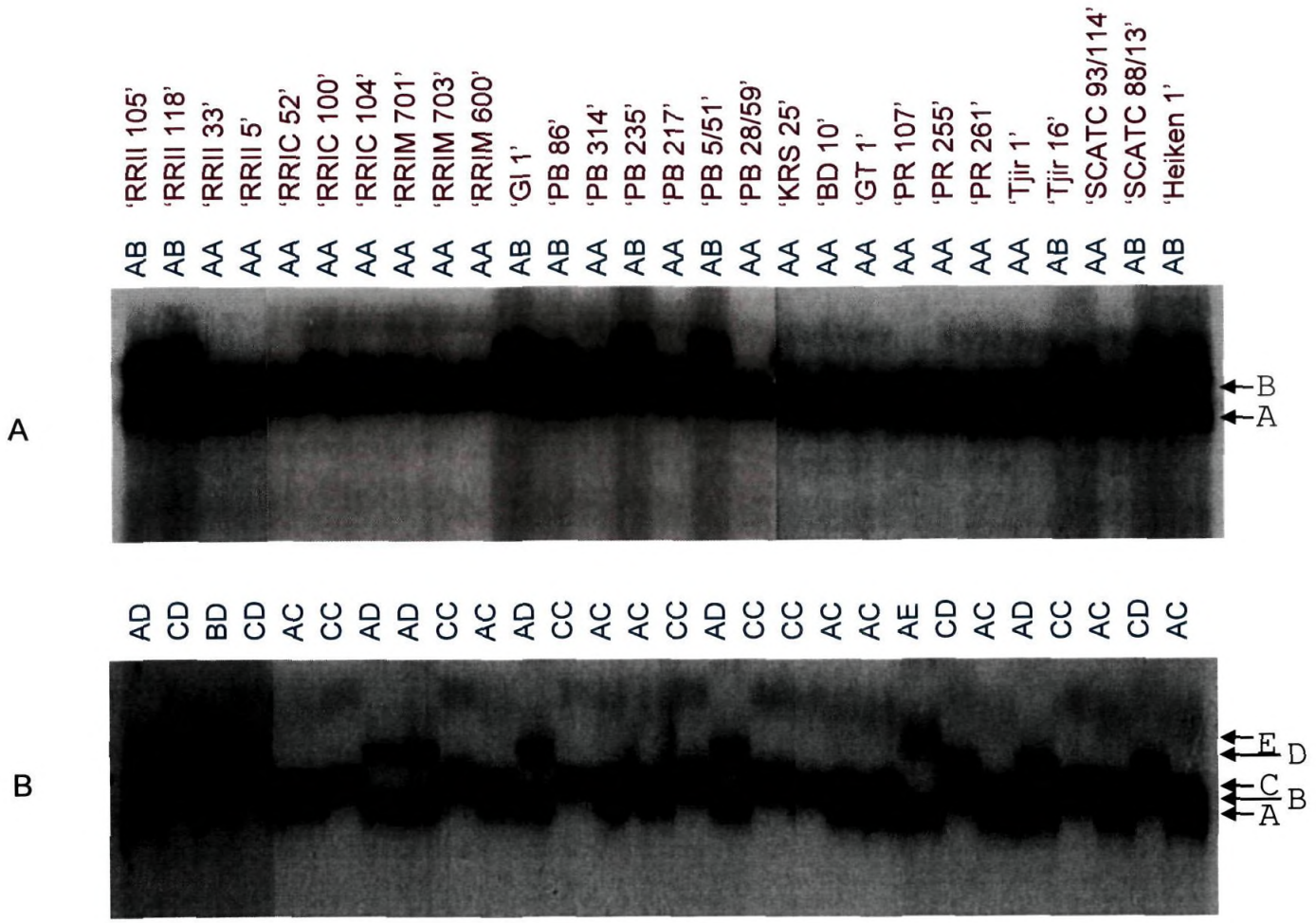


Fig. 1.13. Representative autoradiogram showing allelic profiles of 28 *H. brasiliensis* clones generated with HBE-003 (upper) and HBE-010 (lower) primer pairs. Two microsatellite alleles, designated as A and B (upper) were identified among *H. brasiliensis* clones at HBE-003 locus. Similarly, five alleles (A-E) were detected at the locus HBE-010 among these clones. Genotypes are indicated on each lane

1.4.2.3.4 Trinucleotide microsatellite marker generation

1.4.2.3.4.1 Development of an enriched trinucleotide repeat library

Among the restriction enzymes used, *Tsp* 509I was selected for the digestion of DNA for library construction which gave intense smearing in the region of selected size fraction (200-800 bp) The blot was hybridized with three different repeat motifs and it was observed that $(AAG)_n$ was the most abundant repeat present in *Hevea* genome with uniform hybridization signals. The PCR amplified enriched fragments were ligated on to Lambda ZAP II vector (Promega, USA) for the phage library and pGEM-T vector for the plasmid library. The average plaque forming units (pfu) for the phage library was found to be 2.8×10^6.

1.4.2.3.4.1.1 Screening of library

The primary screening of the lambda phage library (>4000 plaques) was performed. The positive plaques were detected through autoradiography and they were recovered individually. An average of 565 plaques were obtained and 32 plaques were found to be positive (Fig. 1.14) resulting in 5.6% enrichment for the trinucleotide repeat library in the primary screening. These positive plaques were subjected to second round of screening (Fig. 1.15). The recombinant lambda vector from the positive plaques was converted to pBluescript phagemids by *in vivo* excision.

Plasmid screening was done through a different technique as described in materials and methods. since the blotting of colonies directly for screening is not always reliable because there are chances of false positives. Even after thorough washing of blots the problem of adherence of debris resulting in false signals with radioisotope still remains. In brief, the recombinant colonies were streaked individually and PCR amplified using vector directed primers (T7 and SP6). Southern blotting was performed using standard procedures and hybridization was carried out with the DIG-labeled trinucleotide repeat probes (AAG, AAT and GTG). Positive colonies supposed to have the trinucleotide repeat were detected using the DIG DNA Labeling and Detection Kit (Roche) using the colourimetric (Fig.1.16) or CSPD substrate (Fig.1.17). Out of the 263 recombinant clones obtained, 63 were found to have trinucleotide repeat sequences, resulting from an enrichment process (23%). The schematic representation of screening of library is shown in Fig. 1.18.

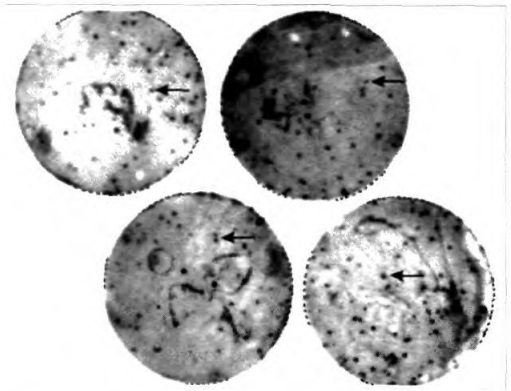


Fig. 1.14. Autoradiogram showing primary screening of plaques (genomic library in lambda phage vector). The black dots represent the positive plaques bearing genomic inserts with trinucleotide repeats, each plaque was isolated individually for second round of selection

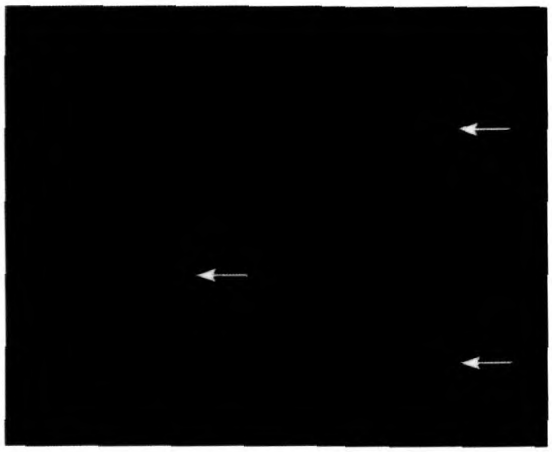


Fig. 1.15. Autoradiogram showing secondary screening of positive plaques detected after primary screening. Four plaques were screened in a single plate. Arrowheads show dark spots representing the positive plaques. The picture shows that out of the four plaques screened, three are showing positive signals

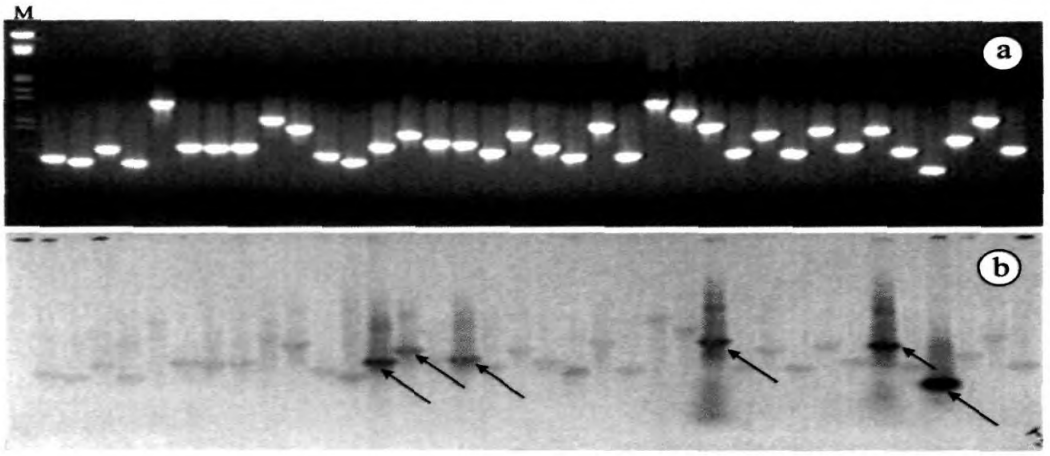


Fig. 1.16. Screening of positive clones bearing repeats from plasmid library using colourimetric method. a) Colony PCR of recombinant clones obtained from the plasmid library, b) Southern hybridization of the colony PCR products with trinucleotide repeat probes and DIG detection through the colourimetric method. Arrowheads show the putative positive colonies enriched with trinucleotide repeats. 'M' represents the lambda DNA marker

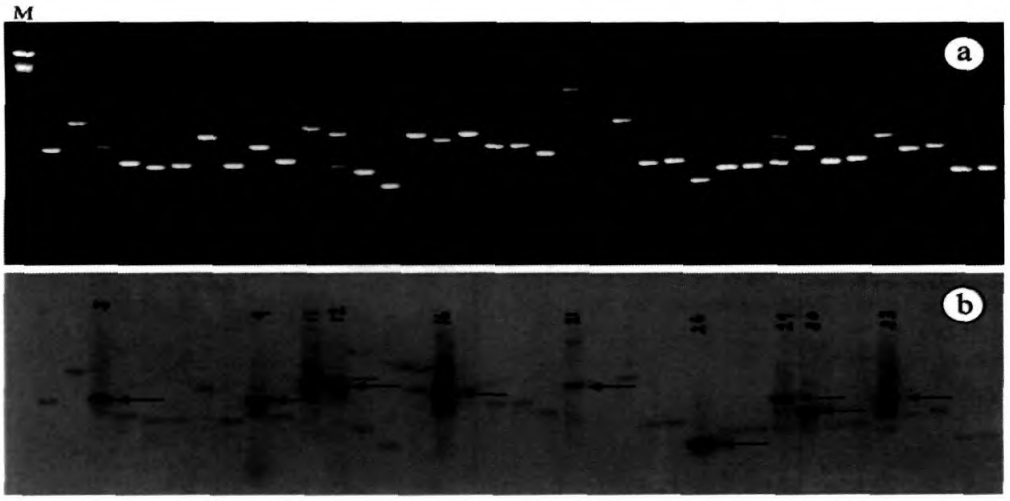


Fig. 1.17. Screening of positive clones bearing trinucleotide repeats from plasmid library using chemiluminiscent detection method. a) Colony PCR of recombinant colonies from the plasmid library, b) Southern hybridization of the colony PCR products with trinucleotide repeat probes and DIG detection through the colourimetric method. Arrowheads show the putative positive colonies enriched with trinucleotide repeats. 'M' represents the lambda DNA marker

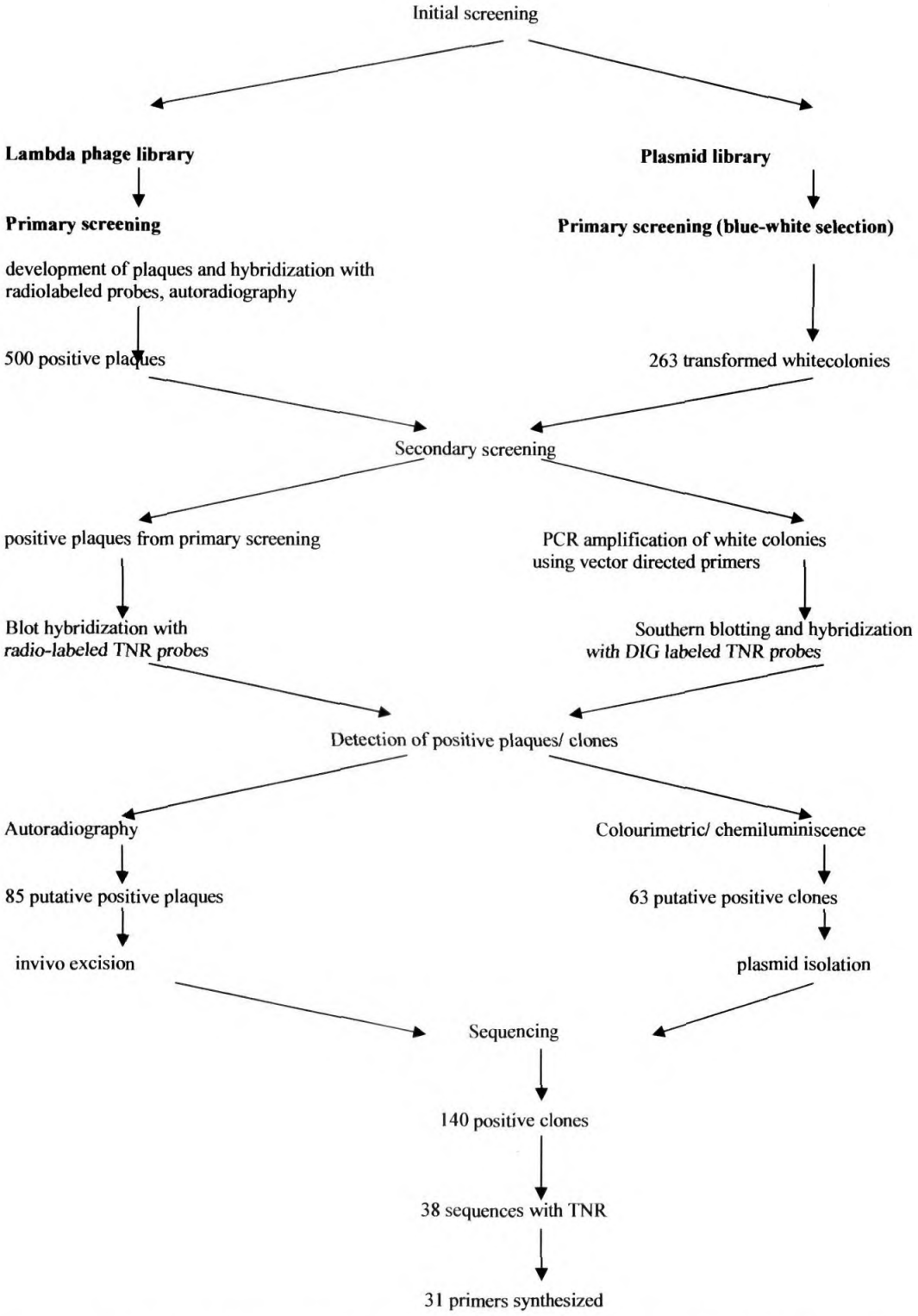


Fig.. 1.18. Schematic representation of screening of lambda phage and plasmid library for the development of trinucleotide repeat markers in rubber

Mass secondary screening of plaques

Another method to screen large number of plaques in a single plate following primary screening was also tried, where the LB plates were incubated with XL-1 Blue cells first and then individual clone in the form of phage was inoculated as droplets (1 µl). Around 30 plaques were screened on a single plate after blot hybridization with trinucleotide repeats following standard procedures (Fig. 1.19). Ten positive plaques were selected based on the hybridization signal and they were sequenced.

1.4.2.3.4.1.2 Sequencing of enriched clones

AAG/CTT repeats were found to be more frequent and longer in repeat length (Fig. 1.20) indicating abundance of the same in *Hevea* genome followed by TGG/ACC. Although long repeat motifs were absent in all the sequences, it was observed that all these sequences were enriched for the repeat motifs selected which were interrupted in between by other sequences. Association of trinucleotide repeats with dinucleotides were also detected (Fig. 1.20). The enrichment for the repeat type (AAT) was less.

1.4.2.3.4.1.3 Primer synthesis and testing for polymorphism

All the thirty-one primers synthesized were checked for amplification in two parental clones of the progeny population for developing markers and further profiling was done on PAGE to identify the informative primer (Fig. 1.21). Two markers, Htnr-90 and Htnr-21 were found polymorphic.

1.4.3 Genetic diversity analysis

1.4.3.1 RAPD markers

Genetic relationship was established by constructing a dendrogram through cluster analysis using the similarity matrix data derived through the application of 138 RAPD markers. The similarity coefficient ranged from 0.67 to 0.91. At a similarity coefficient of 0.75, four clusters were observed. The cluster I comprised of 25 genotypes and the other three clusters comprised of single genotypes each *ie.* PR 107, Tjir 16 and BD 10. Maximum diversity was observed for the clone PR 107 at a similarity coefficient of 0.67. In the cluster I, the Sri Lankan clones: RRIC 52, RRIC 100 and RRIC 104 were grouped together as a sub-cluster at a similarity coefficient of 0.83, which evolved from a common pedigree (Fig. 1.22). The robustness of this node was indicated by the bootstrap value of 55%. The clone RRII 105 was also seen clustered together with one of its

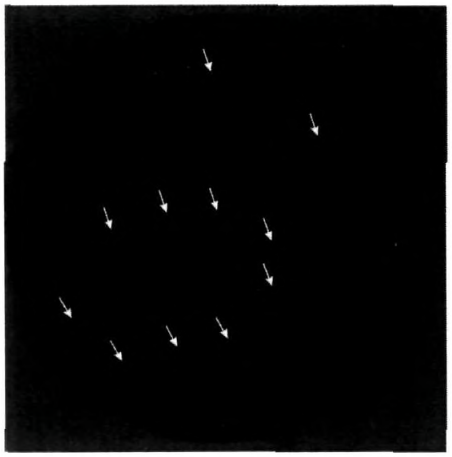


Fig. 1.19. Photograph depicting the mass screening of lambda phage library containing genomic trinucleotide repeats. The positive clones were selected (marked) based on the intensity of the hybridization signals on X-ray film

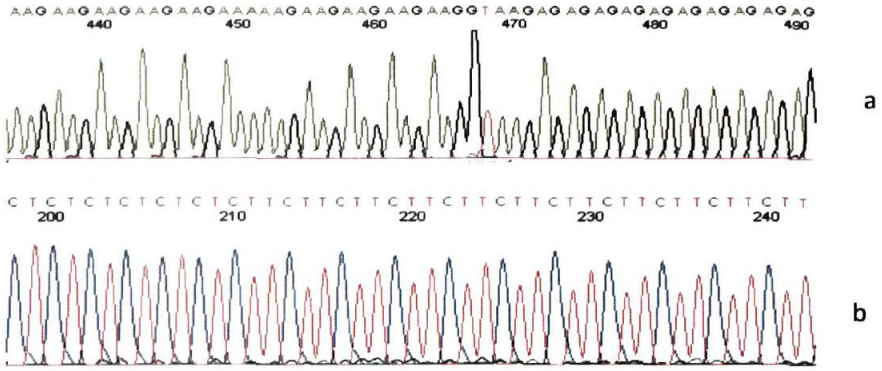


Fig. 1.20. Sequencing of trinucleotide repeats containing genomic clones identified from enriched library. a) Chromatogram showing the (AAG) repeat motifs in the cloned genomic fragments. b) Chromatogram of dinucleotide repeats (CT) associated with the trinucleotide repeat (CTT) motifs

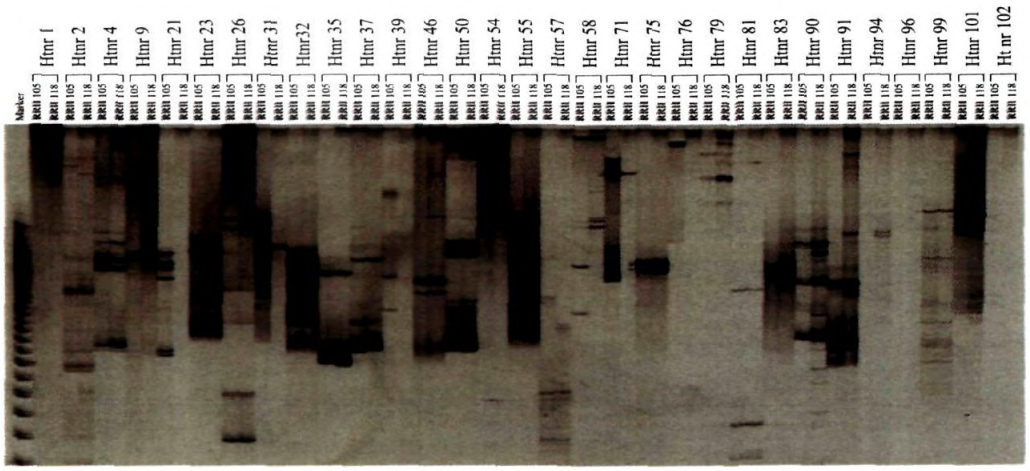


Fig. 1.21. Silver stained PAGE depicting the screening of trinucleotide microsatellite markers in parental clones RRII 105 and RRII 118. Microsatellite markers were developed through the screening of enriched genomic library for trinucleotide repeats

progenitor, Tjir 1 but with a low bootstrap value. The RAPD marker was not able to distinguish the clones, RRII 118 and PB 86. The bootstrapping of clusters indicated ≥ 40% only for two main cluster nodes and three subcluster nodes.

1.4.3.1.1 Principal Coordinates Analysis (PCoA)

The PCoA for RAPD markers further validated the results of the dendrogram. The three dimensional PCoA plot separated all the accessions into four clusters, with the clones Tjir 1 and PR 107 showing maximum diversity (Fig. 1.23). The majority of genotypes were grouped in a single cluster which was similar to that of UPGMA cluster analysis.

1.4.3.2 AFLP markers

The similarity coefficient of the dendrogram generated through cluster analysis using the similarity matrix of 435 AFLP marker data ranged from 0.76 to 0.92 (Fig. 1.24). Three clusters were observed at a similarity coefficient of 0.8. The clone Tjir 16 showed maximum dissimilarity. Cluster I contained almost all the genotypes as in the case of RAPD marker data, even though the genotypes under same group differed from that of RAPD based clusters. The Sri Lankan clones were grouped together which was in conformity with its pedigree relationships. The bootstrap value of the main node (55%) revealed the robustness of this particular grouping. The clone RRII 105 grouped along with one of its maternal parent Gl1 and the grouping was also robust with a bootstrap value of 46%. The clone RRII 118 was seen in close association with RRII 33. AFLP marker could not distinguish the clones, Gl 1 and PB 5/51 and also the clones PB 86 and KRS 25. The bootstrapping of clusters indicated ≥ 40% for four main nodes and seven sub nodes.

1.4.3.2.1 PCo analysis

The three dimensional diagram of principal coordinate analysis revealed the same grouping pattern noticed as in the UPGMA cluster diagram (Fig. 1.25). The clone Tjir 16 was separated while all other clones were clustered together. The Sri Lankan clones were also grouped together as in the case of cluster diagram.

1.4.3.3 SSR markers (genomic and genic)

The simple matching coefficient for the 17 genomic SSR markers ranged from 0.57 to 1, while that of EST-SSR markers ranged from 0.6 to 0.93. The clone PR 107 had the

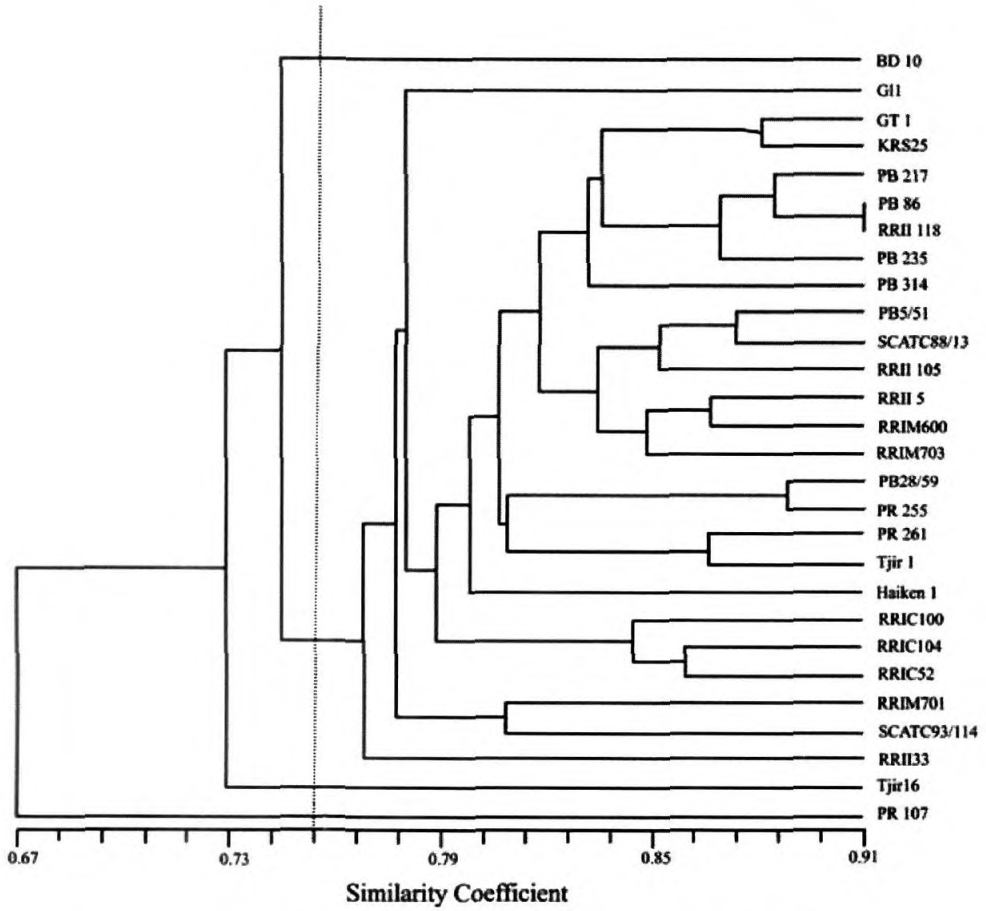


Fig. 1.22. Dendrogram showing the cluster analysis of the popular clones using RAPD markers. The dashed line represents the four main clusters obtained at the similarity coefficient of 0.75

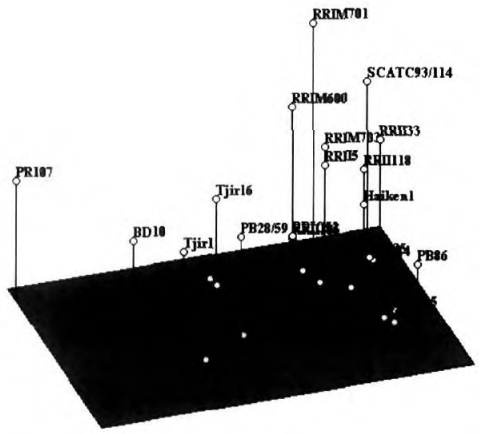


Fig. 1.23. Principal coordinates analysis (PCoA) of the popular clones using RAPD markers

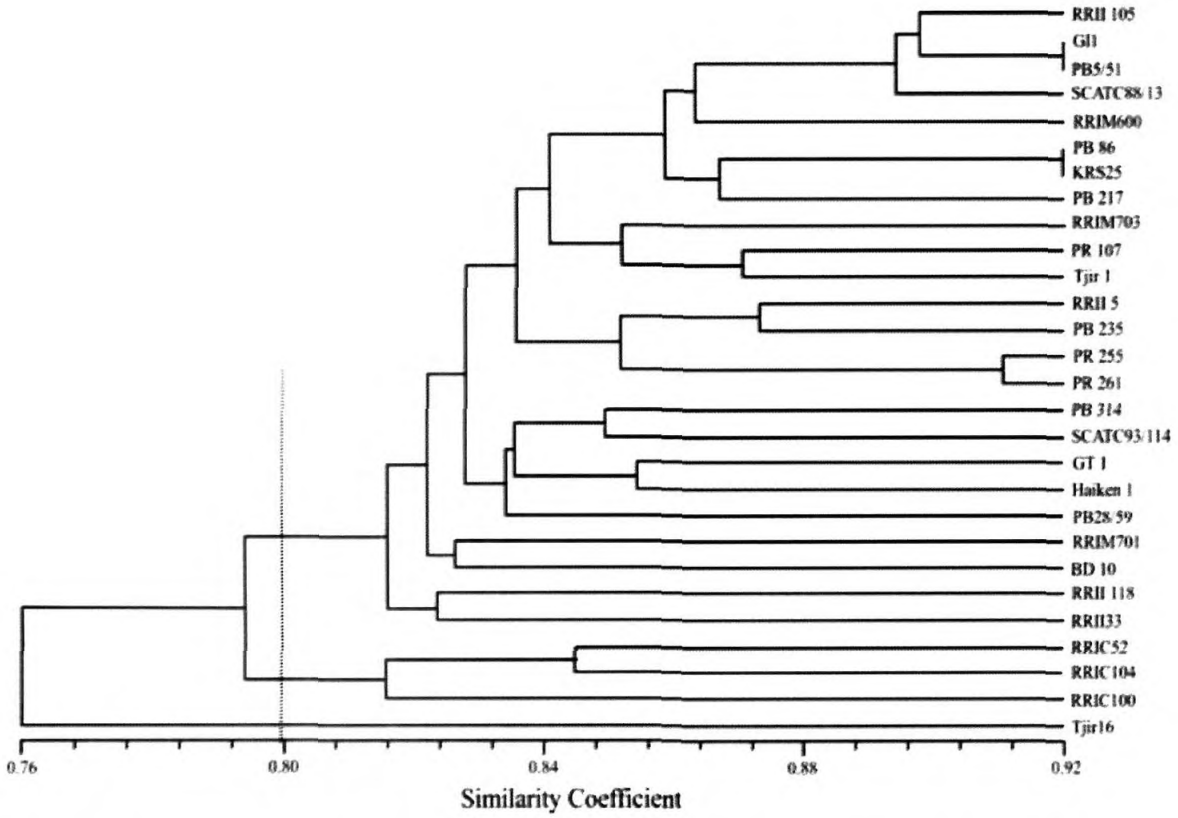


Fig. 1.24. Dendrogram showing the cluster analysis of the popular clones using AFLP markers. The dashed line represents the three main clusters obtained at the similarity coefficient of 0.8

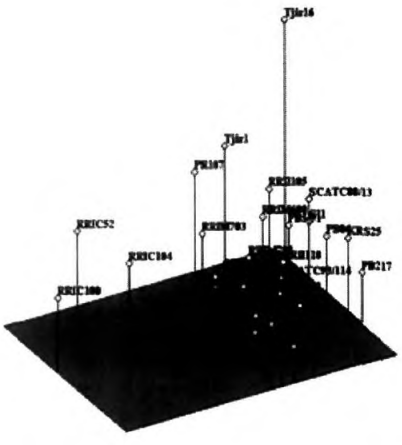


Fig. 1.25. PCo analysis of the popular clones using AFLP markers

maximum diversity and the genotypes were divided in to three clusters with the cluster 1 containing almost all the genotypes. In the second cluster the Sri Lankan clones were grouped together showing the pedigree relationships. This node was robust with the bootstrap value of 47%, for the cluster RRIC 100 and RRIC 52. Six sub clusters have the bootstrap values greater than 40%. EST-SSR markers were not able to differentiate the clones SCATC 93/114 and PB 314 (Fig. 1.26).

1.4.3.3.1 PCo analysis

PCo analysis was performed for the genomic and EST-SSR markers revealing the same clustering pattern as observed in the UPGMA cluster analysis. The Sri Lankan clones were grouped together into a separate cluster while all the other clones were clustered together with out any specific grouping pattern (Fig. 1.27).

1.4.3.4 Combining all RAPD, AFLP and SSR markers

A total of 698 bands/ alleles from the RAPD, AFLP and SSR markers were used for genetic diversity analysis. The similarity coefficient resulting from the analysis based on simple matching coefficient varied from 0.74 to 0.9. No specific grouping was observed for the clones of different origin using these markers except the Sri Lankan clones. Sri Lankan clones were seen clustered together as in the case of other individual marker analysis (Fig. 1.28) using all marker combinations. Clones PB 86 and KRS 25 could not be differentiated. The clone RRII 105 was found closely associated with one of its parent Gl 1.

1.4.3.4.1 PCo analysis

Principal coordinate analysis was performed for all the markers together. All the genotypes were grouped together with the same clustering pattern as exhibited by the cluster diagram. The Sri Lankan clones were grouped together (Fig. 1.29).

The comparison of various markers for the genetic diversity analysis in the present study indicates that SSR markers are the best markers among others indicating a high PIC value (Table. 1.15). Genomic SSR markers showed slightly higher PIC value than the genic SSR markers.

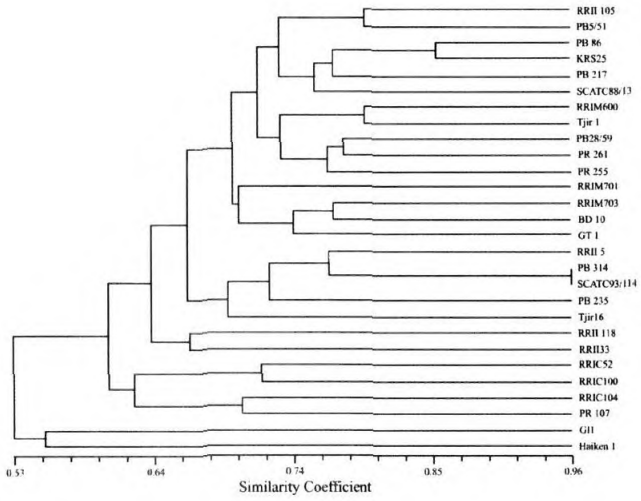


Fig. 1.26. Dendrogram showing genetic relationships of the popular clones using genomic and EST-SSR markers

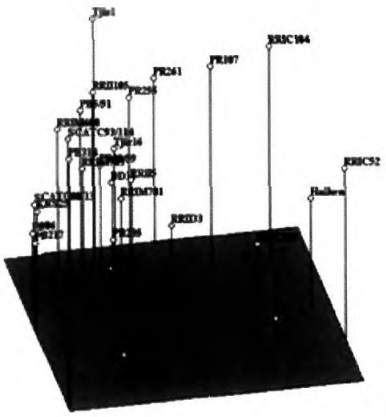


Fig. 1.27. PCo analysis of the popular clones using genomic and EST-SSR markers

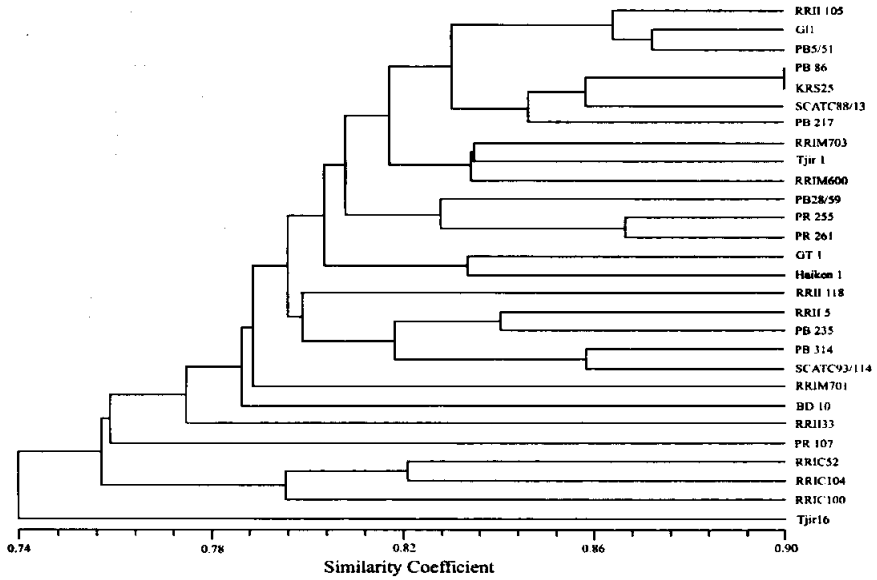


Fig. 1.28. Dendrogram showing genetic relationships among popular clones using RAPD, AFLP, genomic and SSR-SSR markers

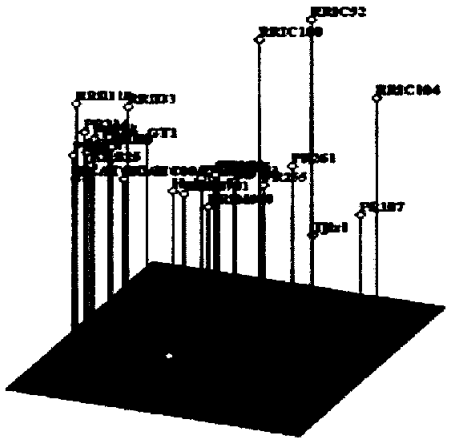


Fig. 1.29. PCo analysis of the popular clones using RAPD, AFLP, genomic and EST-SSR markers

1.4.3.4.2 Mantel's test

Mnatel's test statistic (Z) (Mantel, 1967) is a test for matrix correspondence and it reveals the similarity between two matrices. The mantel test was performed between AFLP and SSR markers, AFLP and RAPD markers, RAPD and SSR markers and genomic and genic SSR markers. The 'r' value showing the matrix correlation between RAPD and SSR markers was r = 0.22709 (p = 0.9828) while 'r' values were very low showing an insignificant correlation between the two matrices for other two comparisons. The corresponding 'p' values was also low (p = 0.37) for matrix correlation between AFLP and RAPD markers and p = 0.158 for matrix correlation between AFLP and SSR markers). The correlation coefficient between the genic and genomic SSR markers was found to be 0.4, but the correlation was insignificant with a 'p' value of 1.

1.4.2.4 Single Nucleotide Polymorphism (SNP)

1.4.2.4.1 Identifying SNPs through sequencing

With the aim of developing SNP markers for *Hevea* genome mapping, partial DNA sequence of the genes/cDNA responsible for complex biochemical traits was PCR amplified using 32 synthesized primers for 12 loci in 16 *Hevea brasiliensis* clones/genotypes as mentioned in the materials and methods. Initially the amplifications were tested in eight selected genotypes mainly the primary clones and the most popular *Hevea* clone RRII 105 (Fig. 1.30). Twelve loci were amplified in each of the 16 genotypes (12 loci X 16 genotypes = 192 amplicons) (Table 1.16.). From the amplification profile of each fragment it was clear that for some of the genes, the amplicon size was more than the expected size due to the presence of introns as the template was genomic DNA. Amplification product of the locus 'ubiquitin precursor' gave more than one band on a gel and the lower strong band was eluted for sequencing.

1.4.2.4.2 Sequence analysis

PCR fragments amplified in each primer-genotype combinations were sequenced from both the directions. SNPs were detected only in five out of 12 loci sequenced from all the 16 genotypes. These five loci were 1) geranylgeranyl diphosphate synthase (*GGDPS*), 2) farnesyl diphosphate synthase (*FDPS*), 3) mevalonate kinase (*MK*), 4) ubiquitin precursor (*UBQ*) and 5) latex patatin homolog (*LPH*). Initially, nucleotide variations were identified using the phred/phrap suite and consed, and later by aligning the sequences

Table 1.16. Details of the loci/genes used in the SNP studies in *Hevea*

Gene designation	GenBank Accn. No.	Region of interest	Expected amplicon size	Actual amplicon size	Coding region	Non-coding regions		No. of SNPs identified	Frequency of SNP
						Introns	UTR		
Geranylgeranyl diphosphate synthase	AB 055496	3'-UTR	500 bp	500 bp	341 bp	-	159 bp	5	1/100 bases
Farnesyl diphosphate synthase	AY 135188	3'-UTR	500bp	1500 bp	335 bp	1045 bp	120 bp	14	1/ 107 bases
Mevalonate kinase	AF 429384	3'-UTR	500 bp	831 bp	256 bp	389 bp	186 bp	13	1/ 64 bases
Ubiquitin precursor	AF 193438	3'-UTR	500 bp	494 bp	286 bp	-	208 bp	7	1/ 71 bases
Latex patatin homolog	U 80598	3'-UTR	500bp	603 bp	284 bp	103 bp	216 bp	6	1 / 100 bases
				3928 bp	1502 bp	1537 bp	889 bp	45	

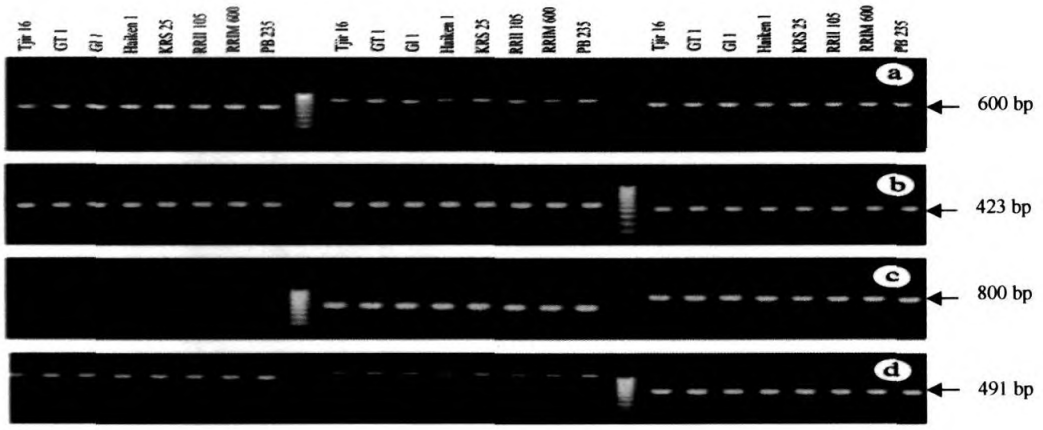


Fig. 1.30. Agarose gel photograph showing the amplification of 12 loci from eight *H. brasiliensis* clones/ genotypes selected for SNP studies a) geranylgeranyl diphosphate synthase, ubiquitin precursor, latex patatin homolog, b) transcript associated with TPD, latex abundant protein, rubber elongation factor, c) farnesyl diphosphate synthase, latex plastidic aldolase, glutathione peroxidase and d) mevalonate kinase, hydroxy methyl glutaryl-CoA synthase, hydroxy methyl glutaryl-CoA reductase

using the software 'DNASIS Max'. If both alleles of an SNP were found at homozygous state in a population the SNP could be spotted very effectively. If only one of the alleles is present in homozygous state, heterozygous samples may be distinguished from the homozygous samples before an SNP can be identified (Fig. 1.31).

1.4.2.4.3 SNP detection in popular *Hevea* clones

Sequencing of five loci components of 3928 bp from each rubber clone resulted in (3928 x 16 = 62848 bp) sequence data. Sequencing for each rubber clone include 1502 bases of coding regions and 2426 bases of non-coding regions containing six introns. A total of 45 SNPs were detected from the gene sequences and out of these, 25 SNPs (55.5%) represented transitions while 14 (28.8 %) represented transversions and the rest were six indels. The details of SNPs identified are given in Table 1.17. Among the identified SNPs, 29 SNPs were present in the intronic sequences and UTR regions, while only 16 SNPs were detected in the coding sequences. SNPs detected in these five genes namely geranylgeranyl diphosphate synthase, farnesyl diphosphate synthase, mevalonate kinase, ubiquitin precursor and latex patatin homolog are described below:

1.4.2.4.3.1 SNPs in geranyl geranyl diphosphate synthase

The expected amplicon size was 500 bp based on the primer site designed on the candidate gene fragment. Actual PCR product was the same as expected. In 500 bp span of the sequence derived from 3'-UTR, five SNPs were identified in the sequences derived from the 16 popular *H. brasiliensis* clones. Four of them were in the coding region and one SNP was detected in the UTR region. The frequency of SNP for this sequence was calculated as 1/100 bp.

1.4.2.4.3.2 SNPs in farnesyl diphosphate synthase

Actual size of the amplicon (1.5 kb) was three times more than the expected size (500 bp) indicating the presence of intron in genomic sequences. The amplicons were sequenced from both directions to get full-length sequence of the same. There were four long intronic regions of size 1165 bp in between the coding regions, and the coding region was found to be only 335 bp (Table 16). Fourteen SNPs were detected in the resulting sequence. Out of these, ten SNPs were in intronic region, three were in the coding region and one SNP was present in the UTR region. The frequency of SNP in this sequence was calculated as 1/107 bp.

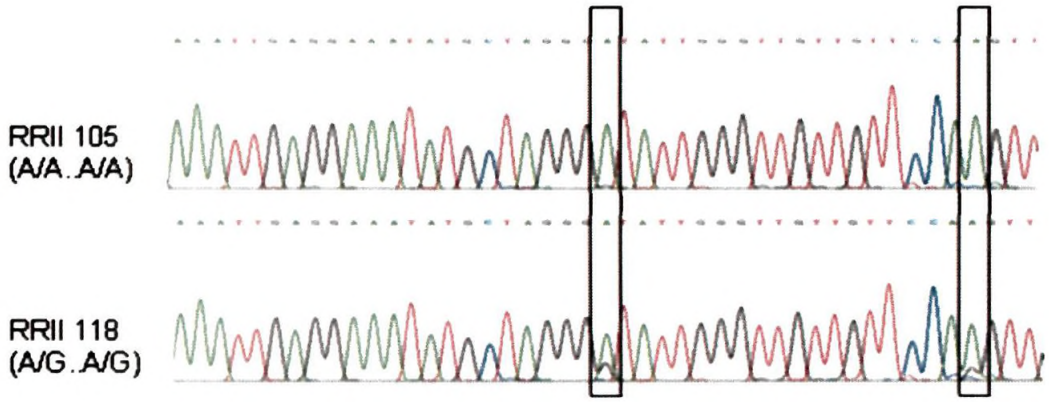


Fig. 1.31. Representative aligned chromatograms showing nucleotide variation in geranylgeranyl diphosphate synthase gene in RRII 105 and RRII 118. Two SNPs were identified at 88 and 107 nucleotide positions. RRII 105 is homozygous and RRII 118 is heterozygous for these SNPs

Table 1.17. SNPs identified in the popular clones of *Hevea brasiliensis*

Geranyl geranyl diphosphate synthase													
Polymorphic base position													
88			107			164				240			400
A/G			A/G			A/G				A/C			T/A
Farnesyl diphosphate synthase													
Polymorphic base position													
195	239	260	296	346	472	510	525	730	916	1155	1301	1310	1426
C/T	G/T	A/T	A/T	A/T	C/T	G/A	T/G	T/C	A/T	G/A	G/A	**G/A**	G/A
Mevalonate kinase													
Polymorphic base position													
118	197	275	332	431	432	433	437	438	455	474	475	476	
A/T	C/T	C/T	A/G	T/-	T/-	T/A	T/-	T/-	A/G	A/-	T/-	C/T	
Ubiquitin precursor													
Polymorphic base position													
112		142	205		290		325		389			450	
G/A		C/T	G/A		T/G		C/G		A/C			A/G	
Latex patatin homolog													
Polymorphic base position													
		244	245		288		313			334		415	
		G/A	A/G		C/T		C/T			G/A		A/C	

1.4.2.4.3.3 SNPs in mevalonate kinase

Actual size of the amplicon was 830 bp whereas the expected one was 499 bp. Therefore existence of an intron having more than 300 bases was evident. Thirteen SNPs were identified within the sequence. (Table 1.16). Out of these only one was in the coding region, which is only 256 bp long, while all others were in the intronic regions. Six SNPs were identified as indels. The frequency of SNP was calculated as 1/64 bp.

1.4.2.4.3.4 SNPs in ubiquitin precursor

The expected amplicon size was 500 bp based on the primer site designed on the candidate gene fragment. Actual PCR product was the same as expected. There were no intronic sequences detected in the amplified product and a UTR region of 208 bp was observed (Table 1.16). Seven SNPs could be identified along the sequences. Three SNPs were detected in the coding region and four were present in the UTR region. The frequency of SNP for this sequence was calculated as 1/71 bp.

1.4.2.4.3.5 SNPs in latex patatin homolog

The expected amplicon size was 500 bp based on the primer site designed on the candidate gene fragment but the amplicon had 103 bp intronic regions resulting in a 603 bp fragment (Table 1.16). Six SNPs were detected, five were in the coding region and one SNP was in the UTR region. The frequency of SNP in this sequence was found to be 1/ 83 bp.

1.4.2.4.4 Identification of SNPs in the parents of the progeny population

The above mentioned five genes /loci were amplified for SNP identification in the parents (RRII 105 and RRII 118) of the progeny for their subsequent use in segregation analysis in the progeny population. Nucleotide variations were identified by aligning the sequences of each gene separately using the software 'DNASIS Max' for the parental clones. Number of SNPs identified was less in the parents with 30 SNPs as base substitution frequency increases with increase in number of genotypes under study. Chromatograms were visually examined to detect the heterozygous SNPs which could not be detected by the software.

1.4.2.4.4.1 SNPs in geranylgeranyl diphosphate synthase

Four SNPs were identified in the sequences derived from parental clones RRII 105 and RRII 118 for the gene geranyl geranyl diphosphate synthase. Three of them were in the coding region and one SNP was detected in the UTR region. All the SNPs were in heterozygous condition in RRII 118, while all were in homozygous condition in RRII 105 (Table 1.18). The sequence details are shown in Fig. 1.32.

1.4.2.4.4.2 SNPs in farnesyl diphosphate synthase

Nine SNPs were detected in the resulting sequence with the parents. Out of these, eight SNPs were in intronic region and one SNP was present in the UTR region. Four SNPs were found in heterozygous state in RRII 18, whereas all existed in homozygous state in RRII 105. (Table 1.18). The sequence details are given in Fig. 1.33.

1.4.2.4.4.3 SNPs in mevalonate kinase

Eleven SNPs including six indels were identified within the sequences of the parental population. Out of these only one was in the coding region and all others were in the intronic regions. All SNPs including indels, were in homozygous state in RRII 118 and heterozygous in RRII 105 (Table 18).The sequence details are given in Fig. 1.34.

1.4.2.4.4.4 SNPs in latex patatin homolog

Four SNPs were detected, of which three of them were in the coding region and one SNP was detected in the UTR region. All the SNPs were in heterozygous state in the clone RRII 105, whereas all SNPs exhibited homozygosity in the clone RRII 118 (Table 1.18).

1.4.2.4.4.5 SNPs in ubiquitin precursor

Two SNPs were identified within the parental sequences and both were homozygous for the concerned allele in both the parents. In the clone RRII 105 a 45 bp deletion was also onserved (Fig. 1.35) (Table 1.18).

1.4.2.4.5 Haplotype construction

Haplotype (haploid genotype) is a particular pattern of sequential SNPs (or alleles) found on a single chromosome and these SNPs tend to be inherited together as a block. The haplotypes were detected for the parental clones RRII 105 and RRII 118. Cloning and sequencing of amplicons were performed as described earlier for haplotype detection.

Table 1.18. SNPs identified in the parental clones RRII 105 and RRII 118

	Geranylgeranyl diphosphate synthase			
	Polymorphic base position			
Genotype	88	107	164	400
RRII 105	A	A	A	T
RRII 118	A/G	A/G	A/G	T/A

	Farnesyl diphosphate synthase								
	Polymorphic base position								
Genotype	260	296	346	730	916	1155	1301	1310	1426
RRII 105	T	A	A	T	A	G	G	A	G
RRII 118	A	A/T	A/T	C	T	A	G/A	G	G/A

	Mevalonate kinase										
	Polymorphic base position										
Genotype	118	197	275	332	432	433	437	438	474	475	476
RRII 105	T/A	C/T	C/T	A/G	T/-	A/-	T/-	T/-	A/-	T/-	C/T
RRII 118	T	C	T	A	-	-	T	T	A	T	C

	Latex patatin homolog				Ubiquitin precursor	
	Polymorphic base position				Polymorphic base position	
Genotype	288	313	334	415	112	205
RRII 105	T/C	C/T	G/A	A/C	G	G
RRII 118	C	C	G	A	A	A

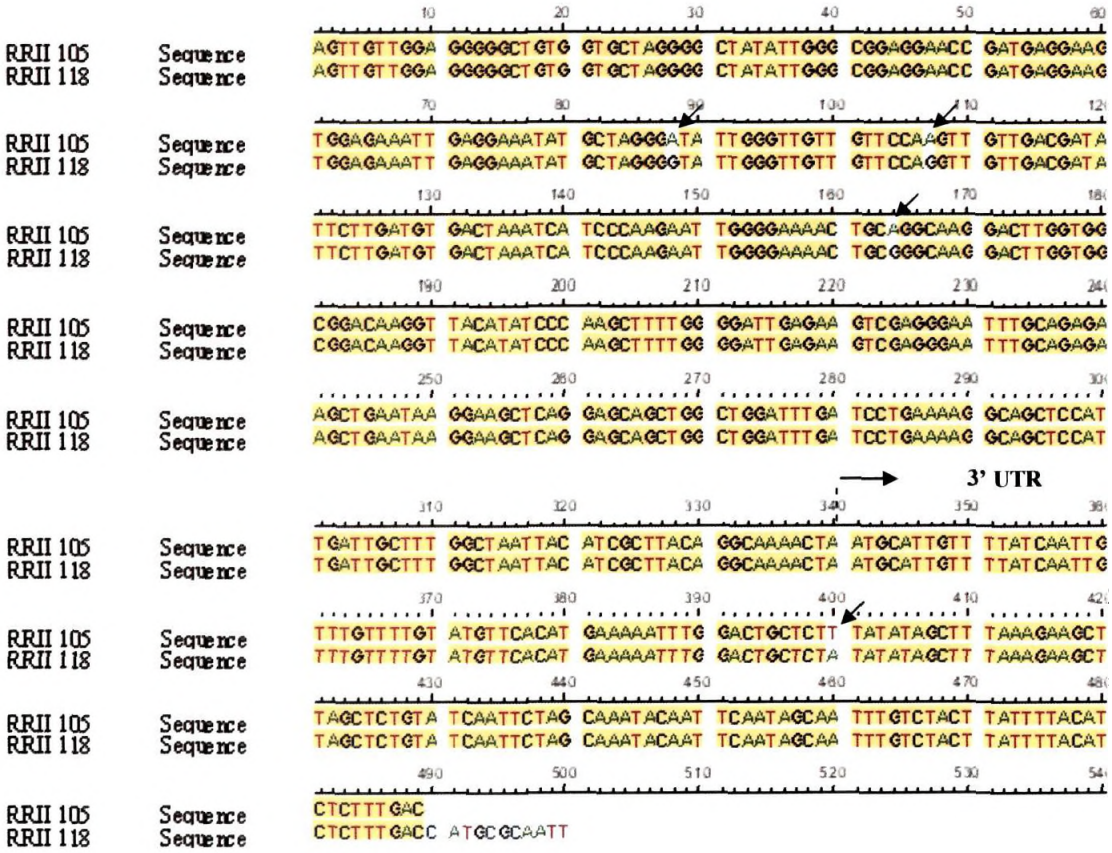
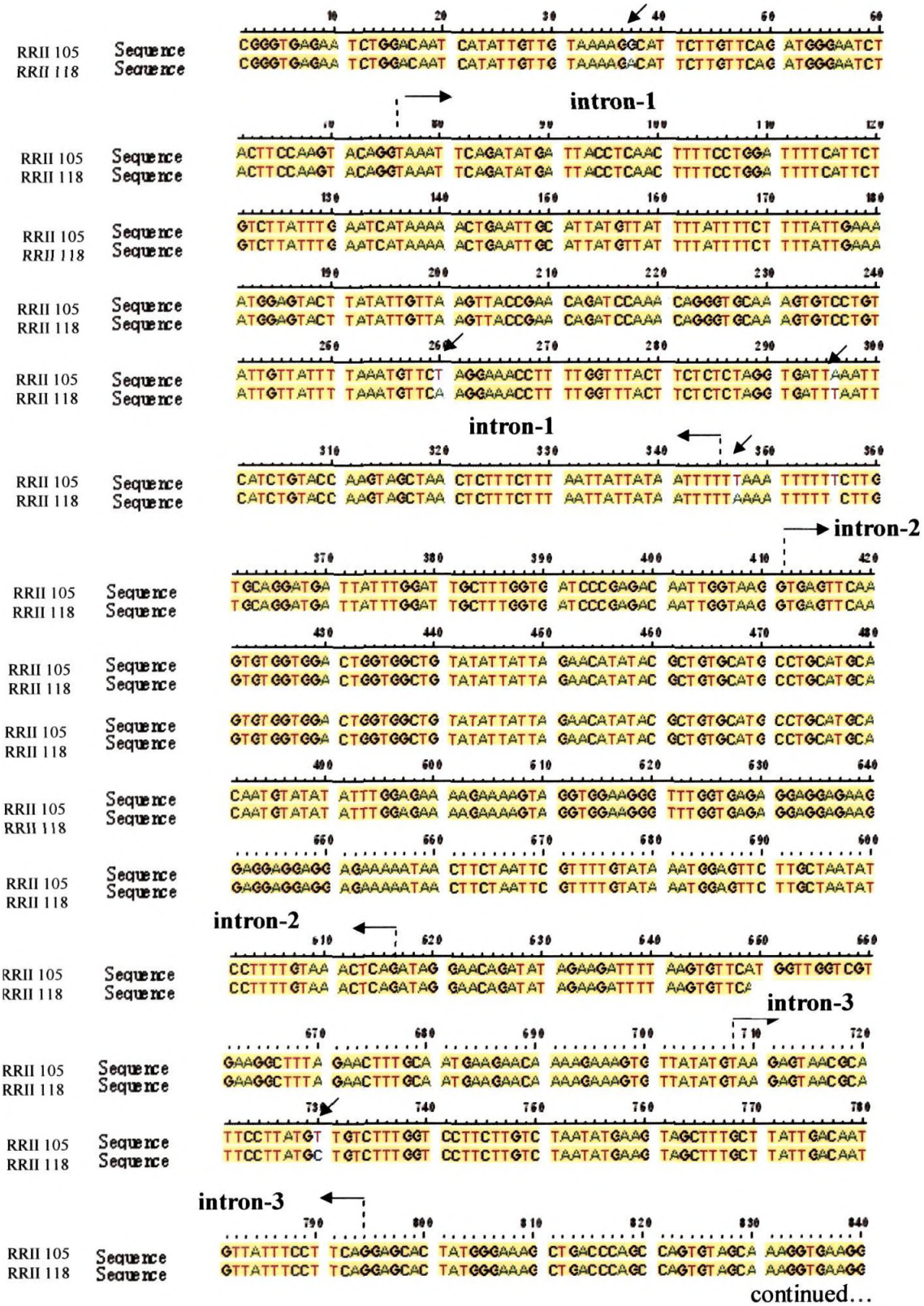


Fig. 1.32. Sequence alignment of geranylgeranyl diphosphate synthase gene from the clo
105 and RRII 118 showing SNP positions (marked) along with the coding sequence and
regions

.


RRII 105 Sequence
RRII 118 Sequence
CGGGTGAGAA TCTGGACAAT CATATTGTTG TAAAAGGCAT TCTTGTTCAG ATGGGAATCT
CGGGTGAGAA TCTGGACAAT CATATTGTTG TAAAAGACAT TCTTGTTCAG ATGGGAATCT
intron-1
ACTTCCAAGT ACAGGTAAAT TCAGATATGA TTACCTCAAC TTTTCCTGGA TTTTCATTCT
ACTTCCAAGT ACAGGTAAAT TCAGATATGA TTACCTCAAC TTTTCCTGGA TTTTCATTCT
GTCTTATTTG AATCATAAAA ACTGAATTGC ATTATGTTAT TTTATTTTCT TTTATTGAAA
GTCTTATTTG AATCATAAAA ACTGAATTGC ATTATGTTAT TTTATTTTCT TTTATTGAAA
ATGGAGTACT TATATTGTTA AGTTACCGAA CAGATCCAAA CAGGGTGCAA AGTGTCCTGT
ATGGAGTACT TATATTGTTA AGTTACCGAA CAGATCCAAA CAGGGTGCAA AGTGTCCTGT
ATTGTTATTT TAAATGTTCT AGGAAACCTT TTGGTTTACT TCTCTCTAGG TGATTAAATT
ATTGTTATTT TAAATGTTCA AGGAAACCTT TTGGTTTACT TCTCTCTAGG TGATTTAATT
intron-1
CATCTGTACC AAGTAGCTAA CTCTTTCTTT AATTATTATA ATTTTTTAAA TTTTTTCTTG
CATCTGTACC AAGTAGCTAA CTCTTTCTTT AATTATTATA ATTTTTAAAA TTTTT CTTG
intron-2
TGCAGGATGA TTATTTGGAT TGCTTTGGTG ATCCCGAGAC AATTGGTAAG GTGAGTTCAA
TGCAGGATGA TTATTTGGAT TGCTTTGGTG ATCCCGAGAC AATTGGTAAG GTGAGTTCAA
GTGTGGTGGA CTGGTGGCTG TATATTATTA GAACATATAC GCTGTGCATG CCTGCATGCA
GTGTGGTGGA CTGGTGGCTG TATATTATTA GAACATATAC GCTGTGCATG CCTGCATGCA
GTGTGGTGGA CTGGTGGCTG TATATTATTA GAACATATAC GCTGTGCATG CCTGCATGCA
GTGTGGTGGA CTGGTGGCTG TATATTATTA GAACATATAC GCTGTGCATG CCTGCATGCA
CAATGTATAT ATTTGGAGAA AAGAAAAGTA GGTGGAAGGG TTTGGTGAGA GGAGGAGAAG
CAATGTATAT ATTTGGAGAA AAGAAAAGTA GGTGGAAGGG TTTGGTGAGA GGAGGAGAAG
GAGGAGGAGG AGAAAAATAA CTTCTAATTC GTTTTGTATA AATGGAGTTC TTGCTAATAT
GAGGAGGAGG AGAAAAATAA CTTCTAATTC GTTTTGTATA AATGGAGTTC TTGCTAATAT
intron-2
CCTTTTGTAA ACTCAGATAG GAACAGATAT AGAAGATTTT AAGTGTTCAT GGTTGGTCGT
CCTTTTGTAA ACTCAGATAG GAACAGATAT AGAAGATTTT AAGTGTTCA
intron-3
GAAGGCTTTA GAACTTTGCA ATGAAGAACA AAAGAAAGTG TTATATGTAA GAGTAACGCA
GAAGGCTTTA GAACTTTGCA ATGAAGAACA AAAGAAAGTG TTATATGTAA GAGTAACGCA
TTCCTTATGT TGTCTTTGGT CCTTCTTGTC TAATATGAAG TAGCTTTGCT TATTGACAAT
TTCCTTATGC TGTCTTTGGT CCTTCTTGTC TAATATGAAG TAGCTTTGCT TATTGACAAT
intron-3
GTTATTTCCT TCAGGAGCAC TATGGGAAAG CTGACCCAGC CAGTGTAGCA AAGGTGAAGG
GTTATTTCCT TCAGGAGCAC TATGGGAAAG CTGACCCAGC CAGTGTAGCA AAGGTGAAGG

continued…

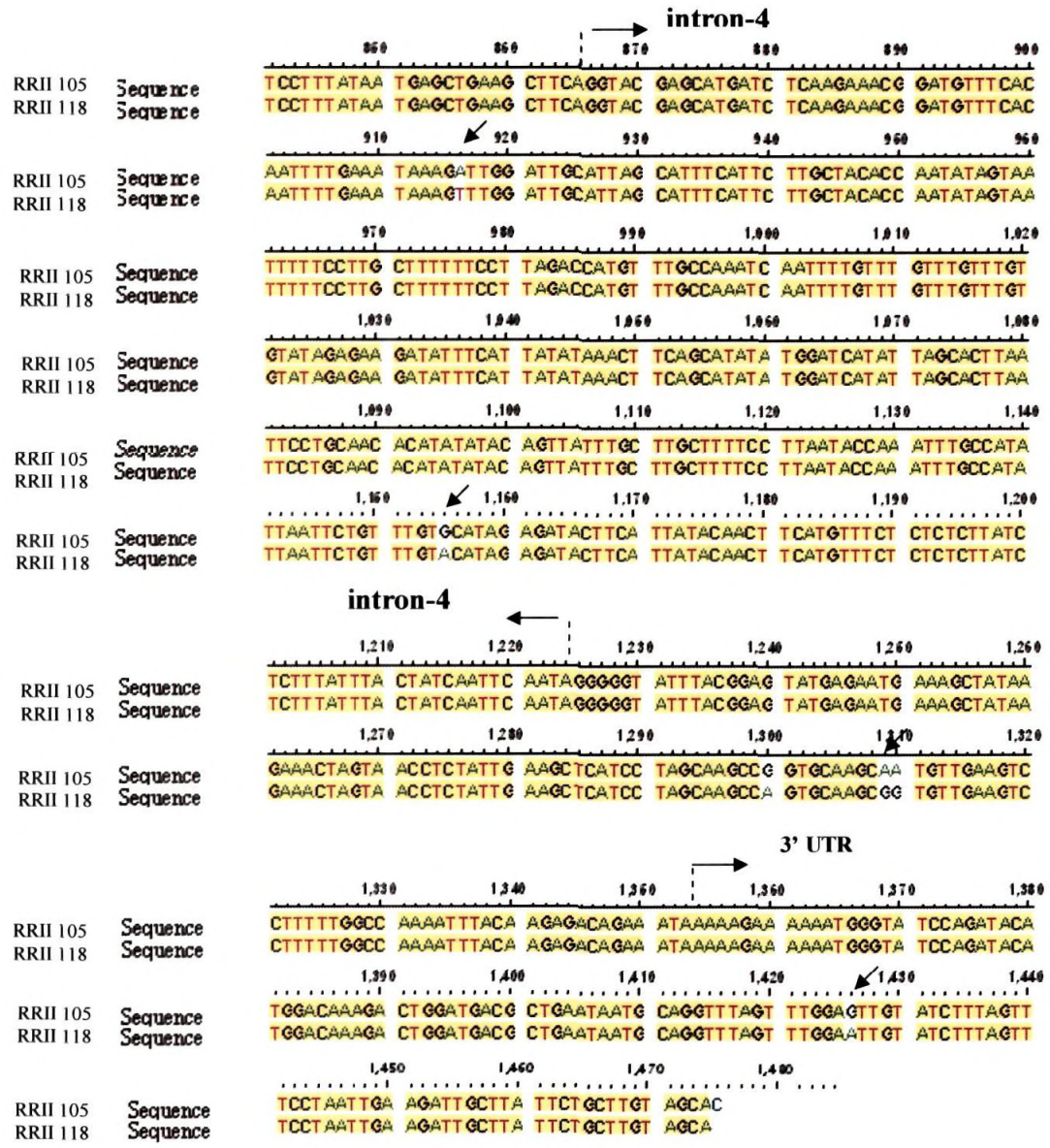


Fig. 1.33. Sequence alignment of farnesyl diphosphate synthase gene from the clone RRII 105 and RRII 118 showing SNP positions (marked) along with the partial coding sequence, 4 introns and 3'UTR regions

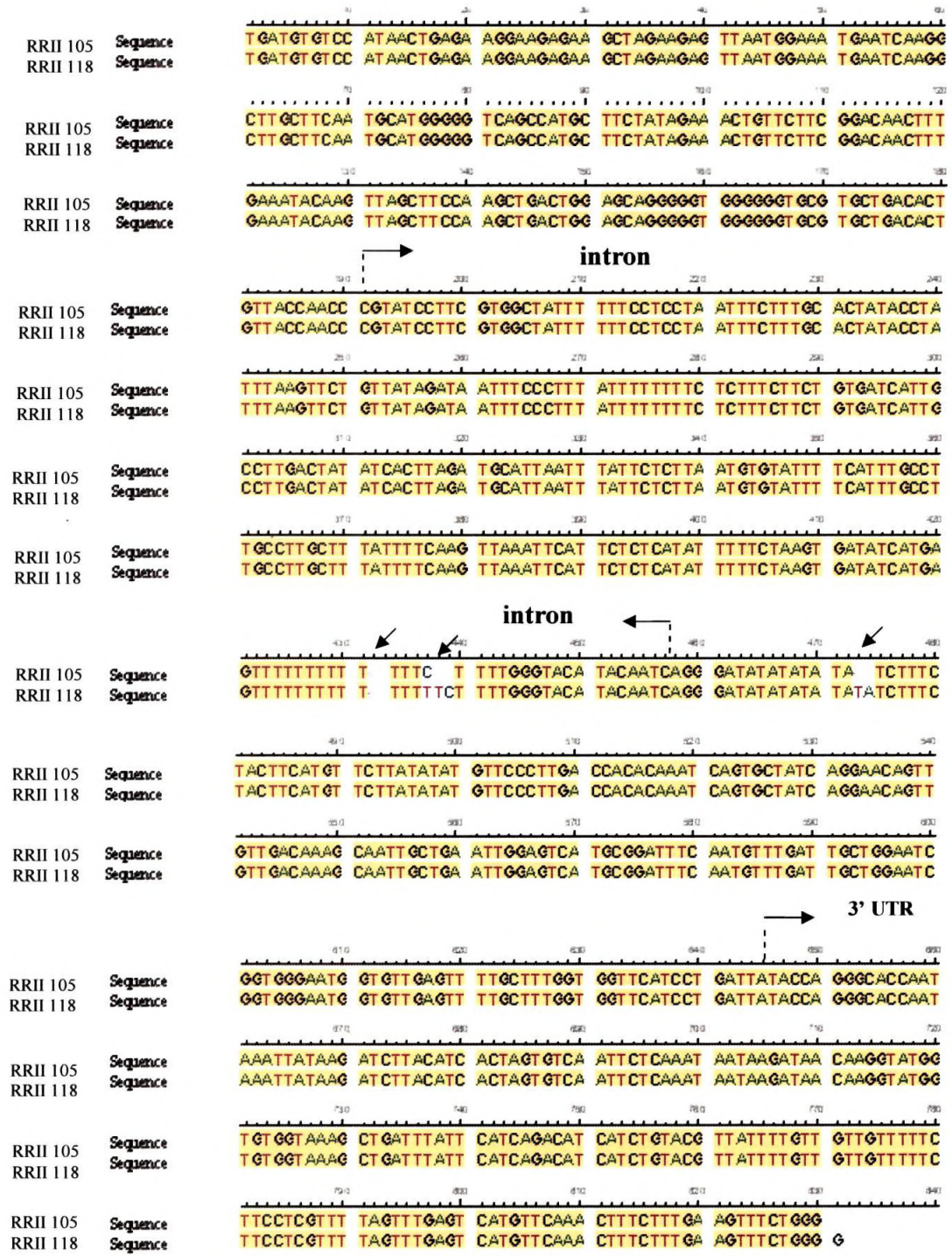


Fig. 1.34. Sequence alignment of mevalonate kinase gene from the clone RRII 105 and RRII 118 showing SNP positions including indels (marked) along with the partial coding sequence, intron and 3'UTR regions

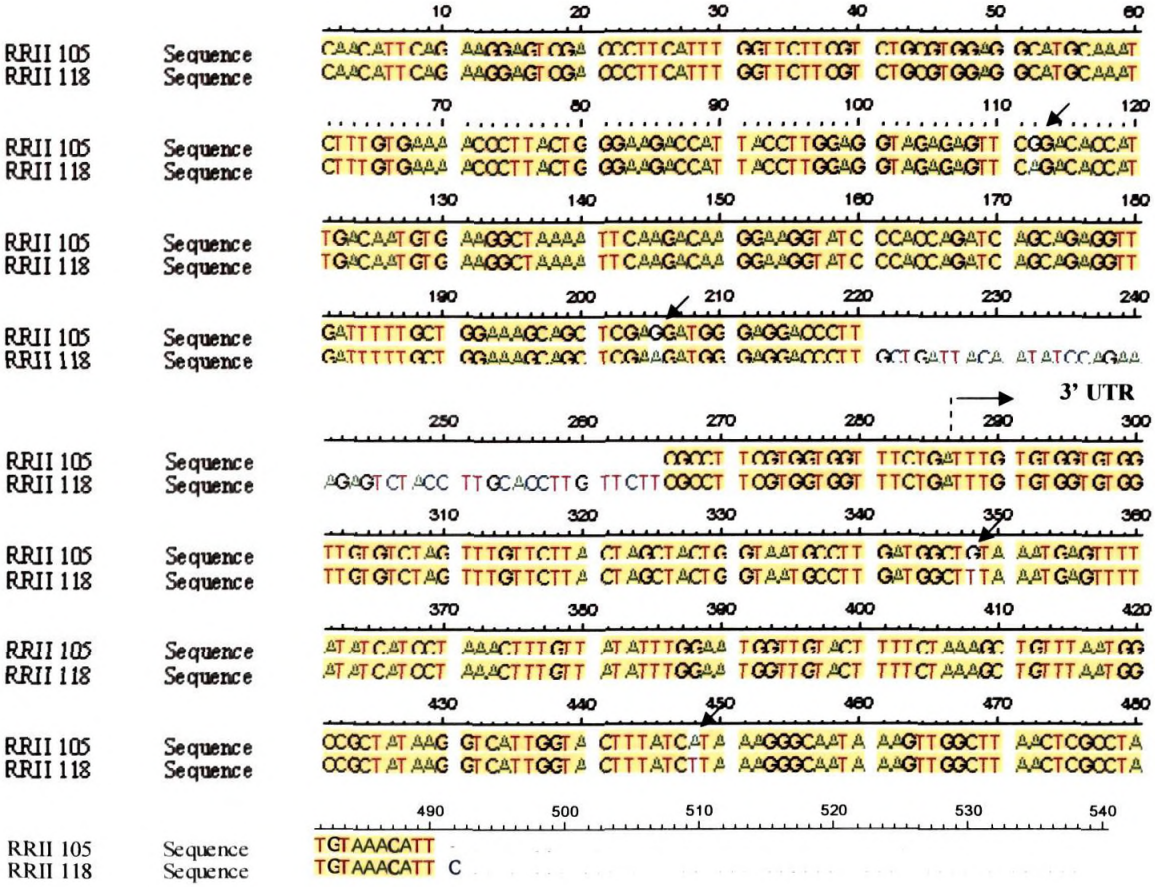


Fig. 1.35. Sequence alignment of ubiquitin precursor gene from the clone RRII 105 and RRII 118 showing SNP positions (marked) along with the partial coding sequence and 3'UTR regions. A 45 bp deletion in the clone RRII 105 was also identified

.

Table 1.19. Haplotype detection in the parental clones (RRII 105 and RRII 118)

a) Geranylgeranyl diphospahte synthase

Clone	88	107	164	400
RRII 105	A	A	A	T
RRII 118	G A	G A	G A	A T

b) Farnesyl diphospahte synthase

Clone	260	296	346	730	916	1155	1301	1310	1426
RRII 105	T	A	T	T	A	G	G	A	G
RRII 118	A A	T A	A T	C C	T T	A A	A G	G G	A G

c) Mevalonate kinase

Clone	118	197	275	332	432	433	437	438	474	475	476
RRII 105	T A	C T	T C	A G	- T	- A	- T	- T	- A	- T	C T
RRII 118	T	C	T	A	-	-	T	T	A	T	C

d) Ubiquitin precursor

e) Latex patatin homolog

Clone	112	205		288	313	334	415
RRII 105	G	G		T C	T C	A G	C A
RRII 118	A	A		C	C	G	A

The number of haplotypes among the parental genotypes containing two or more SNPs was determined by visual inspection. The details of haplotypes identified are given in Table 1.19.

1.4.2.4.6 SNP genotyping

1.4.2.4.6.1 Allele-specific PCR

Allele-specific primers developed for the identified SNPs in parental clones by sequencing various genes were tested on these clones for their validation and to test their utility as SNP markers. The three primer approach was followed successfully for the genotyping of SNPs. Allele specific PCR amplifications were carried out for the three genes 1) geranylgeranyl diphosphate synthase, 2) mevalonate kinase, and 3) latex patatin homolog.

1.4.2.4.6.1.1 Geranylgeranyl diphosphate synthase

For the genotyping of the concerned loci, the SNP position selected was the heterozygous allele A/G (88 bp) for the clone RRII 118 and the corresponding allele was homozygous (A/A) for the clone RRII 105. The allele specific primer used was the forward one (GGDPS_F88G) and size of the allele specific fragment was 412 bp and the size of the gene locus was 500 bp which was distinguished in an agarose gel. The allele selected was 'A' at the 3' end of the allele specific primer. The clone having the homozygous allele for 'A' (RRII 1105) gave the amplification profile where the allele fragment was preferentially amplified with the intense allele specific fragment, compared to the gene loci. For the clone RRII 118, the gene fragment was preferentially amplified and almost 50% reduction in amplification was noticed for the allele specific fragment for that clone explaining the heterozygosity of the allele (Fig. 1.36).

1.4.2.4.6.1.2 Mevalonate kinase

The SNP selected for the gene loci mevolonate kinase, was the heterozygous allele C/T for the clone RRII 105 and the corresponding base was homozygous C for the clone RRII 118. The allele specific primer used was the forward one (MKF_197C) and the size of the allele specific fragment was 633 bp, and the size of the gene locus was 830 bp. The parental genotypes were first standardized for PCR conditions where the heterozygous allele was clearly distinguishable from the homozygous allele. The clone RRII 105 possessing the heterozygous allele showed an amplification pattern where both the gene

fragment and the allele-specific fragment were amplified, for the homozygous allele only the allele-specific fragment was amplified preferentially (Fig. 1.37).

1.4.2.4.6.1.3 Latex patatin homolog

For the gene loci latex patatin homolog, the SNP selected for genotyping was heterozygous for the allele C/T in RRII 105 and homozygous for the allele C/C in RRII 118. The size of the allele specific fragment was 313 bp and the size of the gene specific loci was 600 bp. The allele specific primer used was in the reverse direction (LPH5R_313G). The clone RRII 105 possessing the heterozygous allele showed an amplification pattern where both the gene fragment and the allele specific fragment were amplified for the homozygous allele, only the allele specific fragment was amplified (Fig. 1.38).

1.4.2.4.6.2 Cleaved Amplified Polymorphic Sequence (CAPS)

For the gene ubiquitin precursor, SNPs were homozygous at all the base positions in the parental clones. Hence the allele specific PCR methodology cannot be used here. The standardization of restriction digestion were performed initially with the parental clones where the recognition sequences were identified and after confirming the specificity of the allele, genotyping was performed in the progeny. The amplification of the ubiquitin precursor gene generated three fragments and out of these, the lower fragment was analyzed for SNP detection. The clone RRII 105 was found to be heterozygous with the lower fragment showing a deletion of 45 bases for one chromosome and without deletion for the other homologue. The male parent RRII 118 was homozygous. Aligned sequences of the Ubiquitin precursor loci under study derived from RRII 105 and RRII 118 were subjected to the computer program SNP2CAPS. The restriction enzyme *Xho* I was selected for the development of 'CAPS' marker at the SNP base position 205. The recognition sequence of the enzyme was C/TCGAG. In the female parent RRII 105 the recognition site was functional, whereas in the male parent RRII 118 the sequence was C/TCGAA which was non functional. The expected size of digested products was 200 and 290 bp for RRII 105 and for RRII 118 the locus remained undigested (Fig. 1.39). Additional amplicons were also obtained on digestion with the restriction enzyme, since three fragments were generated in PCR amplification of the gene.

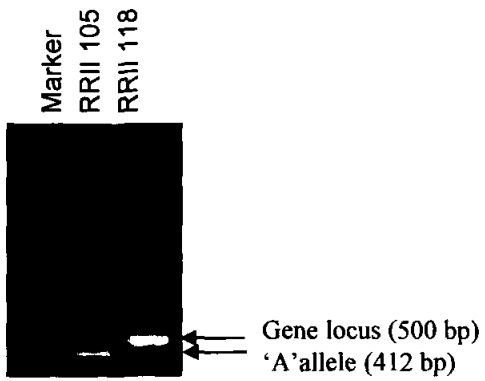


Fig. 1.36. EtBr stained agarose gel photograph showing the allele-specific PCR marker generated from the gene geranylgeranyl diphosphate synthase using allele-specific primer GGDPS_88A. The SNP detected was in heterozygous allelic state in RRII 118 (A/G) and homozygous allelic state (A/A) in RRII 105. The gene loci was preferentially amplified in the clone RRII 118, whereas allelic fragment was preferentially amplified in RRII 105. The ΦX 174 DNA/*Hae* III digest was used as molecular weight marker

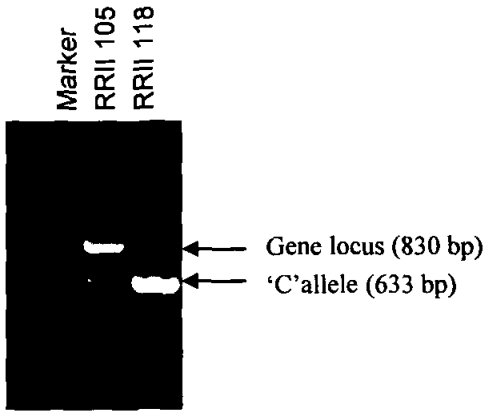


Fig. 1.37. EtBr stained agarose gel photograph showing the allele-specific PCR marker generated from mevalonate kinase gene using allele-specific primer MKF_197C. The SNP detected was in heterozygous state (C/T) in RRII 105 and homozygous state (C/C) in RRII 118. 'C' allele only was amplified in RRII 118 and both the allele-specific fragment and the gene loci were amplified in RRII 105. The ΦX 174 DNA/*Hae* III digest was used as molecular weight marker

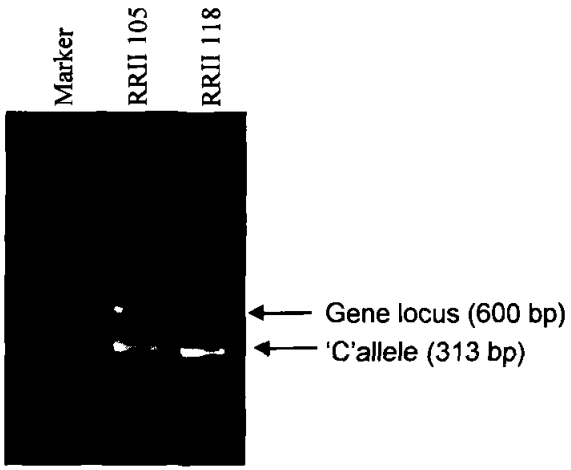


Fig. 1.38. EtBr stained agarose gel showing allele-specific PCR marker, generated from latex patatin homolog gene using allele-specific reverse primer LPH5R_313G. The SNP detected was in heterozygous state (C/T) in the clone RRII 105 and homozygous allelic state (C/C) in RRII 118. 'C' allele was amplified only in the clone RRII 118 and both the allele-specific fragment and the gene loci were amplified in RRII 105. The ΦX 174 DNA/*Hae* III digest was used as molecular weight marker

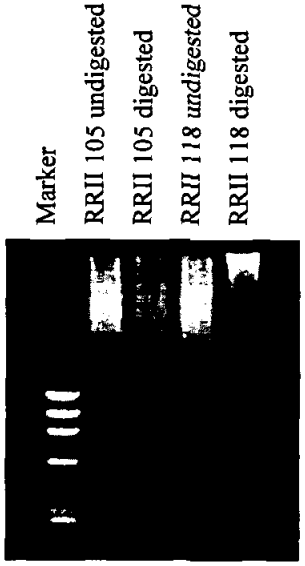


Fig. 1.39. EtBr stained agarose gel photograph showing CAPS marker generaton from the gene ubiquitin precursor using the restriction enzyme *Xho* I. Both the digested and undigested samples were electrophoresed. The gene sequence from RRII 105 having the *Xho* I restriction site, generated two fragments of 290 and 200 bp after complete digestion, which discriminated the clone RRII 105 from RRII 118. The ΦX 174 DNA/*Hae* III digest was used as molecular weight marker

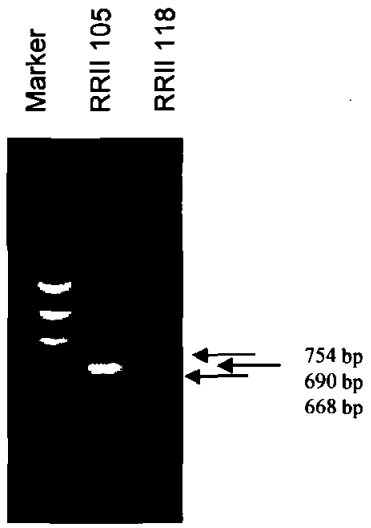


Fig. 1.40. EtBr stained agarose gel photograph showing the sequence tagged site (STS) marker generation from the gene farnesyl diphosphate synthase. Two fragments were generated in RRII 118, whereas one fragment was generated in RRII 105 discriminating the clones. Three bands were distinguished as upper fragment having an addition of 64 bp from the middle common fragment and a lower fragment of 22 bp deletion. Arrowheads show the STS markers. The ΦX 174 DNA/*Hae* III digest was used as molecular weight marker

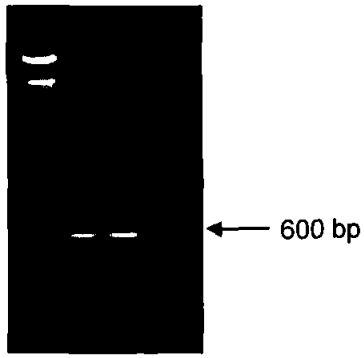


Fig. 1.41. EtBr stained agarose gel photograph showing the amplification of the RGA marker. The λ DNA/*Eco* RI + *Hin*d III digest was used as molecular weight marker

1.4.2.4.6.3 Sequence Tagged Site (STS) marker

The genotyping of the gene farnesyl diphosphte synthase was based on a sequence tagged site marker generated from the amplification of around 700 bases at the 5' end of the gene sequence. The initial screening was done with five selected popular clones based on diversity, where three distinct fragments of 754 bp, 690 bp and 668 bp were amplified. The clone RRII 118 was distinct from other clones as it had generated two fragments of size 0.69 kb and 0.67kb while other clones generated only a single fragment of size 0.76 kb. The three bands were distinguished as upper fragment having an addition of 64 bases to the middle common fragment and a lower fragment of 22 base deletion. The parental clone RRII 105 amplified only the middle allele of 0.69 kb fragments (Fig. 1.40). Since RRII 118 was one of the parents of the progeny population this data was utilized for genotyping the entire progeny for mapping.

1.4.2.5 Resistance gene analogues (RGAs)

Selected degenerated primer combination (RGF3+RGR15) was chosen since our previous studies (Saha *et al.*, 2010) indicated that this primer combination was successful in amplifying RGAs from rubber with all the motifs. The PCR amplified products were run on the agarose gel and they were of the expected size, around 600 bp (Fig. 1.41). The different amplicons derived from the gene family could not be resolved through the agarose gel and hence denaturing PAGE was run to detect polymorphic RGA products. Out of the nine loci identified in PAGE, only one locus was found to be polymorphic between the clones RRII 105 and RRII 118 which was present in the clone RRII 105. This RGA marker developed was used for segregation analysis in the progeny. To characterize the amplicon, the same was cloned and sequenced and the sequence confirmed the RGA with all the conserved motifs (Fig. 1.42).

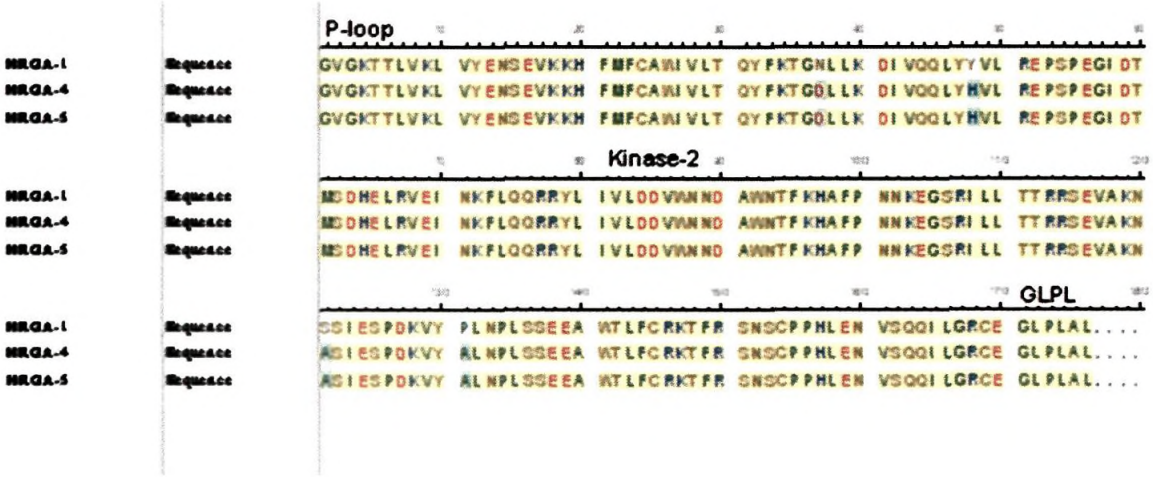


Fig. 1.42. Multiple alignment of aminoacid sequences (conceptual translation) of resistance gene analogues (RGA) isolated from rubber showing the characteristic conserved motifs present in R genes

1.5 Discussion

The history of rubber cultivation started from the collection of a few seeds by Sir Henry Wickham from the Tapajos River basin in Brazil in the year 1896. Continuous breeding efforts in different rubber growing countries have tremendously improved the yield of the crop. However, the breeding of *Hevea* has some major constraints. Being a tree crop the crossing is laborious and time consuming and the long breeding cycle makes the evaluation of hybrids very difficult. Another major constraint is the narrow genetic base of the 'Wickham collections'. Further, intensive selection only for yield neglecting the secondary characters resulted in its erosion (Wycherly, 1968). At present, breeding for biotic and abiotic stress tolerance, development of latex timber clones and improvement for other secondary traits without compromising on yield is gaining importance. Molecular breeding approaches can play a significant role in attaining this goal. Molecular markers are being used in rubber mostly for diversity studies, genetic mapping and identification of genetic loci for agronomical traits. But still extensive genetic investigations are lacking in rubber.

The molecular techniques employed in *Hevea* by several researchers are isozymes, RFLP, RAPD, AFLP and SSR markers. Even though genetic linkage maps were constructed in rubber using different marker systems (Lespinasse *et al.*, 2000; Feng *et al.*, 2010; Souza *et al.*, 2011), it is essential to develop our own genetic markers for population studies of the cultivated *Hevea* clones, diversity analysis and construction of genetic linkage maps.

Availability of genomic resources in the public domain is limited. Due to these limitations, SSR marker development was initiated with the conventional methods. The developments of microsatellite markers have a major impact on genetic analysis and implementation of marker assisted breeding programs in *Hevea* through the generation of marker saturated genetic linkage maps. Genomic microsatellite enriched libraries can be an efficient alternative for marker development involving low cost and technology. In *Hevea*, enriched microsatellite libraries were constructed for marker generation (Atan *et al.*, 1996; Souza *et al.*, 2009; Le Guen *et al.*, 2010; Mantello *et al.*, 2012). Efficient method to produce highly enriched microsatellite libraries in plants was described (Edwards *et al.*, 1996; Tenzer *et al.*, 1999). Most of the microsatellite markers, so far developed in plants have been dinucleotide repeats, since they are more abundant in the genome, and thus easier to isolate (Morgante and Olivieri, 1993). Dinucleotide

microsatellite markers suffer from technical drawbacks for their use in genotyping as 2-bp differences have to be resolved. Moreover, complex banding patterns (shadow bands) arise on gels, making them sometimes difficult to score. For these reasons tri- and tetra-nucleotide repeats have become the markers of choice for population, linkage and forensic studies in humans as well as in other animal species (Gastier *et al.*, 1995; Sheffield *et al.*, 1995; Tozaki *et al.*, 2000). Trinucleotide repeats are three times more frequent in transcribed than in non-transcribed regions in *Arabidopsis thaliana* as well as in maize (Morgante *et al.*, 2002). In plants, no extensive study has been carried out so far on the distribution of trinucleotide microsatellites across the genome. Trinucleotide microsatellite markers were reported in *Citrus* (Kijas *et al.*, (1994), wheat (Bryan *et al.*, 1997) and *Pinus taeda* (Elsik *et al.*, 2000).

1.5.1 SSR marker generation

An enriched library for trinucleotide repeats was constructed both in the lambda phage and in plasmid vectors following the principle of magnetic separation of DNA fragments bearing repeats (Kijas *et al.*, 1994). The libraries were screened for the detection of positive clones bearing the repeats and the positive clones were sequenced, primers were designed based on the sequences flanking the repeat regions and finally analyzed for polymorphism.

In the present study, enrichment percentage for the lambda phage library was found to be 5.6% and that for the plasmid library was detected as 24%. Out of the 140 clones sequenced, trinucleotide repeats were present in 30 clones (21.4%) and similar reports of enrichment of repeats in *Hevea* was published by Atan *et al.* (1996) with 39% enrichment and Mantello *et al.* (2012) with 21.5% enrichment. It was reported that the enrichment levels between species vary from 10% to over 90% (Cordeiro *et al.*, 1999). The reason for the low efficiency of enrichment for the trinucleotide repeats can be attributed to the low frequency of these repeats in the *Hevea* genome. In *Hevea*, Feng *et al.* (2009) found that dinucleotide repeats are three times more abundant than other repeat motifs. In hazelnut, an enriched trinucleotide repeat library was constructed along with the dinucleotide repeats and it was found that the trinucleotide repeat library was less informative than dinucleotide repeat library (Bassil *et al.*, 2005). The same observation was made in tea (Rossetto *et al.*, 1999).

1.5.1.1 Prevalence of repeat motifs

It was observed that the AAG/CTT repeats were the most common repeat identified in the present study. Prevalence of AG/CT and AAG/CTT repeats in *Hevea* genome was also reported (Roy *et al.*, 2004; Feng *et al.*, 2009; Triwitayakorn *et al.*, 2011). A similar bias to AG repeats also found in soybean (Gao *et al.*, 2003), citrus (Chen *et al.*, 2006) and apple (NewComb *et al.*, 2006). Enrichment for the repeat type (AAT) was less, which might be attributed to the self-annealing nature of the probe that gave similar results with AT repeats (Atan *et al.*, 1996; Roy *et al.*, 2004).

Thirty-one primer pairs were designed from the sequences containing trinucleotide repeats and out of these only two markers were found to be polymorphic within the popular clones. It was reported that trinucleotide repeats were found more in the coding regions than in the non-coding regions (Metzgar *et al.*, 2000). This may be the reason for less polymorphism for the trinucleotide derived SSR markers compared to other repeat types.

1.5.2 Genetic diversity analysis

1.5.2.1 RAPD markers

In the present study, out of the 138 amplified fragments, 108 bands were found polymorphic (78%) with an average number of seven polymorphic fragments per primer. Nakkanong *et al.* (2008) used eight RAPD markers for genetic diversity studies of early introduced clones of *Hevea* and obtained 78.5% polymorphic markers. Varghese *et al.* (1997) observed 72% polymorphisms in the evaluation of RAPD markers in *Hevea*.

Among the polymorphic primers, 16 random primers were selected for the genetic diversity analysis of the popular *Hevea* clones. The PIC values of these primers ranged from 0.15 to 0.3 and the average PIC for the data set was found to be 0.22. The similarity coefficients obtained from the genetic diversity analysis of popular clones by the RAPD markers ranged from 0.67 to 0.91. Diversity analysis for the South American, Central American and Asian clones of *Hevea* with RAPD markers resulted in a similarity coefficient ranging from 0.55 to 1.0 (Hernandez *et al.*, 2006). Twenty-four *Hevea* clones were analyzed with 43 decamer primers by Varghese *et al.* (1997) to estimate the genetic diversity between them and it was observed that genetic distance ranged from 0.05 - 0.75. Nakkanong *et al.* (2008), in their attempt to genetically discriminate 87 clones of early introduced and current cultivars of rubber tree with 8 RAPD primers, observed that cultivated clones were less variable than early introduced clones exhibiting a narrow

genetic diversity with a mean genetic similarity of more than 0.6. Genetic diversity analysis by Oktavia *et al.* (2011) in 45 *Hevea* clones using 12 RAPD markers observed that the similarity matrix ranged from 59.18 - 94.23%. All these results support the narrow genetic diversity of popular *H. brasiliensis* clones as determined in the present analysis.

In this study, mainly four clusters were obtained through UPGMA analysis, and the Sri Lankan clones were found to be clustered together with a high bootstrap value which had a common pedigree. RRII 105, the most popular clone developed by the Rubber Research Institute of India (RRII) was grouped with only one of its progenitor Tjir 1 but the bootstrap value was low (7%). The clones RRII 118 and PB 86 could not be discriminated separately by the RAPD analysis.

1.5.2.2 AFLP markers

A total number of 434 loci with an average of 48.3 loci per primer combination, were generated by nine primer combinations in *Hevea* in the present study. Out of these, 262 loci were polymorphic (60.4%) with an average of 29.11 polymorphic bands per primer combination. A low polymorphism was noticed with the AFLP markers in our study compared to the reports of An-Ding *et al.* (2001) regarding fingerprinting analysis in *Hevea* using AFLP markers. They observed 98.6% polymorphism for the amplified fragments. In cultivars of apricot, the genetic diversity analysis was performed using AFLP markers and 69% of the fragments were found polymorphic (Krichen *et al.*, 2008). AFLP markers were used in genetic diversity analysis of olive and estimated polymorphism was found to be 49.5% (Owen *et al.*, 2005) and 59.8% (Sensi *et al.*, 2003). In comparison with the above findings, the polymorphism observed among the popular *Hevea* clones is almost same in the present investigation. It is understood that the primer combinations differed in their ability to detect polymorphism within populations. The average PIC values were found to be 0.17 for the AFLP markers ranging from 0.10 to 0.23. The similarity coefficient of the *Hevea* clones, assessed using AFLP markers ranged from 0.76 to 0.92 with three major clusters at a similarity coefficient of 0.8 indicating a low genetic diversity. An-Ding *et al.* (2001) estimated a much higher genetic diversity for the *Hevea* clones ranging from 0.25 to 0.81. The Sri Lankan clones were clustered together with the AFLP markers showing a robust bootstrap value of 55%, which was in conformity with its pedigree relationship. The clone RRII 105 was grouped along with Gl 1, the maternal parent with a bootstrap value of 46%. The clone RRII 118

was found closely associated with RRII 33. AFLP marker could not discriminate the clones Gl 1 from PB 5/51 and also PB 86 from KRS 25. Only a few main nodes (4) and sub-nodes (7) indicated ≥ 40% bootstrap values.

1.5.2.3 Simple Sequence Repeat (SSR) markers - genomic and genic

A total of 74 alleles were generated with 17 genomic SSR markers (gSSR) with an average of 4.34 alleles per marker. For the genic/ EST-SSR markers, a total of 51 alleles were generated with an average of 3.64 alleles per marker. Genomic SSR markers generated more number of alleles compared to EST-SSR markers in the present analysis. Maximum information was provided by the genomic SSR marker *Hev*-glu with seven alleles, while a maximum of 5 alleles were detected with EST-SSR marker. In other studies, mean allele number reported by Gouvêa *et al.* (2010) was 5.88. Lekawipat *et al.* (2003) reported average of 5.92 alleles per marker in a set of cultivated genotypes using gSSR. The results of the diversity analysis in 12 cultivated popular clones of *Hevea* and four related species by Feng *et al.* (2009), showed an average of 2.47 alleles using the EST-SSRs.

In the present analysis, PIC values varied from 0.42 to 0.80 with an average of 0.59 for the genomic SSR markers and for EST-SSR markers it ranged from 0.2 to 0.76 with an average value of 0.56. Similar PIC values (0.5) were obtained by Souza *et al.* (2009) in *H. brasiliensis* using genomic SSR markers. Perseguini *et al.* (2012) reported an average PIC value of 0.59 in genetic diversity analysis of 46 accessions of *H. brasiliensis* using thirty EST-SSR markers, which was close to the average PIC values of EST-SSR markers used in the present study. It was observed that a slightly higher PIC value was obtained with the genomic SSRs in the present analysis and the same was supported by the studies of Gouvêa *et al.* (2010) in *Hevea.*

The similarity coefficient resulting from the genetic diversity analysis of *Hevea* clones with 17 genomic SSR markers ranged from 0.57 to 1.00, while that of EST-SSR markers ranged from 0.60 to 0.93 in the present study. Low genetic diversity was also observed with this marker system and the same was supported by other reports in *Hevea,* such as Gouvêa *et al.* (2010) estimated a low genetic diversity among the Asian clones and they were clustered into a single group when compared to the African clones using SSR markers. Nakkanong *et al.* (2008) revealed a low genetic diversity with the similarity coefficients ranging from 0.54 - 1.0 using SSR markers. These results were on par with the findings of Saha *et al.* (2007) indicating a much higher genetic diversity in

wild *Hevea* accessions compared to the cultivated popular clones. A low genetic diversity among the Asian clones were reported by Perseguini *et al.* (2012) in rubber while analyzing the genetic structure of Asian, African and Amazonian clones with EST derived SSR markers. Lekawipat *et al.* (2003) used twelve microsatellite markers to assess DNA polymorphism among 108 accessions of *H. brasiliensis* inclusive of 40 cultivated (Wickham) clones and 68 wild accessions collected from Amazon forest (1981 Amazonian accessions) and they found wild accessions were more polymorphic than cultivated Wickham clones.

1.5.3 Genetic relationship among cultivated *Hevea* clones

Mainly three clusters were generated in the present analysis and most of the genotypes were grouped in the first cluster using SSR markers. Only six sub-clusters showed the bootstrap value greater than 40% and the EST-SSR markers were not able to differentiate the clones SCATC 93/114 and PB 314. Even though all the Sri Lankan clones were clustered together, the pedigree relationships could not be established for other clones with RAPD, AFLP and SSR markers. The ancestry could not be considered as a suitable criterion for characterizing most of the groups as showed in the studies of Gouvêa *et al.* (2010) in *Hevea*. Feng *et al.* (2012) observed that the descendants of the cross PB 5/51 × PR 107 were not clustered together. The same was also reported by Nurhaimi-Haris *et al.* (1998) and Toruan-Mathius *et al.* (2002) in *H. brasiliensis* revealing genealogy relationships in few clones. The reasons may be attributed to the cross-pollinated nature of the rubber tree and the F_1 hybrids generated from them are highly heterozygous. They are fixed vegetatively and hence, the proportion of marker alleles in the F1 hybrid from each parent can vary considerably. Thus, in highly heterozygous species with a common ancestry, pedigree information may not always reveal the exact nature of genetic relationships (Varghese *et al.*, 1997).

When all the markers were analyzed together with the UPGMA cluster analysis, it was observed that the similarity coefficient ranged between 0.74 to 0.9. The narrow genetic diversity could be explained by the fact that most rubber breeding materials used in Asia are believed to have originated from 22 seedlings surviving from Sir Henry Wickham's original collection, known as the 'Wickham base', and consequently, the original genetic base of cultivated rubber in Asia is low (Cesar *et al.*, 2006) and also may be attributed to selection of clones for latex yield only in a unidirectional manner.

In the present study, maximum information about the polymorphism was rendered by the SSR markers and among these the genomic SSRs exhibited slightly higher PIC values compared to genic SSRs. The efficiency of microsatellite markers in detecting the genetic diversity was proved in the studies of Lekawipat. (2003); Roy *et al.* (2004); Saha *et al.* (2005); Feng *et al.* (2009); Gouvea *et al.* (2010) and Feng *et al.* (2012) in *Hevea.* The SSR markers generated in popular clones may be used to resolve any clone dispute and varietal protection.

The RAPD markers were not able to distinguish the clones, RRII 118 and PB 86 and similarly, AFLP markers could not able to discriminate the clones Gl 1 from PB 5/51 and PB 86 from KRS 25. EST-SSR markers were not able to differentiate the clones SCATC 93/114 and PB 314. Similar results were shown by Feng *et al.* (2009; 2012). They could not distinguish the clones GT 1 from Reken 525 using 87 EST-SSR markers and Haiken 2 from Dafeng 99. Hence it can be concluded that combination of various types of markers is an excellent way of identifying genotypes as reported by Meszaros *et al.* (2007) in barley.

1.5.3.1 Principal Coordinates Analysis

The principal coordinates analysis was performed with all the marker systems separately and it was observed that the UPGMA clustering was similar to that of the clustering by the principal coordinates analysis. Varghese *et al.* (2007) identified same clustering pattern for the UPGMA and PCoA in *Hevea.*

The mantel test between the three marker systems under study revealed an 'r' value which is insignificant. The study does not represent any correlation between the groupings generated by different marker systems. It has been shown that different markers might reveal different classes of variation (Powell *et al.*, 1996; Russell *et al.*, 1997).

DNA-based marker generation is rather easy now due to the availability of next generation sequencing (NGS) technologies. When this work was initiated, the EST sequences were less than a thousand in the NCBI GenBank. Therefore, classical method was adopted for generating SSR markers in rubber. Later, 12,365 ESTs from *H. brasiliensis* became available in the public domain. Out of these ESTs, 11,256 ESTs were from latex, 1091 ESTs from bark and 18 ESTs from leaves (Han *et al.*, 2000; Ko *et al.*, 2003; Chow *et al.*, 2007). Recently, Triwitayakorn *et al.* (2011) reported 17,819 ESTs from transcriptome sequencing of young shoots and used for the development of SSR

markers. So there is immense scope for further genetic investigations in *Hevea* using the newly developed markers especially the SSRs and SNPs.

1.5.4 SNP markers

1.5.4.1 Identification of SNPs

Single nucleotide polymorphism (SNP) is the most abundant form of genetic variation among individuals within a species. SNPs are used as markers for gene discovery and to assess biodiversity. Currently there is great interest in using SNPs for gene mapping. Success of SNP discovery has been demonstrated in crop species where large numbers of sequences are available in the database in the public domain. However, for SNP discovery in rubber in the present study, 12 genes were selected and the 3' untranslated regions (3'UTR) along with partial coding regions were amplified from a panel of 16 popular *Hevea* clones. The 3' UTR region was selected as the UTR regions have a higher frequency of polymorphism than other parts of the genome and sequencing of this region increases the likelihood of finding multiple SNPs (Zhu *et al.*, 2001; Ching *et al.*, 2002).

SNP analysis in the present study was conducted mainly in the latex biosynthesis genes along with other genes involved in complex biosynthetic pathways *viz.*, ubiquitin precursor, latex patatin homolog, ethylene inducible protein, transcript associated with tapping panel dryness, latex plastidic aldolase, thioredoxin, glutathione peroxidase and hevamin B. Out of the 12 genes SNPs could be identified only in five genes *viz.*, geranylgeranyl diphosphate synthase, farnesyl diphosphate synthase, mevalonate kinase, ubiquitin precursor and latex patatin homolog. In the present study, 44 SNPs were detected in five genes and among these, 29 SNPs were present in the intronic sequences and UTR regions, while only 16 SNPs were detected in the coding sequences. The indels were also present in the non-coding region. The reason may be attributed to the fact that insertion and deletion of nucleotides are likely to produce changes in the open reading frames disrupting correct translation, therefore selection against them is expected in coding regions.

The frequency of SNPs for the gene geranylgeranyl diphosphate synthase was calculated as one SNP in 100 bases, in farnesyl diphosphate synthase one SNP in 107 bases, in mevalonate kinase one SNP in 64 bases, in ubiquitin precursor one SNP in 71 bases and in latex patatin homolog one SNP in 100 bases. The frequency of SNPs was found to be more in the present study, compared to other report in *Hevea* (Pootakham *et al.* (2011) showing 1 SNP in every 1.5 kb. In some crops, the frequency was found to be

much higher with one putative SNP per 223 bp in non-coding regions and one SNP per 71 bp in the coding regions in cacao (Lima *et al.*, 2009). In maize, the frequency of polymorphisms in the US elite inbreeding lines was one SNP per 31 bp in non-coding regions, and one SNP per 124 bp in the coding regions (Ching *et al.*, 2002). In soybean, a study on sequence diversity in 22 diverse genotypes revealed 1.64 SNPs per kb in coding regions and 4.85 SNPs per kb in non-coding regions (Zhu *et al.*, 2003).

Out of the 44 SNPs detected from the five gene sequences, 24 SNPs (55.5%) represented transitions while 14 (28.8 %) represented transversions and 6 (13.6%) were indels. The observed transition: transversion ratio was found to be 1.7. Transition SNPs are the most frequent SNP type reported in both plant and animal genomes and assumed to result from hypermutability effects of CpG dinucleotide sites and deamination of methylated cytosines (Morton *et al.*, 2006). Our results were in conformity with that of Pootakham *et al.* 2011 in *Hevea* detecting 56.3% transitions and 33.7% transversions and the remaining 10% were insertions/deletions. The transition : transversion ratio was found to be 1.67. In grapes, the transition : transversion ratio was 1.56 (Salmaso *et al.*, 2005) and apple 1.27 (Micheletti *et al.*, 2011). Similarly Aranzana *et al.* (2012) on evaluation of the levels of sequence variability in commercial peach, they found that 64% SNPs were due to transitions and 36% were transversions. In cacao, the transition type appeared with a higher frequency than did the transversion , 62% of transitions and 38% of transversions (Lima *et al.*, 2009). The reasons for higher rate of transitions can be due to the fact that as the transversion generates a distortion of DNA molecule width causing serious functional damage in the cell. It was suggested that this kind of mutation was more difficult to be corrected by the DNA system repair (Calcagnotto, 2001).

1.5.4.2 SNP haplotypes

A set of linked SNPs across a long stretch of DNA of the same chromosome is defined as haplotypes which tend to be inherited together. Haplotype is much more informative than single SNPs and also it is essential for developing SNP tags for the gene of interest. For the gene geranylgeranyl diphosphate synthase only one haplotype was detected in the clone RRII 105 and two haplotypes in RRII 118. Similarly, in farnesyl diphosphate synthase one haplotype in RRII 105 and two haplotypes in RRII 118, in mevalonate kinase two haplotypes in RRII 105 and one haplotype in RRII 118, in ubiquitin precursor one haplotype each for both the clones, in latex patatin homolog two haplotypes in RRII 105 and one haplotype in RRII 118. The number of haplotypes detected depends on the

homozygosity/ heterozygosity of the alleles generated. If the neighboring alleles are heterozygous more haplotypes are expected.

1.5.4.3 SNP genotyping

Once SNPs are identified, they could be genotyped using various techniques for their use in linkage map construction and marker assisted selection (MAS). Many SNP detection methods and chemistries were developed based on allelic discrimination, such as denaturing high performance liquid chromatography (DHPLC), oligonucleotide ligation, primer extension, DNA sequencing, PCR primer mismatch, pyrosequencing and heteroduplex assays (Kuppuswami *et al.,* 1991; Ronaghi *et al.*, 1997; Hoogendorn *et al.*, 1999; Pastinen *et al.*, 2000; Vigna *et al.*, 2002; Sobrino *et al.*, 2005). In the present study, SNP genotyping was performed using a simple allele discriminating PCR and CAPS marker for small scale genotyping and mapping studies. The sequence tagged site (STS) marker was also developed for the gene *FDPS* for genotyping. The methodology followed was the three primer approach for the allele-specific PCR amplification using two outer primers amplifying the gene sequences as a control and an allele-specific primer discriminating the SNP. For the gene geranylgeranyl diphosphate synthase, the heterozygous alleles were present in the clone RRII 118. Hence, the gene locus was preferentially amplified along with the SNP allele depending on the allele-specific primer. In case of homozygous state of the allele in the other clone, the allele was amplified preferentially. From the amplification profiles homozygosity and heterozygosity of the alleles were clearly identified. Similarly for the genes mevalonate kinase and latex patatin homolog, the amplicons clearly defined the homozygosity and heterozygosity of the alleles in both the clones with this three primer approach. All these results adhered to the results of the SNPs identified through direct sequencing. Hence, these allele specific-primers could effectively be used for progeny analysis. In plants, the tetra primer ARMS-PCR was followed in crops like barley (Chiapparino *et al.*, 2004). This technique was used in the identification of SNPs in various human genes (Sommer *et al.*, 1989; Duta-Cornescu *et al.*, 2009). By ARMS-PCR method the identification of particular genotypes was accomplished in a single step PCR, excluding other post-PCR manipulation.

The CAPS marker is a simple and accurate genotyping assay in laboratories which does not require any sophisticated equipments (Parani *et al.*, 2001; Kaundun and Matsumoto, 2003). CAPS marker development for the gene ubiquitin precursor was

based on the creation of a restriction site (CTCGAG) by the presence of a SNP, which could discriminate two parents of the mapping population. The involvement of an SNP at the recognition sequence in one of the parents, resulted in the generation of fragments of different sizes during its digestion. The digestion of the clone RRII 105 generated two fragments of 290 bp and 200 bp respectively whereas the clone RRII 118 did not give any digested products as expected with the SNP analysis through direct sequencing of the clone. This CAPS marker which could detect the SNP at the particular base position was successfully used in segregation analysis. CAPS markers were generated in other crops for genotyping assays such as barley (Thiel *et al.*, 2004), strawberry (Kunihisa *et al.*, 2003) for varietal identification and protecting breeders right. In pear (Cho *et al.*, 2009), CAPS marker was identified which could be linked with the scab resistance gene.

Another genotyping method followed in the present study was the use of sequence tagged site marker (STS) based on indels identified in the *FDPS* sequence. Primer sequences were developed for genotyping which could clearly differentiate the parental clones RRII 105 and RRII 118.

1.5.5 Resistance gene analogues (RGAs)

Majority of the cloned R genes (disease resistance genes) encode cytoplasmic receptor-like proteins characterized by an N-terminal nucleotide-binding site (NBS) and a leucine-rich repeat (LRR) domain (Hammond-Kosack and Jones, 1997). Isolation and characterization of the NBS-LRR-encoding genes are of significance in understanding plant-pathogen interactions for effective management of diseases. Conserved motifs of the NBS-LRR protein have enabled rapid isolation of NBS-LRR genes or RGAs from different plant species through PCR-based approach using degenerate primers designed from these domains. RGA approach was adopted to amplify the disease resistance gene in rubber. In this study an effort was made to place an RGA marker on the linkage map and successfully integrate the same into a linkage group in RRII 105 map (the same was described in Chapter 2).

Chapter II

Construction of genetic linkage map

Chapter II

2.1 Introduction

Genetic linkage mapping is a method by which location of genes/markers on a chromosome are determined based on recombination frequencies observed in pedigrees or progeny population (Clark and Wall, 1996). The concept of linkage between genetic markers in heredity was first proposed a century ago by Morgan (1911). The first genetic linkage map of X-chromosome of *Drosophila* was published by Sturtevant in 1913 and the first partial genetic map of maize was published by Emerson *et al.* (1935).

Genetic mapping is a powerful approach for analyzing the structure of genome and its evolution (Yin *et al.*, 2001). Such studies are a prerequisite for the development of markers for marker assisted breeding through the identification of quantitative trait loci (Bradshaw and Stettler, 1995) and positional cloning of genes (Kumar, 1999). A genetic linkage map tagged with agronomically important traits enables to assess the genetic potential of seedlings at the early stages of development without actually measuring the phenotypic trait reducing the requirements of growing area, maintenance and generation cycles to produce new cultivars (Gardiner *et al.*, 2004). Genetic linkage maps were first constructed using morphological markers which possess major limitations, such as delayed expression for markers of interest, dominance, epistasis, pleiotropy, less polymorphism and mainly the influence of environment on the expression of morphological characters. Later more comprehensive genetic maps were constructed using biochemical markers but they are limited in number especially for tree crops. The limitations of morphological markers and biochemical markers were overcome with the advent of DNA based molecular markers.

2.1.1 Requirements for a genetic linkage map

Construction of genetic linkage map requires selection of an appropriate population, informative markers or genes, calculation of pair wise recombination frequencies, establishment of linkage groups, estimation of map distances and determination of gene or marker order (Staub *et al.*, 1996). The development of an appropriate mapping population is the first step in the construction of a genetic linkage map. Generally F_2, backcross, recombinant inbred lines (RILs), double haploids (DH), near isogenic lines (NILs) have been employed in mapping programs for most of the plant species (Semagn

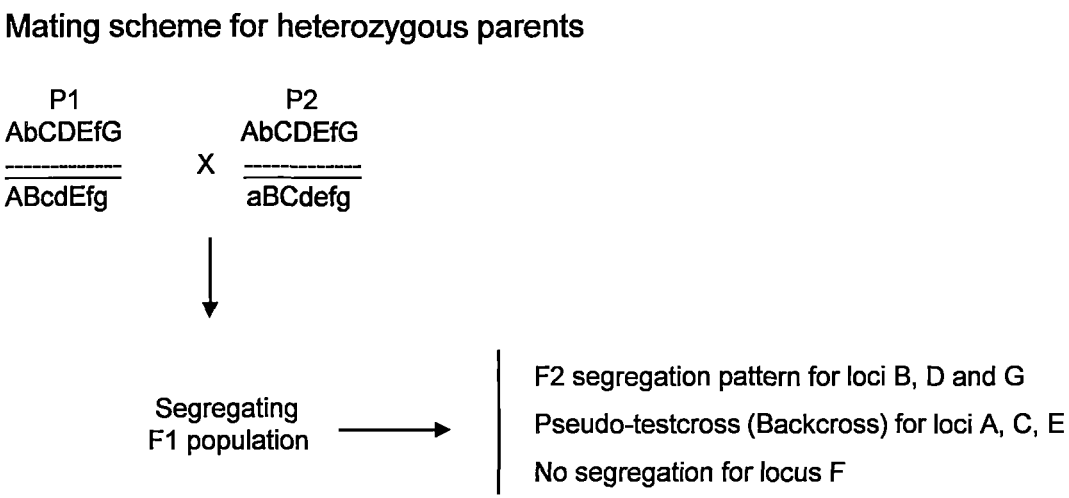


Fig. 2.1. Schematic representation of the pseudo-test cross strategy adopted in rubber. Allelic status at each locus in segregating F_1 population is shown.

et al., 2006). But in tree crops, pure inbred lines are not available due to high genetic load and long generation time. To circumvent this limitation, a 'two-way pseudo-test cross' approach is generally used to construct linkage maps and for QTL analysis in tree crops (Ritter *et al.*, 1990, Grattapaglia and Sederoff, 1994) (Fig. 2.1). This strategy has been successfully used in a number of linkage map constructions such as: *Eucalyptus* (Grattapaglia and Sederoff, 1994), apple (Conner *et al.*, 1997), cassava (Fregene *et al.*, 1997) Prunus (Joobeur *et al.*, 1998), sour cherry (Wang *et al.*, 2000), chestnut (Cassoli *et al.*, 2001), olive (Wu *et al.*, 2004), *Vitis* (Gokabayarak *et al.*, 2006) citrus (Çevik and Moore, 2007) and many other plants.

Genotyping of progeny population is the second step in the development of a genetic linkage map involving the identification of polymorphism in the population. Genotyping is the assessment of polymorphism across the entire population (Collard *et al*, 2005). The genetic basis of linkage map is recombination between homologous chromosomes, a phenomenon that occurs via sexual reproduction during meiosis (Liu, 1988). Such recombination is measured by the recombination fraction, the proportion of recombinant individuals to the total number of individuals within a population (Jones *et al.*, 1997). Many statistical procedures have been developed to calculate recombination frequency and estimate genetic linkage between marker loci (Haldane 1919., Kosambi, 1944., Allard, 1956., Lorieux *et al.*, 1995). Mapping functions convert the recombination fraction (r) into mapping units (cM). One unit of map distance (1 cM) is equivalent to one percent crossing over (Paterson *et al.*, 1988) named after the geneticist Thomas Hunt Morgan by his student Alfred Henry Sturtevant.

Based on the type of the algorithm used for arranging the markers by calculation of maximum likelihood or logarithm of odds ratio (LOD) and the type of the progeny considered, there are several software such as LINKAGE (Suiter *et al.*, 1983), MAPMAKER (Lander *et al.*, 1987), GMENDEL (Echt *et al.*, 1992), MapManager QTX (Manly *et al.*, 2001) and JoinMap (Van Ooijen and Voorrips, 2001) for the construction of genetic linkage maps. The LOD score, to the base 10 (log), hypothesises the likelihood of no linkage existing ($r = 0.5$ null hypothesis) to the linkage existing ($r < 0.5$ alternate hypothesis) based on expected and observed marker class frequencies in the progeny of a cross (Weising *et al.*, 2005). LOD scores ≥ 3, which indicates that the linkage between two markers is 1000 times more likely than no linkage are considered to be strong evidence for linkage and have been used to construct linkage maps (Stam, 1993). If there

are enough markers, the number of linkage groups obtained corresponds to the haploid chromosome number of the organism concerned.

In the present study, an attempt was made to develop genetic linkage map following the pseudo-test cross strategy. Different genetic markers developed in rubber (as indicated in Chapter 1) were used to construct linkage map. Five different marker systems were used in the construction of the linkage maps as selection of the most suitable genetic markers is essential for high quality genetic linkage maps.

2.2 Review of literature

In tree crops, inbreeding depression precludes the genetic analysis of inbred lines and hence it is not possible to utilize testcross or backcross populations regularly for genetic analysis. Instead, "double pseudo-test cross" strategies followed where genetic analysis is performed on both the parents in a controlled cross keeping track on heterozygous loci for each parent. The marker genotype in the F_1 progeny population results from the independent meioses and crossovers in the maternal and paternal parents, thus individual maps are often constructed for each parent (Grattapaglia *et al.*, 1994; Groover *et al.*, 1994). This strategy was successfully adopted in rubber and many other tree crops for the construction of genetic linkage maps.

2.2.1 Genetic linkage maps in *Hevea brasiliensis*

Since rubber tree is an out crossing species, it exhibits inbreeding depression which makes it difficult to develop progeny for classical genetic studies. The first saturated genetic linkage map in *Hevea* was constructed by Lespinasse *et al.* (2000) using RFLP, AFLP, microsatellite and isozyme markers. The pseudo-test cross strategy was applied on F_1 progeny of 106 individuals derived from an interspecific cross between PB260, a *H. brasiliensis* cultivated clone and RO38, an interspecific hybrid between *H.brasiliensis* and *H. benthamiana*. A consensus map was generated and markers were assembled into 18 linkage groups. The same population was used for quantitative trait loci (QTL) analysis for South American leaf blight (SALB) resistance in natural conditions (Lespinasse *et al.*, 2000a) and under controlled infestation of the fungal pathogen, *M. ulei* (Le Guen *et al.*, 2003) detecting eight QTLs mainly on the resistant clone RO38. Feng *et al.* (2010) constructed a genetic linkage map in a F_1 segregating population of 94 progenies from the cross between Reyan 88-13 x IAN873 using 91 EST and genomic SSR markers. The map comprised of 18 linkage groups covering a genetic distance of 937.06 cM with an average of 21.29 cM between adjacent markers. The largest linkage group consisted of 16 marker loci, while the smallest one contained only 2 marker loci. Souza *et al.* (2011) developed a genetic linkage map based on a mapping population derived from a controlled cross between the cultivars PB 217 and PR 255 with genomic and genic SSR markers. Twenty three linkage groups (LG) were obtained spanning over 2,471.2 cM in length with an average genetic distance of 11 cM between adjacent markers using the OneMap software (Margarido *et al.*, 2007). An integrated parental

linkage map was developed by Triwitayakorn *et al.* (2011) using novel EST–SSRs through transcriptome sequencing along with other microsatellite markers available in the GenBank. The analysis was based on a F_1 mapping population, using the double pseudo-test cross strategy by the software JoinMap 3.0. The markers were distributed on 23 linkage groups and covered 842.9 cM with a mean interval of 11.9 cM. QTL mapping for SALB resistance was performed by Le Guen *et al.* (2011) in a population generated from the crossing of a tolerant clone MDF 180 and a susceptible cultivar using AFLP and SSR markers. The experiments were conducted both under natural infestations and controlled conditions and in the resistant parent a major QTL was identified in the linkage group LG-15.

2.2.2 Genetic linkage map in economically important crop of family Euphorbiaceae

The Euphorbiaceae family is a large and diverse family of flowering plants. It includes several economically important plants of the world including rubber tree (*H. brasiliensis*) the main resource of natural rubber, cassava (*Manihot esculenta*) a primary staple food and industrial crop and physic nut (*Jatropha curcas* L.) a high oil content crop with important applications in biodiesel production.

In cassava, the first genetic linkage map was constructed with RFLP, RAPD, SSR and isozyme markers in a full sib intraspecific cross with 20 linkage groups (Fragene *et al.*, 1997) spanning 931.6 cM. Mba *et al.*, (2001) enriched an existing frame work genetic linkage map of RFLP markers with newly characterized SSR markers. Finally 18 LGs were enriched which indicated a broad coverage of the cassava genome and also led to the joining of a few small groups and the creation of one new group. Marker assisted selection (MAS) was performed by Akano *et al.* (2002) using the Bulk Segregant Analysis (BSA) strategy to identify SSR markers linked to cassava mosaic disease (CMD) resistance gene. Consequently two CMD resistance genes designated as CMD1 and CMD2 were placed on the map (Fregene *et al.*, 2001; Akano *et al.*, 2002). The first SSR based molecular genetic map was constructed in cassava in an F_2 population (Okogbenin *et al.*, 2006) and the mapping was performed using MAPMAKER spanning a total genetic distance of 1236.7 cM. A pseudo-test cross population was used in cassava for genetic linkage map construction using SSR markers and subsequent QTL analysis for early yield was conducted by Okogbenin *et al.* (2008). Colorado *et al.* (2009) performed QTL analysis for total carotene content in cassava with a reference linkage map using SSR markers. Kunkeaw *et al.* (2010) generated a genetic linkage map of

cassava using AFLP and SSR markers in segregation analysis of F_1 progenies. Another linkage map of cassava was generated with AFLP, SSR and sequence-related amplified polymorphisms (SRAPs) markers by Chen *et al.* (2010), and 18 linkage groups are identified. Recently another genetic map with 20 linkage groups was generated by Kunkeaw *et al.* (2011) using EST and genomic SSR markers covering 1,178 cM, with an average distance of 5.6 cM between the markers. EST-SSR markers from an enriched library was used for linkage map construction by Sraphet *et al.* (2011) in a F_1 population encompassing 1,420.3 cM, distributed on 23 linkage groups with a mean distance of 4.54 cM. between markers.

In *Jatropha* a first-generation linkage map comprising of 11 linkage groups was constructed using EST-SSR markers and SNPs in backcross population (Wang *et al.*, 2011). Liu *et al.* (2011) performed QTL analyses for the traits, fatty acid composition, total oil content and its expression levels with the help of a gentic linkage map of 95 DNA markers in *Jatropha.*

2.2.3 Genetic linkage maps in other crops

Genetic map was constructed by Chaparro et al (1994) in peach using an intraspecific F2 population. This map consisted of 83 RAPDs, one isozyme and four morphological traits. Hemmat *et al.* (1994) constructed a genetic linkage map in apple based on the pseudo-test cross strategy. Separate maps were constructed for each parent using isozymes and RAPD markers which assembled into 24 and 21 LGs for the female and male parents respectively. The pseudo-test cross mapping strategy was followed by Kenis and Keulemans (2005) for the development of genetic linkage maps for two apple cultivars using AFLP and SSR markers. Two separate maps were obtained for each parent with 17 LGs. The total map length was 1039 cM for the male parent and 1245 cM for the female parent. Celton *et al.* (2009) constructed a framework genetic map in a cross between two apple rootstocks with SSR, sequence-characterized amplified regions (SCARs), SNP and RAPD markers with 17 linkage groups. A high-density genetic map of apple was constructed recently by Han *et al.* (2011) using SSR and STS markers.

Joobeur et al. (1998) constructed the first saturated linkage map for *Prunus* using an almond x peach F_2 progeny with isozymes and RFLPs covering a distance of 491 cM with the average map density of 2.0 cM/marker which was accepted as a reference map for *Prunus* species. The first genetic linkage map for apricot using backcross population

that segregates for resistance to PPV (plum pox virus) was constructed by Lalli *et al.* (2008) and the BSA strategy was applied for QTL identification.

Genetic maps of *Vitis* were constructed from an interspecific hybrid population using RAPD markers exploiting the double-pseudotestcross strategy generating 20 LGs (Dalbó *et al.*, 2000). Gokabayarak *et al.* (2006) constructed a genetic linkage map in grape using 114 RAPD markers. QTL analysis was also performed for flowering time and resistance to powdery mildew. A framework linkage map based on microsatellite markers was developed for *Vitis* using EST SSR markers. The map consisted of 20 linkage groups spanning over 1,728 cM and with an average distance of 11 cM. between markers (Riaz et al.,2004). This map has been adopted as a reference map for the International Grape Genome Program.

A genetic linkage map obtained from a complex intergeneric cross of *Citrus* was constructed with 111 RAPD markers. The map consisted of 63 markers distributed into nine linkage groups with an average map distance of 5.07 cM between markers (Çevik and Moore, 2007). De Oliveira *et al.* (2007) used fluorescent amplified fragment length polymorphism (fAFLP) markers in a backcross population of *Citrus* for map construction.

A linkage map for European hazelnut (*Corylus avellana* L.) was constructed using RAPD and SSR markers. Eleven groups were identified for each parent, corresponding to the haploid chromosome number of hazelnut (Mehlenbacher *et al.*, 2006). Viruel *et al.* (1995) constructed the first map for almond using RFLPs in a cross between two almond varieties and eight linkage groups were constructed for both the parents. Using the pseudo-test cross strategy, individual tree genetic-linkage maps have been constructed in *Eucalyptus grandis* and *E. urophylla* using RAPD markers (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996) and in *E. globulus* and *E. tereticornis* based on AFLP markers (Marques *et al.*, 1998). QTL analysis for wood, fibre and floral loci traits was performed for *E. globulus* based on the segregation of 249 codominant loci in an outbred F1 population of 148 individuals (Thamarus *et al.*, 2002). A linkage map for *E. camaldulensis* was constructed using segregation data from a 92 full-sib progeny based on a two-way pseudo-test cross strategy. Linkage analysis resulted in 168 markers covering 1236 cM of the genome. An integrated genetic linkage map for *E. nitens* was constructed with 210 RFLP, 125 RAPD and 4 isozyme loci resulted in 12 linkage groups covering 1462 cM.

The first linkage map of the olive genome has been constructed by la Rosa *et al.* (2003) using RAPD and AFLP, RFLP and SSR markers applying the pseudo test-cross strategy. Twenty two and twenty seven LGs were obtained for each parent respectively spanning 2,765 cM and 2,445 cM. (Khadari *et al.*, 2010) developed a linkage map of the olive with a full-sib population in a two-way pseudo-test cross mapping configuration. Maternal and paternal maps were constructed using AFLP, SSR and ISSR markers. The female map consisted of 36 linkage groups (LGs) covering 2210.2 cM of total genome length with an average marker spacing 11.2 cM and the male map contained 31 LGs and covered a distance of 1966.2 cM with an average distance between two adjacent markers of 10.3cM.

Linkage maps were constructed for *Populus deltoides*, *P. nigra*, and *P. trichocarpa*e species by analyzing progeny of two controlled crosses sharing the same female parent, based on the two-way pseudo-test cross strategy. AFLP markers that segregated in 1:1 ratio were used to construct the four parental maps. SSR and STS markers were used to merge linkage groups within the individual maps. Linkage analysis and alignment of the homoeologous groups resulted in 19 groups for *P. deltoids* and *P. trichocarpa* (Cervera *et al.*, 2001). A nearly complete linkage maps of *P. sylvestris* was constructed using AFLP markers in a full-sib family. In the maternal parent, 188 markers were mapped in 12 linkage groups, equivalent to the *Pinus* haploid chromosome number, with a total coverage of 1,695.5 cM. In the paternal parent, 245 markers established a map with 15 linkage groups, spanning a genome length of 1,718.5 cM. (Yin *et al.*, 2003). A genetic linkage map was constructed in tea using the pseudo-test cross approach using AFLP and RAPD markers (Hackett *et al.*, 2000). The linkage was estimated using JoinMap 2.0

The first molecular linkage map generated by Hendre *et a.l* (2007) using pseudo-test cross strategy for coffee with RAPD, AFLP and SSR markers in F_1 mapping population. The linkage map comprised of eleven major linkage groups and the total map length was1230.8 cM with an average marker spacing of 3.57 cM.

The first genetic linkage map of European beech was constructed by Scalfi *et al.* (2004) with RAPD, AFLP and SSR markers using both MAPMAKER and JoinMap 3.0 softwares Two parental maps were constructed adopting a "two-way pseudo-test cross" mapping strategy. In the male map 119 markers could be clustered in 11 major groups (971 cM), while in the female map 132 markers were distributed in 12 major linkage groups (844 cM). QTL analysis was also performed in the same pedigree.

A single-tree genetic linkage maps of European larch and Japanese larch (Arcade *et al.*, 2000) was constructed using AFLP, RAPD and ISSR loci in a progeny population 'F2 backcross' showing a test cross configuration. The maternal parent map consisted of 117 markers partitioned within 17 linkage groups (1152 cM) and the paternal parent map (*L. kaempferi*) had 125 markers assembled into 21 linkage groups (1206 cM).

2.3 Materials and methods

2.3.1 Mapping population and DNA extraction

Mapping population was developed from a controlled intraspecific cross between two clones of *Hevea brasiliensis* RRII 105 and RRII 118, which comprised of 61 full sib F_1 progenies. The hybridization programme was conducted in 1995 at Central Experimental Station, Chethackal as a part of the crop improvement programme by the Botany Division of Rubber Research Institute of India, Kottayam (Dr. Kavitha K. Mydin, personal communication). The population was maintained at the experimental fields of RRII campus. Total genomic DNA was extracted from fresh young leaves of parents and progeny as described in previous chapter.

2.3.2 Genetic markers used in genotyping of progenies

Five different types of DNA based molecular markers were used to construct the present genetic linkage map. They include random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), microsatellites/simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs) and resistance gene analogue (RGA) markers.

2.3.2.1 Random Amplified Polymorphic DNA (RAPD) markers

A total of 520 'Operon' arbitrary decamer primers (as mentioned in Chapter 1, Table 1.10) were screened for polymorphism between parental clones RRII 105 and RRII 118 (Fig. 2.2). Selected informative primers were used for segregation analysis in the progeny. The amplification reactions and PCR cycling were carried out as per the protocols described in Chapter 1. Only those primers, which were consistent and gave clear amplification patterns were selected for segregation analysis (Table 2.1). The genotyping data were scored and coded according to the coding system suggested by JoinMap 3.0 (Van Ooijen and Voorrips, 2001).

2.3.2.2 Amplified Fragment Length Polymorphism (AFLP) markers

Preamplified products with *Eco*RI+A and *Mse* I+C from two parental clones were used as template for selective amplifications using 78 primer combinations based on *Eco*RI +ANN and *Mse* I+CNN (Fig. 2.3)(Tables 2.2, 2.3). Reaction procedures and thermal cycles were the same as mentioned in Chapter 1. Selected primer combinations were used

Table 2.1. List of 'Operon' primers selected based on polymorphism following screening of the parents (RRII 105 and RRII 118) of progeny population for segregation analysis

Sl.No.	Operon primer	Primer sequences (5'-3')	Total bands	Polymorphic bands
1	OPA - 01	CAGGCCCTTC	5	3
2	OPC - 05	GATGACCGCC	6	1
3	OPE - 03	CCAGATGCAC	7	1
4	OPF - 10	GGAAGCTTGG	4	1
5	OPG - 17	ACGACCGACA	5	2
6	OPH - 03	AGACGTCCAC	8	2
7	OPJ - 19	GGACACCACT	6	2
8	OPO - 12	CAGTGCTGTG	5	2
9	OPO - 15	TGGCGTCCTT	6	1
10	OPX - 03	TGGCGCAGTG	4	1
11	OPX - 07	GAGCGAGGCT	9	2
12	OPX - 17	GACACGGACC	5	3
13	OPX - 18	GACTAGGTGG	7	3
14	OPAA - 07	CTACGCTCAC	6	3
15	OPAA - 10	TGGTCGGGTC	6	3
16	OPAA - 17	GAGCCCGACT	7	2
17	OPAB - 01	CCGTCGGTAG	7	1
18	OPAB - 07	GTAAACCGCC	9	2
19	OPAB - 14	AAGTGCGACC	6	1
20	OPAD - 14	GAACGAGGGT	6	1
21	OPAE - 04	CCAGCACTTC	4	2
22	OPAE - 08	CTGGCTCAGA	6	2
23	OPAI - 02	AGCCGTTCAG	6	1
24	OPAI - 09	TCGCTGGTGT	5	1
25	OPAJ - 11	GAACGCTGCC	8	2
26	OPAJ - 20	ACACGTGGTC	8	5
27	OPAL - 03	CCCACCCTTG	4	2
28	OPAL - 05	GACTGCGCCA	7	2
29	OPAL - 06	AAGCGTCCTC	7	3
30	OPAL - 12	CCCAGGCTAC	8	2
31	OPAM - 10	CAGACCGACC	9	5
32	OPAO - 01	AAGACGACGG	5	2
33	OPAO-19	GTTCTCGGAC	5	2
34	OPAP - 05	GACTTCAGGG	5	2
35	OPAQ - 07	GGAGTAACGG	4	2
36	OPAQ - 12	CAGCTCCTGT	6	3
37	OPAR - 18	CTACCGGCAC	7	1
38	OPAR - 19	CTGATCGCGG	4	3
39	OPAS - 06	GGCGCGTTAG	5	2
40	OPAS - 10	CCCGTCTACC	8	3
41	OPAS - 14	TCGCAGCGTT	6	4
42	OPAT - 20	ACATCAGCCC	4	2
43	OPAU - 10	GGCGTATGGT	9	3
44	OPAV -19	CTCGATCACC	9	4
45	OPAX - 06	AGGCATCGTG	3	2
46	OPAX - 16	GTCTGTGCGG	4	1
	Total		430	100

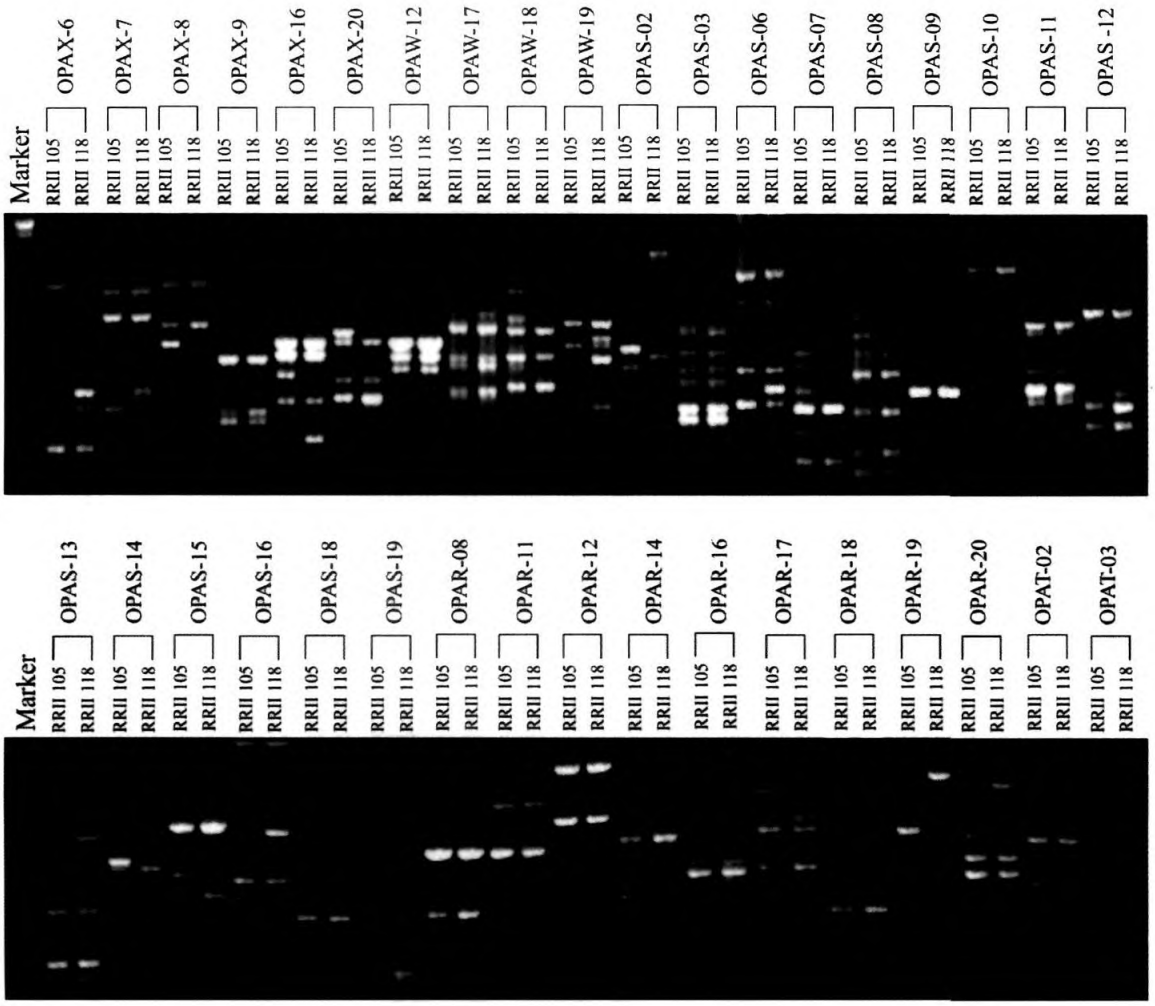


Fig. 2.2. Representative agarose gel photograph showing the primary screening of the random Operon primers in parents (RRII 105 and RRII 118) of the progeny population. The λ DNA/ *Eco* RI + *Hin*d III digest is used as molecular marker

Table 2. 2. Sequence details of adapters and primers for the pre and selective amplification reactions used in the AFLP analysis of progeny population

Adapters	Sequence (5'-3')
Eco R I adapter Oligo-1	CTCGTAGACTGCGTACC
Oligo-2	AATTGGTACGCAGTC
Mse I adapter Oligo-1	GACGATGAGTCCTGAG
Oligo-2	TACTCAGGACTACT
Pre-selective primers	
*Eco*R I –A (E-A)	GACTGCGTACCAATTCA
Mse I- A (M-C)	GATGAGTCCTGAGTAAC
Selective primers	
E-AAC	GACTGCGTACCAATTCAAC
E-AAG	GACTGCGTACCAATTCAAG
E-ACA	GACTGCGTACCAATTCACA
E-ACT	GACTGCGTACCAATTCACT
E-ACC	GACTGCGTACCAATTCACC
E-ACG	GACTGCGTACCAATTCACG
E-AGC	GACTGCGTACCAATTCAGC
E-AGG	GACTGCGTACCAATTCAGG
E-AGT	GACTGCGTACCAATTCAGT
E-AGA	GACTGCGTACCAATTCAGA
M-CAC	GATGAGTCCTGAGTAACAC
M-CAG	GATGAGTCCTGAGTAACAG
M-CAT	GATGAGTCCTGAGTAACAT
M-CTA	GATGAGTCCTGAGTAACTA
M-CTC	GACTGCGTACCAATTCCTC
M-CTG	GACTGCGTACCAATTCCTG
M-CTT	GACTGCGTACCAATTCCTT

'E' and 'M' denotes the adapter sequences of *Eco*R I and *Mse* I respectively

Table 2.3. List of AFLP primer combinations selected based on the number of polymorphic markers generated, following the screening of the parents (RRII 105 anc RRII 118) of the progeny population

Sl. No.	Primer pairs	Total bands	Polymorphic bands
1	E-AGT (C-17)+ M-CTT(MC-5)	40	4
2	E-AGT (C-17)+M-CTC(MC-7)	33	8
3	E-AGC(C-18)+M-CAC(MC-2)	36	4
4	E-AGC(C-18)+M-CTA(MC-6)	34	1
5	E-AGC(C-18)+M-CTC(MC-7)	32	2
6	E-AGC(C-18)+M-CTG(MC-8)	35	3
7	E-AGA(C-19)+M-CAT(M-C-3)	50	5
8	E-AGA(C-19)+M-CTA(MC-6)	57	2
9	E-AGA(C-19)+M-CTC(MC-7)	43	10
10	E-AGG(C-20)+M-CAT(MC-3)	46	3
11	E-AGG(C-20)+M-CAG(MC-4)	37	6
12	E-AAG(EC-7)+M-CAC(MC-2)	45	3
13	E-ACA(EC-5)+M-CAT(MC-3)	1	0
14	E-ACA(EC-5)+M-CAG(MC-4)	30	4
15	E-ACA(EC-5)+M-CTT(MC-5)	44	4
16	E-ACA(EC-5)+M-CTC(MC-7)	31	10
17	E-ACT(EC-9)+M-CAT(MC-3)	38	7
18	E-ACT(EC-9)+M-CTT(MC-5)	37	6
19	E-ACT(EC-9)+M-CTC(MC-7)	22	4
Total		691	86

for segregation analysis in the progeny following the same methods as performed with the parental clones for polymorphism studies.

2.3.2.3 Microsatellite/ SSR markers

2.3.2.3.1 Dinucleotide repeat derived SSR markers

A total of 104 dinucleotide SSR markers were screened for polymorphism between two parents (Fig. 2.4) as described in Chapter 1. Polymorphic microsatellite markers were selected and used for segregation analysis in the progeny population. Protocols are mentioned elsewhere.

2.3.2.3.2 EST derived microsatellite markers

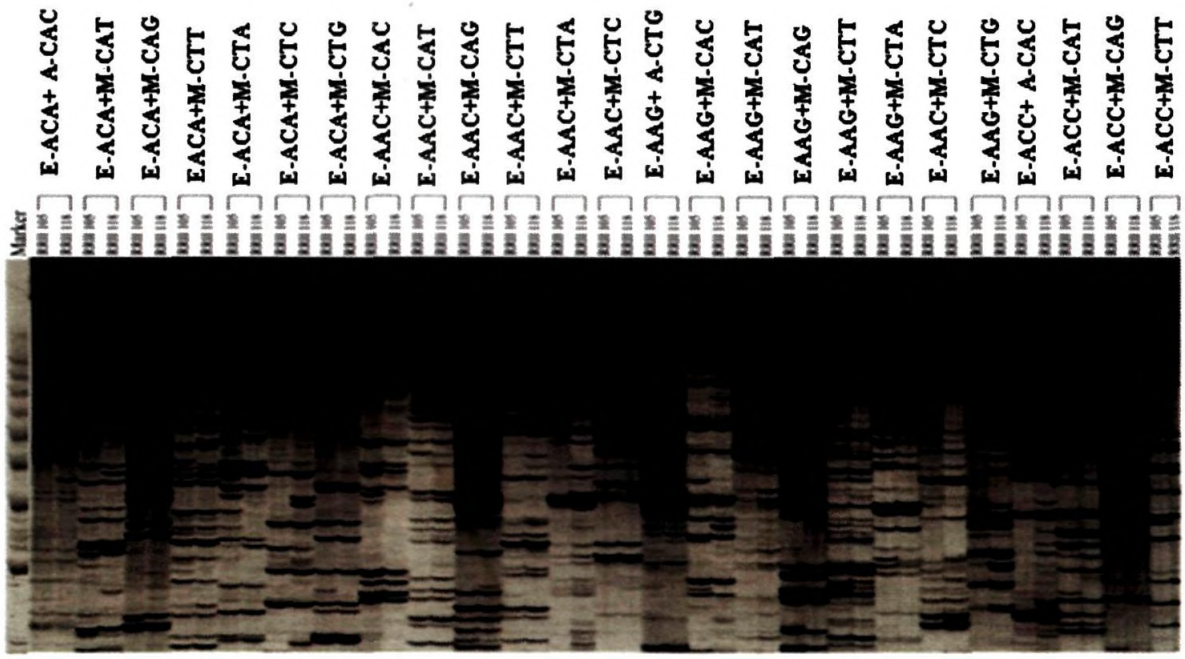
A total of 67 EST derived SSRs for *Hevea* (Feng *et al*, 2009) were subjected to screening for polymorphism between the two parents of the mapping population (Fig. 2.5) and selected primers were used for segregation analysis. Designing, synthesis and screening of primer pairs were mentioned in Chapter 1.

2.3.2.3.3 SSR markers from the genetic linkage map developed by French group

The 'M' series of SSR markers were identified in several linkage groups in the saturated genetic linkage map developed by Lespinasse *et al.* (2000) in *H. brasiliensis*. The details of primers were discussed in Chapter 1. The markers were screened for polymorphisms with the parents along with a few progenies to understand the segregation of alleles. The markers showing simple inheritance pattern were selected for further analysis. PCR reaction set up and PCR thermal cycling programmes for all the microsatellite marker types were the same as mentioned in Chapter 1.

2.3.2.3.4 Trinucleotide derived microsatellite markers

Thirty-one primer pairs were synthesized based on the flanking sequences of repeat bearing positive clones from the enriched trinucleotide repeat genomic library of rubber developed as described in Chapter 1. These primers were assessed for polymorphism between parental clones (Fig. 2.6) before being applied to the entire population to detect the segregation of alleles. Only the markers which were polymorphic and segregating were utilized in the segregation analysis.



ig. 2.3. Representative silver stained PAGE photograph showing screening of AFLP primer ombinations in parents (RRII 105 and RRII 118) of the mapping population. The molecular ıarker used was the 30-330 bp AFLP DNA ladder (Invitrogen)

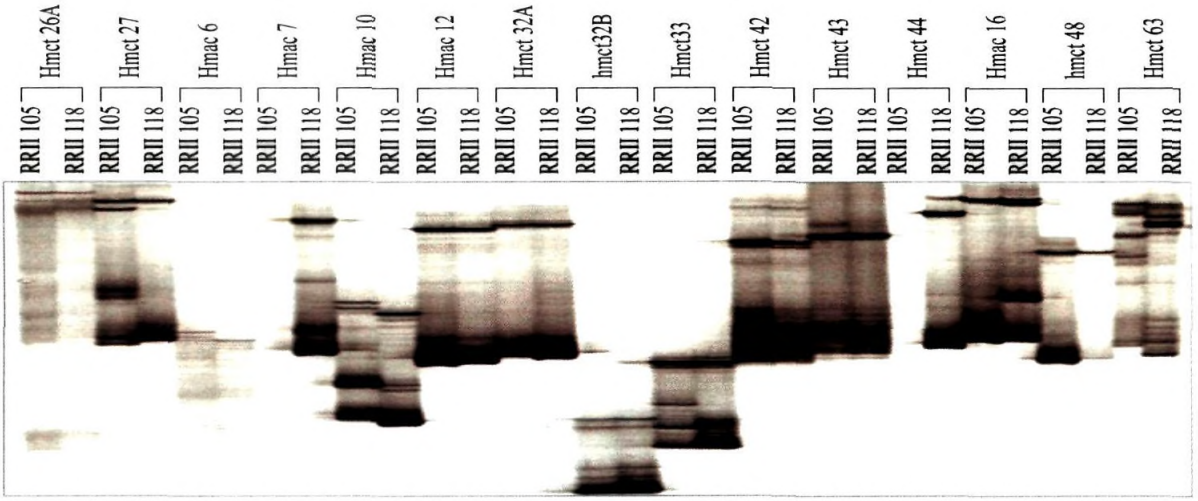


Fig.2.4. Representative silver stained PAGE photograph showing screening of dinucleotide microsatellite/SSR markers in parents (RRII 105 and RRII 118) of the mapping population

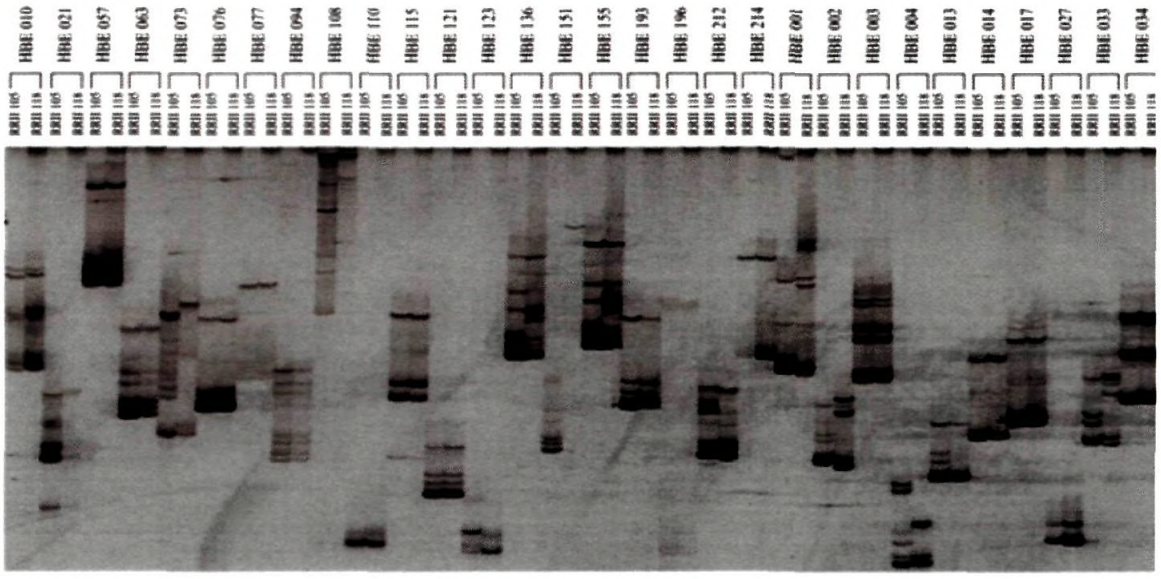


Fig.2.5. Silver stained PAGE photograph showing screening of EST-SSR markers in parents (RRII 105 and RRII 118) of the mapping population

2.3.2.4 Single Nucleotide Polymorphism (SNP)

SNPs were identified from gene sequences as described in Chapter 1 and genotyping of progenies was done using Sequence Tagged Site (STS), allele specific PCR amplification and CAPS markers. The STS marker was developed based on the indels existing in the gene farnesyl diphosphate synthase and genotyping was performed for the same. The allele specific PCR was utilized for the genes, geranylgeranyl diphosphate synthase, mevalonate kinase and latex patatin homolog and the CAPS marker was utilized for the gene ubiquitin precursor. The same PCR amplification reactions and temperature profiling were followed for STS and allele specific primers in the progeny population also as described in the Chapter 1. CAPS marker was used for genotyping following the procedure as mentioned in Chapter 1

2.3.2.5 Resistance Gene Analogue (RGA)

The RGA marker generated with the primer combination RGA 3 + RGA 15 showing polymorphism between parents was used for the segregation analysis in the progeny population. The same PCR conditions and thermal cycling programme, as described in Chapter 1 was followed for the progeny analysis.

2.3.3 Scoring of markers

Both the MAPMAKER/EXP 3.0 (Lincoln *et al.* 1992) and JoinMap 3.0 (Van Ooiijen and Voorrips, 2001) algorithms were used for the construction of the genetic linkage map. Initially only those dominant and codominant markers segregating in a Mendelian ratio of 1:1 were analyzed using MAPMAKER software. Later all the markers with different types of segregation ratios were analyzed with JoinMap software for linkage group assignments of markers.

2.3.3.1 MAPMAKER

In MAPMAKER, the present mapping population was considered as the F_2 backcross and encoded as follows: Code 'A' was given for the homozygote and 'H' was given for the heterozygote. The missing data was indicated by the symbol '-'.

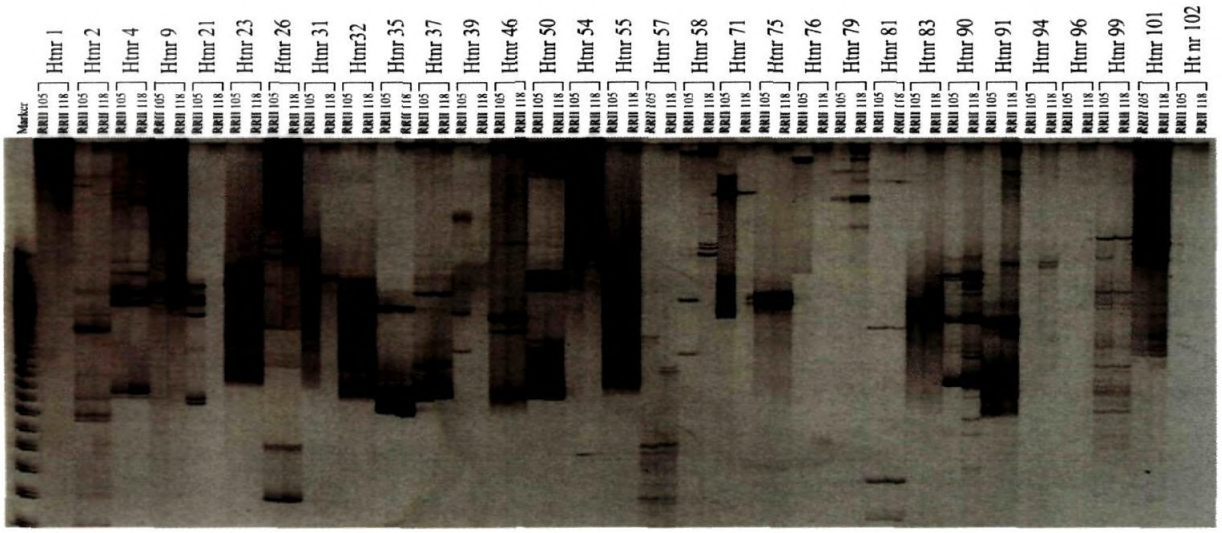


Fig. 2.6. Representative silver stained PAGE photograph showing screening of trinucleotide SSR markers in parents (RRII 105 and RRII 118) of the mapping population. Molecular maker was a 20 bp PCR low ladder (Sigma)

2.3.3.2 JoinMap

Mapping data were obtained by visual scoring of the banding pattern for all types of markers. Only unambiguous and distinct bands were scored. For dominant markers (RAPD, AFLP and RGA), a single band appeared for the presence of the band (both homozygous and heterozygous conditions). Therefore, parental genotypes were determined by the segregation ratio of the progeny, where the heterozygosity of amplicon was expressed as 3:1 ratio and homozygosity as 1:1 ratio in the progeny. In the case of co-dominant markers (SSR and SNP), segregation ratios of 1:1, 1:2:1 and 1:1:1:1 were obtained based on the number of parental alleles. All scorings were performed according to the standard coding systems described in JoinMap 3.0 using different characters to denote the different alleles of the marker loci.

For the dominant markers, the scoring lm x ll was done for the marker present only in the female parent RRII 105, nn x np for the marker present only in the male parent RRII 118 and hk x hk for heterozygous markers in both the parents. For co-dominant markers, four types of scoring were performed. The scoring lm x ll was done for the heterozygous marker present only in the female parent, nn x np for the heterozygous marker present only in the male parent. If both the parents were heterozygous for the same alleles the scoring was given as hk x hk. Initially, for heterozygous parents with three alleles, the scoring was performed as ef x eg. The scoring ab x cd was given for heterozygous parents with four different alleles (Fig. 2.7). The missing genotypes were indicated using the symbol '--'. All the score types, ab x cd and ef x eg which maintained segregation ratios of 1:1:1:1 were reduced to the segregation type 1:1 with the codes given as lm, ll, nn and np in accordance with the genotypes of each progeny following the instructions of JoinMap format.

2.3.4 Preparation of data files for linkage analysis

2.3.4.1 MAPMAKER

Separate data files were prepared for both the parents and each file contained the markers specific to that particular parent. The raw data file was prepared in text file with the above mentioned codes for the genotypes. The first line of the data file indicated the type of the data as F_2 backcross. The second line in the raw file consisted of three numbers, the first number representing the number of progeny, second number indicating the number of loci and the third for any quantitative characters. Out of the 61 progenies analyzed for marker segregation, one progeny was found to be off type (not the true progeny), which

	Parent genotypes	Progeny genotypes	Segregation ratio	Type of segregation	Locus
Single band	—	—	1:1	lm x ll, nn x np	Dominant
	— —	—	3:1	hk x hk	
Alleles One allele	— — —	— — —	1:1	lm x ll, nn x np	Co-dominant
Two alleles	— — — —	— — — —	1:2:1	hk x hk	Co-dominant
Three alleles	— — — —	— — — — — — —	1:1:1:1	ef x eg	Co-dominant
Four alleles	— — — —	— — — — — — — —	1:1:1:1	ab x cd	Co-dominant

Fig. 2.7. Schematic representation of the segregation of dominant and co-dominant markers of a pseudo- test cross progeny of the cross RRII 105 x RRII 118

showed abnormal allelic inheritance of SSR markers. Hence, this individual was omitted from further linkage analysis and finally 60 progenies were considered.

2.3.4.2 JoinMap

Simple text files were used for the JoinMap data file preparation called as 'locus genotype files'. Two separate data sets were prepared for each parent. Each parental file contained the markers specific to the respective parents and the bridge markers common to both the parents.

2.3.5 Linkage analysis and map construction

The markers generated were separately analyzed using MAPMAKER and JoinMap software.

The data analysis was initiated with the 'prepare data' command by which the MAPMAKER can extract the details given in the header row. The 'sequence' command was used for 'two-point' data analysis and markers were associated into linkage groups with the 'group' command. The most likely order of markers were analyzed using the 'compare ' command and the map was generated with the best order of markers using the 'map' command. Other markers were added with the 'try' command into the known order of five markers one by one. After adding all the markers, the 'map' command was given to see the complete map. The marker groups were determined using a maximum likelihood distance of 40 and a minimum LOD score of 2.0. The Kosambi mapping function was (Kosambi, 1994) was used to convert the recombination frequencies into genetic distances in centimorgan (cM).

In JoinMap separate data sets were organized for each parent: one for the female parent RRII 105 and another for the male parent RRII 118. Individual genotype frequency was calculated for all individuals with the available markers, from which individuals missing for large number of markers were excluded from the linkage analysis to avoid any probable distortion. Chi-Square test for goodness of fit ($P \geq 0.05\%$) to the expected segregation ratios was performed for all markers so that the highly distorted markers could be initially excluded from the data set. The similarity of loci was determined using the similarity of loci tab-sheet, where we could exclude one of the loci, showing similarity value 1, since identical loci will map exactly at the same position. The same Kosambi mapping function was used for the calculation of genetic distances in

centimorgan (cM). Marker groups were determined using a minimum LOD score of 3.0 and a recombination frequency of 0.4 and the other parameters kept as default.

2.3.6 Map integration

An attempt was made to integrate those linkage groups having markers common between them, following the ratios of 3:1, 1:2:1 and 1:1:1:1 using the map integration function. The original map data were compiled for graphical presentation of linkage map using MapChart 2.2 (Voorrips, 2002). Map chart reads the linkage information from text files.

2.3.7 Cluster analysis of the progeny population along with the parents

The SSR marker data (52 markers), scored for the segregation analysis was used for the cluster analysis also to determine the genetic distance between the parents and the population. The data obtained were scored as binary unit character as presence or absence (1 and 0) of bands. Genetic relationships were inferred by Treecon 1.3B software (Van de Peer and de Wachter, 1997) using unweighted paired group method with arithmetic averages (UPGMA) cluster analysis.

2.4 Results

2.4.1 Segregation analysis of genetic markers

2.4.1.1 Random Amplified Polymorphic DNA (RAPD)

Among the 520 random arbitrary primers, 46 primers were selected based on the polymorphism between the parents and its reproducibility. These primers provided altogether 100 scorable polymorphic fragments with an average of 2.1 markers per primer. Out of the of 100 RAPD markers obtained, 79 markers followed a segregation ratio of 1:1 and 21 markers followed a segregation ratio of 3:1 (Fig. 2.8). Forty markers were found to be originated from RRII 105 (50.6%), 39 markers from RRII 118 (49.3 %) following a segregation ratio of 1:1 and the remaining 21 markers were intercross markers, common to both the parents expected to follow the segregation ratio of 3:1. The list of segregating loci generated are given in Table 2.4.

2.4.1.2 Amplified Fragment Length Polymorphism (AFLP)

Out of 78 primer combinations, 19 combinations were used for the progeny analysis based on their ability to produce large number of polymorphic markers. Only highly reproducible markers with a strong intensity were retained for mapping. A total of 86 markers, giving an average of 4.7 markers per primer combination were recorded on the mapping population and out of these, 38 (44%) were found to be specific to RRII 105 and 31 (36 %) were specific to RRII 118. Seventeen markers were intercross markers, common to both the parents (Fig. 2.9). Sixty-nine markers following a segregation ratio of 1:1 and 17 following a ratio of 3:1 were selected. Details of AFLP markers generated are given in Table 2.5.

2.4.1.3 Microsatellite marker analysis

2.4.1.3.1 Dinucleotide SSR markers

Among the 104 dinucleotide microsatellites markers screened for polymorphism, 27 were selected and scored for segregation analysis in the progeny. Among these, 15 markers followed the segregation ratio of 1:1, four followed a segregation ratio of 1:2:1 and eight followed the segregation ratio of 1:1:1:1. Out of the 15 markers showing 1:1 segregation, 11 were specific to the female parent RRII 105 and four specific to the male parent RRII 118. Out of the eight markers with a segregation ratio of 1:1:1:1, six were heterozygous generating three alleles and two were heterozygous generating four alleles (Table 2. 6).

Table 2.4. List of RAPD markers generated through the segregation analysis of the progeny population. Markers are designated as Operon primer name used in amplification followed by the fragment size in base pairs

Sl.No.	lmxll		nnxnp		hkxhk
1	OPA-01-1332	1	OPA-01-1966	1	OPA-01-2378
2	OPG-17-588	2	OPC-05-877	2	OPF-10-564
3	OPJ19-563	3	OPE – 03 – 506	3	OPH3-1375
4	OPAB07-589	4	OPG17-661	4	OPAA07-1037
5	OPAB07-512	5	OPH3-1025	5	OPAA07-569
6	OPAJ11-788	6	OPAA07-1558	6	OPAI2-1488
7	OPAJ20-1170	7	OPAB14-1073	7	OPAJ20-1817
8	OPAJ20-498	8	OPAJ11–2482	8	OPAA10-670
9	OPAA10-954	9	OPAJ20-1096	9	OPAA10-1048
10	OPAA17-595	10	OPAJ20-986	10	OPX18-523
11	OPX7-1143	11	OPAA17-1728	11	OPAL-12-613
12	OPX17-793	12	OPX3-1058	12	OPO12-1401
13	OPX-17-1090	13	OPX7-1000	13	OPAL-3-2346
14	OPAL5-881	14	OPX-17-947	14	OPAP-5-2957
15	OPAL5-305	15	OPAL-6-1253	15	OPAP-5-2164
16	OPX18-2269	16	OPAL-6-1110	16	OPAX-16-788
17	OPAL-6-1029	17	OPAL-12-2300	17	OPAU-10-1008
18	OPO15-1232	18	OPO12-1558	18	OPAV-19-934
19	OPAL-3-1180	19	OPJ19-886	19	OPAS-10-2660
20	OPAM-10-1153	20	OPAI9-1048	20	OPAS-14-831
21	OPAM-10-998	21	OPAM-10-1904	21	OPAX-6-2319
22	OPAE-4-578	22	OPAM-10-1359		
23	OPAO-1-1360	23	OPAM-10-1216		
24	OPAO-1-1021	24	OPAD-14-1335		
25	OPAQ-7-422	25	OPAE-4-688		
26	OPAR-18-1152	26	OPAR-19-1813		
27	OPAR-19-1322	27	OPAR-19-327		
28	OPAV-19	28	OPAS-6-820		
29	OPAS-10-1438	29	OPAV-19-518		
30	OPAS-14-1729	30	OPAV-19-1258		
31	OPAS-14-1197	31	OPAS-10-1109		
32	OPAQ-12-914	32	OPAS-14-752		
33	OPAQ-12-778	33	OPAS-6-809		
34	OPAT-20-1015	34	OPAX-6-895		
35	OPAT-20-692	35	OPAQ-12-697		
36	OPA0-19-1508	36	OPX-18-1077		
37	OPA0-19-1139	37	OPAE-08-1400		
38	OPAB1-531	38	OPAE-08-1584		
39	OPAU-10-1935	39	OPAQ-07-644		
40	OPAU-10-483				

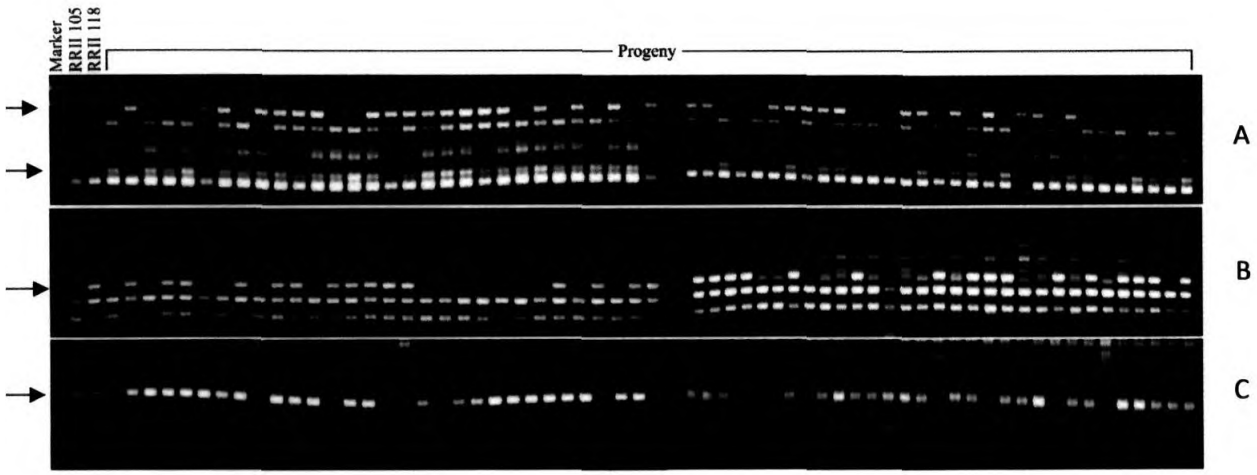


Fig. 2.8. Representative ethidium bromide (EtBr) stained agarose gel photograph showing segregation of RAPD markers among mapping population. A) Marker specific to the parent RRII 105 (allelic configuration: lm), generated with primer OPA-01 segregating in 1:1 ratio among the progenies, B) Marker specific to the parent RRII 118 (allelic configuration: np) generated with the primer OPC-05 segregating in 1:1 ratio among the progenies, C) Marker common to both the parents (allelic configuration: hk x hk) developed with primer OPF-10 segregating in 3:1. Arrowheads indicate the segregating bands. λ DNA/ *Eco* RI + *Hin*d III digest was used as molecular marker

Table 2.5. List of AFLP markers used in the segregation analysis of the progeny population along with their segregation types

Sl.No.	RRII 105 (lmxll)	Sl.No.	RRII 118 (nnxnp)	Sl.No.	Both parents (hkxhk)
1	EC-5+MC7-2	1	EC-5+MC-5-5	1	EC-9+MC-5-2-680
2	EC-5+MC-7-5	2	EC-5+MC-7-1	2	EC-9+MC-5-5-141
3	EC-5+MC-7-7	3	EC-5+MC-7-3	3	EC-9+MC-5-6-105
4	EC-5+MC-7-9	4	EC-5+MC-7-4	4	EC-9+MC-3-5-168
5	EC-5+MC-7-10	5	EC-5+MC-7-6	5	EC-9+MC-7-1-310
6	C-18+MC-2-2	6	EC-5+MC-7-8	6	EC-9+MC-7-2-248
7	C-18+MC-2-3	7	EC-5+MC-4-3-175	7	EC-9+MC-7-4-180
8	C-17+MC-7-3	8	C-18+MC-2-1	8	EC-5+MC-4-1
9	C-17+MC-7-4	9	C-18+MC-2-4	9	C-17+MC-7-8
10	C-17+MC-7-6	10	C-17+MC-7-1	10	C-17+MC-7-2
11	C-17+MC-5-4	11	C-17+MC-7-5	11	C-17+MC-5-1
12	C-19+MC-6-1	12	C-17+MC-7-9	12	C-17+MC-5-3
13	C-19+MC-6-4	13	C-17+MC-5-2	13	C-19+MC-7-4
14	EC-7+MC-2-1	14	EC-7+MC-2-3	14	C-19+MC-7-7
15	EC-5+MC-5-1	15	C-19+MC-7-3	15	C-19+MC-3-2-380
16	EC-5+MC-5-2	16	C-19+MC-7-5	16	C-19+MC-3-4-280
17	EC-5+MC-5-3	17	C-19+MC-7-6	17	C-20+MC-4-3-200
18	EC-7+MC-2-2	18	C-19+MC-7-9		
19	C-19+MC-7-1	19	C-18+MC-7-1		
20	C-19+MC-7-2	20	C-18+MC-7-2		
21	C-19+MC-7-8	21	C-18+MC-6-1		
22	C-19+MC-7-10	22	C-19+MC-3-3-370		
23	C-18+MC-8-1	23	C-19+MC-3-5-135		
24	C-18+MC-8-2	24	C-20+MC-3-3-210		
25	C-18+MC-8-3	25	C-20+MC-4-1-800		
26	C-20+MC-3-2-260	26	EC-9+MC-3-2-300		
27	C-20+MC-4-2-310	27	EC-9+MC-3-6-142		
28	C-20+MC-4-5-138	28	EC-5+MC-4-2		
29	C-20+MC-4-6-110	29	EC-9+MC-7-3-238		
30	EC-9+MC-5-1-720	30	C-20+MC-4-4-150		
31	EC-9+MC-5-3-210	31	EC-9+MC-3-1-500		
32	EC-9+MC-5-4-143				
33	EC-8+MC-3-1-270				
34	EC-9+MC-3-3-295				
35	EC-9+MC-3-4-215				
36	EC-5+MC-4-120				
37	C-19+MC-3-1-470				
38	C-20+MC-3-1-350				

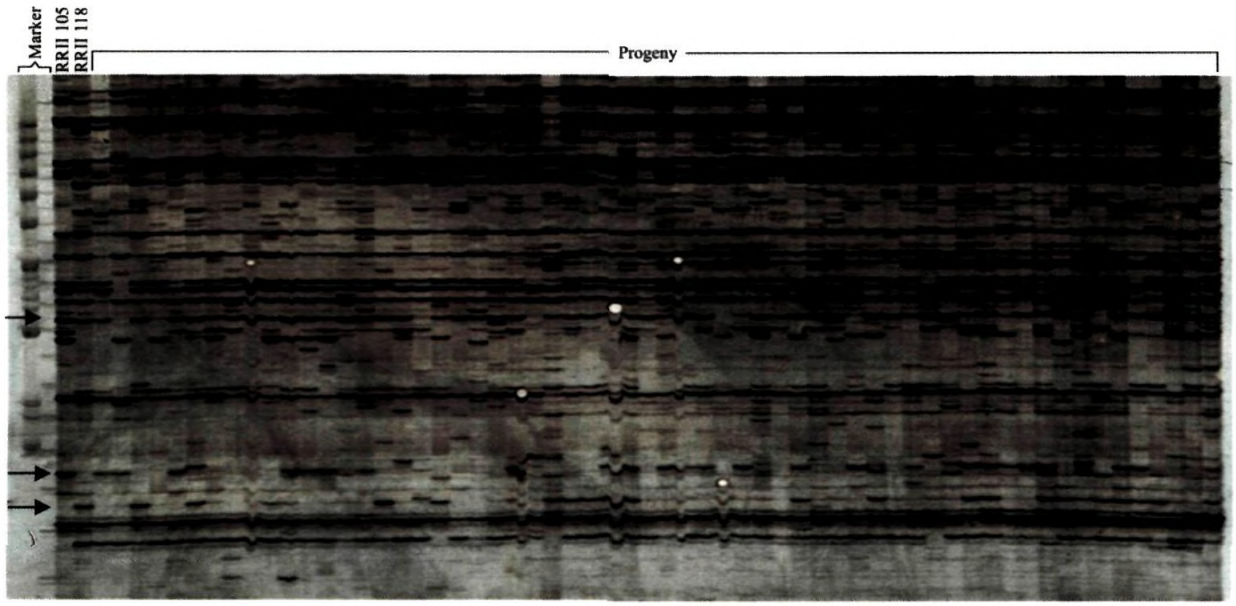


Fig. 2.9. Representative silver stained PAGE photograph of AFLP markers showing segregation among mapping population, primer combination E-AGT+M-CTC was used to generate the AFLP profiles. a) Marker specific to the parent RRII 118 (allelic configuration: np), segregating in 1:1 ratio among the progenies, b) Marker specific to the parent RRII 105 (allelic configuration: lm), segregating in 1:1 ratio among the progenies and c) Marker common to both the parents (allelic configuration: hk x hk), segregating in 3:1 ratio among the progenies. Arrowheads represent the segregating AFLP markers/bands. Molecular weight marker is 30-330-bp DNA ladder (Invitrogen)

2.4.1.3.2 EST derived microsatellite markers

Twenty-five EST derived SSR markers were selected following screening as described in the previous chapter for segregation analysis in the progeny. Out of these 25 markers, 17 markers segregated in 1:1 ratio, four markers followed a segregation ratio of 1:2:1 and four markers segregated in 1:1:1:1 ratio. Out of the 17 markers segregating in 1:1 ratio, eight markers were specific to the female parent and nine markers were specifically heterozygous in the male parent. Four markers segregated in 1:2:1 ratio with common alleles in both the parents. Out of the five markers segregating in 1:1:1:1 ratio, four markers were heterozygous generating three alleles with the configuration of ef x eg and one marker was heterozygous generating four alleles with a configuration of ab x cd (Table 2.6).

2.4.1.3.3 SSR markers from the genetic linkage map developed by French group

The 'M' series of markers were identified in several linkage groups in the saturated genetic linkage map developed by Lespinasse *et al.* (2000) in *H. brasiliensis*. In order to link the SSR markers generated by the French group with our markers, primer pairs were designed based on sequences from the GenBank for twelve 'M' series microsatellite markers, details of which are given in previous chapter. Among these, only two markers were selected for segregation analysis in the progeny and only the marker (M 574) was included in the mapping data, which could be scored unambiguously. The marker was heterozygous in both the parents generating three alleles (ef x eg), resulting in a 1:1:1:1 segregation ratio (Table 2.6).

2.4.1.3.4 Trinucleotide microsatellite markers

Two trinucleotide repeat markers (Htnr-21 and Htnr-90) were selected for progeny analysis from 31 markers developed through the screening of an enriched library. Out of these, the marker Htnr-21 segregated in 1:1ratio and the marker Htnr-90 amplified two loci with segregation ratios of 1:2:1and 1:1 respectively (Table 2.6; Fig. 2.10).

A total number of 52 SSR markers segregating at 1:1, 1:2:1 and 1:1:1:1 ratios were used for the construction of the present linkage map. Representative gel photographs of SSR marker segregation profiles are given in Fig. 2.11.

Table 2.6. Details of microsatellite markers selected along with their parental types used in segregation analysis in the progeny

Sl. No.	Clone	GenBank accession No.	SSR Motif	Parental genotype	Expected Segregation ratio	Forward primer	T_m (°C)	Reverse primer (5`-3`)	T_m (°C)
1	HBE-010	EC609548	(AG)14	ab x cd	1:1:1:1	GGTGAAACTTCGCACGC	55.96	TTTTGCTGCCTGTTGTCG	53.68
2	HBE-063	EC607362	(GA)16	hk x hk	1: 2: 1	GGTTTTGGGTTTAACAGAGATC	56.52	GAGGAAAAGAAGAGTGGCTATG	58.39
3	HBE-115	EC605557	(AT)10	nn x np	1: 1	ACAACCAGGCGTTCTTCAA	56.52	ATACACGGGAATACCATTTGAT	58.39
4	HBE-001	EC609907	(CTT)13	nn x np	1: 1	TATTTGCTGTTGGAGGAAGAA	53.96	CACCAATCAATAATCACCAAGA	54.66
5	HBE-002	EC609891	(TCT)8	nn x np	1: 1	AACTTCTTGTTTGCTTCCCTT	53.96	CGGGCGATACAAGCAGTC	58.24
6	HBE-003	EC609774	(TTC)6	hk x hk	1: 2: 1	AGCCAGTCCCTTGTCTTTCC	59.34	TGTTTGGCTTTCTTTAGTTTTTAC	54.17
7	HBE-033	EC608405	(TTC)9	ef x eg	1: 1: 1: 1	TTCATCAATCATCTTGCCTTTT	52.80	CCCAAAACTGAGCTTGAAAGT	55.91
8	HBE-034	-	(TTC)6	hk x hk	1: 2: 1	AGCTTCTGCATGGGTTTGG	56.66	CTAAATCCAGGAGGCAGTCTAA	58.39
9	HBE-043	EC608110	(CTT)8	ef x eg	1: 1: 1: 1	GCCAAATGGGTCGTCCTT	55.96	AAAATGGAGCCCCAATTCTA	53.14
10	HBE-054	EC607575	(TC)6	ef x eg	1: 1: 1: 1	TTCAGATCCACACGCTAAAAT	54	TCTGCCGTTGGCGGTTAT	56
11	HBE-044	(TTC)8	(TTC)8	ab x cd	1: 1: 1: 1	TTAAAGAGCTAGAATTGGGGC	55.91	GCACACCGTTCTTGACAAAAT	55.91
12	HBE-067	EC607289	(GGC)7	hk x hk	1: 2: 1	TCGCACCTCTTTCCCATC	55.96	CAGAAGGCAGAGCGGAGA	58.24
13	HBE-068	EC607281	(GAA)8	lm x ll	1: 1	TCGGGTCACCAACAGCAA	55.96	CTCCCAATCGCCGTATCG	58.24
14	HBE-080	EC606832	(TGC)7	lm x ll	1: 1	GGTGGAAGCATTCTTGTATTTG	56.52	GAGAAGCCTTGATGAACTGGA	57.87
15	HBE-092	EC606292	(ACC)6	lm x ll	1: 1	TTCTTGTGAGCGAGTTGCC	56.66	TGCCGTTACCAGGGACAA	55.96
16	HBE-098	EC606169	(GGA)6	nn x np	1: 1	GTGGTTGTGGTGGTTCTGG	58.82	CCTGGGACCATCACCTTG	58.24
17	HBE-101	EC606085	(TTA)7	lm x ll	1: 1	CCAGAAACTGACGATGAAGAA	55.91	AACAAGAGAGGACATTGACCC	57.87
18	HBE-122	EC605312	(ATA)8	ab x cd	1: 1: 1: 1	GATGCCCGATGGACCTCT	58.24	CTACAAAACCCGACCCAATA	55.25
19	HBE-139	EC604443	(TCT)6	nn x np	1: 1	AGTCTTCAGCAACATCATCTTCT	57.08	GCAGTTCTCCATTAGGGTTTAT	56.52
20	HBE-140	-	(CAT)7	ab x cd	1: 1: 1: 1	GGTTTTGGTGCTTGTGAGTATG	58.39	CTTGATGGGAAGTTGGGC	55.96
21	HBE-170	EC602995	(AAG)10	nn x np	1: 1	ATTTGGGTGACCTGTTTGG	54.51	CATGACCTCTTGTACTCGTTGC	60.25
22	HBE-190	EC601511	(CTT)11	ef x eg	1: 1: 1: 1	CCCTTCCTTCTGTCCCTCTC	61.40	GGGTAGATTCTGGAGGTCGG	61.40
23	HBE-192	EC601354	(GAA)9	lm x ll	1: 1	GGTCGCCATCTCATCATAATACT	58.86	CAAAGCCAAGATACTCCACATA	56.52
24	HBE-207	EC600469	(AGA)10	nn x np	1: 1	CCCCTCACCTCCACTACCT	60.98	CGCTTGATTCCCCGATTA	53.68
25	HBE-225	CB376545	(AGA)9	nn x np	1: 1	CCAAAACAAGGGAAATCACA	53.19	GGACCGAGACGCTTGATT	55.96
26	Hmct-1	AY135651	$(CT)_{12}AG(GT)_8$	lm x ll	1: 1	AACCAGAAGGGTGTCATGCT	63.5	GGAATCCCATGACAATCCAC	64.0
27	hmct-16	AY439296	(AG)16	lm x ll	1: 1	CATGCAAATAACGAACCAGA	53.2	TTCCGTTCACTATATCGCTCA	55.9
28	Hmct-19	AY439301	(CA)5(TG)2AAA(AG)13	lm x ll	1: 1	CAGATGGGCATTGGTCTTTT	55.3	TCAGAAAATGGATATGGGGC	55.3
29	Hmct-27	AY439314	(AG)3AC(AG)8	lm x ll	1: 1	TGAGCAACGGAGGAGAGAAC	59.4	AAACACCCAAACCCAATTCA	53.2
30	Hmac-4	AY135656	(CT)17(CA)3(CT)2(GT)16	lm x ll	1: 1	GTTTTCCTCCGCAGACTCAG	63.9	ATCCACCAAA TAAGGCATGA	61.7
31	Hmac-10	AY439311	(TCA)6TCT(TCA)4	lm x ll	1: 1	CAGTTGAGGCAGGTTCCAGT	59.4	GGAGCAGGTAGCAGAGTTGG	61.4
32	Hmac-13	AY962198	(AC)11	hk x hk	1: 2: 1	CATCCCTGCATTCCCCTAAT	64.7	ATGGGATGGGACCGTAACAA	66.1
33	Hmct-53	AY962223	(CT)13	hk x hk	1: 2: 1	GCCCCTCTCCCTCTTCAGAT	65.8	CATTCGGGTGGGAAAAAGAA	65.7
34	Hmac-2	AY439299	(TG)4AG(TG)6	lm x ll	1: 1	AGCATGCTCATGCATCATTT	53.2	TGCTACYCACCCCAACTTCC	59.4
35	Hmct-17	AY439298	$(CT)_6$ $(CA)_{14}(TA)_4$	hk x hk	1: 2: 1	GGAAAATGAAGCTTTAACGG	53.2	AATGAGGACGTGGATGGAAT	55.3
36	Hmct-14D	-		ab x cd	1: 1: 1: 1	ACGCATGTCCACAGACCAAA	66.6	TCTGGTTGCAATTCCATGTCA	66.4
37	Hmct-29	AY962191	(AG)16	ab x cd	1: 1: 1: 1	TCCAGGCTTCAGAATCACCA	65.8	TCAGCTGGAAAACAAAGTTCCA	65.6
38	Hmac-14	AY962202	$(CA)_{13}$	ef x eg	1: 1: 1: 1	CTCCACTTGTGCAAGCGTTC	65.9	GCAATCAAACGTGCCAAGAA	65.9
39	Hmac-17	AY962212	$(CT)_6$ $(CA)_{14}(TA)_4$	ef x eg	1: 1: 1: 1	CGCTGTGCCTTGTTGATCTT	65.2	CACTGGCACGTGAATCCCTA	66.0
40	*Hev*-glu	J Q 650524	$(CTT)_{18}$	ef x eg	1: 1: 1: 1	TAGTTCCCTGCCCTCAAGAA	57.3	TAGTTCCCTGCCCTCAAGAA	59.3

	(glucanase)								
41	Hmct-9	AY439289	$(CT)_{10}GT(CT)_{11}$	ef x eg	1: 1: 1: 1	CCCTTTTCATTTAGCTTCATCC	56.5	GTGTACCTGTTCGGCCTTTA	57.3
42	Hmct-2	AY439287	$(AG)_{11}$	ef x eg	1: 1: 1: 1	AAGTTGTAGCGACAGCAGCA	57.3	CCCTGGGTTTTACTTTTGCC	57.3
43	Hmct-58	AY962229	$(AG)_{10}$	ef x eg	1: 1: 1: 1	GGAGGCAGTGAGCACTGAAA	65.6	TTCTTTGTTGACGTCCCCTTC	65.1
44	Hmct-45	AY962213	$(CTT)_9$	nn x np	1: 1	CGCCTACCTGCAGCTCTCTT	65.8	GAGACCCAGCTGCCAAAAAC	65.8
45	Hmac-16	AY962211	$(TA)_5(CA)_8$	nn x np	1: 1	GAATTTTTATGTGGTAGGCAAA	59.4	GCTGCCATTGATGAATCTGG	65.5
46	Hmct-5	AY135653	$(GAT)_4(GTT)_3(GA)_5A(AG)_{10}$	nn x np	1: 1	ATGTATGTGTGCGCAGGAAG	63.7	CTGTAGTCATGGCAGCAGGA	64.1
47	Hmct-2A-1	AY439287	$(AG)_{11}$	lmxll	1:1	AGCAGGCAGAAAAAGAAA	55.3	GCGTACTAACCTCTCCCTGG	61.4
48	Hmct-2A-2	AY439287	$(AG)_{11}$	lmxll	1:1	AGCAGGCAGAAAAAGAAA	55.3	GCGTACTAACCTCTCCCTGG	61.4
49	Htnr-90 lous 1	FJ 160580	$(ACC)_2ACA(ACC)_8$	hkxhk	1:2:1	ACTGGGCCACTATTGCTATTA	60.6	GGTAGTGGTGGCGACAGTGA	61.4
50	Htnr-90 locus 2	FJ 160580	$(ACC)_2ACA(ACC)_8$	lmxll	1:1	ACTGGGCCACTATTGCTATTA	60.6	GGTAGTGGTGGCGACAGTGA	61.4
51	Htnr- 21	FJ 160557	$(TTC)_9$	lmxll	1:1	GCCTTCAAAACAAAGACCAAAGACC	61.3	TGAAGGGAAGAGGCTGGGTGT	61.8
52	M-574	AF 221706	-	ef x eg	1: 1: 1: 1	GCCTTGTGTCCTCTACTTGTC	60.6	GCTGGTTCACCGATCTACTTTT	58.4

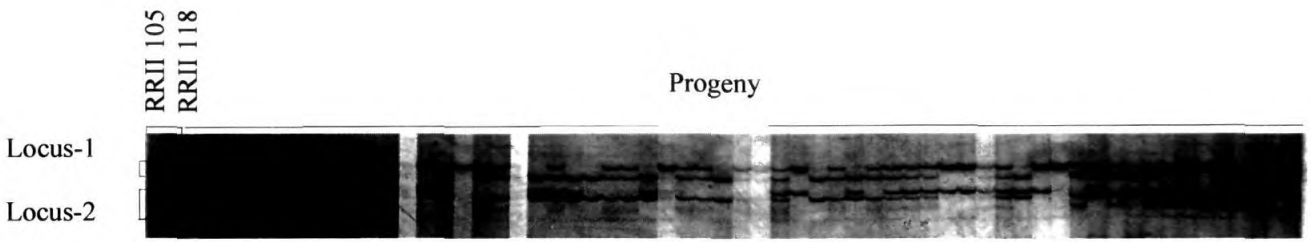


Fig. 2.10. Silver stained PAGE photograph showing the amplification of two loci with the trinucleotide SSR marker Htnr-90. The locus-1 was in hk alleleic configuration, which segregated in 1:2:1 (hk:hh:kk) ratio among the progenies and the locus-2 in np configuration, which segregated in 1:1 (nn:np) ratio, was scored as presence/absence of the band.

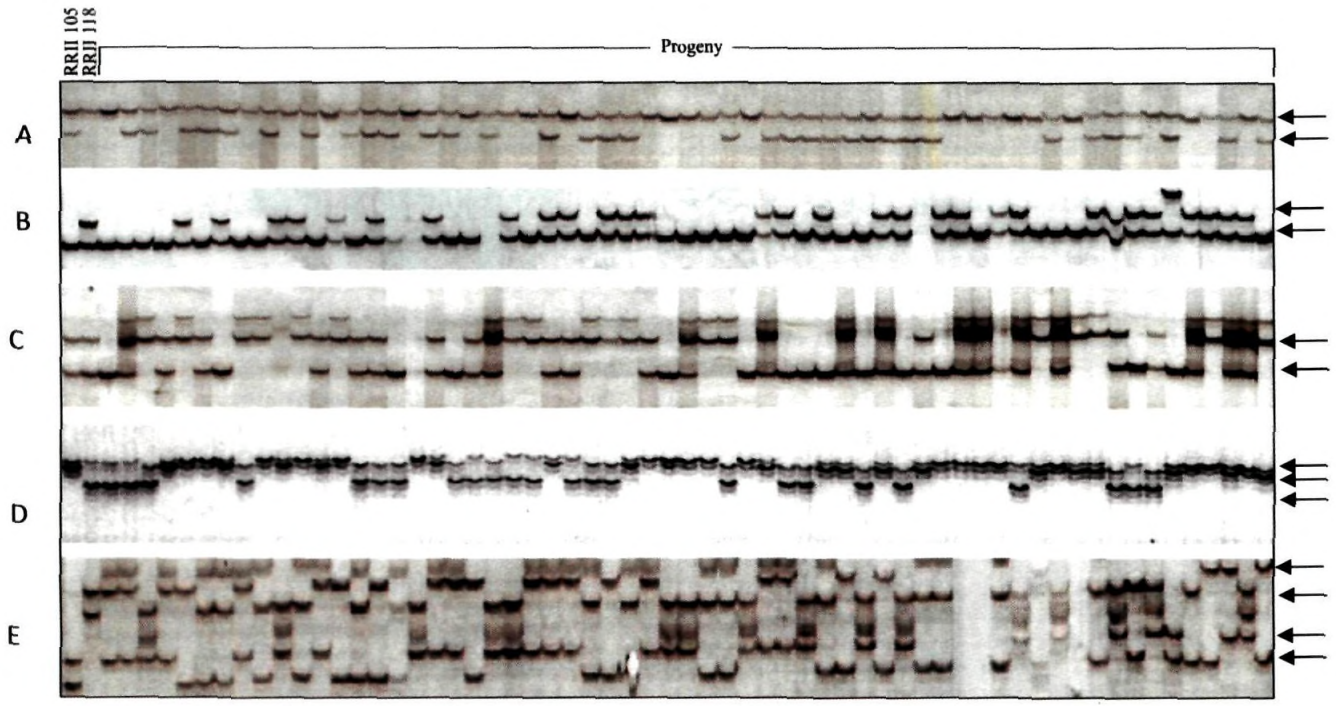


Fig. 2.11. Representative silver stained PAGE photograph showing segregation of SSR markers. A) Hmct-27 locus in allelic configuration 'lm' present in parent RRII 105 segregated in 1:1 (lm:ll) ratio among progenies, B) HBE-207 locus in allelic configuration 'np' present in parent RRII 118 segregated in 1:1 (nn:np) ratio among progenies, C) Hmac-13 locus in 'hk' allelic configuration in both the parents segregated in 1:2:1 (hh:hk:kk) ratio among progenies, D) Hmac-14 locus in 'ef' allelic configurations in parent RRII 105 and 'eg' in parent RRII 118 segregated in 1:1:1:1 ratio (ef:eg:ee:fg) and E) HBE-044 locus in 'ab' allelic configurations in parent RRII 105 and 'cd' in parent RRII 118 segregated in 1:1:1:1 (ab:cd:ac:bd) ratio among progenies. Arrowheads show the segregating alleles

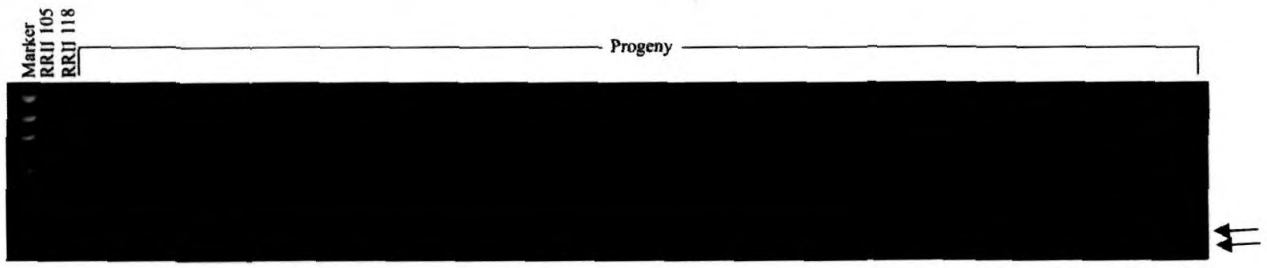


Fig. 2.12. EtBr stained agarose gel photograph showing the segregation of CAPS marker in 1:1 ratio among the progenies, developed from the ubiquitin precursor gene. Presence/absence of the bands like dominant marker was considered in segregation analysis. ΦX 174/ *Hae* III digest was used as molecular weight marker. The arrowheads shows the segregating bands

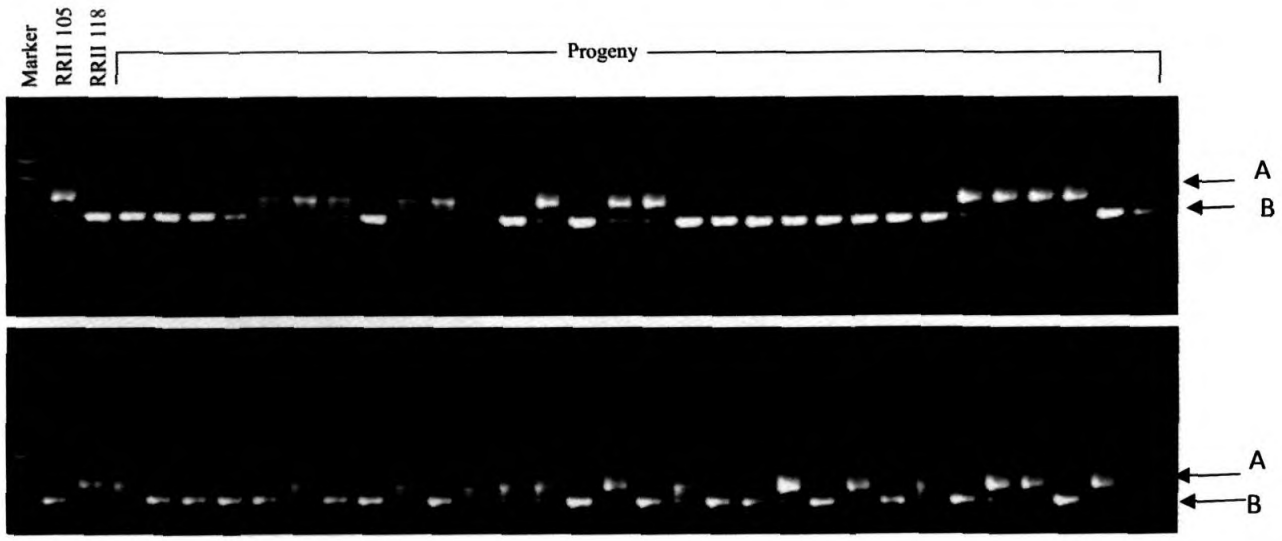


Fig. 2.13. EtBr stained agarose gel photograph showing segregation of the SNP alleles in mapping population, which was generated from the mevalonate kinase gene using allele-specific primer MKF_197C. The SNP detected was in heterozygous state in RRII 105 (C/T) and homozygous state in RRII 118 (C/C). Heterozygosity was represented by two fragments in the segregating progeny (both gene locus and allele specific products) and homozygosity was represented by a single fragment (allele specific only). The arrowhead 'A' shows the gene locus and the arrow head 'B' shows the allele. ΦX 174/ *Hae* III was used as molecular weight marker

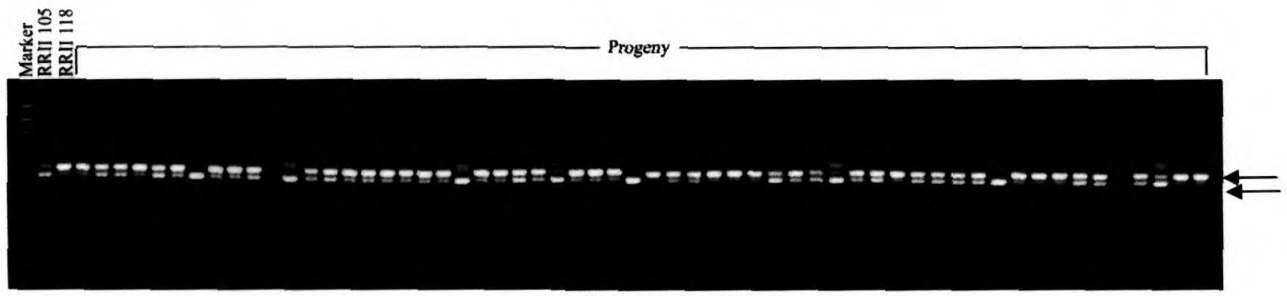


Fig. 2.14. EtBr stained agorose gel photograph showing the segregation of the allele-specific SNP marker developed from geranylgeranyl diphosphate synthase gene using allele-specific primer GGDPS_F88A. The amplified SNP allele is more intense for the homozygote (A/A) and gene locus is preferentially amplified when the allele is in heterozygous state (A/G). The arrowhead 'A' shows the gene locus and 'B' shows the SNP allele. ΦX 174/ *Hae* III digest was used as molecular weight marker

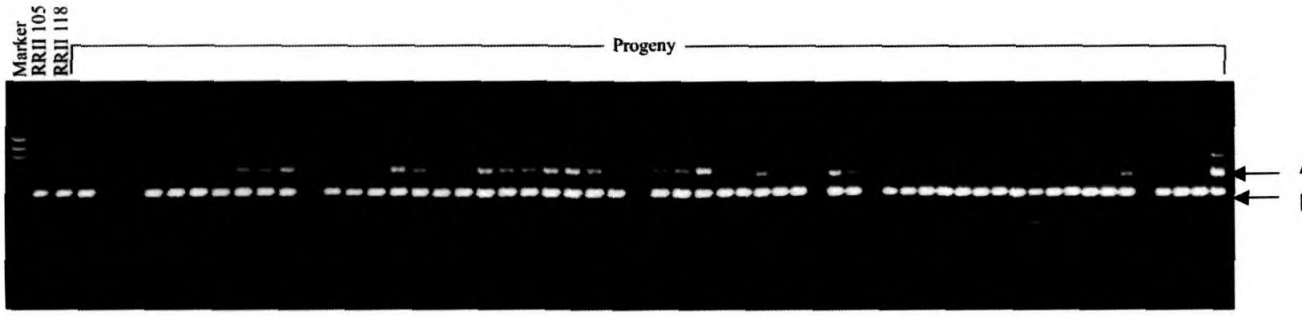


Fig. 2.15. EtBr stained agarose gel photograph representing the segregation of allele-specific PCR marker generated from latex patatin homolog gene using the allele specific reverse primer LPH5R_313G. The SNP detected here was in hetrozygous state for the clone RRII 105 (C/T), homozygous allele (C/C) was amplified as single band and heterozygous as two fragments: allele-specific product and the gene locus. The arrowhead 'A' shows the gene locus and 'B' shows the SNP allele. ΦX 174/ *Hae* III was used as molecular weight marker

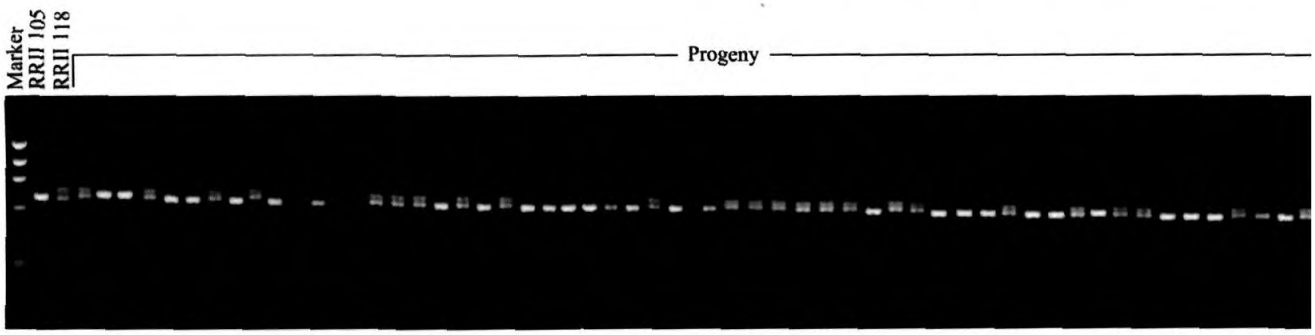


Fig. 2.16. EtBr stained agarose gel photograph representing the segregation of the sequence tagged site (STS) marker generated from farnesyl diphosphate synthase (*FDPS)* gene based on indels. The arrowheads represents the STS marker alleles generated. ΦX 174/ *Hae* III digest was used as molecular weight marker

2.4.1.4 Single nucleotide polymorphisms (SNPs)

SNP genotyping was done using allele-specific primers as well as CAPS markers. Three SNP-based markers from the genes: mevalonate kinase, latex patatin homolog and ubiquitin precursor were specific for the female parent RRII 105 and the SNP from the gene geranylgeranyl diphosphate synthase was specific to RRII 118. The segregation ratio was 1:1 in all the cases. CAPS marker was utilized for the gene ubiquitin precursor where the recognition sequence was absent for the particular enzyme *Xho* I in the case of RRII 118. Scoring of the CAPS marker was performed as presence and absence of band segregating in 1:1 ratio and two bands at 290 bp and 200 bp were scored in the progeny (Fig. 2.12). For the gene mevalonate kinase, SNP was identified at 197 bp of the amplified fragment (denoted as MKF197C). The allele was heterozygous for RRII 105 only and progeny genotyping was done based on this SNP (Fig. 2.13). SNPs were identified in geranylgeranyl diphosphate synthase (designated as GGDPSF_88A) and genotyping of SNP was performed based on the polymorphic base at 88 bp, where heterozygosity of alleles was observed for the clone RRII 118 only (Fig. 2.14). For latex patatin homolog (designated as LPH5R_ 313G), heterozygous SNP was identified for the clone RRII 105 at 313 bp (Fig. 2.15). Scoring was performed as lm x ll for heterozygous alleles in RRII 105 and nn x np in RRII 118. The segregation ratio of 1:1 was followed by all the allele specific markers.

2.4.1.4 .1 Sequence Tagged Site marker (STS)

The genotyping for the gene farnesyl diphosphte synthase was based on a sequence tagged site (STS) marker generated from the amplification of around 700 bases at the 5' end of the gene sequence. Heterozygosity was detected for the gene in the parental clone RRII 118, where two fragments were amplified, whereas the clone RRII 105 was homozygous amplifying a single band. The genotyping of progenies showed 1:1 segregation ratio (Fig. 2.16).

2.4.1.5 Resistance gene analogue (RGA)

The RGA marker developed segregated in a 1:1 ratio which was specific for the female parent RRII 105.

A total of 244 markers comprising of the 100 RAPD markers, 86 AFLP markers, 52 SSR markers, four gene specific SNP markers, one STS marker and one RGA marker was used for the construction of linkage map (Table 2.7).

2.4.1.6 Cluster analysis of the progeny population

The UPGMA algorithm was used for grouping all the progenies and their parents based on their genetic distances. Dendrogram representing the genetic relationship between parents and the progeny is presented in (Fig. 2.17). Seven clusters were observed at a distance coefficient of 0.33. One of the progeny was found to be in a single cluster. It was observed that the two parents (RRII 105 and RRII 118) were in separate clusters and the overall diversity was found to be 40%, which was more than the parents. Each progeny could be distinguished from the other. The dendrogram was drawn to provide additional information regarding the segregation ratios followed by the progeny population.

2.4.2 Linkage map construction

2.4.2.1 Segregation distortion

A Chi-Square test was performed to test the null hypothesis of segregation ratios of 1:1, 1:2:1 and 3:1 markers in the progenies derived from a cross between RRII 105 and RRII 118. Significant segregation distortion (χ^2 value > 10 for the present analysis, $P \leq 0.05$) was detected with eight markers out of 164 loci analyzed (4.8 %) in RRII 105, and with six markers, out of 145 loci analyzed (4%) in RRII 118. The distorted markers belonged to RAPD and AFLP markers only in both the parents. Most of the distorted markers followed a segregation ratio of 3:1. Microsatellite and SNP based markers did not show high distortion from the expected values. The distorted markers were first removed from the linkage analysis for establishing an order of markers.

2.4.2.2 Genetic linkage analysis with MAPMAKER

Linkage analysis was first performed with the MAPMAKER software. The data set was separated into two different groups corresponding to each parent. For MAPMAKER only the markers following a 1:1 segregation ratio were generally considered for linkage analysis. However, segregation ratio of 1:1:1:1 generated in mapping population of the parents with three and four alleles were reduced to 1:1 following the methods described by Dr. Van Ooiijen (JoinMap developer; personal communication) and genetic analysis was performed using the F_2 backcross option. For the female parent RRII 105, 115

Table 2.7. Total number of dominant and co-dominant markers generated for linkage analysis along with their segregation types and ratios. The specificity of the marker to the particular parent is also mentioned

Type of marker	Co-dominant markers					Dominant markers			Total
	(1:1:1:1) ab x cd	ef x eg	(1:2:1) hk x hk	(1:1) lm x ll	nn x np	(1:1) lm x ll	nn x np	(3:1) hk x hk	
RAPD	-	-	-	-	-	40	39	21	100
AFLP	-	-	-	-	-	38	31	17	86
SSR	6	11	8	16	11				52
SNP	-	-	-	3	1				4
RGA	-	-	-	-	-	1			1
STS	-	-	-	1	-				1
Total	6	11	8	20	12	79	70	38	244

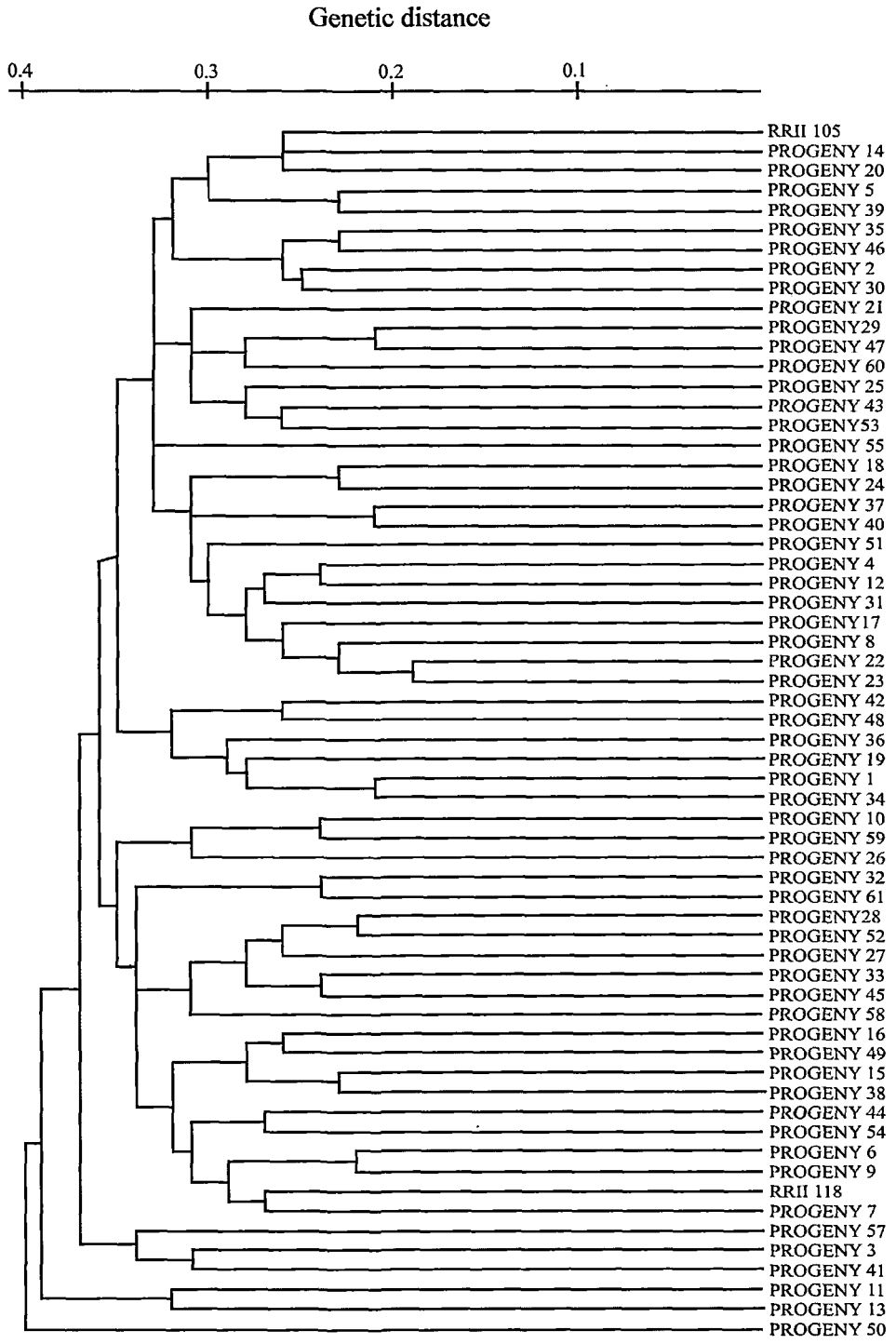


Fig. 2.17. Dendrogram showing genetic distance among the individuals of a mapping population along with their parents RRII 105 and RRII 118. It is shown that each progeny is distinct from the other

segregating markers were analyzed for linkage using a maximum likelihood distance of 40 and a minimum LOD score of 2.0. Twenty three linkage groups (LGs) were obtained from 89 mapped markers and 26 markers remained unlinked at the minimal LOD score. The largest group was found to be the LG-7 with 18 markers. Ten linkage groups were detected with more than two markers per linkage group, while all other groups possessed only two markers each. The largest linkage group covered a genetic distance of 408.7 cM. The details of linkage groups are given in Table 2.8. Total distance estimated for female map was 1384.7 cM using MAPMAKER. The average distance between markers was calculated as 15.5 cM.

In the male parent RRII 118, linkage analysis was performed with 99 markers following the segregation ratios of 1:1, resulting in 23 linkage groups with 66 markers, while 33 markers remained unlinked. The largest linkage group was the LG-17 with seven markers spanning a distance of 134.1 cM. Eleven linkage groups were detected having more than two markers while all other linkage groups contained only two markers. The lowest genetic distance was detected in group 18, spanning a distance of 6.9 cM only. The total distance covered by the male linkage map using the MAPMAKER was 768.5 cM. The average distance between markers was calculated as 11.64 cM. The details of linkage groups are given in Table.2.9.

2.4.2.3 Genetic linkage analysis with JoinMap

In the analysis with JoinMap, all the markers with different segregation ratios were taken into consideration using the JoinMap software. Separate linkage maps were constructed for each parent following the pseudo-test cross model analysis (Grattapaglia and Sederoff, 1994). A total of 164 markers were utilized for the linkage map construction in RRII 105 and 145 markers were utilized for genetic map construction in RRII 118.

Twenty four linkage groups were generated for the female map (RRII 105) spanning 762 cM, comprising of 103 markers in linkage with an average distance between two framework loci of 7.39 cM. Sixty-one markers were unlinked in RRII 105. The size of the linkage groups ranged from 7 cM (LG 23) with two markers to 121 cM (LG 1) with 14 linked markers with an average of 31.75 cM. Thirteen linkage groups were found to be with three or more markers and 11 linkage groups possessed only two markers each. Sixty-two markers were unlinked in RRII 105. Approximately 62.8% of the total markers analyzed (103 loci) were located on the female genetic linkage map (Fig. 2.18).

Table 2.8. Details of linkage groups RRII 105 generated by MAPMAKER

Marker	Genetic distance	Marker	Genetic distance
LG -1(9 markers)		LG-2 (5 markers)	
capub200	11.6 cM	EC5+MC7-5	24.0 cM
OPAJ201170	21.5 cM	OPAQ-12-778	38.5 cM
A11332	33.4 cM	capub290	23.6 cM
AM101153	22.0 cM	OPJ19-563	34.1 cM
AM10998	8.4 cM	C19+MC7-8	----------
AU101935	42.8 cM		120.2 cM
AR181152	19.5 cM		
AV19	13.1 cM		
CT_1	----------		
	172.2 cM		
LG-3		LG-4 (3 markers)	
OPAB-07-589	15.5 cM	EC9+MC3-4	27.2 cM
Hm ac-17		M-574	7.4 cM
		OPAJ-20-498	----------
			34.6 cM
LG-5 (5 markers)		LG-6	
HMct_14D	27.5 Cm	EC5+MC-4	32.4 cM
OPAO1-1360	18.8 cM	OPAA17-595	2.4 cM
OPAA-10954	11.9 cM	OPAL-5305	3.9 cM
HBE_054	23.9 cM	RGA3_15	----------
C19+MC7-1	----------		38.7 cM
	82.2 cM		
LG-7	30.0 cM	LG-8	
C19+MC-6-1	16.6 cM	OPX17-1090	0.0 cM
HBE_044	2.1 cM	OP-O15-1232	
Hmac_4	37.6 cM		
Htnr_21	22.9 cM		
Hmct_58	28.4 cM	LG-9	13.9 cM
OPA1-1021	63.8 cM	OPX18-2269	
C19+MC-7-2	26.8 cM	EC9+MC3-3	
OPA019-1508	20.9 cM		
OPAQ-7_422	25.1		
OPAB-11531	22.0 cM		
Hmac-2	39.0 cM	LG-10	
OPAS-14-1197	11.3 cM	OPAE4-578	12.6 cM
Hev glu	31.8 cM	EC5+MC-7-2	30.8 cM
EC5+MC-7-7	0.8 cM	OPAL6-1029	26.2 cM
OPAL-5-881	29.6 cM	EC9+MC5-1 -	---------
OPAQ-12-914	----------		69.6 cM
HBE_101	408.7 cM		
LG-11	24.4 cM	LG-12	20.3 cM
OPAL3-1180	26.6 cM	C20+MC3-1	12.9 cM
EC7+MC-2-2	32.5 CM	EC5+MC7-9	28.2 cM
HBE_080	---------	EC7+MC-2-1	26.9 cM
CT_9	83.5 cM	OPAS-14-1729	----------
		HBE_092	88.3 cM
LG- 13	17.7 cM	LG-17	26.9 cM
OPAT-20-1015		C17+MC-7-4	
C20+MC3-2		C19+MC7-10	
LG-14	22.0 cM	LG-18	
OPA019-1139		C20+MC4-6	20.5 cM
C18+MC8-2		C17+MC-7-6	23.9 cM
LG-15	15.6 cM	Hmac_10	5.1 cM
C5+MC7-10		C20+MC4-2	----------
EC8+MC3-1			49.5 cM
LG-16	25.8 cM		
C7+MC-7-3			
C18+MC8-3			

Continued...

LG-19 C19+MC6-4 C19+MC3	 24.2 cM	LG-20 HBE_033 Hmct_2 hmct2A2	 22.1 cM 4.3 cM ---------- 26.4 cM
LG-21 EC9+MC5-4 HBE_068	 10.9 cM ---------- 10.9 cM	LG-22 HBE_043 MKF197C LG-23 Hmct_16 HBE_122	 12.1 cM ---------- 12.1 cM 26.2 cM

Table 2.9. Linkage groups generated in RRII 118 by MAPMAKER

MARKER	GENETIC DISTANCE	MARKER	GENETIC DISTANCE
LG-1 C18+MC2 -4 OPA01-1966 HBE-115	6.7 cM 29.4 cM --------- 36.1 cM	LG-2 OPC-05-877 OPJ-19-866 LG-3 OPE- 03-506 OPAE-04-688	20.8 13.4 cM --------- 13.4 cM
LG-4 OPAL-06-1253 OPG17-661 Hmct-14D	8.3 cM 0.0 cM --------- 8.3 cM	LG-5 OPH-03-1025 OPAI-09-1048 C18+MC7-2 EC5+MC4-2	16.7 cM 14.5 cM 30.7 cM --------- 61.9 cM
LG-6 EC5+MC4-3-175 OPAB-14-1073 C19+MC-3-5-135 C20+MC4-1-800	25.5 cM 22.8 cM 18.9 cM --------- 67.2 cM	LG-7 Htnr locus 2a OPAR-19	23.4
LG-8 OPAJ-11-2482 HBE-122	24.1	LG-9 OPAA17-1728 EC-9+MC3-6	12.1
LG-10 OPX-7-1000 HBE-001	28.0	LG-11 OPAS-06-809 OPAS6-820 EC5MC7_8 OPX-17-947	3.0 cM 30.3 cM 0.0 cM --------- 33.3
LG-12 OPAL-6-1110 OPAL-12-230	34.7	LG-13 OPO-12-1558 EC5+MC-7-6	20.4
LG-14 OPAE-08-1400 Hmct-58	34.4	LG-15 OPAE-08-1584 HbFDP2 HBE-002	3.4
LG-16 EC-5+MC-7-3 EC-9+MC-3-2 OPAQ-7-644 EC-5+MC-7-4 Hev-Glu	0.0 0.0 9.8 42.3 --------- 52.1	LG-17 HBE_207 HBE_225 C18+MC6-1 C20+MC3-3-210 C17+MC7-1 OPAX6-895 OPAS14-752	0.0 cM 34.7 cM 20.1 cM 36.5 cM 26.6 cM 16.2 cM --------- 134.1 cM
LG-18 C-17+ MC-7-9 EC-7+MC-2-3	6.9	LG-19 HBE-139 HBE-140 C17+MC5-2	0.0 cM 24.1 cM --------- 24.1 cM
LG-20 C19+MC7-6 HBE-170 Hmct-5	32.8 cM 0.1 cM --------- 32.9 cM	LG-21 EC9+MC3-1-500_ Hmct-45	14.2 cM --------- 14.2 cM
LG-22 EC-9+ MC-7-3 HBE-098	34.4	LG-23 HBE 033 Hmct-2	18.4

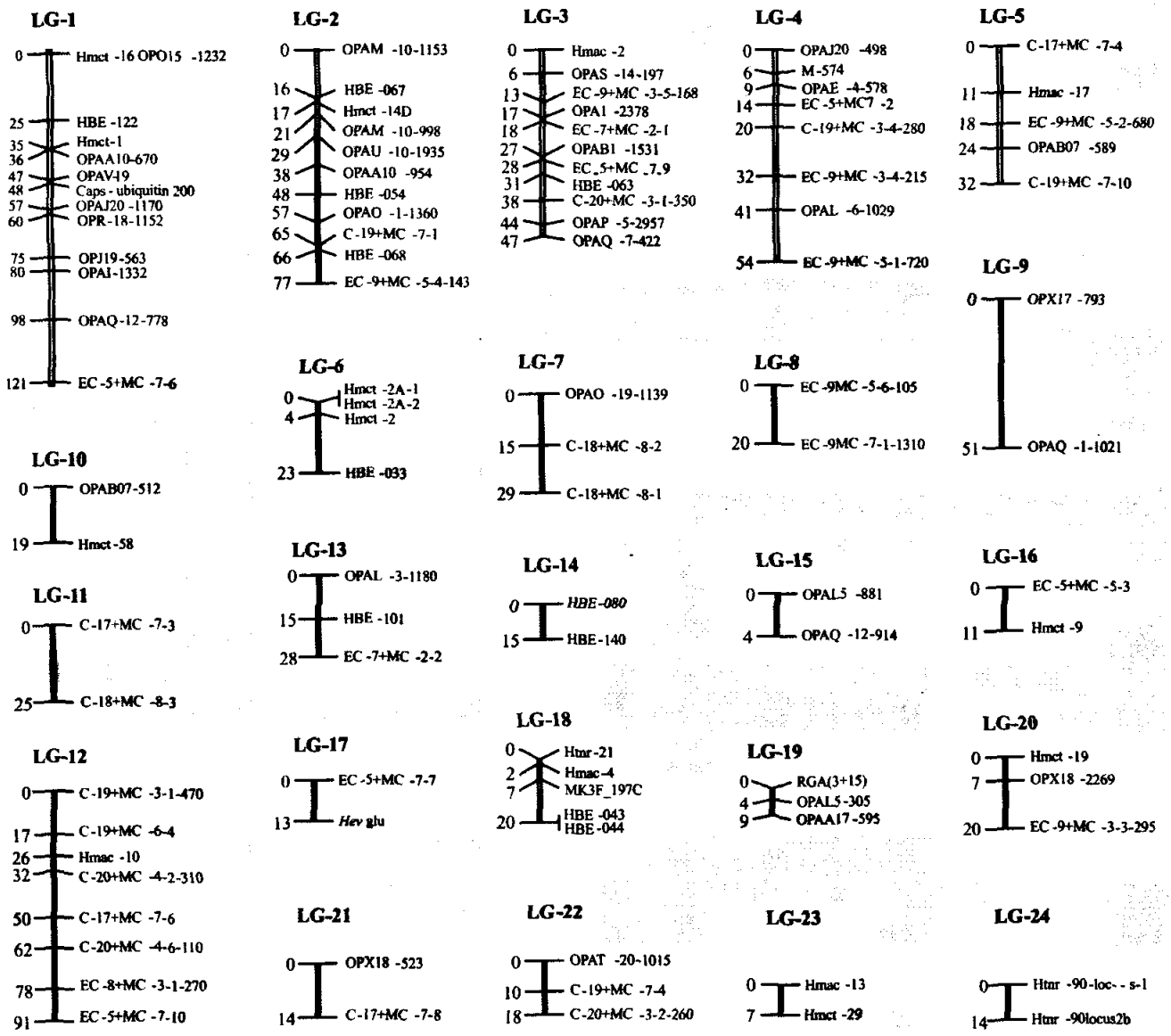


Fig. 2.18. Linkage map of the female parent RRII 105 generated by the software JoinMap

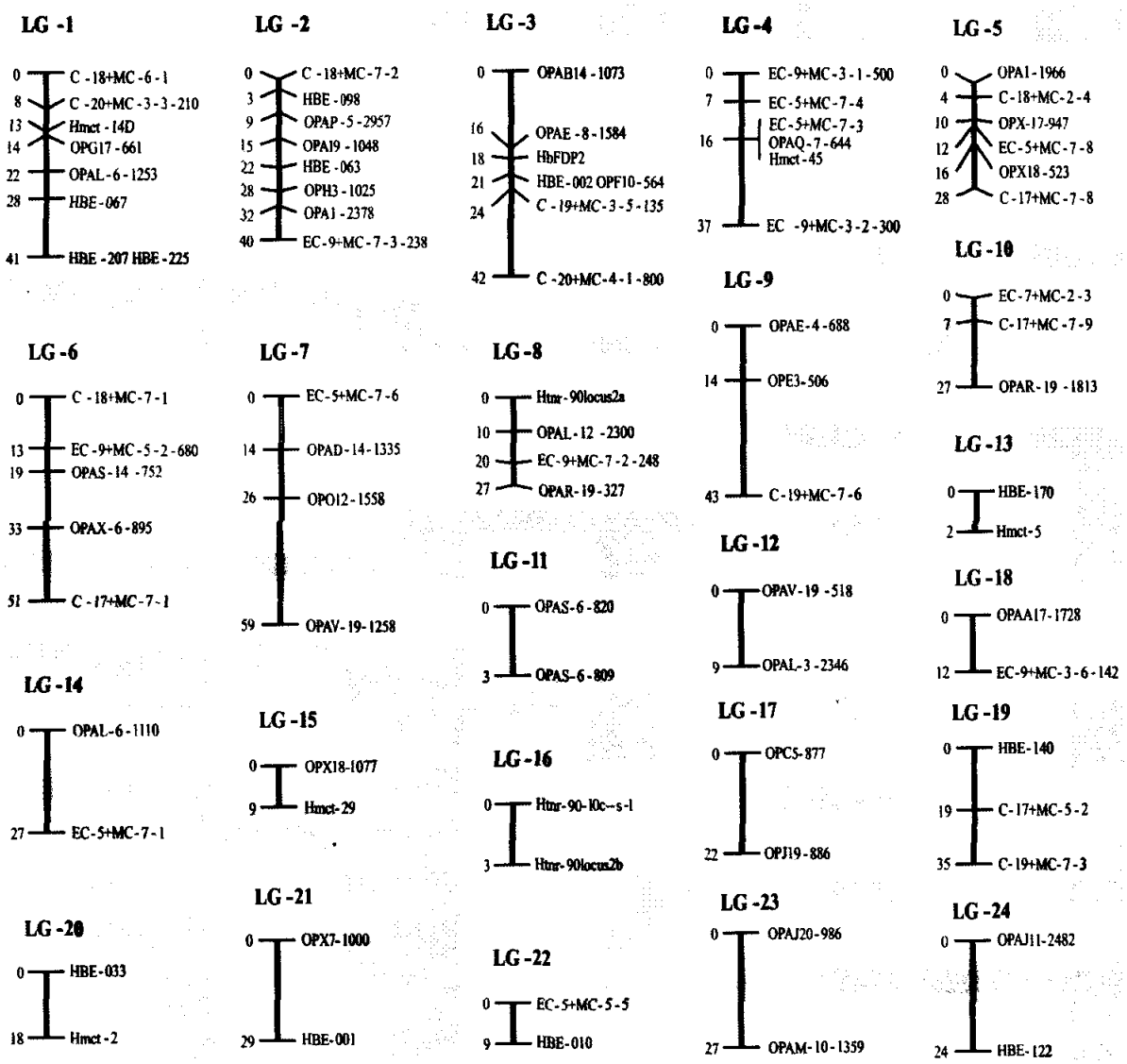


Fig. 2.19. Linkage map for the male parent RRII 118, generated by the software JoinMap

Twenty-four LGs were obtained for the male parent RRII 118 also, covering a distance of 634 cM comprising of 83 markers in the linkage, with an average distance between two loci of 7.63 cM. Sixty-two markers were unlinked in RRII 118. The size of the linkage groups ranged from 2 cM (LG 13) to 59 cM (LG 7) with an average of 26.4 cM . Eleven linkage groups were found to be with three or more markers and 13 linkage groups possessed only two markers each. Sixty two markers were unlinked in RRII 118. Approximately 57.2% of total markers analyzed in the male parent (83 loci) were located on the linkage map (Fig. 2.19). The details of marker segregation and their respective Chi-Square values including assigned linkage groups are given in Tables 2.10, 2.11, 2.12, 2.13 and 2.14.

2.4.2.4 Linkage analysis of SNPs in latex biosynthesis genes and other genes

In the present study, the SNPs anchoring mevalonate kinase locus, involved in rubber biosynthesis pathway was mapped on LG-18 in the parent RRII 105. The STS marker developed from the gene farnesyl diphosphate synthase, was mapped on LG-3 of the male parent RRII 118. The SNPs from ubiquitin precursor gene was mapped on LG-1 of the female parent RRII 105 (Fig. 2.20).

2.4.2.5 Map integration

Only two groups could be merged to form an integrated map. The first integrated map (LG-2 in RRII 105 and LG-1 in RRII 118) comprised of 17 markers linked by two orthologous microsatellite markers. The common markers present in these two linkage groups were Hmct-14D and HBE-068. The genetic distance covered by the map was 82 cM. The second integrated map comprised of four microsatellite markers, which was integrated by joining the LG-6 of RRII 105 and LG-20 of RRII 118. The common markers present were the SSR markers: Hmct-2 and HBE-033. The integrated map was found to be the same as LG-6 in RRII 105 and the same genetic distance of 23 cM was obtained for the integrated map (Fig. 2.21).

2.4.3 Comparison between linkage maps generated by MAPMAKER and JoinMap

It was observed that the markers which were linked through the MAPMAKER software were same as those linked in the JoinMap. The order of markers also remained same in both the cases. Some of the markers which were unlinked in JoinMap analysis were found linked in MAPMAKER. The total genetic distance calculated through MAPMAKER was 1384 cM in the female parent while through JoinMap it was 762 cM.

Table 2.10. Segregating RAPD loci along with their Chi-Square (χ^2) values and linkage group assignments

Sl. No.	RAPD	Parental genotypes	Expected ratio	Observed ratio	χ^2 values	Linkage group RRII-105	Linkage group RRII-118
1	OPA-01-1332	lm x ll	1:1	25:32	0.9	LG-1	-
2	OPG-17-588	lm x ll	1:1	30:27	0.2	unlinked	unlinked
3	OPJ19-563	lm x ll	1:1	31:26	0.4	LG-1	-
4	OPAB07-589	lm x ll	1:1	26:31	0.4	LG-5	-
5	OPAB07-512	lm x ll	1:1	35:22	3.0	LG-10	-
6	OPAJ11-788	lm x ll	1:1	35:22	3.0	unlinked	unlinked
7	OPAJ20-1170	lm x ll	1:1	28:29	0.0	LG-1	-
8	OPAJ20-498	lm x ll	1:1	26:31	0.4	LG-4	-
9	OPAA10-954	lm x ll	1:1	31:25	0.6	LG-2	-
10	OPAA17-595	lm x ll	1:1	33:24	1.4	LG-19	-
11	OPX7-1143	lm x ll	1:1	33:24	1.4	unlinked	unlinked
12	OPX17-793	lm x ll	1:1	23:34	2.1	unlinked	unlinked
13	OPX-17-1090	lm x ll	1:1	27:30	0.2	unlinked	unlinked
14	OPAL5-881	lm x ll	1:1	30:27	0.2	LG-15	-
15	OPAL5-305	lm x ll	1:1	32:25	0.9	LG-19	-
16	OPX18-2269	lm x ll	1:1	33:24	1.4	LG-20	-
17	OPAL-6-1029	lm x ll	1:1	28:28	0.0	LG-4	-
18	OPO15-1232	lm x ll	1:1	27:30	0.2	LG-1	-
19	OPAL-3-1180	lm x ll	1:1	26:28	0.1	LG-13	-
20	OPAM-10-1153	lm x ll	1:1	28:26	0.1	LG-2	-
21	OPAM-10-998	lm x ll	1:1	23:31	1.2	LG-2	-
22	OPAE-4-578	lm x ll	1:1	33:24	1.4	LG-4	-
23	OPAO-1-1360	lm x ll	1:1	30:27	0. 2	LG-2	-
24	OPAO-1-1021	lm x ll	1:1	12:45	19.1**	LG-9	-
25	OPAQ-7-422	lm x ll	1:1	13:44	16.9**	LG-3	-
26	OPAR-18-1152	lm x ll	1:1	25:28	0.2	LG-1	-
27	OPAR-19-1322	lm x ll	1:1	31:25	0.6	unlinked	unlinked
28	OPAV-19	lm x ll	1:1	27:29	0.1	LG-1	-
29	OPAS-10-1438	lm x ll	1:1	36:21	4.0*	unlinked	unlinked
30	OPAS-14-1729	lm x ll	1:1	32:23	1.5	unlinked	unlinked
31	OPAS-14-1197	lm x ll	1:1	31:24	0.9	LG-3	-
32	OPAQ-12-914	lm x ll	1:1	31:24	0.9	LG-15	-
33	OPAQ-12-778	lm x ll	1:1	22:33	2.2	LG-1	-
34	OPAT-20-1015	lm x ll	1:1	31:22	1.5	LG-22	-
35	OPAT-20-692	lm x ll	1:1	18:35	5.5*	unlinked	unlinked
36	OPA0-19-1508	lm x ll	1:1	20:35	4.1*	unlinked	unlinked
37	OPA0-19-1139	lm x ll	1:1	32:23	1.5	LG-7	-
38	OPAB1-1531	lm x ll	1:1	25:32	0.9	LG-3	-
39	OPAU-10-1935	lm x ll	1:1	30:27	0.2	LG-2	-
40	OPAU-10-483	lm x ll	1:1	38:17	8.0*	unlinked	unlinked
41	OPA-01-1966	nn x np	1:1	26:33	0.8	-	LG-5
42	OPC-05-877	nn x np	1:1	28:31	0.1	-	LG-17
43	OPE – 03 – 506	nn x np	1:1	32:27	0.4	-	LG-9
44	OPG17-661	nn x np	1:1	27:32	0.4	-	LG-1
45	OPH3-1025	nn x np	1:1	25:34	1.4	-	LG-2
46	OPAA07-1558	nn x np	1:1	32:27	0.4	unlinked	unlinked
47	OPAB14-1073	nn x np	1:1	27:32	0.4	-	LG-3
48	OPAJ11–2482	nn x np	1:1	32:27	0.4	-	LG-24
49	OPAJ20-1096	nn x np	1:1	19:40	7.5*	unlinked	Unlinked
50	OPAJ20-986	nn x np	1:1	22:37	3.8	-	LG-23

51	OPAA17-1728	nn x np	1:1	39:20	6.1*	-	LG-18
52	OPX3-1058	nn x np	1:1	26:33	0.8	unlinked	unlinked
53	OPX7-1000	nn x np	1:1	30:29	0.0	-	LG-21
54	OPX-17-947	nn x np	1:1	38:21	4.9*	-	LG-5
55	OPAL-6-1253	nn x np	1:1	28:29	0.0	-	LG-1
56	OPAL-6-1110	nn x np	1:1	22:36	3.4	-	LG-14
57	OPAL-12-2300	nn x np	1:1	29:30	0.0	-	LG-8
58	OPO12-1558	nn x np	1:1	22:37	3.8*	-	LG-7
59	OPJ19-886	nn x np	1:1	18:41	9.0*	-	LG-17
60	OPAI9-1048	nn x np	1:1	25:34	1.4	-	LG-2
61	OPAM-10-1904	nn x np	1:1	33:24	1.4	unlinked	unlinked
62	OPAM-10-1359	nn x np	1:1	27:30	0.2	-	LG-23
63	OPAM-10-1216	nn x np	1:1	16:40	1.6	unlinked	unlinked
64	OPAD-14-1335	nn x np	1:1	36:23	2.9	-	LG-7
65	OPAE-4-688	nn x np	1:1	24:35	2.0	-	LG-9
66	OPAR-19-1813	nn x np	1:1	33:25	1.1	-	LG-10
67	OPAR-19-327	nn x np	1:1	35:23	2.5	-	LG-8
68	OPAS-6-820	nn x np	1:1	32:27	0.4	-	LG-11
69	OPAV-19-518	nn x np	1:1	38:21	4.1	-	LG-12
70	OPAV-19-1258	nn x np	1:1	22:36	3.4	-	LG-7
71	OPAS-10-1109	nn x np	1:1	21:38	4.1	unlinked	unlinked
72	OPAS-14-752	nn x np	1:1	28:27	0.0	-	LG-6
73	OPAS-6-809	nn x np	1:1	30:29	0.0	-	LG-11
74	OPAX-6-895	nn x np	1:1	25:32	0.9	-	LG-6
75	OPAQ-12-697	nn x np	1:1	25:31	0.6	unlinked	unlinked
76	OPX-18-1077	nn x np	1:1	29:29	0.0	-	LG-15
77	OPAE-08-1400	nn x np	1:1	31:28	0.1	-	LG-
78	OPAE-08-1584	nn x np	1:1	35:24	2.0	-	LG-3
79	OPAQ-07-644	nn x np	1:1	34:25	1.4	unlinked	unlinked
80	OPA-01-2378	hk x hk	3:1	17:30	0.1	LG-2	LG-2
81	OPF-10-564	hk x hk	3:1	14:43	0.0	unlinked	unlinked
82	OPH3-1375	hk x hk	3:1	17:40	0.7	unlinked	unlinked
83	OPAA07-1037	hk x hk	3:1	48:9	106.1***	unlinked	unlinked
84	OPAA07-569	hk x hk	3:1	17:40	0.7	unlinked	unlinked
85	OPAI2-1488	hk x hk	3:1	14:43	0.0	unlinked	unlinked
86	OPAJ20-1817	hk x hk	3:1	16:41	0.3	unlinked	unlinked
87	OPAA10-670	hk x hk	3:1	18:39	1.3	LG-1	-
88	OPAA10-1048	hk x hk	3:1	11:46	1.0	unlinked	unlinked
89	OPX18-523	hk x hk	3:1	21:36	0.0	LG-21	LG-5
90	OPAL-12-613	hk x hk	3:1	17:40	0.7	unlinked	unlinked
91	OPO12-1401	hk x hk	3:1	16:41	0.3	unlinked	unlinked
92	OPAL-3-2346	hk x hk	3:1	14:42	1.0	-	LG-12
93	OPAP-5-2957	hk x hk	3:1	13:40	0.0	LG-3	LG-2
94	OPAP-5-2164	hk x hk	3:1	10:43	1.1	unlinked	unlinked
95	OPAX-16-788	hk x hk	3:1	15:42	0.1	unlinked	unlinked
96	OPAU-10-1008	hk x hk	3:1	20:36	3.6	unlinked	unlinked
97	OPAV-19-934	hk x hk	3:1	10:49	2.0	unlinked	unlinked
98	OPAS-10-2660	hk x hk	3:1	13:45	0.2	unlinked	unlinked
99	OPAS-14-831	hk x hk	3:1	15:40	0.1	unlinked	unlinked
100	OPAX-6-2319	hk x hk	3:1	19:37	2.4	unlinked	unlinked

* slightly distorted segregation from Mendel's law

** significantly distorted

*** highly distorted

Table 2.11. Segregating AFLP loci along with their Chi-Square (χ^2) values and linkage group assignments

Sl. No.	AFLP	Expected genotypes	Expected ratio	Observed ratio	χ^2 values	Linkage group RRII-105	Linkage group RRII-118
1	EC-5+MC7-2	lm x ll	1:1	30:25	0.5	LG-4	-
2	EC-5+MC-7-5	lm x ll	1:1	20:36	4.6*	LG-1	-
3	EC-5+MC-7-7	lm x ll	1:1	31:24	0.9	LG-17	-
4	EC-5+MC-7-9	lm x ll	1:1	32:24	1.1	LG-3	-
5	EC-5+MC-7-10	lm x ll	1:1	35:20	4.1*	LG-12	-
6	C-18+MC-2-2	lm x ll	1:1	23:29	0.7	unlinked	unlinked
7	C-18+MC-2-3	lm x ll	1:1	28:24	0.3	unlinked	unlinked
8	C-17+MC-7-3	lm x ll	1:1	24:33	1.4	LG-11	LG-19
9	C-17+MC-7-4	lm x ll	1:1	27:29	0.1	LG-5	-
10	C-17+MC-7-6	lm x ll	1:1	31:25	0.6	LG-12	-
11	C-17+MC-5-4	lm x ll	1:1	35:20	4.1*	unlinked	unlinked
12	C-19+MC-6-1	lm x ll	1:1	15:36	8.7*	unlinked	unlinked
13	C-19+MC-6-4	lm x ll	1:1	26:24	0.1	LG-12	-
14	EC-7+MC-2-1	lm x ll	1:1	36:21	4.0*	LG-3	-
15	EC-5+MC-5-1	lm x ll	1:1	23:30	0.9	unlinked	unlinked
16	EC-5+MC-5-2	lm x ll	1:1	16:38	9.0*	unlinked	unlinked
17	EC-5+MC-5-3	lm x ll	1:1	23:30	0.9	LG-16	-
18	EC-7+MC-2-2	lm x ll	1:1	24:31	0.9	LG-13	-
19	C-19+MC-7-1	lm x ll	1:1	21:36	4.0*	LG-2	-
20	C-19+MC-7-2	lm x ll	1.1	10:47	24.0*	unlinked	unlinked
21	C-19+MC-7-8	lm x ll	1:1	25:32	0.9	unlinked	unlinked
22	C-19+MC-7-10	lm x ll	1:1	31:24	0.9	LG-5	-
23	C-18+MC-8-1	lm x ll	1:1	18:35	5.5*	LG-7	-
24	C-18+MC-8-2	lm x ll	1:1	34:23	2.1	LG-7	-
25	C-18+MC-8-3	lm x ll	1:1	31:25	0.6	LG-11	-
26	C-20+MC-3-2-260	lm x ll	1:1	36:18	6.0*	LG-22	-
27	C-20+MC-4-2-310	lm x ll	1:1	23:31	1.2	LG-12	-
28	C-20+MC-4-5-138	lm x ll	1:1	36:17	1.8	unlinked	unlinked
29	C-20+MC-4-6-110	lm x ll	1:1	34:20	3.6	LG-12	-
30	EC-9+MC-5-1-720	lm x ll	1:1	36:19	5.3*	LG-4	-
31	EC-9+MC-5-3-210	lm x ll	1:1	29:27	0.1	unlinked	unlinked
32	EC-9+MC-5-4-143	lm x ll	1:1	30:24	0.7	LG-2	-
33	EC-8+MC-3-1-270	lm x ll	1:1	27:26	0.0	unlinked	unlinked
34	EC-9+MC-3-3-295	lm x ll	1:1	35:21	3.5	LG-20	-
35	EC-9+MC-3-4-215	lm x ll	1:1	26:31	0.4	LG-4	-
36	EC-5+MC-4-120	lm x ll	1:1	32:22	1.9	unlinked	unlinked
37	C-19+MC-3-1-470	lm x ll	1:1	16:37	8.3*	LG-12	-
38	C-20+MC-3-1-350	lm x ll	1:1	34:20	3.6	LG-3	-
39	EC-5+MC-5-5	nn x np	1:1	21:34	3.1	-	LG-22
40	EC-5+MC-7-1	nn x np	1:1	32:26	0.6	-	LG-14
41	EC-5+MC-7-3	nn x np	1:1	31:25	0.6	-	LG-4
42	EC-5+MC-7-4	nn x np	1:1	29:37	0.1	-	LG-4
43	EC-5+MC-7-6	nn x np	1:1	23:34	2.1	-	LG-7
44	EC-5+MC-7-8	nn x np	1:1	38:20	5.6*	-	LG-5
45	EC-5+MC-4-3-175	nn x np	1:1	31:28	0.1	unlinked	unlinked
46	C-18+MC-2-1	nn x np	1:1	25:28	0.2	unlinked	unlinked
47	C-18+MC-2-4	nn x np	1:1	21:33	2.7	-	LG-5
48	C-17+MC-7-1	nn x np	1:1	29:28	0.0	-	LG-6
49	C-17+MC-7-2	nn x np	1:1	9:50	3.0	unlinked	unlinked
50	C-17+MC-7-9	nn x np	1:1	31:27	0.3	-	LG-10
51	C-17+MC-5-2	nn x np	1:1	29:28	0.0	-	LG-19
52	EC-7+MC-2-3	nn x np	1:1	29:29	0.0	-	LG-10
53	C-19+MC-7-3	nn x np	1:1	31:38	0.1	unlinked	unlinked
54	C-19+MC-7-5	nnx np	1:1	28:30	0.1	unlinked	unlinked
55	C-19+MC-7-6	nn x np	1:1	22:36	3.4	-	LG-9
56	C-19+MC-7-9	nn x np	1:1	20:39	6.1*	unlinked	unlinked
57	C-18+MC-7-1	nn x np	1:1	26:31	0.4	-	LG-6
58	C-18+MC-7-2	nn x np	1:1	19:38	6.3*	-	LG-2
59	C-18+MC-6-1	nn x np	1:1	34:24	1.0	-	LG-1
60	C-19+MC-3-3-370	nn x np	1.1	32:25	0.9	unlinked	unlinked

62	C-20+MC-3-3-210	nn x np	1:1	29:26	0.2	-	LG-1
63	C-20+MC-4-1-800	nn x np	1:1	33:25	1.1	-	LG-3
64	EC-9+MC-3-2-300	nn x np	1:1	33:23	1.8	-	LG-4
65	EC-9+MC-3-6-142	nn x np	1:1	31:27	0.3	-	LG-18
66	EC-5+MC-4-2	nn x np	1:1	19:39	6.4*	unlinked	unlinked
67	EC-9+MC-7-3-238	nn x np	1:1	27:29	0.1	-	LG-2
68	C-20+MC-4-4-150	nn x np	1:1	33:23	1.8	unlinked	unlinked
69	EC-9+MC-3-1-500	nn x np	1:1	31:27	0.3	-	LG-4
70	EC-9+MC-5-2-680	hk x hk	3:1	16:41	0.3	LG-5	LG-6
71	EC-9+MC-5-5-141	hk x hk	3:1	27:25	20.1**	unlinked	unlinked
72	EC-9+MC-5-6-105	hk x hk	3:1	21:29	7.5*	LG-8	-
73	EC-9+MC-3-5-168	hk x hk	3:1	17:39	0.9	LG-3	-
74	EC-9+MC-7-1-310	hk x hk	3:1	16:41	0.3	LG-8	-
75	EC-9+MC-7-2-248	hk x hk	3:1	15:41	0.1	-	LG-8
76	EC-9+MC-7-4-180	hk x hk	3:1	18:38	1.5	unlinked	unlinked
77	EC-5+MC-4-1	hk x hk	3:1	14:40	0.1	unlinked	unlinked
78	C-17+MC-7-8	hk x hk	3:1	26:30	13.7**	LG-21	LG-5
79	C-17+MC-7-2	hk x hk	3:1	8:49	3.6*	unlinked	unlinked
80	C-17+MC-5-1	hk x hk	3:1	23:31	8.9*	unlinked	unlinked
81	C-17+MC-5-3	hk x hk	3:1	7:45	3.7	unlinked	unlinked
82	C-19+MC-7-4	hk x hk	3:1	23:33	7.7*	LG-22	
83	C-19+MC-7-7	hk x hk	3:1	27:28	17.0**	unlinked	unlinked
84	C-19+MC-3-2-380	hk x hk	3:1	20:36	3.4	unlinked	unlinked
85	C-19+MC-3-4-280	hk x hk	3:1	19:35	3.0	LG-4	-
86	C-20+MC-4-3-200	hk x hk	3:1	23:30	9.6*	unlinked	unlinked

* slightly distorted segregation from Mendel's law

** significantly distorted

Table 2.12. Segregating genomic SSR loci along with their Chi-Square (χ^2) values and linkage group assignmemnts

Sl. No.	Genomic-SSR	Expected genotypes	Expected ratio	Observed ratio	Chi-Square value	Linkage group of RRII-105	Linkage group of RRII-118
1	hmct-1	lm x ll	1:1	23:33	1.8	LG-1	-
2	hmct-16	lm x ll	1:1	30:23	0.9	LG-1	-
3	hmct-19	lm x ll	1:1	24:33	1.4	LG-20	-
4	hmct-27	lm x ll	1:1	24:33	1.4	unlinked	unlinked
5	hmac-4	lm x ll	1:1	23:34	2.1	LG-18	-
6	hmac-10	lm x ll	1:1	29:28	0.0	LG-12	-
7	hmac-13	hk x hk	1:2:1	14:27:16	0.3	LG-23	-
8	hmct-53	hk x hk	1:2:1	8:29:15	2.6	unlinked	unlinked
9	hmac-2	lm x ll	1:1	27:28	0.0	LG-3	-
10	hmct-17	hk x hk	1:2:1	15:27:15	0.2	unlinked	unlinked
11	hmct-14D	lm x ll, nn x np	1:1; 1:1	31:23; 27:29	1.2, 0.1	LG-2	LG-1
12	hmct-29	lm x ll, nn x np	1:1; 1:1	22:34; 30:28	2.6, 0.9	LG-29	LG-15
13	hmac-14	lm x ll, nn x np	1:1; 1:1	26:31; 27:32	0.4	unlinked	unlinked
14	hmac-17	lm x ll, nn x np	1:1; 1:1	25:32; 25:34	0.9, 1.4	LG-5	-
15	*Hev*-glu (glucanase)	lm x ll, nn x np	1:1; 1:1	31:26; 31:28	0.4, 0.1	LG-17	-
16	hmct-9	lm x ll, nn x np	1:1; 1:1	26:22; 26:22	0.3, 0.2	LG-16	-
17	hmct-2	lm x ll, nn x np	1:1; 1:1	27:24; 27; 25	0.2, 0.1	LG-6	LG-20
18	hmct-58	lm x ll, nn x np	1:1; 1:1	16:39, 17:39	9.6, 8.6*	LG-10	-
19	hmct-45	nn x np	1:1	25:33	1.1	unlinked	unlinked
20	hmac-16	nn x np	1:1	25:21	0.3	unlinked	unlinked
21	hmct-5	nn x np	1:1	27:23	0.1	-	LG-13
22	hmct-2A-1	lm x ll	1:1	17:37	1.2	LG-6	-
23	hmct-2A-2	lm x ll	1:1	28:26	0.1	LG-6	-
24	M-574	lm x ll	1:1	22:29, 23:29	1.0, 0.7	LG-4	-
25	TNR-90locus2a	nn x np	1:1	24:30	0.7	LG-24	LG-8
26	TNR-90-locus-s-1	hk x hk	1:2:1	14:34:9	3.0	LG-24	LG-16
27	Htnr-21	1m x ll	1:1				

*slightly distorted segregation from Mendel's law

Table 2.13. Segregating EST - SSR loci along with their Chi-Square (χ^2) values and their linkage group assignments

Sl. No.	EST-SSR	Observed genotypes in parents	Expected ratio	Observed ratio	Chi-Square value	Linkage group in RRII-105	Linkage group in RRII-118
1	HBE-010	(nn x np)	1:1	37:22	3.8	-	LG-22
2	HBE-063	(hk x hk)	1:2:1	12:31:12	0.9	LG-3	LG-2
3	HBE-115	(nn x np)	1:1	21:35	3.5	unlinked	unlinked
4	HBE-001	(nn x np)	1:1	31:27	0.3	-	LG-21
5	HBE-002	(nn x np)	1:1	33:26	0.8	-	LG-3
6	HBE-003	(hk x hk)	1:2:1	14:38:7	6.6*	unlinked	unlinked
7	HBE-033	(nn xnp), (lm x ll)	1:1; 1:1	28:30; 28:28	0.1, 0.0	LG-6	LG-20
8	HBE-034	(hk x hk)	1:2:1	14:34:9	3.0	unlinked	unlinked
9	HBE-043	(lm x ll)	1:1	30:27	0.2	unlinked	unlinked
10	HBE-054	(lm x ll)	1:1	27:23	0.3	LG-2	-
11	HBE-044	(nn xnp), (lm x ll)	1:1; 1:1	36:23; 27:30	2.9, 0.2	LG-18	-
12	HBE-067	(hk x hk)	1:2:1	12:28:10	0.9	LG-2	LG-1
13	HBE-068	(lm x ll)	1:1	31:26	0.4	LG-2	-
14	HBE-080	(lm x ll)	1:1	29:26	0.82	LG-14	-
15	HBE-092	(lm x ll)	1:1	33:23	1.8	unlinked	unlinked
16	HBE-098	(nn xnp)	1:1	34:25	1.4	-	LG-2
17	HBE-101	(lm x ll)	1:1	30:25	0.5	LG-13	-
18	HBE-122	(nn xnp), (lm x ll)	1:1	29:29; 35:22	0.0; 3.0	LG-1	LG-24
19	HBE-139	(nn x np)	1:1	34:23	2.1	unlinked	unlinked
20	HBE-140	(nn xnp), (lm x ll)	1:1; 1:1	35:24; 25:32	2.0, 0.9	LG-14	LG-19
21	HBE-170	(nn xnp)	1:1	32:27	0.4	-	LG-13
22	HBE-190	(nn xnp), (lm x ll)	1:1; 1:1	25:33; 24:32	0.1	unlinked	unlinked
23	HBE-192	(lm x ll)	1:1	27:27	0.0	unlinked	unlinked
24	HBE-207	(nn x np)	1:1	26:30	0.3	-	LG-1
25	HBE-225	nn x np)	1:1	26:32	0.6	-	LG-1

*slightly distorted segregation from Mendel's law

Table 2.14. Segregating SNP loci along with the Chi-Square (χ^2) values and their linkage group assignments

Sl. No.	SNP marker	Parental genotypes	Expected ratio	Observed ratio	Chi-Square value	Linkage group
1	Farnesyl diphosphate synthase (STS)	nn x np	1:1	32:25	0.9	LG-3 (RRII 118)
2	Geranylgeranyl diphosphate synthase	nn x np	1:1	23:33	1.8	unlinked
3	Mevalonte kinase	lm x ll	1:1	31:25	0.6	LG-18 (RRII 105)
4	Latex patatin homolog	lm x ll	1:1	26:24	0.1	unlinked
5	Ubiquitin - 200 (CAPS)	lm x ll	1:1	27:29	0.1	LG-1 (RRII 105)
1	RGA(3+15)	lm x ll	1:1	29:25	0.3	LG-19 (RRII 105)

Table 2.15. Details of linkage analysis using MAPMAKER and JoinMap

MAPMAKER analysis						JoinMap analysis					
RRII 105 (Female parent) (115 markers)			RRII 118 (Male parent) (99 markers)			RRII 105 (Female parent) (164 markers)			RRII 118 (Male parent) (145 markers)		
Linkage group	No. of markers	Map distance (cM)	Linkage group	No: of markers	Map distance (cM)	Linkage group	No: of markers	Map distance (cM)	Linkage group	No: of markers	Map distance (cM)
LG-1	9	172.2	LG -1	3	36.1	LG -1	14	121	LG -1	8	41
LG -2	5	120.2	LG -2	2	20.8	LG -2	11	77	LG -2	8	40
LG -3	3	15.5	LG -3	2	13.4	LG -3	11	47	LG -3	7	42
LG -4	3	34.6	LG -4	3	8.3	LG -4	8	54	LG -4	6	37
LG -5	5	82.2	LG -5	4	61.9	LG -5	5	32	LG -5	6	28
LG -6	5	38.7	LG -6	4	67.2	LG -6	4	23	LG -6	5	51
LG -7	17	408.7	LG -7	2	23.4	LG -7	3	29	LG -7	4	59
LG -8	2	0.0	LG -8	2	24.1	LG -8	2	20	LG -8	4	27
LG -9	2	13.9	LG -9	2	12.1	LG -9	2	51	LG -9	3	43
LG -10	4	69.6	LG -10	2	28.0	LG -10	2	19	LG -10	3	27
LG -11	4	83.5	LG -11	4	33.3	LG -11	2	25	LG -11	2	3
LG -12	5	88.3	LG -12	2	34.7	LG -12	8	91	LG -12	2	9
LG -13	2	17.7	LG -13	2	20.4	LG -13	3	28	LG -13	2	2
LG -14	2	22.0	LG -14	2	34.4	LG -14	2	15	LG -14	2	27
LG -15	2	15.6	LG -15	3	3.4	LG -15	2	14	LG -15	2	9
LG -16	2	25.8	LG -16	5	52.1	LG -16	2	11	LG -16	2	13
LG -17	2	26.9	LG -17	7	134.1	LG -17	2	13	LG -17	2	22
LG -18	4	49.5	LG -18	2	6.9	LG -18	5	20	LG -18	2	12
LG -19	2	24.2	LG -19	3	24.1	LG -19	3	9	LG -19	3	35
LG -20	3	26.4	LG -20	3	32.9	LG -20	3	20	LG -20	2	18
LG -21	2	10.9	LG -21	3	44.1	LG -21	2	14	LG -21	2	29
LG -22	2	12.1	LG -22	2	34.4	LG -22	3	18	LG -22	2	9
LG -23	2	26.2	LG -23	2	18.4	LG -23	2	7	LG -23	2	27
						LG -24	2	14	LG -24	2	24
Total	89	1384		66	768.5		103	762		83	634

Similarly the genetic distance calculated through MAPMAKER in the male parent was 768.5 cM and calculated as 634 cM in JoinMap analysis. Hence the genetic distance calculated by JoinMap was less compared to that calculated by MAPMAKER (Table 2.15).

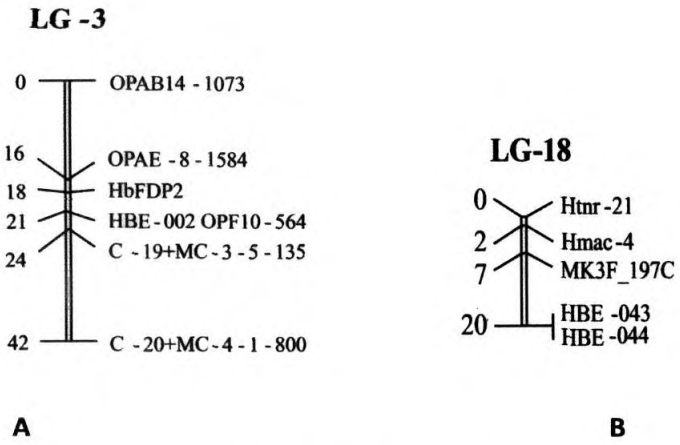


Fig. 2.20. Linkage assignment of latex biosynthesis genes. A) Linkage of farnesyl diphosphate synthase gene (denoted as HbFDP2) to the LG-3 of RRII 118 and B) Linkage of mevalonate kinase gene (denoted as MK3F_197C) to the LG-18 of RRII 105

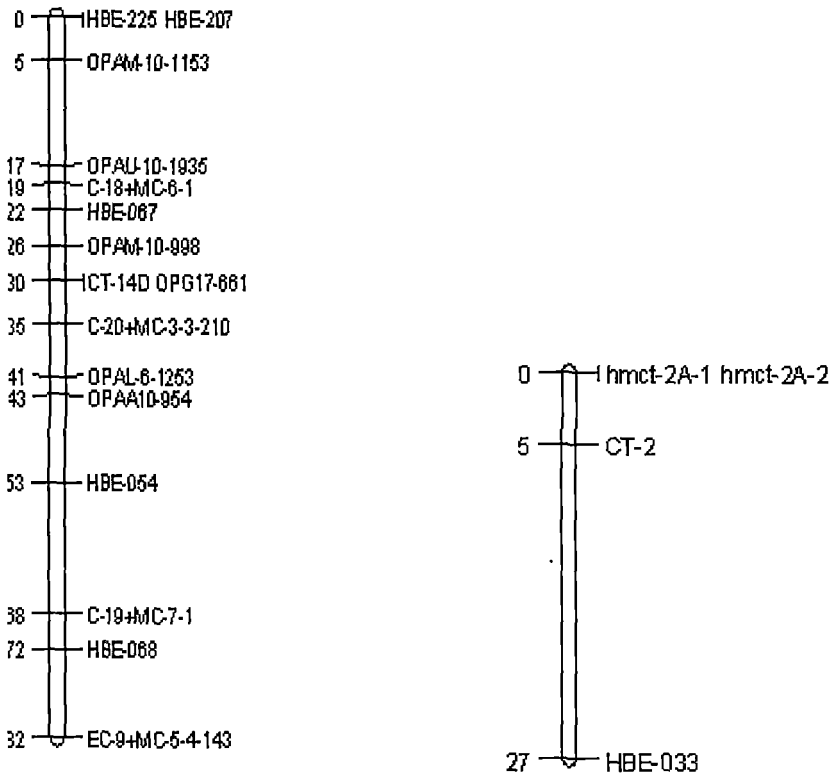


Fig. 2.21. Two integrated linkage groups were constructed using common SSR markers present in both the parents (heterozygous condition in both the parents), A) Integrated map of LG-1 and LG-2 of RRII 105 and RRII 118 respectively; B) Integrated map of LG-20 and LG-6 of RRII 105 and RRII 118 respectively

2.5 Discussion

2.5.1 Mapping strategy in *Hevea*

The "pseudo-test cross" mapping strategy was used to construct the present linkage maps for *Hevea brasiliensis*. The mapping population consisted of 61 F_1 individuals obtained from the cross between RRII 105 and RRII 118. The pseudo-test cross mapping strategy is conceptually simple to implement and can be applied with any type of molecular marker. The name "pseudo-test cross" was applied for this strategy because the testcross mating configuration of the markers is not known *apriori*, as in a conventional testcross where the tester is homozygous recessive for the locus of interest. Instead, the configuration is inferred *aposteriori* in the case of dominant markers like RAPD and AFLP after analyzing the genetic segregation of the marker in the progeny of a cross between highly heterozygous parents with no prior genetic information. When this inference is done for both parents involved in the cross, the term "two-way pseudo-test cross" is more appropriately used (Grattapaglia and Sederoff, 1994). In out-crossing tree crops, individual plants are highly heterozygous (Hamrick and Godt, 1990). Hence, the F_1 mapping population had been widely adopted in trees due to the high heterozygosity (Wu *et al.*, 2000; Achere *et al.*, 2004). In the present work the same mapping strategy was applied to rubber.

In *Hevea brasiliensis* pseudo-test cross strategy was adopted for genetic map construction in F_1 progeny populations obtained from PB 260 x RO 38 cross combination (Lespinasse *et al.*, 2000). In the reports of Feng *et al.* (2010), F_1 segregating population derived from the cross between Reyan 88-13 x IAN 873 was considered for mapping. Another genetic linkage map was constructed in *Hevea* using the full-sib progeny derived from a controlled cross between the cultivars PB 217 and PR 255 (Souza *et al.*, 2011). An integrated parental linkage map of *H. brasiliensis* based on F_1 mapping population of RRIM 600 x RRII 105 was constructed by Triwitayakorn (2011). In other tree crops like *Eucalyptus* (Grattapaglia and Sederoff, 1994; Brondani *et al.*, 2004), apple (Conner *et al.*, 1997), olive (Wu *et al.*, 2004) and *Pinus* (Yin *et al.*, 2003; Chen *et al.*, 2010), this mapping strategy was followed.

2.5.2 Genetic markers used in linkage mapping

In the present study, dominant markers: RAPD, AFLP and RGA were employed along with other co-dominant markers SSRs and SNPs for the construction of genetic linkage map. Single-dose polymorphic markers like RAPD and AFLP segregated in a 1:1 ratio in F_1 population for the presence and absence of the band, thus facilitating the occurrence of pseudo-test cross configurations (Grattapaglia and Sederoff, 1994). AFLP analysis is an efficient and economic technique for identifying markers for mapping due to its higher multiplex ratio and reproducibility than RAPDs (Vos *et al.*, 1995). A reliable framework map constructed with random markers could be used as the starting point for the addition of new co-dominant markers. The dominant AFLP markers were utilized by Lespinasee *et al.* (2000) along with other co-dominant markers for the construction of the saturated genetic linkage map in rubber. Genetic linkage maps of *Eucalyptus grandis*, *E. urophylla*, apple and *Vitis* were generated using the dominant RAPD markers Grattapaglia and Sederoff, 1994; Conner *et al.*, 1994; Gokabayarak *et al.*, 2006).

Genomic and EST-derived SSR markers were also used for the construction of the present linkage map in *Hevea*. In *H. brasiliensis*, microsatellite markers were utilized for genetic linkage map construction by Lespinasse *et al.* (2000), which was later used to identify the QTL variants conferring resistance to the South American leaf blight (Lespinasse *et al.* 2000a; Le Guen *et al.*, 2003; 2007). Recently, genetic linkage maps solely based on microsatellite markers were developed by Feng *et al.* (2010) and Souza *et al.* (2011) in rubber. Microsatellite markers were utilized along with other dominant markers and further QTL analysis was performed in European beech (Scalfi *et al.*, 2005). In *Vitis*, Riaz *et al.* (2004) developed a genetic linkage map using microsatellite markers with an aim of developing a high-density consensus map. Genetic linkage maps of two apple cultivars were generated based on AFLP and microsatellite markers (Kenis *et al.*, 2005). Khadari *et al.* (2010) used AFLP, SSR and ISSR markers for the construction of the genetic linkage map in olive. Dondini *et al.* (2006) constructed linkage map in apricot using SSR markers. In cassava, the first microsatellite based linkage map was constructed by Okogbenin *et al.* (2006) and further QTL analysis was performed. In grapes, QTL analysis was performed based on a linkage map generated with dominant and co-dominant markers (Dalbo *et al.*, 2000).

2.5.3 Segregation analysis of genetic markers

In the present study, 100 scorable segregating RAPD markers were generated from 46 random primers with an average of 2.1 markers/ primer. Almost similar results were obtained with other crops like *Citrus* spp. (Cristofani *et al.*, 1999), *Eucalyptus* spp. (Verhaegen and Plomion, 1996), *P. thunbergii* (Kondo *et al.*, 2000), European beech (Scalfi *et al.*, 2005), *P. pinaster* (Costa *et al.*, 2000), *Eucalyptus* (Grattapaglia and Sederoff, 1994) and tea (Hackett *et al.*, 2000). AFLP analysis resulted in 86 scorable amplicons from nine primer combinations yielding 4.35 markers per primer combination. AFLP markers generally reveal a larger number of polymorphic bands than other markers like RAPDs. The less number of polymorphic loci obtained in the present study is due to the less genetic divergence between the two parental clones RRII 105 and RRII 118. Lespinasse *et al.* (2000) could generate on an average 13 segregating bands/primer combination with AFLP markers in their linkage map construction in rubber. The parental combination of the progeny was highly divergent as the same was derived from a cross between *H. brasiliensis* (clone PB 260) and an interspecific hybrid of *H. brasiliensis* and *H. benthamiana* (RO 38, alias FX 3899). With AFLP markers, on an average 10.5 bands per primer combination was achieved in tea (Hackett *et al.*, 2000) and 15 polymorphic markers per combination in maritime pine (Chagné *et al.*, 2002). In the present analysis, 27 SSR markers with two alleles showing heterozygosity in any of the parent (lm x ll, nn x np), eight markers with two alleles - heterozygous in both parents (hk x hk), 11 markers with three alleles - heterozygous in both the parents with a common allele (ef x eg) and six markers with four alleles - heterozygous in both the parents (ab x cd) were generated.

2.5.4 Linkage analysis of SNPs in latex biosynthesis genes and other genes

In the present study, the SNPs anchoring mevalonate kinase locus, involved in rubber biosynthesis pathway was mapped on LG-18 in the female parent RRII 105. STS marker developed from the gene farnesyl diphosphate synthase was mapped on LG-3 in the male parent RRII 118. The SNPs from ubiquitin precursor gene was mapped on LG-1 of the female parent RRII 105. The ubiquitin tag directs proteins to the proteasome which degrades and recycles unwanted proteins. The genes geranylgeranyl diphosphate and latex patatin homolog were remained unlinked. Even though these findings could not be

utilized at this moment because of unsaturation of the map, in future this linkage map can be useful for QTL analysis for yield related traits. In almond, Wu *et al.* (2009) mapped 12 gene anchored SNPs to various linkage groups. In *Vitis*, SNPs were identified from EST sequences and they were mapped to linkage groups (Troggo *et al.*, 2007).

2.5.5 Segregation distortion

Segregation distortion is referred to as a departure from an expected Mendelian ratio for a given genotype within a segregating population (Dufour *et al.*, 2001). Distorted segregation is a common feature noticed among different crops. In the present study, the segregation distortion ($P \leq 0.05$) was detected only with eight markers out of 164 loci analyzed (6%) in the female parent and distortion of six markers out of 145 loci (4%) was noticed in the male parent. Only the dominant markers RAPDs and AFLPs showed distorted segregation. Lorieux *et al.* (1995) reported that segregation distortion had less impact on the recombination frequencies of co-dominant markers compared to the dominant markers. Segregation distortion may occur due to several reasons such as deleterious genes segregating in the mapping population, heterogeneity within the parents (Grandillo and Tanksley, 1996), chromosomal rearrangements (Tanksley, 1984), reproductive differences between the two parents (Foolad *et al.*, 1995) and pre or post zygotic selection of genes flanking the distorted loci (Joober *et al.*, 2000).

2.5.6 Cluster analysis of progeny population

The cluster analysis was also performed to get an idea into the segregation pattern of the progeny. It was observed that no two progenies were similar which indicated that the population was generated through random selection without any bias. The less number of skewed markers as revealed through segregation analysis may be due to symmetric behavior of the progeny population. In other reports of *Hevea*, segregation distortion was only 1.4% (Lespinasse *et al.*, 2000) and 7.38% of loci were distorted from the Mendelian ratio in the genetic linkage map constructed by Feng *et al.* (2010), which was based on genomic and EST derived SSR markers. In almond 17.9% of markers had significantly skewed genotype ratios at the level of $P < 0.05$ (Tavassolian *et al.*, 2010). Among these, only 5.3% of the co-dominant markers showed deviation. In *Eucalyptus*, only 5% of the markers departed from the Mendelian expectation of 1:1 (Grattapaglia and Sederoff, 1994) and a low segregation distortion of 1.5% was noticed in Japanese larch (Arcade *et al.*, 2002).

2.5.7 Linkage groups

In the preliminary linkage analysis with 1:1 markers using MAPMAKER, 23 linkage groups (LGs) were obtained for both the female and male parents, spanning a distance of 1384.7 cM with 89 markers and 768.5 cM with 66 markers respectively. The average distance between markers in the female map was calculated as 15.5 cM and that of male map was 11.64 cM. The JoinMap analysis of all the markers segregating in different ratios, resulted in 24 LGs for both the female and male maps; the female map covering 762 cM, comprising of 103 markers with an average distance between two of 7.39 cM and the male map spanning a distance of 634 cM comprising of 83 markers, with an average distance between two loci of 7.63 cM. The length of both the maps remained almost the same. Similar results were obtained with almond (Tavassolian *et al.*, 2010) and yellow passion fruit (Carneiro *et al.*, 2002). Sixty one and 62 markers remained unlinked in the female and the male parent respectively of which the majority of them were SSR markers in the present linkage analysis. Other linkage maps developed in *Hevea* spanned over a distance of 2144 cM with a total of 717 loci assembled into 18 linkage groups (Lespinasee *et al.*, 2000a). Linkage map constructed by (Feng *et al.*, 2010) covered a total distance of 1937.06 cM with an average genetic distance of 21.29 cM between adjacent markers in 18 linkage groups. In the studies of Souza *et al.* (2011), 225 markers were distributed in 23 linkage groups and total genetic distance was 2471.2 cM with an average genetic distance of 11 cM between adjacent markers using the OneMap software (Margarido *et al.*, 2007).

The expected number of eighteen linkage groups for a comprehensive linkage map of *Hevea* (2n = 36) could not be achieved in the present study and was exceeded by five linkage groups. This excess of linkage groups was due to incomplete coverage of the genome with the available marker loci. Furthermore, the high percentage of polymorphic SSR markers that remained unmapped resulted in unsaturation of the map. It is anticipated that the smaller linkage groups will be brought together as new markers are introduced in the segregation analysis. Similar results were obtained with *Hevea* and other crops also, where the linkage groups exceeded the basic chromosome number. Souza *et al.* (2011) identified 23 linkage groups with 225 markers in rubber. In cassava (Okogbenin *et al.*, 2006; Kunkeaw *et al.*, 2010), 25 and 33 linkage groups were obtained respectively, although the haploid chromosome number is 18. In *Vitis* (2n=38) 20 linkage groups were obtained by Gokabayarak *et al.* (2006).

2.5.8 Map integration

In the present analysis, only two parental groups could be merged together to form an integrated map using the map integration function of JoinMap. This is due to the less number of linked markers which are common to both the parents. Many informative microsatellite markers which could be used as bridge markers remained unlinked in the present linkage analysis. Byrne *et al.* (1995) suggested that the number and distribution of fully informative markers greatly influenced the generation of an integrated map. Also, the bridge markers linked in the female parent differ from that of the male parent and hence groups cannot be integrated. In *Hevea*, Lespinasse *et al.* (2000) constructed a synthetic map by merging the individual parental maps. In *Eucalyptus*, Thamarus *et al.* (2002) constructed a combined parental map using the loci which segregated in both the parents.

The total genetic map length of rubber, calculated by MAPMAKER (1384 cM) was more than that of calculated by JoinMap (762 cM). The comparison between the total map lengths obtained with JoinMap and MAPMAKR showed difference, even though the same mapping function (Kosambi) was used in both the software, which was repoeted also reported in barley (Qi *et al.*,1998), maritime pine (Chagne *et al.*, 2002), loblolly pine (Sewell *et al.*,1999) and in sugi (Tani *et al.*, 2003). The difference between map lengths obtained with JoinMap and MAPMAKER results from the different methods of computing map lengths. MAPMAKER calculates the map length as the sum of adjacent distances, *i. e.* using adjacent marker pairs only. JoinMap on the other hand uses all pair-wise estimates (above a pre-defined LOD threshold) for calculating the total map length. Whenever the assumed level of interference does not exactly reflect the true interference, the two methods will produce slightly different total map lengths. The likelihood method applied in MAPMAKER assumes an absence of interference and recombination frequencies are simply translated into centimorgans according to the chosen mapping function. The JoinMap package, however, does take interference into account. Therefore, when there is interference, JoinMap will produce shorter maps than MAPMAKER, even when both programmes use the same Kosambi mapping function (Stam 1993).

Although the present maps are relatively unsaturated in terms of the number of markers incorporated, this may be considered as an initial step towards the genetic dissection of agronomically important characters in rubber.

Conclusion and future prospects

Molecular markers were generated from different marker systems such as RAPDs, AFLPs, SSRs, RGA and SNPs. These markers were used for polymorphism analysis and also for the generation of molecular markers for the segregation analysis towards the construction of genetic linkage maps. The genetic diversity assessed using the RAPD, AFLP and SSR markers revealed a narrow genetic diversity of approximately 35% within the 28 popular *H. brasiliensis* clones. Genetic linkage maps were constructed with MAPMAKER and JoinMap softwares using all the developed markers in the parents of the progeny population (RRII 105 and RRII 118). Twenty four linkage groups were generated for both the parents using the JoinMap software and twenty three linkage groups were generated with the MAPMAKER software.

The molecular markers developed are major tools for cultivar identification, germplasm preservation and protecting the varietal right of farmers. The knowledge of the extent of diversity between various clones can be effectively used in Hybridization programmes for the development of diverse elite cones. The linkage maps created in this study are preliminary, but provide a stepping stone towards a high density genetic map incorporating large number of high throughput markers in near future. The mapping of the latex biosynthesis gens in various linkage groups in the present study can be used for marker assisted selection and map based cloning which is possible through fine mapping with a saturated genetic linkage map. This also provides a base for QTL mapping studies in which phenotypic traits of importance can be identified and associated to a specific location of the genome. Further, the linkage maps generated with the markers developed in the present initiative can be merged with other segregation datas of a different progeny population also and thereby a saturated genetic linkage map can be developed.

References

Achere V, Faivre-Rampant P, Jeandroz S, Besnard G, Markussen T, Aragones A, Fladung M, Ritter E, Favre JM (2004) A full saturated linkage map of Picea abies including AFLP, SSR, ESTP, 5S rDNA and morphological markers. Theor Appl Genet 108: 1602-1613

Adams M. D, Kelley J M, Gocayne J D, Dubrick M, Polymeropoulos M H, Xiao H, Merril C R, Wu A, Olde B, Moreno R F, Kerlavog A R, McCombie W R and Venter J C (1991) Complimentary DNA sequencing: expressed sequence tags and human genome project. Science 252: 1651-1656

Agwanda C O, Philippe L, Pierre T, Marie-Christine C, Andr´e C (1997) entification of RAPD markers for resistance to coffee berry disease *Colletotrichum kahawae* in arabica coffee. Euphytica 97: 241-248

Aishwarya Vand Sharma P C (2007) UgMicroSatdb:database for mining microsatellites from unigenes. Nucleic Acids Res 1: 1-4

Akano A O, Dixon A G O, Mba C, Barrera E, Fregene M (2002) Genetic mapping of a dominant gene conferring resistance to the cassava mosaic disease (CMD). Theor Appl Genet 105: 521–525

Aljorna L, Bueno M A, Aguinagalde I, Martín J P (2007) Fingerprinting and genetic variability in cork oak (*Quercus suber* L.) elite trees using ISSR and SSR markers Ann Forest Sci 64: 773-779

Allard R W (1956) Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia 24: 235-278

An Ding L, Shou-cai C, Kun-Xin W, Shao-Ping F (2001) AFLP fingerprinting analysis of elite *Hevea brasiliensis* germplasm. Acta Bot Sin 43: 941-947

Aranzana M J, Illa E, Howad W, Arús P (2012) A first insight into peach [*Prunus persica* (L.) Batsch] SNP variability. Tree Genet Genomes 8: 1359-1369

Arcade A, Anselin F, Faivre P R (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. Theor Appl Genet 100: 299-307

Arnold M L, Buckner C M and Robinson J J (1991) Pollen-mediated introgression and hybrid speciation in Louisiana. Proc Natl Acad Sci 88: 1398–1402

Ashworth V E T M and Clegg M T (2003) Microsatellite markers in avocado (*Persea americana* Mill.): genealogical relationships among cultivated avocado genotypes. J Heredity 94: 407-415

Ashworth V E T M, Kobayashi M C, De La Cruz M and Clegg M T (2004) Microsatellite markers in avocado (*Persea americana* Mill.): development of dinucleotide and trinucleotide markers. Scientia Horticuturae 101: 255-267

Atan S, Low F C and Saleh N M (1996) Construction of a microsatellite–enriched library from *Hevea brasiliensis*. J Nat Rubber Res 11: 247-255

Baes P, Custsem V (1993) Electrophoretic analysis of eleven isozyme system and their possible use as biochemical markers in breeding chicory (*Chychorium intybus* L.). Plant Breed 110: 16-23

Barbazuk W B, Emrich S J, Chen H D, Li L and Schnable P S (2007) SNP discovery via 454 transcriptome sequencing. Plant J.51: 910–918

Bassil N V, Botta R, Mehlenbacher S A (2005) Microsatellite markers in the hazelnut: isolation, characterization and cross-species amplification in *Corylus*. J. Amer. Soc. Hort.Sci. 130: 543-549

Beckman J S, Soller M (1986) Restriction fragment length polymorphisms in genetic improvement of agricultural species. Euphytica 35: 111-124.

Bergmann F, Gregorius H R, Scholz F (1989) Isoenzyme: Indicators of environmental impact on plants or environmentally stable gene markers? In: Scholz F, Gregorius H R, Rudin D (eds) Genetic effects of air pollutants in forest tree populations. Springer-Verlag Heidelberg, New York, Tokyo, pp: 17-28

Besse P, Lebrun P, Seguin M, Lanaud C (1993) DNA fingerprints in *Hevea brasiliensis* (rubber tree) using human minisatellite probes. Heredity 70: 237-244

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

Bhattramakki D, Dolan M, Hanafey M, Wineland R, Vaske D, Register J C, Tingey S V, Rafalski A (2002) Insertion deletion polymorphisms in 3 regions of maize genes occur frequently and can be used as highly informative genetic markers. Plant Mol Biol 48: 539–547

Bicalho K C, Oliveira L E M, de Santos J B, Mesquita A C (2008) Similaridade genética entre clones de seringueira (*Hevea brasiliensis*) por meio de marcadores RAPD. Ciência Agrotécnica 32: 1510-1515

Botstein D, White R L, Skolnick M, Davis R W (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphism. Am J Hum Genet 32: 314-331

Bradshaw H D, Settler R F (1995) Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, form, and phenology traits in a forest tree. Genetics 139: 963–973

Brown S, Kresovich S (1996) Molecular characterization for plant genetic resources conservation. In: A.H. Paterson (ed) Genomic mapping in plants. R.G. Landes Austin, TX, USA pp.

Byrne M, Murrell J C, Allen B, Moran G F (1995) An integrated genetic linkage map for Eucalypts using RFLP, RAPD and isozyme markers. Theor Appl Gen 91: 6-7

Calcagnotto D (2001) Taxas de Evolução e o Relógio Molecular. In: (Matioli SR ed.). Biologia Molecular e Evolução Holos Ribeirão Preto 51-63

Cesar A H, Lucia A K, Rafael A I, Mario L A (2006) Analysis of genetic variation in clones of rubber (*Hevea brasiliensis*) from Asian, South and Central American origin using RAPD markers. Revista Colombiana de Biotech 8: 29-34

Cassoli M, Mattioni C, Cherubini M, Villani F (2001) A genetic linkage map of European chestnut (*Castanea sativa* Mill.) based on RAPD, IISR and isozyme markers. Thoer Appl Genet 102: 1190-1199

Celton J M, Tustin D, Chagné D, Gardiner S (2009) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. Tree Genet Genomes 5: 93-107

Cervera T M, Storme R V, Ivens B, Gusmao J, Liu B H, Hostyn V, Slycken V J, Montagu V M, Boerjan W (2001) Dense genetic linkage maps of three *populus* species (*Populus deltoides*, *P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. Genetics 158: 787–809

Cevik S M, Moore A G (2007) Construction of a genetic linkage map of citrus with Random Amplified Polymorphic DNA (RAPD) markers using a progeny population from a complex intergeneric cross. Turk J Bot 31: 79-86

Cevik S M, Moore A G (2007) Construction of a genetic linkage map of citrus with Random Amplified Polymorphic DNA (RAPD) markers using a progeny population from a complex intergeneric cross. Turk J Bot 31: 79-86

Chagné D, Lalanne C, Madur D, Kumar S, Frigerio J M, Krier C, Decroocq S, Savouré A, Bou-Dagher-Kharrat M, Bertocchi E, Brach J, Plomion C (2002) A high density genetic map of Maritime pine based on AFLPs. Ann For Sci 59: 627–636

Chagne D, Gasic K, Crowhurst R N, Han Y, Bassett H C, Bowatte D R, Lawrence T J, Rikkerink E H A, Gardiner S E, Korban S S (2008) Development of a set of SNP markers present in expressed genes of the apple. Genomics 92 :353-358

Chaparro J S, Werner P J, O Malley D, Sederoff R R (1994) Targetted mapping and linkage analysis of morphological, isozyme and RFLP markers in peach. Theor Appl Genet 87: 805-815

Chase M, Kesseli R, Bawa K (1996) Microsatellite markers for population and conservation genetics of tropical trees. Am J Bot 83: 51-57

Chase M, Kesseli R, Bawa K (1996) Microsatellite markers for population and conservation genetics of tropical trees. Am. J. Bot. 83: 51-57

Chen C X, Zhou P, Choi Y A, Huang S, Gmitter F G, Jr (2006) Mining and characterizing microsatellites from citrus ESTs. Theor Appl Genet. 112: 1248–1257.

Chen M M, Feng F, Sui X, Han S, (2010) Genetic linkage maps of Pinus koraiensis Sieb. et Zucc based on AFLP markers. Afr. J. Biotech. 9(35): 5659-5664

Chiapparino E, Lee D, Donni P (2004) Genotyping single nucleotide polymorphisms in barley by tetra-primer ARMS-PCR Genome 47: 414-420

Ching A, Caldwell K S, Jung M, Dolan M, Smith O S, Tingey S (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet 3:19

Cho Y G, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch S R, Park W D, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100: 713-722

Cho KH, Shin IS, Kim KT, Suh EJ, Hong SS, Lee HJ (2009) Development of AFLP and CAPS markers linked to the scab resistance gene, *Rvn2*, in an inter-specific hybrid pear (*Pyrus* spp.). J Hortic Sci Biotechnol 84:619–624

Chow K S, Wan K L, Isa M N (2007) Insights into rubber biosynthesis from transcriptome analysis of *Hevea brasiliensis* latex. J Exp Bot 58: 2429–40

Clark M S, Wall W J, (1996) Chromosomes: The Complex Code. New York: Chapman and Hall p 345

Clément Demange A, Legnate H, Seguin M, Carron M P, Le Guen V, Chapuset T, Nicolas D (2000) Rubber Tree. In: Charrier A, Jacquot M, Hamon S, Nicolas D (eds).Tropical Plant breeding, Collection Repères, ISBN 978-1-57808-144-8, Montpellier, France pp. 455-480

Clement Demange A, Priyadarshan P M, Hoa T T T, Venkatachalam P (2007) *Hevea* rubber breeding and genetics. Plant Breed Rev 29

Collard B C Y, Jahufer M Z Z, Brouwer J B, Pang E C K (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement. The basic concepts. Euphytica 142: 169-196

Colorado M, Alberto J, Hernando R, Martin F (2009) Genetic mapping and QTL analysis for carotenes in a s1 population of cassava. Acta Agron Palmira 58

Concibido V C, Young N D, Lange D A, Denny R L, Danesh D, Orf J H, (1996) Targeted comparative genome analysis and qualitative mapping of a major partial-resistance gene to the soybean cyst nematode. Theor Appl Genet 93: 234-241

Conditt R, Hubbell S P (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34: 66-71

Conner P J, Brown S K, Weeden (1997) Randomly amplified polymorphic DNA-based genetic linkage maps of three apple cultivars. J Am Soc Hort Sci 122: 350–359

Cordeiro G M, Maguire T L, Edwards K J, Henry R J (1999) Optimisation of a microsatellite enrichment technique in *Saccharum* spp. Plant Mol. Biol. Reporter 17: 225-220

Costa P, Pot D, Dubos C, Frigerio J M, Pionneau C, Bodénès C, Bertocchi E, Cervera M T, Remington D L , Plomion C (2000) A genetic map of maritime pine based on AFLP, RAPD and protein markers. Theor. Appl. Genet. 100,39-48.

Cristofani M, Machado M A, Grattapaglia D (1999) Genetic linkage maps of *Citrus sunki* Hort. ex. Tan. and *Poncirus trifoliata* (L.) Raf. and mapping of *citrus tristeza* virus resistance gene. Euphytica 109:25-32

Dalbo M A, Ye G N, Weeden N F, Steinkellner H, Sefc K M, Reisch B I, (2000) A gene controlling sex in grapevines placed on a molecular marker-based genetic map. Genome 43: 333–340

De Oliveira A C, Bstianel M, Cristofani-Yaly M, do Amaral A M, Machado M A (2007) Development of genetic maps of the citrus varieties 'Murcott' tangor and 'Pera' sweet orange by using fluorescent AFLP markers. J Appl Genet 48: 219-231

Dean W (1987) Brazil and the struggle for rubber: A study in environmental history. Cambridge University Press,Cambridge p 234

De Riek J, Calsyn E, Everaert I, Van Bockstaele E, De Loose M (2001) AFLP based alternatives for the assessment of distinctness,uniformity and stability of sugar beet varieties. Theor Appl Genet 103: 1254-1265

Despres L, Gielly L, Redoutet B, Taberlet P (2003) Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability. Mol Phylogenet Evol 27:185–196

Dondini L, Lain O, Geuna F, Banfi R, Gaiotti F, Tartarini S, Bassi D, Testolin R (2007). Development of a new SSR based linkage map in apricot and analysis of synteny with existing prunus maps. Tree Genet Genomes 3: 239-249

Dorgac C, Ozgen M, Simsek O, Kakar Y A, Kiyga Y, Celebi S, Gunduz K, Serce S, (2008) Molecular and morphological diversity among pomegranate (*Punica granatum*) cultivars in Eastern Mediterranean region of Turkey. Afr J Biotechnol 7: 1294-1301

Doyle J J, Doyle J L (1990) A rapid total DNA preparation procedure for fresh plant tissue. Focus 12:13-15

Dow B D, Ashley M V (1998) Factor influencing male mating success in bur oak *Quercus macrocarpa*. New Forest 15: 161-180

Dufour P, Johnsson C, Antoine-michard S, Cheng R, Murigneux A, Beckert M (2001) Segregation distortion at marker loci: variation during microspore embryogenesis in maize. Theor Appl Genet 102: 993-1001

Duta-Cornescu G, Simon-Gruita A, Constantin N, (2009) A comparative study of ARMS – PCR and RFLP – PCR as methods for rapid SNP identification. Rom Biotech Let 14: 4845-4850

Dutton C, Sommer S S (1991) Simultaneous detection of multiple single-base alleles at a polymorphic site. BioTechniques 11:700–702.

Dvornik V, Mikkonen M, Sirvio A, Savolainen O (2000) Plant and Animal Genome VIII Conference, 9–12 January San Diego, USA

Echt C S, May-Marquardt, Hseih M, and Zahorchak R (1996) Characterization of microsatellite markers in eastern white pine. Genome 39:1102-1108

Echt C, Knapp S, Liu B-H (1992) Genome mapping with noninbred crosses using GMendel 2.0. Maize Genet Coop Newsl 66: 27–29

Edwards K J, Barker J H A, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTech 20: 758-760

Elsik C G, Minihan V T, Hall S E, Scarpa A M, Williams C G (2000) Low-copy microsatellite markers for *Pinus taeda* L. Genome 43: 550-555

Emerson R, Beadle G, Fraser A (1935) A summary of linkage studies in maize. Cornell Univ Agric Esp Stn Memoir 180: 1-83

Etienne C, Rothan C, Moing A, Plomion C, Bodénè C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E (2002) Candidate genes and QTLs for sugar and organic acid contents in peach [*Prunus persica* (L.) Batsch]. Theor Appl Genet 105:145-159

Feng S P, Li W G, Huang H S, Wang J Y, Wu Y T (2009) Development, characterization and cross-species/genera transferability of EST-SSR markers for rubber tree (*Hevea brasiliensis*). Mol Breed 23: 85-97

Feng S P, Li W G, Yu F, Wang J Y, Wu Y T (2010) Construction of genetic linkage map for rubber tree (*Hevea brasiliensis*) based on SSR markers. Hereditas 8: 857-863

Fischer D, Bachman K (1998) Microsatellite enrichment in organisms with large genome (*Allium cepa* L.) BioTech 24: 796-802

Flint-Garcia S A, Thuillet A, Yu J M, Pressoir G, Romero S M, Mitchell S E, Doebley J, Kresovich S, Goodman M M, Buckler E S (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44: 1054–1064

Foolad M R, Arulsekar S, Becerra V, Bliss F A (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theor Appl Genet 91: 262-269

Fregene M, Angel F, Gomez R, Rodriguez F, Chavarriaga P, Roca W, Tohme J, Bonierbale M (1997) A molecular map of cassava (*Manihot esculenta* Crantz). Theor Appl Genet 95:431–441

Fregene M, Okogbenin E, Mba C, Angel F, Suarez M C, Guitierez J, Chavarriaga P, Roca W, Bonierbale M, Tohme J (2001) Genome mapping in cassava improvement. Challenges, achievements and opportunities. Euphytica 120: 159-165.

Ganal M W, Altmann T, Roder M S (2009) SNP identification in crop plants. Current Opinion Plant Biol 12:211-217

Gao L F, Tang J, Li H, Jia J (2003). Analysis of microsatellites in major crops assessed by computational and experimental approaches. Mol Breed 113:163–185.

García R I A, González S S M, Montoya C D, Aristizabal F (2011) Identification in silico of SSR markers for genotyping *Hevea* sp. clone gardens in Colombia. Agr Colomb 29: 359-366

Gardiner J S, Schroeder M L, Polacco H, Sanchez-Villeda H, Fang Z , Morgante M, Landewe T, Fengler K, Useche F, Hanafey M, Tingey S, Chou H, Wing R, Soderlund C, Coe E H Jr.(2004) Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization. Plant Physiol 134: 1317–1326

Garg K, Green P, Nickerson D A (1999) Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. Genome Res 9: 1087-1092

Gastier J M, Pulido J C, Sunden S, Brody T, Buetow K H, Murray J C, Weber J L, Hudson T J, Sheffield V C, Duyk G M (1995) Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. Hum Mol Genet 4: 1829-1836

Gerlach H K, Stosser R (1997) Patterns of random amplified polymorphic DNAs for sweet cherry (*Prunus Avium* L.) cultivar identification. J Appl Bot 71: 212–218

Germer S, Higuchi R (1999) Single-tube genotyping without oligonucleotide probes. Genome Res 9: 72–78

Glaszmann J C, Fautret A. Noyer J L, Feldmann P, Lanaud C (1989) Biochemical genetic markers in sugarance. Theor Appl Genet 78: 537-543

Glaubitz J C, Embiri L C, Moran G F (2001) Dinucleotide microsatellite from *Eucalyptus sieberi*: inheritance diversity and improved scoring of single-base differences. Genome 44: 1041-1045

Gokabayarak Z, Ozer C, Soylemezolu G (2006) Preliminary results on genome mapping of an Italia x Mercan grapevine population. Turk J Agric For 30: 273-280

Gokabayarak Z, Ozer C, Soylemezolu G (2006) Preliminary results on genome mapping of an Italia x Mercan grapevine population. Turk J Agric For 30: 273-280

Gouvea L R L, Rubiano L B, Chioratto A F, Zucchi M I, Gonçalves P S (2010) Genetic divergence of rubber tree estimated by multivariate techniques and microsatellite markers. Genet Mol Biol 33: 308-318

Grandillo S, Tanksley S D (1996) Genetic analysis of RFLPs, GATA microsatellites and RAPDs in a cross between *L. esculentum* and *L. pimpinellifolium*. Theor Appl Genet 92:957–965

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: Mapping strategy and RAPD markers. Genetics 137: 1121–1137

Grattapaglia D, Silva Jr O B, Kirst M, Lima B M, Faria D M and Pappas Jr G (2011) High-throughput SNP genotyping in the highly heterozygous genome of Eucalyptus: assay success, polymorphism and transferability across species.BMC Plant Biology 2011, 11:65

Grodzicker T, Williams J, Sharp P, Sambrook J (1974) Physical mapping of temperature sensitive mutations Cold spring Harbor Symp Quart Biol 39:439-446

Guimarães P E M, Costa M C R (2002) SNPs: sutis diferenças em um código. Biotecnol Cienc Desenvolv 26: 24-27

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

Gupta P K, Roy G K, Prasad M (2001) Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Sci 80: 525-535

Gut I G (2004) An overview of genotyping and single nucleotide polymorphisms (SNP). In: Rapley R Harbron S (eds) Molecular analysis and genome discovery. Wiley Chichester pp 43–64.

Hackett A C, Wachira F N, Paul S, Powell W, Waugh R (2000) Construction of a genetic linkage map for *Camellia sinensis* (tea). Heredity 85: 346-355

Haldane J B S (1919) The combination of linkage values, and the calculation of distances between the loci of linked factors. J Genet 8: 299–309

Halle and Comb 1974 hevea

Hamilton M B, Pincus E L, Di Fiore A, Fleischer R C (1999) Universal linkers and ligation procedures for the construction of genomic DNA libraries enriched for microsatellites. BioTech 27: 500-507

Hammond-Kosack K E, Parker J E (2003) Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. Curr Opin Biotechnol. 14:177-193

Hamrick J L, Godt M J (1990) Allozyme diversity in plant species. In: Brown A H D, Clegg M T, Kahler A L, Weir B S (eds.) Plant population genetics, breeding, and genetic resources, Sinauer, Sunderland, M A pp. 43–63

Han K H, Shin D H, Yang J, Kim I J, Oh S K, Chow K S (2000) Genes expressed in the latex of *Hevea brasiliensis*. Tree Physiol 20:503–510

Han Y, Zheng D, Vimolmangkang S, Khan A M, Beever E J, Korban S S (2011) Integration of physical and genetic maps in apple confirms whole-genome and segmental duplications in the apple genome. J Exp Bot 62(14): 5117–5130

Hansson B, Kawabe A (2005) A simple method to score single nucleotide polymorphisms based on allele-specific PCR and primer-induced fragment-length variation. Mol Ecol Notes 5:692–696

Hemmat M, Weeden N F, Manganaris A G, Lawson D M (1994) Molecular marker linkage map for apple. J Hered 85: 4-11

Hendre S P, Singh L, Aggarwal K R (2007) Construction of framework molecular linkage map of Robusta Coffee, *Coffea canephora* with RAPD, AFLP, g-SSRs And EST-SSRs using pseudo-testcross strategy. Plant & Animal Genomes XV Conference, January 13-17, 2007,Town and Country Convention Center San Diego, CA P651 : Other Species.

Hermosa M R I, Grondona J M, Diaz-Minguez E A, Iturriaga Monte E (2001) Development of a strain-specific SCAR marker for the detection of *Trichoderma viride*, a biological control agent against soil borne fungal plant pathogens. Curr Genet 38: 343-50

Hernandez C R L, Afanador R A, Isaza, Lobo A M (2006) Analysis of genetic variation in clones of rubber (*Hevea brasiliensis*) from Asian, south and central American origin using RAPD markers. Rev Colomb Biotecnol 8: 29-34

Hinten G N, Hale M C, Gratten J, Mossman JA, Lowder BV, Mann M K, Slate J (2007). SNP-SCALE: SNP scoring by colour and length exclusion. Mol Ecol Notes 7:377–388

Hoogendoorn B, Owen M J, Oefner P J, Williams N, Donovan M C (1999) Genotyping single nucleotide polymorphisms by primer extension and high performance liquid chromatography. Hum Genet 104(1): 89-93

Hospital F, Charcosset A (1997) Marker-assisted introgression of quantitative trait loci. Genetics 147:1469-1485

Hurtado A, Romero C, Abbott G V, Llácer G, Badenes L (2002) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. Theor Appl Genet 108: 1120-1130

Ishiguro A, Kubota T, Soya Y, Sasaki H, Yagyu O, Takarada Y (2005) High-throughput detection of multiple genetic polymorphisms influencing drug metabolism with mismatch primers in allele-specific polymerase chain reaction. Anal Biochem; 337:256–61

Iwaki K, Nishida J, Yanagisawa T, Yoshida H, Kato K (2002) Genetic analysis of Vrn-B1 for vernalization requirement by using linked dCAPS markers in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 104:571–576

James C M, Wilson F, Hadonou A M, Tobutt K R (2003) Isolation and characterization of polymorphic microsatellites in diploid strawberry (*F. vesca* L.) for mapping diversity studies and clone identification. Mol Ecol 3: 171-173

Jeffreys A J (1979) DNA sequence variants in the GA and B-globin genes of man. Cell. 18: 1-10

Jeffreys A J, Wilson V, Thein S L (1985) Individual-specific 'fingerprints' of human DNA. Nature 316: 76–79

Jones C J, Edwards K J, Castaglione S, Winfield M O, Sala F, van de Wiel C, Bredemeijer G, Vosman B, Matthes M, Daly A, Brettschneider R, Bettini P, Buiatti M, Maestri E, Malcevschi A, Marmiroli N, Aert R, Volckaert G, Rueda J, Linacero R, Vazquez A, Karp A (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breed 3: 381–390

Joobeur T, Viruel M A, de Vicente M C, Jáuregui B, Ballester J, Dettori M T, Verde I, Truco M J, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arús P. (1998) Construction of a saturated linkage map for *Prunus* using an almond × peach F2 progeny. Theor Appl Genet 97: 1034–1041

Joobeur T, Periam N, de Vicente M C, King G J, Arús P (2000) Development of a second generation linkage map for almond using RAPD and SSR markers. Genome 43: 649-655

Kaundun S S, Matsumoto S (2003) Development of CAPS markers based on three key genes of the phenylpropanoid pathway in tea, Camellia sinensis (L.) O. Kuntze, and differentiation between assamica and sinensis varieties. Theor. Appl. Genet. 106: 375–383

Kenis K, Keulemans J (2005) Genetic linkage maps of two apple cultivars (Malus x domestica Borkh.) based on AFLP and microsatellite markers. Mol Breed 15: 205–219

Khadari B, Aabidine E Z A, Grout C A (2010) Genetic linkage map of olive based on amplified fragment length polymorphism, intersimple sequence repeat and simple sequence repeat markers. J Am Soc Hort Sci. 135: 6548-555

Kijas J H M, Fowler J C S, Garbett C A, Thomas M R (1994) Enrichment of microsatellites from citrus genome using biotinylated oligonucleotides sequences bound to streptavidin-coated magnetic particles. BioTech 16: 656-662

Ko J H, Chow K S, Han K H (2003) Transcriptome analysis reveals novel features of the molecular events occurring in the laticifers of *Hevea brasiliensis* (para rubber tree). Plant Mol Biol 53: 479–92

Koblizkova A, Dolezel J, Macas J (1998) Subtraction with 3' modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. BioTech 25: 32-38

Kondo T, Terada K, Hayashi E, Kuramoto N, Okamura M, Kawasaki H (2000) RAPD markers linked to a gene for resistance to pine needle gall midge in Japanese black pine (*Pinus thunbergii*). Theor Appl Genet 100:391-395.

Kosambi D (1944) The estimation of map distances from recombination values. Ann Eugen 12: 172–175

Kostia S, Sirkaa-Liisa V, Vakkari P, Pulkkinen P (1995) Microsatellite sequence in a conifer *Pinus sylvestris*. Genome 38: 1244-1248

Kota R, Rudds S, Facius A, Kolesov G, Thicl T, Zhang H, Stein N, Mayer K and Graner A (2003) Snipping polymorphisms from large EST collection in barley. Mol Genet Genomics 270: 24-33

Krichen L, Martins J M S, Lambert P, Daaloul A, Trifi-Farah N, Marrakchi M, Audergon J M (2008) Using AFLP markers for the analysis of the genetic diversity of apricot cultivars in Tunisia. J Am Soc Horti Sci 133: 204–212

Kumar L S (1999) DNA markers in plant improvement. Biotechnol Adv 17: 143-182

Kumar P, Gupta V K, Misra A K, Modi D R, Pandey B K (2009) Potential of molecular markers in plant biotechnology. Pl Omics J 2: 141-162

Kumpatla S P, Mukhopadhyay S (2005) Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. Genome 48: 985-998

Kunert K J, Baaziz M, Cullis C A (2003) Techniques for determination of true to type date palm. Emir J Agric Sci 15: 01-16

*Kunkeaw S, Tangphatsornruang S, Triwitayakorn K (2010) Genetic linkage map of cassava (*Manihot esculenta *Crantz) based on AFLP and SSR markers. Plant Breed 129: 112–115*

Kuppuswamy M N, Hoffman J W, Kasper C K, Spitzer S G, Groce S L, Bajaj S P (1991) Single nucleotide primer extension to detect genetic diseases: Experimental application to hemophilia B (factor IX) and cystic fibrosis genes. Proc Natl Acad Sci USA 88: 1143-1147

Kwok S, Kellogg D E, McKinney N, Spasic D, Goda L,Levenson C and Sninsky J J (1990) Effects of primertemplate mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res 18: 995-1005

Kwok P Y (2001) Methods for genotyping single nucleotide polymorphisms. Ann Rev Genom Hum Genet 2: 235-258

la Rosa R, Angiolillo A, Guerrero C, Pellegrini M, Rallo L, Besnard G, Bervillé A, Martin A, Baldoni L (2003) A first linkage map of olive (*Olea europaea* L.) cultivars using RAPD, AFLP, RFLP and SSR markers. Theor Appl Genet. 106:1273-82

Lalli D A, Abbott A G, Zhebentyayeva T N, Badenes M L, Damsteegt V, Polák, J, Krška B, Salava J (2008) A genetic linkage map for an apricot (*Prunus armeniaca* L.) BC_1 population mapping plum pox virus resistance. Tree Genet Genomes 4: 481–493

Lam L V, Thanh T, Chi V T Q, Tuy L M (2009) Genetic Diversity of *Hevea* IRRDB'81 Collection Assessed by RAPD Markers. Mol Biotechnol 42: 292–298

Lambert P, Hagen L S , Arus P, Audergon J M (2004) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas x peach earlygold reference map for Prunus. Theor App Genet 108: 1120-1130

Landegren U, Nilson M and Kwok P Y (1998) Reading Bits of Genetic information: methods for single nucleotide polymorphism Analysis. Genome Res. 8: 769-776

Lander E S) , Green P (1987 Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci USA 84: 2363–2367

Lavi U, Akkaya M, Bhagwat A, Lahav E, Cregan P B (1994) Methodology of generation and characteristics of simple sequence repeat DNA markers in avocado (Persea americana M.) Euphytica 80: 171-177

Le Guen V, Lespinasse D, Oliver G, Rodier-Goud M, Pinard F, Seguin M (2003) Molecular mapping of genes conferring field resistance to South American Leaf Blight (*Microcyclus ulei*) in rubber tree. Theor Appl Genet 108(1): 160-167

Le Guen V, Garcia D, Mattos C R, Doare F, Lespinasse D, Seguin M (2007). Bypassing of a polygenic Microcyclus ulei resistance in rubber tree, analyzed by QTL detection, New Phytol. 173: 335–45.

Le Guen V, Gay C, Xiong T C, Souza L M, Rodier-Goud M, Seguin M (2010) Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*). Plant Breed. DOI: 10.1111/j.1439-0523.2010.01774.x

Le Guen V, Garcia D, Doaré F, Mattos R R C, Condina V, Couturier C, Chambon A, Weber C, Espéout S, Seguin M (2011) A rubber tree's durable resistance to

Microcyclus ulei is conferred by a qualitative gene and a major quantitative resistance factor. Tree Genet Genomes 7: 877-889

Leitch A R, Lim K Y, Leitch I J, O'Neill M, Chye M, Low F (1998) Molecular cytogenetic studies in rubber, *Hevea brasiliensis* Muell. Arg. (Euphorbiaceae). Genome 41: 464-467

Lekawipat N, Teerawatanasuk K, Rodier– Goud M, Seguin M, Vanavichit A, Toojinda T, Tragoonrang S (2003) Genetic diversity analysis of wild germplasm and cultivated clones of *Hevea brasiliensis* Muell. Arg. by using microsatellite markers. J Rubber Res 6: 36-47

Lespinasse D, Rodier-Goud M, Grivet L, Leconte A, Legnate H, Seguin M (2000) A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite and isozyme markers. Theor Appl Genet 100:127–38

Lespinasse D, Grivet L, Troispoux V, Rodier-Goud M, Pinard F, Seguin M (2000a) Identification of QTLs involved in the resistance to South American leaf blight (*Microcyclus ulei*) in the rubber tree. Theor Appl Genet 100: 975–84

Li D, Deng Z, Qin B, Liu X, Men Z (2012) De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.) BMC Genomics 13:192

Licy J (1997) Variability correlation and heterosis for yield and yield components in certain hybrid clones of para rubber tree *Hevea brasiliensis* Muell. Arg. Ph. D. Thesis University of Kerala India 158

Liebhard R, Gianfranceschi L, Koller B, Ryder C D, Tarchini R, Van de Wag E (2002) Development and charecterisation of 140 new microsatellites in apple (Malus x domestica Borkh.) Mol Breed 10: 217-241

Lima L S, Gramacho K P, Carels N, Novais R, Gaiotto F A, Lopes U V, Gesteira A S, Zaidan H A, Cascardo J C M, Pires J L, Micheli F (2009) Single nucleotide polymorphisms from *Theobroma cacao* expressed sequence tags associated with witches' broom disease in cacao. Genet Mol Res 8: 799-808

Lincoln S E, Daly M, Lander E S (1992) Constructing genetic maps with MAPMAKER/EXP 3.0 (Cambridge, Mass.: Whitehead Institute Technical Report)

Litt M, Luty J A (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44: 397-401

Liu Q, Thorland E C, Heit J A, Sommer S S (1997) Overlapping PCR for bidirectional PCR amplification of specific alleles: a rapid one-tube method for simultaneously differentiating homozygotes and heterozygotes. Genome Res 7:389–398

Liu B-H (1998) Statistical Genomics: Linkage, mapping, and QTL analysis. CRC

Liu P, Wang C, Li L, Sun F, Yue G (2011) Maping QTLs for oil traits and eQTLs for oleosin genes in Jatropha. BMC Plant Biol 11: 132

Lorieux M, Perrier X, Goffinet B, Lanaud C, Gonzalez de Leon D (1995) Maximum likelihood models for mapping genetic markers showing segregation distortion. 2. F_2-populations. Theor Appl Genet 90: 81–89

Low F C, Atan S, Jaafar H, Tan H (1996) Recent advances in the development of molecular markers for *Hevea* studies. J Nat Rubber Res 11: 32-44

Lu Z X, Reighard G L, Baird W V, Abbott A G, Rajapakse S (1996) Identification of peach rootstock cultivars by RAPD markers. Hort Sci 31: 127-129

MacIntosh G C, Wilkerson C, Green P J (2001) Identification and analysis of Arabidopsis expressed sequence tags characteristics of non-coding RNAs. Plant Physiol 127: 765-776

Mantello C C, Suzuki F I, Souza L M, Gonçalves P de S, Souza A P (2012) Microsatellite marker development for the rubber tree (*Hevea brasiliensis*): characterization and cross-amplification in wild *Hevea* species. BMC Res notes 5: 329

Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: software for genetic mapping in outcrossing species. Hereditas 144: 78–79

Markert C L, Moller F (1959) Multiple forms of enzymes: tissue, ontogenetic, and species specific patterns. Proc Natl Acad Sci USA 45: 753-763.

Marques C M, Araújo J A, Ferreira J G, Whetten R, O'Malley D M, Liu B H, Sederoff R (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. Theor Appl Genet 96: 727–737

Martin G B, Williams J G K and Tanksley S D (1991) Rapid identification of markers linked to a Pseudomonas resistance gene in tomato by using random primers and near isogenic lines. Proc Nat Acad Sci USA 88: 2336-2340

Martin A, Baldoni L (2003) A first linkage map of olive (*Olea europaea* L.) cultivars using RAPD, AFLP, RFLP and SSR markers. Theor Appl Genet 106:1273-82

Mathew N M (1992) Physical and technological properties of natural rubber. In: Sethuraj M R and Mathew N M (eds) Natural Rubber: Biology, Cultivation and Technology . Developments in Crop Science 23 Elsevier Amsterdam pp. 399-425

Mathew R, Roy B, Ravindran M, Nazeer M A, Saha T (2005) Phylogenetic relationship of *Hevea brasiliensis* species as revealed through molecular markers. Nat Rub Res 18: 14-24

Mba R E C, Stephenson P, Edwards K, Melzer S, Nkumbira J, Gullberg U (2001) Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta Crantz*) genome: towards an SSR-based molecular genetic map of cassava. Theor Appl Genet 102(1): 21-31

Mehlenbacher S A, Brown R N, Nouhra E R, Gökirmak T, Bassil N V, Kubisiak T L (2006) A genetic linkage map for hazelnut (*Corylus avellana* L.) based on RAPD and SSR markers. Genome 49: 122–133

Melchinger A E (1990) Use of molecular markers in breeding for oligogenic disease resistance. Plant Breed. 104: 1–19

Mendel G 1866. Experiments on plant hybrids (English translation). In: Stern C, Sherwood E R (eds) The origin of genetics: a Mendel source book. San Francisco: W. H. Freeman; 1-48.

Metzgar D, Bytof J, Wills C (2000) Selection against frameshift mutations limits microsatellite expansion in coding DNA.Genome Res 10: 72-80

Michaels S D, Amisino R M (1998) A robust method for the detecting single-nucleotide changes as polymorphic markers by PCR. Plant J 14:381–385

Micheletti D, Troggio M, Zharkikh A, Costa F, Malnoy M, Velasco R, Salvi S (2011) Genetic diversity of the genus *Malus* and implications for linkage mapping with SNPs. Tree Genet Genome doi:10.1007/s11295-011-0380-8

Monte-Corvo L L, Cabrita C M, Oliveira, Leitao J M (2000) Assessment of genetic relationships among *Pyrus* species and cultivars using AFLP and RAPD markers. Gen Res Crop Evol 47: 257-265

Morgan T H (1911) An attempt to analyze the constitution of the chromosomes on the basis of sex-limited inheritance in Drosophila. J Exp Zool 11: 365-413

Morgante M, Olivieri A M (1993) PCR amplified microsatellites as markers in plant genetics. Plant J 3: 175-182

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with non-repetitive DNA in plant genomes. Nature Genetics 30: 194-200

Morton B R, Bi I, McMullen M D, Gaut B S (2006) Variation in mutational dynamics across the maize genome as a function of regional and flanking base composition. Genetics 172: 569–577

Myers R M, Sheffield V C, Cox D R (1988) Detection of single base changes in DNA: ribonuclease cleavage and denaturing gradient gel electrophoresis: A Practical Approach. IRL. In: Davies K.E (ed) Genome Analysis. Oxford pp 95–139

Nakkanong K, Nualsri C, Sdoodee S (2008) Analysis of genetic diversity in early introduced clones of rubber tree (*Hevea brasiliensis*) using RAPD and microsatellite markers. Songklanakarin J Sci Technol 30: 553-560

Nartvaranant P, Nartvaranant K (2011) Analysis based on AFLP markers of the genetic variations and their relationships for pummelo cultivars grown in the central region of Thailand. Songklanakarin J Sci Technol. 33: 499-508

Neff M M, Neff J D, Chory J, Pepper A E (1998) dCAPS a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. Plant J 14: 387-392

Newcomb R D, Crowhurst R N, Gleave A P, Rikkerink E H, Allan A C, Beuning L L, Bowen J H, Gera E, Jamieson K R, Janssen B J (2006) Analysis of expressed sequence tags from apple. Plant Physiol 141: 147–166

Nurhaimi H, Woelan S, Darusamin A (1998) RAPD analysis of genetic variability in plant rubber (*Hevea brasiliensis* Muell. Arg.) clones. Menara Perkebunan 66: 9-19

Okogbenin E, Marin J, Fregene M (2006) An SSR-based molecular genetic map of cassava. Euphytica 147: 433–440

Okogbenin E, Marin J, Fregene M (2008) QTL analysis for early yield in a pseudo F_2 population of cassava. Afr J Biotechnol 7: 131-138

Oktavia F, Kuswanhadi L M (2011) Genetic relationship of Wickham and IRRDB 1981 rubber population based on RAPD markers analysis. HAYATI J Biosci 18: 27-32

Olsen M, L Hood, C Cantor and D Botstein (1989) A common language for physical mapping of the human genome. Science 245: 1434-1435.

Ong S H (1976) Chromosome morphology at the pachytene stage in *Hevea brasiliensis*: A preliminary report. Proc Internatl Rubber Conference, KualaLumbur, Malaysia, p. 1-12

Orita M, Iwahana H, Kanazawa H, Hayashi K, Sekiya T (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphism. Proc Natl Acad Sci USA 86: 2766–2770

Orozco-Castillo C, Chalmers K J, Waugh R, Powell (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor Appl Genet 87: 934-940

Orozco-Castillo C, Chalmers K J, Waugh R, Powell (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor Appl Genet 87: 934-940

Owen C A, Bita E C, Banilas G, Hajjar S E, Sellianakis V, Aksoy U, Hepaksoy S, Chamoun R, Talhook S N, Metzidakis I, Hatzopoulos P, Kalaitzis P (2005) AFLP reveals structural details of genetic diversity within cultivated olive germplasm from the Eastern Mediterranean. Theor Appl Genet 110: 1169–1176

Paetkau D (1999). Microsatellites obtained using strand extension: An enrichment protocol. BioTech 26: 690-697

Panikkar A O N, Nair V K B, Markose V C (1980). Breeding and tree improvement. In: Pillay P N R (ed) Handbook of natural rubber production in India. Rubber Research Institute of India, Kottayam, India. pp. 35-43

Paran I, Michelmore R W (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85: 985-993

Parani M, Rajesh K, Lakshmi M (2001) Species identification in seven small millet species using polymerase chain reaction-restriction fragment length polymorphism of trn S-psb C gene region. Genome 44: 495-499

Pastinen T, Raitio M, Lindroos K, Tainola P, Peltonen L, Syvaries A C (2000). A system for specific high through put genotyping by allele specific primer extension on micro arrays. Genet. Res. 10(7): 1301-1042.

Paterson A H (1996) Making genetic maps In. Paterson A H (ed) Genome mapping in plants. Sam diego, California Academic Press, Austine, Texas. pp: 23-29

Perry D J, Bousquet J (1998) Sequence-tagged-site (STS) markers of arbitrary genes: development, characterization and analysis of linkage in black spruce. Genetics 149: 1089-1098

Perry DJ, Isabela N, Bousquet J (1999) Sequence-tagged-site (STS) markers of arbitrary genes: the amount and nature of variation revealed in Norway spruce. Heredity, 83: 239-248

Perseguini J M K C, Romão L R de C, Briñez B, Junior E J S, Gonçalves P de S, Benchimol L L (2012) Genetic diversity of cultivated accessions and wild species of rubber tree using EST-SSR markers. Pesq agropec bras Brasília 47(8): 1087-1094

Petch T (1914) Notes on the history of the plantation rubber industry in the east. Ann Royal Botanic Gardens Peradeniya 5: 440 – 487

Phan J, Reue K, Peterfy M (2000) MS-IRS: A simple method for isolation of microsatellites. BioTech 28: 18-20

Picoult-Newberg L, Ideker T E, Pohl M G, Taylor S L, Donaldson M A, Nickerson D A and Boyce-Jacino M (1999) Mining SNPs from EST data bases. Genome Res 9: 167-174

Plomion C, Frigerio J M, Ridolfi M, Pot D, Pionneau C, Bodenes C, Kremer A, Hurme P, Savolainen O, Avila C, Gallardo F, Canovas F M, David H, Neutelings G, Campbell M (1999) Developing SSCP markers in two *Pinus* species. Mol Breed 5: 21–31

Pootakham W, Chanprasert J, Jomchai N, Sangsrakru D, Yoocha T, Therawattanasuk K (2011) Single nucleotide polymorphism marker development in the rubber tree, *Hevea brasiliensis* (euphorbiaceae) Amer J Bot 98: e337–e338

Powell W, Morgante M, Andre C (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol Breed 2: 225-238

Qi X, Stam P, Lindhout P (1998) Use of locus specific AFLP markers to construct a high density molecular map in barley. Theor Appl Genet 96: 376-384

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5: 94-100

Ratnaparkhe M B, Tekeoglu M and Muehlbauer F J (1998) Intersimple-sequence-repeat (ISSR) polymorphisms are useful for finding markers associated with disease resistance gene clusters. Theor Appl Genet 97: 515–519

Rebecca C J, Dorothy A, Steane B M, Potts G, Vaillancourt R E (2002) Microsatellite and morphological analysis of *Eucalyptus globulus* populations. Can J For Res 32: 59–66

Riaz S, Dangl G S, Edwards K J, Meredith C P (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theor Appl Genet 108: 864–872

Risch N, Merikangas K (1996) The future of genetic studies of complex human disease. Science 273: 1516-1517

Ritter E, Gebhardt C, Salamini F (1990) Estimation of re-combination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. Genetics 125: 645-654

Rohlf F J 1993. NTSYS-pc. Numerical taxonomical and multivariate analysis system.Exerter Software, Setauket, New York, USA

Ronaghi M, Uhlen M, Nyren P A (1996) A sequencing method based on real time pyrophosphate. Science 281: 363-365

Rossetto M, Mclauchlan A, Harriss F C L, Henry R J, Baverstock P R, Lee L S, Maguire T L, Edwards K J (1999) Abundance and polymorphism of microsatellite markers in the tea tree (Melaleuca alternifolia, Myrtaceae). Theoretical and Applied Genetics 98: 1091-98

Roy B C, Nazeer M A, Saha T (2004) Identification of simple sequence repeats in rubber (*Hevea brasiliensis*). Curr sci 87: 807-811

Roy CB, Ravindran M, Saha T (2012) One hundred microsatellite markers from Hevea genome for characterization of cultivated clones of rubber (*Hevea brasiliensis*). 24th Kerala Science Congress, 29-31, January, Rubber Research Institute of India, Kottayam

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

Saha T, Kumar A , Sreena A S, Joseph A, Jacob C K, Kothandaraman R, Nazeer M A (2002) Genetic variability of *Corynespora cassiicola* infecting *Hevea brasiliensis* isolated from the traditional rubber growing areas in India. Ind J Nat Rub Res 13: 1-10

Saha T, Roy C B, Nazeer M A (2005) Microsatellite variability and its use in the characterization of cultivated clones of *Hevea brasiliensis*. Plant breed 124: 86-92

Saha T, Roy C B, Ravindran M, Bini K, Nazeer M A (2007) Allelic diversity revealed through SSR polymorphisms at locus encoding HMG– CoA reductase in rubber (*Hevea brasiliensis*). Silvae Genet 56: 58-65

Saha T, Priyadarshan P M (2012) In: Schnell R J, Priyadarshan PM (eds) Genomics of *Hevea* rubber. Springer Newyork pp. 261-298

Salmaso M G, Faes C, Segala M, Stefanini I, Salakhutdinov E, Zyprian R, Toepfer M, Grando R, Velasco (2005) Genome diversity and gene haplotypes in the grapevine (*Vitis vinifera* L.) as revealed by single nucleotide polymorphisms. Mol Breed 14: 385–395

Saraswathyamma C K , Markose V C, Licy J and Panikkar A O N (1984) Cytomorphological studies in an induced polyploidy of *Hevea brasiliensis* Muell. Arg. Cytologia 49: 725-729

Sarkhosh A, Zamani Z, Fatahi R, Ebadi A (2006) RAPD markers reveal polymorphism among some Iranian pomegranate (*Punica granatum* L.) genotypes. Sci Hort 111: 24-29

Scalfi M, Troggio M, Piovani P, Leonardi S, Magnaschi G, Vendramin G G, Menozzi P (2005) A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). Theor Appl Genet 108: 433–441

Schlotterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20(2): 211-215

Schmidt S, Schöne N, Ritz E (1995)Association of ACE gene polymorphism and diabetic nephropathy? The Diabetic Nephropathy Study Group. Kidney Int. 1176-81

Schnell R J, Brown J S, Olano C T (2003) Evaluation of avocado germplasm using microsatellite markers. J Am Soc Hort Sci 128: 881-889

Schultes R E (1970) The history of taxonomic studies in *Hevea*. The botanical review, 36: 197-276

Schultes R E (1977) The odyssey of the cultivated rubber tree. Endeavor 1:135-137

Seguin M, Gay C, Xiong T C, Rodier-Goud M (2001) Microsatellite markers for genome analysis of rubber tree (*Hevea* spp.). Biotechnology and rubber tree: Proc IRRDB symp p. 25-28

Semagn K, Bjørnstad A, Ndjiondjop M N (2006) Principles, requirements and prospects of genetic mapping in plants. Afr J Biotechnol 5: 2569–2597

Sensi E, Vignani R, Scali M, Masi E, et al. (2003). DNA fingerprinting and genetic relatedness among cultivated varieties of Olea europaea L. estimated by AFLP analysis. Sci. Hort. 97: 379-388.

Sewell M M, Sherman B K, Neale D B (1999) A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. Genetics 151: 321–330

Shabir A, Wani M A, Bhat G N, Malik Afifa S, Kamili M R (2010) Molecular markers and their role in mulberry improvement. Intntl J Curr Res 4: 20-24

Sharma R K, Gupta P, Sharma V, Sood A, Mohapatra T, Ahuja P S (2008) Evaluation of rice and sugarcane SSR markers for phylogenetic and genetic diversity analysis in bamboo. Genome 51: 91-103

Sharon D, Cregan P B, Muhameed S. (1997) An integrated genetic linkage map of avocado. Theor Appl Genet 95: 911-921

Sheffield V C, D Y Nishimura, E M Stone (1995) Novel approaches to linkage mapping. Curr Opin Genet Dev 5:335-341

Shoucai C S, Hansuang H, Dong Qiong L Sheng , Xueqin Z (1994) Identification of mildew resistant gene from *Hevea* tree by RAPD technique. Chinese J Trop Crops 15: 26

Simmonds N W (1989) Rubber breeding In: Rubber Longman Scientific and Technical, Essex, 85-124.

Simpson E H (1949) Measurement of diversity. Nature 163: 688

Smith D N, Devey M E (1994) Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37: 977-983

Sobrino B, Brio N M. & Carracedo A (2005) SNPs in forensic genetics: A review on SNP typing methodologies. *Forensic Science International,* 154, 181

Soleimani V D, Baum B R, Johnson D A (2003) Efficient validation of single nucleotide polymorphisms in plants by allele-specific PCR, with an example from barley. Plant Mol Biol Rep 21:281–288

Soller M, Beckmann J S (1983) Genetic polymorphism in varietal identification and genetic improvement. Theor Appl Genet 82: 233-241

Sommer S S. Groszbach A R, Bottema C D K(1992) PCR amplification of specific alleles (PASA) is a general method for rapidly detecting known single-base changes.BioTech 12:82–87.

Souza L M, Mantello C C, Santos M O, Goncalves P de S, Souza A P (2009) Microsatellites from rubber tree (*Hevea brasiliensis*) for genetic diversity analysis and cross-amplification in six *Hevea* wild species. Conser Genet Resour 1:75-79

Souza M L, Mantello C C, Suzuki F, Gazaffi R, Garcia D, Le Guen V, Garcia F A A, Souza P A (2011) Development of a genetic linkage map of rubber tree (*Hevea brasiliensis*) based on microsatellite markers. BMC Proc 5: 39

Sraphet S, Boonchanawiwat A, Thanyasiriwat T, Boonseng O, Tabata S, Sasamoto S, Shirasawa K, Isobe S, Lightfoot D A, Tangphatsornruang S, Triwitayakorn K (2011) SSR and EST-SSR-based genetic linkage map of cassava (*Manihot esculenta* Crantz). Theor Appl Genet 122: 1161-70

Sriyadi B, Setiamihardja R, Baihaki A, Astika W (2002) Genetic relatedness among the F_1-tea plant from crossing of clones Tri 2024xPS 1, based on RAPD markers. Zuriat 13: 11-20

Stadler L J (1929) Chromosome number and the mutation rate in *Avena* and *Triticum* Proc Nat Acad Sci 15:876-881

Stam P (1993) Construction of integrated genetic linkage maps by means by a new computer package: JOINMAP Plant J 3: 739–744

Staub J E, Serquen F C, Gupta M (1996) Genetic markers, map construction, and their application in plant breeding. Hort Sci 31: 729–741

Suiter K A, Wendel J F, Case J S (1983). LINKAGE-1: a PASCAL computer program for the detection and analysis of genetic linkage. J Hered 74: 203-204

Talebi Bedaf M, Sharifnabi B, Bahar M (2003) Analysis of genetic diversity in pomegranate cultivars of Iran using Random Amplified Polymorphic DNA (RAPD) markers. Proc 3rd Natl Congr Biotechnol Iran 2: 343-345

Tani N, Takahashi T, Iwata H, Mukai Y, Ujino-Ihara T, Matsumoto A, Yoshimura K, Yoshimaru H, Murai M, Nagasaka K, Tsumura Y (2003) A consensus linkage map for sugi (*Cryptomeria japonica*) from two pedigrees, based on microsatellites and expressed sequence tags. Genetics 165: 1551-1568

Tanksley S D (1984) Linkage relationships and chromosomal locations of enzyme-coding genes in pepper, *Capsicum annuum*. Chromosoma 89:352–360

Tanksley S D, Young N D, Paterson A H and Bonierbale M W (1989) RFLP mapping in plant breeding: New tools for an old science. Nature Biotechnology 7: 257-264

Tautz D, Trick M, Dover G A (1986) Cryptic simplicity in DNA is a major source of genetic variation. Nature 322: 652-656

Tavassolian I, Rabiei G, Gregory D, Mnejja M, Wirthensohn M G, Hunt P W, Gibson J P, Ford C M, Sedgley M and Wu S B (2010) Construction of an almond linkage map in an Australian population Nonpareil × Lauranne *BMC Genomics*11:551

Tavaud M A, Zanetto J L, David F, Laigret, Dirlewanger E (2004) Genetic relationships between diploid and allotetraploid cherry species (*Prunus avium* x *Prunus gondouinii* and *Prunus cerasus*). Heredity 93: 631-638

Tenzer I, Ivanissevich D S, Morgante M, Gessler C (1999) Identification of microsatellite markers and their application to population genetics of *Venturia inaequalis*. Phytopathology 89: 748-753

Testolin R, Marrazzo T, Cipriani G, Verde I, Dettori M T, Pancaldi M, Sansavini S (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. Genome 43: 512-520

Thiel T, Kota R, Grosse I, Stein N, Graner A (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acid Res 32: 238-242

Thomas M R, Scott N S (1993) Microsatellite repeats in grape vine reveal DNA polymorphism when analysed as sequence tagged sites. Theor Appl Genet 86: 985-990

Toruan-Mathius N, Lalu Z, Soedarsono Aswidinnor H (2002) Genetic variation of rubber (*Hevea brasiliensis* Muell. Arg.) clones resistance and susceptible to *Corynespora cassiicola* using RAPD and AFLP markers. Menara Perkebunan 70: 35-48

Tozaki T, Inoue S, Mashima S, Ohta M, Miura N, Tomita M (2000) Sequence analysis of trinucleotide repeat microsatellites from an enrichment library of the equine genome. Genome 43:354-365

Triwitayakorn K (2011) SSR and EST-SSR-based genetic linkage map of cassava (*Manihot esculenta* Crantz). Theor Appl Genet 122: 1161-70

Triwitayakorn K, Chatkulkawin P, Kanjanawattanawong S, Sraphet S, Yoocha T, Sangsrakru D, Chanprasert J, Ngamphiw C, Jomchai N, Therawattanasuk K, Tangphatsornruang S (2011) Transcriptome Sequencing of *Hevea brasiliensis* for development of microsatellite markers and construction of a genetic linkage map. DNA Res. 18: 471–482

Troggio M G, Malacarne G, Coppola C, Segala D A, Cartwright , Pindo M, Useche F J, Gao G, Harafey M, Rafalski (2007) A: High-throughput identification,database storage and analysis of SNPs in EST utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. Theor Appl

Ujino-Ihara T, Taguchi Y, Moriguchi Y, Tsumura Y (2010) An efficient method for developing SNP markers based on EST data combined with high resolution melting (HRM) analysis. BMC Res Notes. 3: 51

Upadhyay A, Kadam U, Chacko S, Priya, Karibasappa G S (2010) Microsatellite and RAPD analysis of grape (*Vitis* spp.) accessions and identification of duplicates/ misnomers in germplasm collection. Indian J Hort 67 : 8-15

Vallejos C E (1983) Enzyme activity staining. In: Tanksley S D, Orton T J (eds) Isozymes in plants genetics and breeding, Part A. Elsevier, Amsterdam, pp: 469-516

Van de Peer Y, de Wachter R (1997) Treecon for windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comp Appl Biosci 10: 569-570

Van de Schoot J, Pospiskova M, Vosman B Smulders M J M (2000) Development and characterization of microsatellite markers in Black Poplar (*Populus nigra* L.) Theor Appl Genet 101: 317-322

Van de Ven W T, McNicol R J (1996) Microsatellites as DNA markers in Sitka spruce. Theor Appl Genet 93: 613-617

Van Der Wurff A W, Chan Y L, Van Straalen N M, Schouten J (2000) TE-AFLP: combining rapidity and robustness in DNA fingerprinting. Nucleic Acids Res 28: E105

Van Ooijen J W, Voorips R E (2001) JoinMap3.0, Software for the calculation of genetic linkage maps. Plant Res Internatl Wageningen The Netherlands

Varghese Y A, Knaak C, Sethuraj M R, Ecke W (1997) Evaluation of random amplified polymorphic DNA (RAPD) markers in *Hevea brasiliensis*. Plant Breed 116: 47-52

Varshney R K, Graner A, Sorrells M E (2005) Genic microsatellite markers in plants: features and applications. Trends Biotech 23: 48-55

Venkatachalam P, Thomas S, Priya P, Thanseem I, Gireesh T, Saraswathyamma C K, Thulaseedharan A (2002) Identification of DNA polymorphism among clones of Hevea brasiliensis Muell. Arg. using RAPD analysis. Indian J. Nat. Rubber Res. 15: 172-181

Venkatachalam P, Priya P, Saraswathyamma C K, Thulseedharan A (2004) Identification, cloning and sequence analysis of a dwarf genome specific RAPD marker in rubber tree (*Hevea brasiliensis*). Plant Cell Rep 23: 327-332

Venkatachalam P, Priya P, Gireesh T, Saraswathy-Amma C K, Thulaseedharan A (2006) Molecular cloning and sequencing of a polymorphic band from rubber tree [*Hevea brasiliensis* (Muell.)Arg.]: the nucleotide sequence revealed partial homology with proline specific permease gene sequence. Current Sci 90: 1510-1515

Venkatachalam P, Geetha N, Priya P and Thulaseedharan A (2010) Identification of a differentially expressed thymidine kinase gene related to tapping panel dryness syndrome in the rubber tree (*Hevea brasiliensis* Muell. Arg.) by random amplified polymorphic DNA screening. Interntl J Plant Biol (Germany) 1:33-39

Verhaegen D, Plomion C (1996) Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers. Genome 39:1051–1061

Vigna A, Milana D, Sancristobla M A, Eggenb A (2002). A review on SNP and other types of molecular markers and their use in animal genetics. Genet. Sel. Evol. 34: 275-305.

Violi H A, Brown J S, Tondo C L, Borrone J W, Schnell R J (2009) Microsatellite markers reveal low breeding system efficacy and pollen contamination can limit production of full-sib avocado progeny. Scientia Hort 120: 360-366

Viruel M A, Messeguer R, Vicente M C, Garcia-Mas J, Puigdomènech P, Vargas F, Arus P (1995) A linkage map with RFLP and isozyme markers for almond. Theor Appl Genet 91: 964-971

Voorrips R E (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. J Hered 93: 77-78

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-4414

Vouillamoz J, Maigre D, Meredith C P (2003) Microsatellite analysis of ancient alpine grape cultivars: pedigree reconstruction of *Vitis vinifera* L. Cornalin du Valais. Theor Appl Genet 107: 448-454

Wang D G, Fan J B, Siao C J, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J (1998) Large scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077–1082

Wang D, Karle R, Lezzoni A F (2000) QTL analysis of flower and fruit traits in sour cherry. Theor Appl Genet 100: 535-544

Wang CM, Liu P, Yi C, Gu K, Sun F, Li L, Lo L C, Liu X, Feng F, Lin G, Cao S, Hong Y, Yin Z, Yue G H (2011) A first generation microsatellite and SNP based linkage map of jatropha. PLoS ONE 6: e23632

Weber D, Helentjaris T (1989) Mapping RFLP loci in maize using B-A. translocations. Genetics 121: 583-590

Weising K, Nybom H, Wolf K, Kahl G (2005) DNA finger printing in plants: Principles, methods and applications, 2nd edn CRC Press Boca Raton Taylor and Francis Group pp. 277

Weller J I, Soller M, Brody T (1988) Linkage analysis of quantitative traits in an interspecific cross of tomato (*Lycopersicon esculentum* x *Lycopersicon pimpinellifolium*) by means of genetic markers. Genetics 118: 329-339

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

White G, Powell W (1997) Isolation and characterization of microsatellite loci in Swietenia humilis (Meliacea): An endangered tropical hardwood species. Mol Ecol 6: 851-860

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

Williamson V M, Ho J V, Wu F F, Miller N, Kaloshian I (1994) A PCR base marker lightly linked to nematode resistance gene Mi in tomato. Theor Appl Genet 87: 757-63

Witsenboer H, Michelmore R W, Vogel J (1997) Identification, genetic localization and allelic diversity of selectively amplified microsatellite polymorphic loci in lettuce and wild relatives (*Lactuca* spp). Genome 40: 923-936

Wu R L, Han Y F, Hu J J, Fang J J, Li L, Li M L, Zeng Z B (2000) An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. Theor Appl Genet 100: 1249–1256

Wu S B, Collins G, Sedgley M (2004) A molecular linkage map of olive (*Olea europaea* L.) based on RAPD, microsatellite, and SCAR markers. Genome 47: 26-35

Wu W M, Tsai H J, Pang J H S, Wang H S, Hong H S, Lee Y S (2005). Touchdown thermo cycling program enables a robust single nucleotide polymorphism typing method based on allele-specific real-time polymerase chain reaction. Anal Biochem 339:290–296

Wu S H, Wirthensohn M, Hunt P, Gibson J, and Sedgley M (2008) High resolution melting analysis of almond SNPs derived from ESTs. Theor Appl Genet 118: 1–14

Wycherly P R (1968) Introduction of Hevea to the orient. The planter 44: 127-137

Yamanaka S, Nakamura I, Watanabe K N, Sato Y I (2004). Identification of SNPs in the waxy gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice. Theor Appl Genet 108:1200–4

Yanagisawa T, Kiribuchi-Otobe C, Hirano H, Suzuki Y, Fujita M (2003) Detection of single nucleotide polymorphism (SNP) controlling the waxy character in wheat by using a derived cleaved amplified polymorphic sequence (dCAPS) marker. Theor Appl Genet 107:84–88

Yap V I, Nelson R J (1996) WINBOOT: A program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI discussion paper series No. 14, Manila, Philippines; International Rice Research Institute

Yin T M, Huang M R, Wang M X, Zhu L H, Zeng Z B, Wu R L (2001) Preliminary interspecific genetic maps of the *Populus* genome constructed from RAPD markers. Genome 44: 602–609

Yin T M, Wang X R, Andersson B, Lerceteau-Kohler E (2003) Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. Theor Appl Genet 106: 1075–1083

Yin T M, Wang X R, Andersson B, Lerceteau-Kohler E (2003) Nearly complete genetic maps of *Pinus sylvestris* L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. Theor Appl Genet 106: 1075–1083

Yuan Z, Yin Y, Qu J, Zhu L, Li Y (2007) Population genetic diversity in Chinese pomegranate (*Punica granatum* L.) cultivars revealed by fluorescent-AFLP markers. J Genet Genomics 34: 1061–1071

Zane I, Bergelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11: 1-16

Zhu Y L, Hyatt S, Quigley C, Song Q J (2001) Single nucleotide polymorphisms (SNPs) in soybean genes cDNAs and random genomic sequence. In: Plant and Animal Genome IX Conference, San Diego.

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