LINKAGE MAPPING AND QTL ANALYSIS OF DROUGHT RELATED TRAITS IN GROUNDNUT (*Arachis hypogaea L.*)

A THESIS

Submitted

in the partial fulfillment of the requirements for

the award of the degree of

DOCTOR OF PHILOSOPHY

in

FACULTY OF BIOTECHNOLOGY

By

B. GAUTAMI
[Reg. No. 0803PH0242]

RESEARCH AND DEVELOPMENT CELL
JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
KUKATPALLY, HYDERABAD – 500 085
INDIA

NOVEMBER, 2012

LINKAGE MAPPING AND QTL ANALYSIS OF DROUGHT RELATED TRAITS IN GROUNDNUT (*Arachis hypogaea L.*)

Thesis / Dissertation submitted in partial fulfillment of the requirement for the award of the degree of Ph.D.

By

B.GAUTAMI
[Reg. No: 0803PH0242]

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
HYDERABAD – 500 085, A.P., INDIA

NOVEMBER, 2012

DECLARATION

I hereby declare that the work described in this thesis, entitled "**Linkage mapping and QTL analysis of drought related traits in groundnut (*Arachis hypogaea L.*)**" which is being submitted by me in partial fulfillment for the award of Doctor of Philosophy (Ph.D.) in the Dept. of Biotechnology to the Jawaharlal Nehru Technological University Hyderabad, Kukatpally, Hyderabad (A.P.) -500 085, is the result of investigations carried out by me under the Guidance of Dr. Dave Hoisington, Deputy Director General -Research, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru and Co-supervisor Prof. M. Lakshmi Narasu, Centre for Biotechnology, Jawaharlal Nehru Technological University Hyderabad, Hyderabad.

The work is original and has not been submitted for any Degree/Diploma of this or any other university.

Place: Hyderabad

Date: Signature

Name of the Candidate: **B.GAUTAMI**

Roll No.: 0803PH0242

CERTIFICATE

This is to certify that the thesis entitled **"Linkage mapping and QTL analysis of drought related traits in groundnut (Arachis hypogaea L.)"** that is being submitted by **Ms B.GAUTAMI** in partial fulfillment for the award of Ph.D. in Faculty of Biotechnology to the **Jawaharlal Nehru Technological University, Hyderabad** is a record of bonafide work carried out by her under our guidance and supervision.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Signature of Co-Supervisor

Prof. M. Lakshmi Narasu

Centre for Biotechnology

JNTUH

Hyderabad-500085

Signature of Supervisor

Dr. Dave Hoisington

DDG-Research

ICRISAT

Hyderabad -500085

CERTIFICATE

This is to certify that the thesis entitled "**Linkage mapping and QTL analysis of drought related traits in groundnut (*Arachis hypogaea* L.)**" that is being submitted by **Ms B.GAUTAMI** in partial fulfillment for the award of Ph.D. in the faculty of Biotechnology to the **Jawaharlal Nehru Technological University, Hyderabad** is a record of bonafide work carried out by her at our organization/institution.

Signature of Head/Director
of Organization/Institution
Name and Designation
Dr. CLL GOWDA
Director of Grain Legumes Program
ICRISAT
Hyderabad- 500085

ACKNOWLEDGEMENTS

With great pleasure I express my esteemed gratitude and deep sense of indebtedness to my respectful supervisor ***Dr. Dave Hoisington****, the Deputy Director General (DDG)-Research, ICRISAT for his invaluable guidance, critical evaluation, emending suggestions and constant encouragement throughout to bring out this thesis successful. I feel it to be a privilege to work under his able guidance.*

I wish to record my profound gratitude and sincere thanks to the personification of generosity and kindness to my co-supervisor ***Prof. M. Laxmi Narasu,*** *Centre for Biotechnology, JNTUH University for her expedient advice, ever-encouraging suggestions, timely help, invaluable support, kind concern and consideration regarding my academic requirements during this tenure of research work.*

This dissertation would not have been possible without the guidance and the kind help of ***Dr. Rajeev K Varshney****, Principal Scientist-Applied Genomics, ICRISAT and Leader Theme1-Comparative and Applied Genomics (CAG), Generation Challenge Programme. I am sincerely obliged and indebted for his peer guidance, constructive comments, enthusiastic discussions, endless support, encouragement and able direction throughout my research project. I wish to express my sincere thanks for the critical evaluation and emending suggestions in the preparation of scientific papers for journals and in achieving the final form of dissertation.*

I wish to express my thanks and gratitude to ***Dr. William D. Dar****, the Director General, ICRISAT for giving me an opportunity to work in this esteemed organization.*

I wish to express my special thanks ***to Dr. CLL Gowda****, Director of Grain Legumes Program,* ***Dr. Vincent Vadez****,* ***Dr. Nigam****,* ***Dr. L Krishnamurthy****,* ***Dr. Senthivel****,* ***Dr. Trusha Shah, Mr. Abhishek Rathore, Dr. Rosana Mula,*** *the LSU- Coordinator from ICRISAT and* ***Asst. Prof. Archana Giri****, Centre for Biotechnology, JNTUH University for their valuable suggestions and help during this work.*

It is my immense pleasure to render sincere regards to all the collaborators of the project especially ***Dr. MVC Gowda*** *(University of Agricultural Sciences, Dharwad),* ***Dr. T Radhakrishnan*** *(Directorate of Groundnut Research (DGR), Junagadh) for their valuable support and co-operation to complete this invaluable piece of work.*

With immense pleasure, I express my cordial thanks ***to Mrs. Aruna****,* ***Mrs. Seetha Kannan****,* ***Mrs. B Manjula****,* ***Mr. KDV Prasad, Mr. Murali Krishna, Mr. Prasad and Mr. Damodar for*** *their administrative help. My sincere and special thanks to* ***Mr. B.J.Moss****,* ***Mr. K Eshwar Mr. G Somaraju*** *and* ***Mr. Ranjan*** *for helping me in laboratory during this work. My sincere and heart felt thanks to all my labmates in AGL, ICRISAT, for their support during my happiness and hard times. I also thank all the administrative staff and research scholars at Department of Biotechnology of JNTUH, Hyderabad for their kind help and co-operation.*

I also thank the Bioinformatics and Biometrics unit for the assistance given at the time of need. Assistance rendered by the members of Central Support Lab, library and Learning Systems Unit at ICRISAT is gratefully acknowledged.

I feel blessed to have my husband **Mr. *A Ravi Chandra*** *in my life, who was the constant inspiration for me to carry out the research and gave emotional support whenever I had difficult times. I feel happy to have my*

beautiful little daughter ***A. Kundana*** *and my husband who have traveled every step of this journey by cheering me up and stood by me through the good and hard times. I feel indebted to my parent's in-law* ***Shri A. Krishna Prasad*** *and* ***Smt. A. Usha Kiran*** *for their support and encouragement throughout the tenure.*

I avail this opportunity to thank my beloved parents who are the Almighty's most treasured gift to me. I feel scanty of words to magnitude the boundless love and tireless sacrifice, affection and encouragement showed on me by my parents, ***late Shri B. Gopala Rao and Smt. B. Rama Devi*** *and affection of my sister* ***Smt. M. Madhavi*** *that I could attain academic heights to accomplish my doctoral degree. And I express my deepest adoration to them for teaching me the etiquettes of life.*

Financial support from National Fund of Indian Council of Agricultural Research (NBFSRA), New Delhi, India and Tropical Legume I- Objective 1 project of CGIAR Generation Challenge Programme (http://www.generationcp.org), Mexico and Bill and Melinda Gates Foundation (BMGF), USA of Consultative Group on International Agricultural Research (CGIAR) are gratefully acknowledged.

I convey my whole hearted thanks to many of my well wishers and friends.

***Date**:*

***Place**: Hyderabad* *B.Gautami*

ABSTRACT

Groundnut (*Arachis hypogaea* L.), commonly called peanut, is one of the most important oilseed crops in the smallholder-farming sectors of the semi-arid tropical regions of the world where drought is the major production constraint. Until recently, the low level of molecular diversity in the cultivated groundnut genome and the scarcity of co-dominant DNA-based molecular markers were critical constraints in using modern genomics in groundnut improvement. To increase the number of molecular markers for groundnut, 23 novel simple sequence repeat (SSR or micro-satellite) markers were isolated from a SSR-enriched genomic library. These new markers, along with 3215 already available markers from different sources were tested for detecting polymorphism among parental genotypes of the two recombinant inbred line (RIL) mapping populations (ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76) to understand the genetic basis and identification of QTLs for drought related traits. As a result, two new genetic linkage maps were developed with 119 (2208 cM) and 82 (831 cM) marker loci. In addition, a consensus map consisting of 293 SSR loci located across 20 linkage groups and spanning a map distance of 2841 cM was constructed using the two new genetic maps (from the present study) and the reference map TAG 24 × ICGV 86031.

The comprehensive QTL analysis detected 153 main effect QTLs (M-QTLs) and 25 epistatic QTLs (E-QTLs) for drought tolerance related traits. Localization of these QTLs revealed 16 genomic regions that contained 125 QTLs in the consensus map. Importantly, no major single QTL for drought tolerance was detected. Therefore, novel breeding approaches like MARS (marker-assisted recurrent selection) and GWS (genome wide selection) are more likely to be required for the introgression of a larger number of QTLs in order to develop drought tolerant groundnut genotypes. As a final objective, an international reference consensus genetic map using the marker segregation data for 10 RILs and one BC population from the international groundnut community has been constructed. This consensus genetic map is composed of 897 marker loci, distributed on 20 linkage groups (a1–a10 and b1–b10) and covering a genetic distance of 3864 cM. The highest numbers of markers (70) were located on linkage group 'a1' and the least number of markers (21) on 'b9'. The marker density was lowest (6.4 cM) on 'a8' and highest (2.5 cM) on 'a1'. The reference consensus genetic map has been divided into 203 BINs, each of 20 cM. PIC (polymorphism information content) value was provided for a total of 526 markers in 190 BINs.

In summary, the newly developed genomic resources such as SSR markers and consensus genetic maps with the localized QTLs for drought

tolerance related traits will be extremely useful for groundnut genetics and breeding applications. Moreover, the international reference consensus map developed will serve as a reliable reference for aligning new genetic and physical maps, accelerate QTL mapping in a multi-population design, and serve other genetic and marker assisted breeding activities in groundnut.

CONTENTS

LISTOF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

%	:	Per cent
°C	:	degree Celsius
ABA	:	Abscisic acid
AFLPs	:	Amplified Fragment Length Polymorphism
ANOVA	:	Analysis of variance
BAC	:	Bacterial Artificial Chromosome
BC	:	Back crossing
BDT v 3.1	:	BigDye Terminator version 3.1
BES	:	BAC-end Sequences
bp	:	base pair
BSA	:	Bulk segregant analysis
CAPS	:	Cleaved amplified polymorphic sequence
cDNA	:	Complementary DNA
CIAT	:	Centro internacional de agricultura tropical
CIM	:	composite interval mapping
CIMMYT	:	Centro internacional de mejoramiento de Maizy Trigo
cM	:	centiMorgan
COS	:	Conserved Orthologous Set
CTAB	:	Cetyl Trimethyl Ammonium Bromide
DAF	:	DNA Amplification Fingerprinting
DAGs	:	directed acyclic graphs
DArT	:	Diversity Array Technology
DH	:	Doubled haploids
DNA	:	Deoxyribonucleic Acid
DW	:	Biomass dry weight
EDTA	:	Ethylenediaminetetraacetic acid
E-QTLs	:	epistatic QTLs
EST	:	Expressed Sequence Tag
GMM	:	Genotype Matrix Mapping
GMN	:	Genotype matrix network
GMs	:	Genotype matrices
GRD	:	Groundnut rosette disease
GS	:	Genomic selection
GSS	:	Genome Survey Sequences
h	:	recombination fraction
h2b.s	:	broad-sense heritability

HI	:	harvest index
HSP	:	Heat shock proteins
ICGM	:	ICRISAT Groundnut Microsatellite
ICRISAT	:	International Crops research Institute for the Semi-Arid Tropics
IPTG	:	Isopropyl β-D-1-thiogalactopyranoside
ISSR	:	Inter simple sequence repeats
ITP	:	inositol triphosphate
kbp	:	kilo base pairs
LA	:	leaf area
LB-amp	:	Luria Broth-ampicillin
LD	:	Linkage Disequilibrium
LDW	:	leaf dry weight
LEA	:	Late Embryogenesis Abundant
LG	:	linkage group
LOD	:	Logarithm of odds (base 10)
MAS	:	Marker-Assisted Selection
MISA	:	MIcroSAtellite
mM	:	milliMolar
MPC	:	magnetic particle concentrator
M-QTLs	:	Main effect QTLs
NaOAc-EDTA	:	Sodium Acetate-EDTA
NCBI	:	National Center for Biotechnology Information
ng	:	nanograms
PAGE	:	Polyacrylamide Gel Electrophoresis
PCR	:	Polymerase Chain Reaction
PIC	:	Polymorphism Information Content
PVE	:	Phenotypic variation explained
PVP	:	Polyvinylpyrolidone
QTLs	:	Quantitative Trait Loci
R2	:	Phenotypic variance
RAPDs	:	Random Amplified Polymorphic DNA
RFLPs	:	Restriction Fragments Length Polymorphisms
RIL	:	Recombinant Inbred Line
RIL	:	Recombinant Inbreed Line
ROS	:	Reactive oxygen radicals
RWC	:	Relative water content
SAT	:	Semi-arid tropics
SCMR	:	SPAD chlorophyll meter reading
ShDW	:	Shoot plus pod dry weight
SIM	:	Simple interval mapping
SLA	:	Specific leaf area

SMA	:	Single marker analysis
SNP	:	Single nucleotide polymorphism
SSRs	:	Simple sequence repeat
STMS	:	Sequence tagged microsatellites sites
STS	:	Sequence tagged sites
T	:	Transpiration
TE	:	Transpiration efficiency
TSWV	:	Tomato spotted wilt virus
VNTR	:	Variations in the number of tandem repeats
VW	:	Vegetative weight at harvest
WS	:	Water stress
WUE	:	Water-use efficiency
WW	:	Well watered
x2	:	Chi-square
X-gal	:	5-bromo-4-chloro-indolyl-β-D-galactopyranoside
µl	:	Microliter

1. INTRODUCTION

Groundnut (*Arachis hypogaea* L.), commonly known as peanut or monkey nut, is an important food and cash crop for millions of smallholder farmers in the semi-arid tropics (SAT). It is native to South America and belongs to the leguminous family *Fabaceae*. Groundnut is a self-pollinated segmental amphidiploid (2n=4x=40) (Stebbins 1957) that is believed to have originated from a single hybridization event between *A. duranensis* and *A. ipaensis* (Halward et al., 1991 and Young et al., 1996), and has a relatively large genome size of 2800 Mb/1C (Guo et al., 2009).

Groundnut is produced in both subsistence and commercial farming systems. The 'nuts' are high in edible oil content (47-53%), dietary protein (25%) and carbohydrates (20%), and are also a good source of a variety of essential vitamins and minerals. The hulms are excellent for fodder and cakes used for animal feed. The plant roots left after harvest serve as an excellent bionutrient to the soil, especially in the less developed countries where the crop is grown under low input conditions.

As a crop of global economic significance, groundnut is the sixth most important oil seed crop and fourth most important source of edible oil in the world. It is cultivated in more than 109 countries across the world on 24 million hectares with a global production of 38 million tons (FAOSTAT 2010). China, India and the USA are the leading producers. India ranks

second in groundnut production after China with an area of 5.5 million hectares and a production of 5.5 million tons in 2009 (FAOSTAT 2011). The average groundnut yield in India is low at 1007 kg/ha compared to the world average of 1522 kg/ha, and far below the average yield in China of 3356 kg/ha.

Groundnut is usually grown under low input conditions particularly in developing countries. Productivity of the crop is restricted primarily by two types of stresses: (i) abiotic and (ii) biotic. Among the abiotic stresses, drought is the most important constraint challenging global groundnut production. Though the improvement of drought tolerance is a major focus of most breeding programmes, breeding for enhanced tolerance has been difficult due to the (i) genetic complexity of the trait, (ii) high genotype by environment interactions, (iii) lack of precise phenotypic evaluation strategies at the field level, and (iv) duration and severity of drought in many locations. In the past, many efforts to improve drought tolerance have been made using conventional breeding; however, these have had limited success because of the complex nature of inheritance and/or the difficulty to measure the trait under field conditions, e.g., drought parameters such as root length, root density, variation in transpiration and water use efficiency. Thus, the improvement of key traits, especially drought tolerance, has become a key challenge for conventional breeding approaches that rely on selection for yield under

stressed environments. The difficulties of controlling the level of water stress under natural conditions and genotype by environment interactions for yield makes direct selection difficult. Therefore, approaches that combine genomics with breeding and physiology, termed genomics-assisted breeding (Varshney et al., 2005), provide strategies for improving component traits of drought tolerance that should prove more effective and efficient than the conventional selection methods.

Construction of a genetic linkage maps has become an essential step for molecular breeders in order to use various molecular breeding strategies for improving abiotic and biotic stress resistance varieties (Azhaguvel et al., 2006) and also in identification of potential regions in the genome which may be further transferred into important cultivar varieties and/or used in map based cloning of the resistance genes. Therefore, appropriate molecular markers and genetic maps integrated with molecular markers are prerequisites for MAS (marker-assisted breeding).

In groundnut, several attempts have been made to construct genetic linkage maps in diploid species using RFLPs (Restriction Fragments Length Polymorphisms), AFLPs (Amplified Fragment Length Polymorphism), RAPDs (Random Amplified Polymorphic DNA) and SSRs (Simple Sequence Repeats) (Halward et al., 1993; Milla et al., 2003; Moretzsohn et al., 2005; Garcia et al., 2005 and Gobbi et al., 2006) but

very few studies have been reported in tetraploid species using RFLPs and AFLPs (Burrow et al., 2001 and Herselman et al., 2004). However, low level of polymorphism in the tetraploid (AABB) groundnut has limited the integration of SSR markers into genetic maps. To date, the number of SSR marker loci integrated into a single cultivated groundnut genetic map has not been sufficiently high (Varshney et al., 2009a; Khedikar et al., 2010; Sarvamangala et al., 2011; Ravi et al., 2011 and Hong et al., 2010).

The paucity of DNA (Deoxyribonucleic Acid) polymorphism in cultivated groundnut may be due to the suspected single event of polyploidization. Further isolation from its wild relatives also poses a considerable obstacle to genetic mapping in groundnut. For instance, earlier studies using RFLPs, RAPDs and AFLPs markers detected limited DNA variation in *Arachis* species (Kochert et al., 1991; Halward et al., 1991; 1992; Paik-Ro et al., 1992; Gimenes et al., 2002; Bhagwat et al., 1997 and Subramaniam et al., 2000). Among different types of marker systems, the simple sequence repeat (SSR) or microsatellite markers that are co-dominant and hyper-variable markers are considered to be the potential markers of choice for application in various breeding programs (Gupta and Varshney 2000). They have detected higher levels of polymorphism in most crops compare to RFLPs, RAPDs and AFLPs (Hopkins et al., 1999; He et al., 2003; 2005; Ferguson et al., 2004 and Mace et al., 2006).

In addition, the availability of more than 4000 SSR markers in both public domain and /or accessed from various collaborators (e.g., Hopkins et al., 1999; He et al., 2003; Ferguson et al., 2004; Moretzsohn et al., 2005; Mace et al., 2007; Cuc et al., 2008; Liang et al., 2009; Bertioli et al., (unpublished) and Knapp et al., (unpublished)) provides the opportunity to integrate these markers into various genetic linkage maps of groundnut.

Genomic studies in cultivated groundnut are quite challenging because of the large genome size, narrow genetic diversity in the primary gene pool, paucity of DNA polymorphism and lack of knowledge on the genetic basis of most important traits. Therefore, developing a dense genetic map such as a "consensus map" that can be used as a reference resource for many genetic studies in different genetic backgrounds would provide the framework for transferring genetic information between different marker technologies. Such a map also allows the rapid localization of markers between various published maps and facilitates the selection of markers for high-density mapping in defined regions.

Consensus maps were developed in several crop species such as *Brassica oleracea* (Kianian and Quiros 1992), maize (Beavis and Grant 1991; Cone et al., 2002 and Falque et al., 2005), soyabean (Song et al., 2004 and Choi et al., 2007), barley (Wenzl et al., 2006; Varshney et al., 2007b and Marcel 2007), and wheat (Somers et al., 2004). However, groundnut is still lagging behind except for a recent report of a

comprehensive genetic map developed by Hong et al., 2010 with 175 loci using three mapping populations. Therefore, one objective of the present study is to construct a high-density genetic linkage map for cultivated groundnut using exclusively SSR markers.

Due to the demand to increase groundnut production under various stresses, several mapping populations have been developed using diverse parents for a combinations of traits. However, most of the studies are focused on biotic stresses such as tomato spotted wilt virus, leaf rust, late leaf spot, aphid vector of groundnut rosette disease, and nematode resistance. Only a few studies focused on abiotic stresses such as drought tolerance (Varshney et al., 2009a and Ravi et al., 2011), even though drought being a major abiotic constraint of groundnut production that weakens the plant making it more vulnerable to disease infestation and insect pests. Now-a-days, developing drought tolerant varieties is the most recommended and sought after strategy to mitigating drought stress in groundnut, and is becoming even more important due to the ever changing weather patterns. Thus, more attention has been paid to drought tolerance by groundnut breeders and physiologists over the past few years.

To assist in the efforts to employ marker-assisted selection in groundnut, there is need to increase the density of markers in the cultivated groundnut genetic maps and to identify the QTLs (Quantitative Trait Loci) for drought tolerance.

In the view of above, the present study employed two mapping populations (ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76) that segregated for drought related traits with the following seven objectives:

1. Development of novel simple sequence repeat (SSR) markers in groundnut;
2. Screening for parental polymorphisms (ICGS 76, CSMG 84-1 and ICGS 44) using SSR markers and genotyping of the respective mapping populations;
3. Construction of two genetic linkage maps using polymorphic microsatellite markers for ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 mapping populations;
4. Phenotyping of two mapping populations for drought related traits;
5. Identification of genes/QTLs associated with tolerance to drought;
6. Construction of consensus genetic map using three Recombinant Inbreed Line (RIL) mapping populations segregating for drought related traits and mapping of several main effect QTLs (M-QTLs) and epistatic QTLs (E-QTLs); and
7. Construction of an international reference consensus genetic map based on eleven mapping populations for tetraploid groundnut.

2. LITERATURE SURVEY

2.1 Groundnut

Groundnut (*Arachis hypogaea* L.) is one of the most important oilseed crops in the world. It occupies 31% of the total cropped area under oilseeds and accounts for 36% percent of total oilseed production in the world (FAOSTAT 2011). Groundnut – the 'king of oilseeds' in India – occupies an area of about 7.8 million ha with a production of 9.0 million tons. Groundnut production in the last three decades in India has increased considerably from 4.6 to 9.0 million tons. However, there has been only a marginal increase in groundnut area over the past five years (FAOSTAT 2011).

Domesticated groundnut is a segmental amphidiploid (2n=4x=40) which is believed to be originated from a single hybridization event between *A. duranensis* (A-genome) and *A. ipaensis* (B-genome), followed by a rare spontaneous duplication of chromosomes (Halward et al., 1991). In contrast, wild diploid *Arachis* species are genetically more diverse (Hilu and Staler 1995; Moretzsohn et al., 2004 and Bravo et al., 2006), providing a rich source of variation for agronomical traits and DNA polymorphisms for genetic and genomic studies (Stalker and Simpson 1995 and Dwivedi et al., 2007).

Groundnut probably originated as a geocarpic form of stylosanthinaes in the southern Bolivia/northwest Argentina region of South America (Krapovikas et al., 2000). Presently, it is grown in six continents but mainly Asia, America and Africa. China, India and USA are the top producers.

2.2 Economic Importance

Groundnut kernels contains high quality edible oil (45-55%), easily digestible protein (25-30%), carbohydrates (20%) (Encyclopaedia of Agricultural Science, 1994), and on average 40% fat, 25% protein and fairly a rich source of calcium, iron and the vitamin B complexes thiamine, riboflavin and niacin. It has multifarious usages. Groundnut oil is not only used as a major cooking medium for various food items but also utilized for manufacture of soap, cosmetics, lubricants, etc. In fact, groundnut plays a pivotal role in oilseed economy of India. It is estimated that the shell represents about 25 percent of the dry weight of unshell groundnut, and the kernel comprises 75 percent. Groundnuts are a reasonable source of dietary minerals especially potassium, phosphorus and magnesium. Groundnut oil is an excellent source of mono- and polyunsaturated fatty acids (Nwokolo 1996).

2.3 Constraints in groundnut Production

Approximately 70% of the world's groundnut production comes from semi-arid regions in which developing countries contribute about 90%. The SAT regions are mainly characterized by extremes of temperature and moisture availability especially during the peak period of crop cultivation. Despite of its economical importance, groundnut productivity, especially in SAT regions of Asia and Africa, is very low (<900 kg/ha) when compared to the world's average (1500 kg/ha) (FAOSTAT 2010). This is due to various abiotic (temperature extremes, frequent drought stress, soil factors such as alkalinity, poor soil fertility and nutrient deficiencies) and biotic (attacks by pests and diseases) constraints. Therefore, scientists have been working to improve the yield of the crop under various biotic and abiotic stresses.

2.4 Studies on Drought stress

Drought stress is one of the major environmental factors that contribute to reduced agricultural productivity and food security worldwide. Drought stress varies spatially and temporally at several different scales. Drought overall affects yield (Suther and Patel 1992) by altering membrane lipids, membrane permeability and photosynthetic responses. The ability of a plant to maintain membrane integrity under drought conditions will not only determine the plants tolerance to stress but also provide mechanisms for adaptation to water and heat stress

including stomatal conductance, osmotic adjustments and paraheliotropism.

2.4.1 Responses to drought stress

Drought stress can be defined as a persistent and abnormal moisture deficiency that causes an adverse impact on plants, and has a tremendous effect on agriculture by limiting the crop production because water limitations causes stress in plants and thereby limits the production of cultivation (Boyer 1982). During crop domestication, plants were selected on the basis of different economically important traits, where water limitation tolerance being unlikely one of them. Presently, drought stress is more severe in SAT regions due to erratically and low availability of rainfall. There were several factors of drought, which include precipitation, evaporation caused due to transpiration, temperature and humidity that occurs individually or in combination (Renu and Suresh 1998). Although selection for genotypes with increased productivity in drought environments has became a major goal and challenge of many plant breeding programs, the biological basis for drought tolerance is poorly understood because drought stress being a highly complex trait and varies with time and space and making it difficult to evaluate the reactions of genotypes to drought in a consistent manner.

Drought resistance can be categorized mainly into two groups: (i) drought avoidance and (ii) drought tolerance. Drought avoidance is a mechanism for avoiding lower water status in tissues by maintaining cell turgor and cell volume either through aggressive water uptake by an extensive root system or through reduction in water loss from transpiration and other non-stomatal pathways. While drought tolerance is a mechanism by which plant maintains metabolism even at low water potential. This trait is considered to be the most difficult one to improve through conventional plant breeding. However, in the recent year's research has been done on the identification of component traits, sources of genes and the field management practices required to approach and solve such a complex trait.

The need for new methodologies for a sustainable agriculture (Khush 1999), such as drought-tolerant plants, may provide a better practical solution to alleviate the problem of water limitation. However, most of these alternatives are based either on accelerating the selection of natural varieties and / or by transferring genes from other plant varieties to provide drought tolerance (CIAT, 2001). In order to achieve this goal, the biological base for drought tolerance needs to be clearly understood.

2.4.1.1 Physiological adaptations

In nature, almost all the terrestrial plants develop different strategies, which are genetically encoded (Monneveux and Belhassen 1996) and one among them is in accumulating water to delay or escape from the stress. Drought tolerant plants are able to overcome the stress by diminishing their metabolic functions, which are resumed once water potential is sufficient (Chandler and Bartels 1999). Other strategies to limit water loss include abscisic acid-mediated regulation of stomatal closure (Blum 1996) which causes the accumulation of gases such as carbon dioxide, that diminish photosynthesis (Bohnert and Sheveleva 1998) resulting in an energy imbalance (Levine 1999). Regarding root development, a general adaptation such as hygrotropism, in which roots detect a water gradient and redirect its growth towards it has been proposed (Lambers et al., 2000).

2.4.1.2 Biochemical responses

The most common biochemical adaptation seen in plants is osmotic adjustment, which is the result of newly synthesized metabolites (Bartels and Sunkar 2005) such as amino acids, glycine-betaine, sugars and sugar alcohols, non-toxic molecules at high concentrations. Sugars together with other macromolecules such as LEA (Late Embryogenesis

Abundant) proteins, accumulated during drought stress are likely to stabilize membranes and thereby prevent membrane fusion. Trehalose, a disaccharide, is also accumulated under drought stress and functions during embryo and flower development, as well as in the regulation of carbon metabolism and photosynthesis (Phillips et al., 2002). While glycine betaine serves as an osmoprotectant, thereby maintaining water equilibrium in plant organs (Chen and Murata 2002).

2.4.1.3 Molecular responses

Drought tolerance as a quantitative trait involves the participation of a complex set of genes and several studies have been performed on model plants as well as in drought tolerant species (Yang et al., 2004 and Montalvo-Hernandez et al., 2008). Whenever drought stress is perceived by the plant, changes in the expression pattern will be monitored ranging from genes whose products are involved in early response such as signal transduction, transcription and translation factors; to late response genes, such as water transport, osmotic balance, oxidative stress and damage repair. (Ramanjulu and Bartels 2002, and Knight and Knight 2001). Sometimes adaptive responses are also observed as a consequence of such changes, which includes early flowering and growth inhibition (Bray 2002). Details regarding mechanism are discussed below.

2.4.1.3.1 Drought Sensing and Signal Transduction

The actual sensor for drought stress is still unknown, although it is accepted that the organ with such ability is the “root”. The plant regulator abscisic acid (ABA) is the key endogenous messenger for this stress response (Raghavendra et al., 2010). However, diverse hypotheses suggest as redox imbalance and changes in cell wall-membrane integrity could trigger the response to drought (Kacperska 2004). Since drought and salinity induce high levels of ABA together with major changes in gene expression and adaptive physiological responses (Christmann et al., 2007), it is considered that ABA plays a key role in early plant response to drought. In *Arachis hypogaea*, in contrast with susceptible plants, drought tolerance is also correlated to PLD accumulation (Guo et al., 2006).

2.4.1.3.2 Induced Genes at Transcriptional Level

A significant number of drought-induced genes appear to be controlled at the transcriptional level. Bioinformatics studies have also identified several transcription factors induced under drought stress (Ashraf 2010). Transgenic plants expressing such identified

transcriptional activators have been developed for the production of drought tolerant plants (Lam and Meisel 1999).

2.4.1.3.3 Drought-Induced proteins

Translational control is another mechanism regulating plant responses to abiotic stress. Synthesized proteins have direct functions in membrane and protein protection and are involved in the acquisition of water and ions, and the transportation and homeostasis maintenance of basal cell functions. Late Embryogenesis Abundant proteins (LEA), highly accumulated in plant embryos (Galau et al., 1986), are expressed at basal levels and induced to high levels during osmotic and drought stress (Barrera-Figueroa et al., 2007). Heat shock proteins (HSP) are found to be highly accumulated during stress and are widely distributed in nature. They are involved in protein folding and assembly and induced by drought and salinity (Alamillo et al., 1995). *In vivo* evidence suggests that HSPs prevent protein thermal aggregation (Lee et al., 1995), thus facilitating the recovery of cell functions after abiotic stress. Cyclophilin, a chaperon protein is also involved in protein folding and highly induced during drought stress; overexpression of cyclophilin-encoding genes confers multiple abiotic stress tolerance (Sekhar et al., 2010).

2.4.1.3.4 Oxidative stress

One of the main effects of the dehydration in plants is the production of reactive oxygen radicals (ROS) (Bartels 2001). ROS are mainly produced in chloroplasts, where the photosynthetic activity is compromised during stress. Drought tolerance is unequivocally related to an efficient antioxidant cellular process (Montero-Tavera et al., 2008).

2.4.2 Approaches used to develop drought tolerant crops

The study of the molecular, physiological and biochemical mechanisms of the plants are mainly employed to respond to drought stress and has provided scientific knowledge for plant breeding. A number of genetically-improved drought tolerant crops have been developed using different approaches, such as (i) conventional breeding, (ii) marker-assisted breeding and (iii) genetic engineering (not discussed in the present study). For optimal success, a combination of the aforementioned techniques will likely be needed to produce new varieties showing drought tolerance in the field (Mittler and Blumwald 2010). Regardless of the approach, an interesting method to prove tolerance in the field was described by Salekdeh et al., 2009 based on yield quantification as a function of the water use and harvest index (HI).

2.4.2.1 Conventional breeding

Conventional breeding focuses on obtaining new individuals based on their genetic variation and uses phenotype-based selection to incorporate better characteristics into the progeny. In this regard, two plants possessing desirable traits are selected and then crossed to exchange their genes, so that the offspring have new genetic arrangements. Individual plants are finally tested for the expression of the desirable characteristic and are maintained in future plant generations (McCouch 2004). In case of drought tolerance, varieties displaying drought tolerance are crossed with susceptible, and resulted in developing the high yielding plants (McCouch 2004).

2.4.2.2 Molecular Breeding

Genetic improvement can be assisted by using recognizable tags (known as molecular markers) linked to target genes. These markers are based on polymorphisms that occur naturally in the DNA sequence. Different methods have been employed to detect markers such as RFLPs, RAPDs, AFLPs, and SSRs (Van Berloo et al., 2008). The genetic factors responsible for the phenotypic variations observed for a quantitative characteristic are named as QTLs (Quantitative Trait Loci). The use of molecular markers to aid in the selection of new varieties has an enormous potential to accelerate the breeding process (Ashraf 2010).

2.4.3 Study of Drought Stress on groundnut

Groundnut plants exposed to drought stress mostly lose moisture from pods that lead to the reduction in seed physiological activity, and thereby increasing the susceptibility to fungal invasion. Drought stress not only affects the food quality but also alters the nutritional quality of seed proteins. Due to lack of desirable genetic variation in groundnut, several conventional and molecular breeding techniques were adopted to improve drought and aflatoxin tolerance varieties (Holbrook et al., 2000).

Moreover during the past few decades, several advanced molecular tools have been developed and used to screen drought tolerance in various groundnut genotypes where effect of drought stress are being studied at the molecular and cellular level. These have generated enormous amount of genomic and proteomic data that help to explain the mechanism by which groundnut plants respond to drought stress. Engineering of groundnuts to withstand drought stress has also been achieved using different strategies, while few of them have succeeded in developing improved groundnut genotypes that withstand drought stress while others are in the process of developing advanced genotypes.

2.4.3.1 Responses to drought in groundnut

Drought stress has adverse affects on water relations (Babu & Rao 1983), mineral nutrition, metabolism, growth and yield of groundnut (Suther & Patel 1992). Parameters like relative water content (RWC), leaf water potential, stomatal resistance, rate of transpiration, leaf temperature and canopy temperature influence water relations in

groundnut during drought (Babu & Rao 1983). Transpiration rates generally correlate with the incident solar radiation under sufficient water availability. However, drought stressed plants transpire less than unstressed plants. Subramaniam & Maheswari 1990 reported that leaf water potential, transpiration rate and photosynthetic rate decreased progressively with increasing duration of water stress indicating that plants under mild stress were postponing tissue dehydration. Stomatal conductance also decreases during the stress period indicating that stomatal conductance is more sensitive than transpiration during the initial stress period. Under water deficit conditions, the leaves show marked diurnal variation in leaf turgor, while the pegs show less variation and maintain much higher turgor levels largely because of their lower solute potentials (Stirling et al., 1989). Marked osmotic adjustment also occurrs in growing leaves but not in mature leaves, allowing them to maintain higher turgor during periods of severe stress. Azam Ali (1984) reported that stomatal resistance of older leaves was greater than that of younger leaves and leaves become thicker under moderate drought stress conditions. Reddy and Rao (1968) reported that severe drought stress reduces leaf area by slowing leaf expansion, affecting the levels of chlorophyll *a*, *b,* and also supply of carbohydrates. Periodic water stress leads to anatomical changes such as a decrease in size of cells and intercellular spaces, thicker cell walls and greater development of epidermal tissue. Moisture stress also delays nodule formation in leguminous crops (Reddi & Reddy 1995). There is considerable evidence

that nitrogen, phosphorus and potassium uptake by groundnut is reduced by drought stress (Kulkarni et al., 1988). Leakage of solutes as a consequence of membrane damage is the most commonly observed response of groundnut tissue when exposed to drought stress. Severe water deficits causes decrease in enzymatic activity and results in breaking of complex carbohydrates and proteins into simpler sugars and amino acids (Pandey et al., 1984). Accumulation of proline is observed in the later stages of drought stress and therefore its concentration is considered to be a good indicator of moisture stress (Reddi & Reddy 1995).

2.4.3.2 Effect of drought during flowering and pod formation

The start of flowering is not delayed by drought stress (Boote & Ketring, 1990); however, the rate of flower production is affected (Gowda & Hegde 1986; Janamatti et al., 1986 and Meisner & Karnok 1991). A significant burst in flowering on alleviation of stress is a unique feature in the pattern of flowering under moisture stress, particularly when drought is imposed just prior to reproductive development (Janamatti et al., 1986). When stress is imposed during 30–45 days after sowing, the first flush of flowers produced up to 45 days do not form pegs; however, flowers produced after re-watering compensated for this loss (Gowda & Hegde 1986).

Groundnut often experiences water stress during pegging and pod formation (Jogloy et al., 1996) and results in a drastic reduction in yield. Peg longation is a turgor dependent, and is delayed due to drought stress (Boote & Ketring 1990). When adequate moisture is supplied to the root zone, it keeps the pegs alive and allows penetration and initiation of pod development (Skelton & Shear 1971). Dry pegging-zone soil delays pod and seed development; and root zone decreases the pod and seed growth rates by 30%. Peg growth during drought stress can be suspended during the period of soil water deficit and reinitiated after the drought stress is relieved (Sexton et al., 1988). It has been reported several times that under water stress, pegging and seed set responses of various groundnut cultivars vary substantially (Nageswara Rao et al., 1989).

(iii) Relationship of drought tolerance and aflatoxin contamination

Drought stress in groundnut has significant effects on phytoalexins, antifungal proteins and phenols that influence the growth of *Aspergillus* spp. and aflatoxin synthesis. Aflatoxin contamination increases with increased seed maturity. When the seed moisture content decreases, seeds lose the tendency to produce the phytoalexins resulting in *Aspergillus* spp. invasion and aflatoxin production. Enzymes such as chitinases, osmotins, peroxidases and proteases also are adversely affected during drought stress.

Drought stress and drought stress mediated-fungal infection compromise groundnut defenses and exacerbate aflatoxin formation in the seeds (Guo et al., 2006). Thus, breeding for drought tolerance has been considered to be one of the important strategies for reducing the aflatoxin content in groundnut cultivars, which would not only reduce water usage but also help in expanding groundnut production in marginal and sub-marginal soils. However, the rate of success in this effort is still slow due to the lack of genetic resources and information on the relationship and interaction between the pathways affected due to drought.

2.4.3.3 Breeding for groundnut improvement

2.4.3.3.1 Breeding towards drought tolerance

Several efforts have been made to improve groundnut cultivars that focus on yield as the only environmental method for screening of tolerance. Currently, more-integrated approaches for groundnut breeding are focused to offer success in developing stress-tolerant varieties. Understanding both physiological and molecular mechanisms of stress responses would help to develop new varieties tolerant to various stresses. Therefore, significant attempts are being made by scientists to improve the performance by selecting the varieties/cultivars that produces high and good quality pod yield even under adverse conditions. By conducting large scale trials, parameters that correlate best with drought tolerance were identified. Water transpired (T), water-use

efficiency (WUE) and harvest index (HI) are low-cost, rapid and easily measured indicators for drought-tolerance and can be used to screen large numbers of breeding populations. The application of this physiological model in groundnut breeding has not been possible because of many practical problems associated with measurement of the traits under field conditions.

A new drought tolerant groundnut variety, ICGV 91114, has become very popular in Anantapur district in Andhra Pradesh, India, replacing TMV 2 (a seven decade old variety). Moreover, the crosses GG-2 x NCAC 17135, GG-2 x PI 259747, J 11 x PI 259747, S 206 x FESR-8, Kisan x FESR-S-PI-B1-B, and the genotypes JB 223 and JB 224 were also termed as drought tolerant genotypes. Therefore the lines/genotypes that could be grown under regions of limited rainfall may be also used as parents in breeding programs for developing drought tolerant groundnut cultivars.

2.4.3.3.2 Limitations to traditional breeding

Crop improvement in terms of production and development of desirable traits and resistance to drought stress has become a pre-requisite in modern day agriculture. However, conventional breeding for developing drought tolerant varieties is labor intensive and time-consuming process because of the quantitative and complex nature of drought tolerance and difficulties in selection for drought tolerance (Ribaut et al., 1997). Combining high levels of resistance varieties into

higher yielding cultivars with acceptable and /or desirable traits that are market preferred is considered to be difficult (Holbrook & Stalker 2003). In addition, several breeding programs that have focused on incorporating resistance genes from wild *Arachis* relatives have been largely unsuccessful due to (i) genetic incompatibility, and (ii) limited gene pool or the restricted range of organisms between which genes can be transferred. Therefore, in addition to traditional conventional methods, new *omics* techniques are to be undertaken to develop new groundnut cultivars/varities with high tolerance to drought.

2.4.3.4 Applications of molecular breeding tools for groundnut improvement

2.4.3.4.1 Genomic approach

Groundnut, a segmental allotetraploid (2n=4x=40) (Stebbins 1957) and has a relatively large genome size of 2800 Mb/1C (Guo et al., 2009). Complete sequencing of the whole genome will be expensive and labor intensive. Research in molecular aspects began in groundnut only in the early 1980's when protein and isozyme variation in *A. hypogaea* was determined to be of less use in characterizing variation within the cultivated genotypes. Over the past five years, a large number of molecular markers have been detected (Stalker et al., 1994), but still the number of is too small to be routinely used in breeding programs.

2.4.3.4.2 Gene expression during drought stress in groundnut

Abiotic stress has become a major growing constrain for groundnut cultivation. Major production areas are in SAT environments that have unreliable rainfall, and global climate changes. Physiological adaptation and selection for drought tolerance have been studied by many researchers (Reddy et al., 2003). Groundnut genomics has been limited by many biological constraints, and many basic tools of genomics have yet to be developed (Gepts et al., 2005). Since, the groundnut genome is large, insertional mutagenesis and sequencing of the whole genome will be expensive, and requires large genomic libraries for physical mapping and positional cloning. To date, 136,901 groundnut sequences, including 87,688 ESTs from cultivated groundnuts and 39,866 nucleotide sequences have been deposited in the NCBI (National Center for Biotechnology Information) EST (Expressed Sequence Tag) database. Out of which only 52 nucleotide sequences and 25,914 EST sequences were available in response to drought treatments. One of the major molecular responses that plants exhibit to drought stress is altered expression of genes, related to different pathways associated with stress perception, signal transduction, regulators and synthesis of a number of compounds (Ramanjulu & Bartels 2002). Differential display reverse transcriptase PCR (Polymerase chain reaction) was used to identify genes induced and suppressed in groundnut seed during drought. A total of 1235 differential display products were observed in irrigated samples, compared to 950 differential display products in stressed leaf samples (Jain et al., 2001). These studies demonstrated

qualitative and quantitative differences in the gene expression during drought stress in groundnuts.

Drought is a complex process and there are certain genes that are expressed at elevated levels whenever plants encountered stress. It is also important to note that tolerance to drought is unlikely to be under the control of a single gene. Therefore, it will be important to employ a combination of conventional screening efforts, marker assisted selection and genetic engineering inorder to switch on a transcription factor regulating the expression of several genes related to drought tolerance.

Although significant progresses have been made to understand the genetic mechanisms that underlie drought tolerance in groundnut, considerable challenges still remain unsolved. Under field conditions, plants are subjected to variable levels of multiple stresses, and hence, the response to a combination of stresses deserves much more attention. Apart from that, the response of plants to multiple stresses cannot be inferred from the response to individual stress. Therefore, it is very important to test newly developed varieties to multiple stresses, and to perform extensive field studies under diverse environments inorder to assess their tolerance.

2.4.3.5 Drought related traits in groundnut

Drought tolerance is likely to be conditioned by many genes under different and high environmental influence and thus the networks

involved in drought tolerance are highly complex in nature. Therefore, selection based on the phenotype will be difficult for such traits (Collins et al., 2008). Water-use efficiency (WUE) is considered to be an important drought avoidance trait that deals with utilization of soil water more efficiently for biomass production (Blum 2005 and Collins et al., 2008). Raising the WUE of both irrigated and rainfed crop production has become an urgent imperative (Nigam et al., 2005). Surrogate traits for TE (transpiration efficiency) such as specific leaf area (SLA) and SPAD chlorophyll meter reading (SCMR) have also been used as proxies for TE (Hubick et al., 1986, Nageswara Rao and Wright 1994); however, some recent studies have raised concern about the use of these surrogates (Krishnamurthy et al., 2007). Nevertheless, SLA can be used as an indicator of leaf thickening processes, which indirectly effect or condition the rate at which a plant uses water, and is thus an important component in understanding drought adaptation (Kholova et al., 2010a and 2010b). SPAD reading can also be used as a proxy for the nitrogen status.

Developing drought tolerant varieties through conventional breeding is considered to be time-consuming, costly and labor intensive due to the quantitative nature of drought tolerance, and the difficulties in selection for drought tolerance traits (Ribaut et al., 1997). Recent advances in crop genomics offer tools to assist breeding through identification and introgression of genomic regions associated with drought tolerance to

develop improved cultivars/ varities with increased drought tolerance using marker-assisted selection (MAS) (Ribaut et al., 1996 and Varshney et al., 2006). Therefore, inorder to identify the genomic regions that are suitable for marker-assisted breeding strategies, it is important to establish accurate phenotyping methods coupled with development of saturated genetic linkage maps and identification of QTLs (quantitative trait loci) for traits of interest.

Several studies in many other crops have reported the identification of QTLs for drought tolerance or related traits. However, in groundnut, QTL studies for drought tolerance traits have been conducted only on one mapping population (TAG 24 × ICGV 86031). Comprehensive QTL analysis resulted in the identification of a total of 117 small M-QTLs and 23 E-QTLs for drought related traits (Ravi et al., 2011). However, from the above study it was inferred that QTLs identified are not suitable for their deployment in marker-assisted selection strategies. Therefore, to confirm the involvement of small effect QTLs from the study by Ravi et al., 2011, it is essential to undertake a similar drought tolerance QTL analysis using other mapping populations. Such a QTL analysis may also yield new QTLs that were not identified based on earlier studied mapping population.

2.5 Groundnut Genomics

Genetics, the study of genes through their variation, has made a major contribution to agriculture. In spite of progress made through genetic enhancement, additional gains in agricultural productivity are in great demand to cope with the continued population growth. The science of molecular biology in recent years has provided tools suitable for rapid analysis of different organisms using DNA markers. The most wide spread application of molecular markers is in the construction of the genetic linkage maps to determine the chromosomal location of genes affecting both qualitative and quantitatively inherited traits. By knowing the map position of a gene, one can use nearby or flanking molecular markers to diagnose the presence of the gene without having to determine the effect of the gene.

Marker-assisted selection offers great scope for improving the efficiency of conventional plant breeding. Molecular markers are especially advantageous for traits with low heritability where traditional selection is difficult, expensive and inaccurate (Crouch 2001). The essential requirements for developing MAS breeding programs include (i) availability of diverse germplasm with useful characteristics, (ii) identification of flanking markers closely linked on either side of the gene/ QTL, (iii) simple robust marker detection technology to facilitate rapid and cost effective screening of large breeding populations, and (iv) highly accurate and precise screening techniques for phenotyping of mapping populations. The molecular markers offer certain advantages over morphological markers as they are phenotypically neutral, occur

throughout the genome, and neither influenced by environments nor by pleiotropic and epistatic interactions.

2.5.1 Molecular markers

A molecular marker is a gene or fragment of DNA that is associated with a known location on a chromosome and may or may not be associated with a trait. Molecular markers offers a powerful tools for the (i) construction of genetic and physical maps, (ii) marker-trait association studies, (iii) marker-assisted selection programmes (MAS), (iv) gene pyramiding, (v) positional cloning of useful genes, (vi) genetic diversity analysis, (vii) DNA profiling and (viii) tagging of genes (Gupta and Rustgi 2004). During the last three decades, a number of molecular marker technologies have been utilized to visualize DNA polymorphisms in plant species (Gupta et al., 2002). Depending on the method of detection of the sequence variation, the molecular markers have been categorized into two classes (i) hybridization based (non-PCR based) molecular markers and (ii) PCR dependent molecular markers including micro-array based molecular markers (Gupta et al., 2002). Hybridization based molecular markers include RFLPs, while PCR-dependent molecular markers include RAPDs, AFLPs, SSRs, and sequence tagged sites (STS) and cleaved amplified polymorphic sequence (CAPS) (Gupta et al., 2002). The micro-array based molecular markers comprise of single nucleotide

polymorphism (SNP) and diversity array technology (DArT) (Gupta et al., 2008).

2.5.2 Molecular markers studies in groundnut

Cultivated groundnut has been analyzed by several markers systems including RFLPs, RAPDs, AFLPs and SSRs; however only 5% of the markers analyzed detect polymorphisms among diverse genotypes, and this is much lower between pairs of *A. hypogaea* lines.

2.5.2.1 Restriction Fragment Length Polymorphism (RFLPs)

RFLPs are first generation molecular marker systems that detected large number of polymorphisms in plant at the sequence level. They are robust, reliable and transferable markers across the mapping populations but at the same time are time consuming, laborious, expensive and require a large amount of genomic DNA. RFLPs are produced by digesting genomic DNA with restriction endonucleases that recognize a specific DNA sequence and then cleave the DNA strand in near recognition sites of the sequence. The fragments produced can be separated by size using gel electrophoresis. Due to large genome size, plants often produce so many fragments that the resulting gel is not interpretable. For such complex genomes, a labeled (radioactive or non-radioactive) probe is designed from cloned DNA homologous to a specific DNA sequence in the species being investigated (Botstein 1980).

Hybridized DNA fragments to the probe are finally visualized by detection of the specific label.

In *Arachis*, Kochert et al., (1991) observed a very low level of RFLP variability among the allotetraploids (U.S cultivars) and *A. monticola*, which is a wild species. RFLPs also revealed very low levels of variability in unadapted germplasm lines though considerable morphological and physiological variability existed among the lines (Halward et al., 1991). Paik-Ro et al., (1992) assessed RFLPs among accessions within six groundnut species of the *Arachis* section and observed significant amount of variation present among the *Arachis* species. *Arachis monticola* was found to be more closely related to *A. hypogaea* subspecies *hypogaea* than to subspecies *fastigiata.* Kochert et al., (1991) observed no variation between *A. hypogaea* and *A. monticola.* RFLPs have also been used to analyze the species in the section *Arachis* and the determined clusters (Kochert et al., 1991) corresponded closely with morphological groups (Stalker 1990). Stalker et al., (1995) used RFLPs to study genetic diversity among eighteen accessions of *A. duranensis* and observed a large amount of variation in the species. RFLP analysis also revealed that the cultivated groundnut resulted from the cross between *A. duranensis* X *A. ipaensis,* and chloroplast analysis indicated that *A. duranensis* was the female progenitor (Kochert et al., (1991).Gimenes et al., (2002) used RFLPs to study genomic relationship between AA genome, BB genomes and AABB genome species. The lowest genetic

variation was detected within accessions of *A. duranensis* (17 accessions), followed by *A. batizocoi* (4 accessions) and *A. cardenasii* (9 plants of accession GKP 10017).

2.5.2.2 Random Amplified Polymorphic DNAs (RAPDs)

The RAPDs assay is using a single arbitrary nucleotide sequence primer. The assay was first developed used to detect nucleotide sequence polymorphisms in DNA by Williams et al., in 1990. RAPDs are quick, simple and inexpensive, can detect multiple loci using a single primer and require a small amount of DNA to carry out but the assay. However, RAPDs are generally not as popular as other markers due to problems like poor reproducibility and transferability, fuzzy products, and difficulty in scoring of bands that lead to inappropriate inferences. Halward et al., (1992) used RAPDs to study *Arachis* species variability and reported very little variation, concluding that the dominant behavior of the markers prevented the differentiation of heterozygotes from homozygotes. However, Lanham et al., (1992) was able to detected nearly 82% variation between *A. hypogaea* and synthetic amphidiploids using RAPDs. Hilu and Stalker (1995) observed maximum variation among accessions of *A. cardenasii* and *A. glandulifera,* whereas in the case of *A. hypogaea* and *A. monticola* less variation was observed using RAPDs. Based on this study, *A. duranensis* was most closely related to the domesticated groundnut and was believed to be the donor of the A genome. Bhagwat et al., (1997) observed 6% polymorphism and were able to detect variation among the

different plant height and pod size mutants using RAPDs. Through single RAPD primers, a high degree of polymorphism among 14 closely related groundnut genotypes was reported by Bhagwat et al., (2001). Subramanian et al., (2000) studied RAPD differences among 70 selected cultivated groundnut genotypes that represent variability for several morphological, physiological and other characters with 48 oligonucleotide primers. Of these, only seven primers (15%) yielded polymorphic amplification products. Dwivedi et al., (2001) assessed genetic diversity among 26 selected groundnut accessions using eight 10-mer primers and reported that the pair-wise genetic similarity (S_{ij}) ranged from 59 to 99%, with an average of 86%, and identified five accessions with diverse profiles for mapping and genetic enhancement studies.

2.5.2.3 Amplified Fragment Length Polymorphism (AFLPs)

AFLPs are mainly based on the selective PCR amplification of restriction fragments from a total digest of genomic DNA. It involves three steps: (i) restriction of the DNA and ligation of oligonucleotide adapters, (ii) selective amplification of sets of restriction fragments, and (iii) gel analysis of the amplified fragments. PCR amplification of restriction fragments is achieved by using the adapter and restriction site sequence as target sites for primer annealing. The selective amplification is achieved by the use of primers that extend into the restriction fragments, amplifying only those fragments in which the primer extensions match the nucleotides flanking the restriction sites (Vos et al., 1995). Using this

assay, even multiple loci can be detected. With AFLPs, it is possible to detect high levels of polymorphism but the major disadvantages are the requirement of large amount DNA and complicated methodology. He and Prakash (1997) used DNA Amplification Fingerprinting (DAF) and AFLP techniques to detect genetic variation among the groundnut cultivars and found that AFLPs were more efficient since 43% of AFLP primers combinations could detect polymorphism in contrast to 3% of DAF primers. He and Prakash (2001) concluded that AFLP approach can detect considerable amount of DNA variation in the cultivated groundnut germplasm. They conducted evolutionary studies demonstrating that the botanical varieties *aequatoriana* and *peruviana* were closer to subspecies *hypogaea* than subspecies *fastigiata Waldr*. to which they belong, and that the wild *A. monticola* was not distinct from the cultivated *A. hypogaea*. Gimenes et al., (2002) used AFLPs to study the genetic relationship among 20 species from seven of the nine sections of genus *Arachis*. The level of polymorphism was evaluated among nine accessions of the cultivated groundnut, *A. hypogaea*. Moreover, this study revealed the genetic relationship assessed using AFLPs agreed with the classification established using morphological and crossability data. The results indicated that AFLPs are good markers, can be used for studying the genetic relationship among *Arachis* species and detect higher levels of polymorphism than RAPDs and RFLPs. Milla et al., (2005) used the AFLP technique to determine intra- and inter-specific relationships among and within 108 accessions of 26 species of *Arachis* section. They determined

that the A-genome accessions KG 30029 (*Arachis helodes*) and KSSc 36009 (*Arachis simpsonii*) and B-genome accession KGBSPSc 30076 (*A. ipaensis*) were most closely related to both *A. hypogaea* and *A. monticola* suggesting their involvement in the evolution of the tetraploid groundnut species.

2.5.2.4 Simple Sequence Repeats (SSRs)

Simple sequence repeats (SSRs), also known as microsatellites, are often chosen as the preferred markers for a variety of applications in plant breeding programmes because of their multi-allelic nature, co-dominance, reproducibility, requirement for small amounts of DNA and extensive genome coverage (Gupta and Varshney 2000). Since they are PCR based markers, SSRs contain short, tandemly repeated DNA sequence motifs that consist of two to six nucleotide core units (Litt and Lutty 1989). SSR detection is technically simple, robust, reliable and transferable between populations. A large amount of time and labour are required to generate primers and polyacrylamide gels are usually required to resolve the fragments. Polymorphisms are detected as variations in the number of tandem repeats (VNTR loci) in a given repeat motif. The high incidence of detectable polymorphism through changes in repeat numbers is caused by an intramolecular mutation mechanism called DNA slippage (Gupta et al., 1996). The regions flanking the microsatellites are generally conserved and PCR primers specific to the flanking regions are used to amplify SSR containing DNA fragments.

Powell et al., (1996) reported that SSR markers show higher level of polymorphism than RFLPs, RAPDs and AFLPs, and have been widely adopted for genetic analysis in plants (Rongwen et al., 1995). Thus, SSRs are considered important markers to facilitate routine diversity analysis and molecular breeding applications (Dwivedi et al., 2003). In 2000 Gupta and Varshney reported that microsatellites are more variable than RFLPs and RAPDs, and have been widely utilized in plant genomic studies.

Groundnut is thought to have evolved relatively recently through a single hybridization event, most likely between the unreduced gametes of two diploid species representing the A and B genomes (Kochert et al., 1991). It is postulated that the resultant amphidiploid plant was then reproductively isolated from diploid wild relatives leading to a very narrow genetic base. Genetic maps have been reported for the genomes of both diploid (Halward et al., 1993) and amphidiploid (Burow et al., 2001) *Arachis*. The number of microsatellite markers published for groundnut has increased considerably in the last 10 years (Hopkins et al., 1999; He et al., 2003; Palmieri et al., 2002; 2005; Fergusson et al., 2004; Moretzsohn et al., 2004; 2005; Nelson et al., 2006; Mace et al., 2007; Proite et al., 2007; Gimenes et al., 2007; Cuc et al., 2008 and Knapp et al., unpublished)), but these are still not sufficient for the construction of saturated linkage maps. The first SSRs to be developed in groundnut detected disappointing levels of polymorphism in cultivated

germplasm (Hopkins et al., 1999), and the first genetic linkage map in cultivated groundnut based on SSR markers (Varshney et al., 2009a) that is used as a reference map as well as the A-genome (Moretzsohn et al., 2005) and B-genome (Moretzsohn et al., 2009) maps were based primarily on SSR markers. Hopkins et al., (1999) isolated 26 microsatellites from a groundnut genomic DNA library and observed 23% polymorphism across a collection of 22 groundnut genotypes representing both cultivated and wild species. Raina et al., (2001) used twenty-one RAPD and 29 ISSR (Inter simple sequence repeats) primers to assess genetic variation and interrelationships among subspecies and botanical varieties of cultivated groundnut and phylogenetic relationships among cultivated groundnut and wild species of the genus *Arachis*. Both random and ISSR primers revealed 48% and 54% polymorphism, respectively. This study strongly supported the view that *Arachis monticola* ($2n = 4x = 40$) and *A. hypogaea* are very closely related, and indicated that *A. villosa* and *A. ipaensis* are the diploid wild progenitors of the tetraploid species. He et al., (2003) isolated 56 different microsatellites by using a SSR enrichment procedure and observed 34% polymorphism among the genotypes suggesting a higher level of DNA polymorphism by these SSRs than other DNA markers in cultivated groundnut. Moretzsohn et al., (2004) screened 67 TTG SSR markers to study polymorphism of seven accessions and observed only 4% polymorphism in cultivated groundnut. Ferguson et al., (2004) generated 110 sequence tagged microsatellites sites (STMS) markers for cultivated

groundnut and in their study, 81% of $(ATT)_n$ and 71% of $(GA)_n$ detected polymorphism in groundnut. Krishna et al., (2004) has shown molecular diversity using microsatellite markers in the cultivated Valencia groundnut (subsp. *fastigiata*) and results indicated that considerable genetic variations was present among the analyzed genotypes. He et al., (2005) have developed 130 simple sequence repeat (SSR) markers in groundnut and observed 29% polymorphism among 24 groundnut accessions. Eight SSR markers were found useful to classify botanical varieties. Mace et al., (2006) screened 23 SSR markers across 22 groundnut genotypes with varying levels of resistance to rust and late leaf spot and detected 52% polymorphism with a PIC (Polymorphism Information Content) value ≥0.5. Bravo et al., (2006) evaluated the transferability of microsatellite primers and the assay of genetic variability between and within the germplasm of some species of the *Arachis* section and reported that 78% were polymorphic. All loci had transferability to all the species analyzed. Upadhyaya et al., (2007) studied genetic diversity in composite collection containing 916 accessions with 21 SSR markers and revealed considerable variation among the accessions (0.819 PIC value; 490 alleles) A total of 101, 50, 11 group-specific unique alleles in wild *Arachis*, *A. fastigiata* and *A. hypogaea*, respectively were identified. Clustering of different genotypes into *fastigiata*, *hypogaea* and wild species was observed and based on common origin, some of the accessions from *fastigiata* grouped with *hypogaea*. Kottapalli et al., (2007) used 73 microsatellite markers to

genotype 72 accessions from the USA groundnut minicore. Moderate levels genetic found and the genetic distance values (D) ranged from 0.88 to 0.25. Nimmakayala et al., (2007) used 96 SSR primers to screen 30 species representing A, B and D genomes of *Arachis* with various ploidy levels (18 diploid, 9 tetraploid and one aneuploid) along with two cultivated groundnut varieties. Of these, 50 (52%) were found to be polymorphic. Tang et al., (2007) assessed the genetic variation from the four sets of 24 accessions each from the four botanical varieties of the cultivated groundnut using 34 microsatellites. Among these accessions, 10 to 16 pairs of microsatellites primers detected polymorphisms. Barkley et al., (2007) studied diversity and phylogenetic relationships among groundnut species using 31 microsatellites with attached M13 tails, which consists of all but one of the 112 accession from the minicore. A total of 477 alleles were detected in this data set with an average of 15.4 alleles per locus. The mean PIC score was 0.687. Gimenes et al., (2007) isolated thirteen microsatellite loci and characterized 16 accessions of *A. hypogaea*. The level of variation detected in *A. hypogaea* using microsatellites was higher than with other markers. Cross transferability of the markers was also high and the same repeated sequence was found in almost all the wild species as in *A. hypogaea* after sequencing of amplified fragments.

Therefore, the studied markers systems in groundnut revealed very low level of molecular polymorphism compared to other crop species (Stalker and Mozingo 2001).

2.5.3 Genetic mapping

Genetic mapping is a method to locate molecular markers, gene loci and QTLs in order, thereby indicating the relative distances among them, and assign them to linkage groups on the basis of their recombination values from all pair wise combinations. Genetic mapping mainly requires two components (i) detectable polymorphic alleles and (ii) recombination or segregation of those alleles. Genetic linkage maps are considered to be a 'route map' of the chromosomes derived from two different parents (Paterson 1996). They serve as structural frameworks to identify chromosomal locations containing genes and QTLs associated with traits of interest (QTL map). 'QTL mapping' is mainly based on the principle that genes and/or markers segregate via chromosome recombination during meiosis and thus allowing their analysis in the progeny (Paterson 1996). If two genes or marker loci are located close to each other on the same chromosome, they will tend to be inherited together and these two loci are said to be 'linked' while markers that have a recombination frequency of 50% are considered to be 'unlinked' and assumed to be located far apart on the same chromosome or on different chromosomes (Hartl and Jones 2001).

Genetic linkage maps are mainly constructed from the analysis of many segregating markers. For linkage map construction, three main steps are required: (1) production of a mapping population, (2) identification of polymorphism between parental genotypes for molecular markers, and (3) linkage analysis of markers. Linkage between markers is usually calculated using an odds ratio (i.e. the ratio of linkage versus no linkage). This ratio is expressed as the logarithm of the ratio, called a logarithm of odds (LOD) value or LOD score (Risch 1992). For constructing linkage maps, LOD *values greater than* 3 are typically used. If the LOD value is 3 between any two markers it indicates that linkage is 1000 times more likely (i.e. 1000:1) than no linkage (null hypothesis). Accepted LOD value threshold may be lowered in order to detect a greater level of linkage or to place additional markers within maps constructed at higher LOD values. The most commonly used mapping software programs are (i) Mapmaker/ EXP (Lander et al., 1987 and Lincoln et al., 1993), (ii) MapManager QTX (Manly et al., 2001), (iii) GMendel (http://cropandsoil.oregonstate.edu/Gmendel), (iv) MSTMap (Wu et al., 2008) and (iv) JoinMap (Stam 1993).

2.5.3.1 Status of genetic mapping in groundnut research

The ability to rapidly construct genetic maps has made possible applications that were unthinkable using conventional mapping techniques. Construction of genetic linkage map is a prerequisite for modern plant breeding programmes. The ease with which a genetic map can be developed and applied to a target crop species depends on the

genetic complexity of the species and the extent of DNA polymorphism present in the species. Genetic mapping in general monomorphic species like groundnut has usually been achieved by using wide crosses between highly divergent parental genotypes, sometimes even using different species (Paterson et al., 1996). The low frequency of DNA polymorphism within a species can also limit the utilization of mapped DNA markers in cross that are of agronomic importance, but involve more genetically monomorphic parents. For these reasons, it is important to establish the frequency of DNA polymorphism within a species before engaging in a plant improvement programme using molecular markers.

In groundnut, it became a challenging task because of its low level of genetic polymorphism due to single event of polyploidization, but with recent explosion of many robust molecular markers methods, significant amount of polymorphism is also observed in this crop.

Halward et al., (1993) constructed the first genetic map in groundnut using a cross between two diploid species A. *stenosperma* and A. *cardenasii*. RFLP markers were used from genomic as well as cDNA libraries of A. *hypogaea* cvGK7. The partial genomic library was constructed using *Pst*I digestion of genomic DNA and cloning of 1-2 Kb fragments. The cDNA libraries were analyzed using seven different restriction enzymes (*Bam*HI, *Dra*I, *Eco*RI, *Hae*III, *Hind*III and *Rsa*I). Out of the 100 genomic and 300 cDNA probes used, 15 and 190 gave

polymorphic profile between the parents respectively. Of the 205 probes that detected polymorphism, 132 were analyzed for segregation with the rest revealing complex patterns and not mapped. Of the 132, 117 segregating loci were distributed on 11 linkage groups. A total map distance of 1400 cM was covered with a 20 cM resolution. This map was able to cover 80% of the groundnut genome (Table 1). Garcia et al., (1995) constructed a linkage map using the combination of one tetraploid parent and one diploid species A. *cardenasii*. In this study, 73 RFLP probes and 70 RAPD markers were used to screen against 46 introgression of A. *cardenasii* Krapovickas and W.C.Gregory (2n=2x=20) for the introgression of *A. cardenasii* chromosome segments. A total of 34 cDNA RFLP probes and 45 RAPD primers identified introgressed chromosomal segments in one or more lines. The introgression segments covered 10 out of the 11 linkage groups, smallest of which were RFLP markers and the largest had 3-4 adjacent markers at a distance of 30-40 cM. Garcia et al., (2005) also used a backcross population *A. stenosperma* x (*A. stenosperma* x *A. cardenasii*) and 39 shared RFLPs and placed 167 RAPD loci onto the RFLP map. The RAPDs were able to cover a total genetic length of 800 cM and mapped onto 11 linkage groups. Herselman et al., (2004) used 60 $F_{2:3}$ lines derived from two *A. hypogaea* (ICGI 2991 and ICGV-SM 93541) genotypes. A total of 308 AFLP primers and 144 primers combinations were used to identify markers associated with aphid resistance and identified 20 putative markers. Of which, 12 are mapped on 5 linkage groups covering a map distance of 139.4 cM.

This study is the first report on the identification of molecular markers linked to aphid resistance to groundnut rosette disease (GRD) and the first partial genetic linkage map of the cultivated groundnut. Burrow et al., (2001) constructed the first molecular map representing the entire tetraploid genome of the groundnut. To introduce variability from diploid wild species into tetraploid cultivated *Arachis hypogaea*, a synthetic amphidiploid TxAG-6 (*A. batizocoi* K 9484 X (*A. cardenasii* GKP 10017 X *A. diogia* GKP 10602)^{4x}) was used as a donor parent to generate backcross population of 78 progenies. A total of 370 RFLP loci were mapped onto 23 linkage groups using the BC1 mapping population. A total of 917 bands were observed, for an average of 4.1 bands per probe. A mean of 1.7 loci per probe were mapped. The total length of the tetraploid map was 2210 cM, which was slightly greater than twice the length (1063 cM) of the diploid map (Garcia et al., 1995). The tetraploid map developed based on an inter-specific cross is useful in locating specific genes of interest in the inter-specific cross and also provides valuable information about genome organization and genome evolution. Milla (2003) constructed a genetic linkage map for an F2 population of *A. kuhlmannii* x *A. diogoi*. This map consisted of 102 AFLP markers mapped on to 12 linkage groups and spanning a map distance of 1068.1 cM. As a first step towards the introgression of resistance genes into cultivated groundnut, Moretzsohn et al., (2005) constructed a linkage map based on microsatellite markers using a F2 population obtained from a cross between two diploid wild species with AA genome (*A. duranensis* and *A.*

stenosperma). A total of 271 new microsatellite markers were developed from SSR enriched genomic libraries, EST and data minning, sequences available in Genbank and another 162 published groundnut microsatellites markers screened against both progenitors. Two hundred and four of these (47%) were polymorphic and were screened across 93 F2s. The resulting linkage map consists of 11 linkage groups covering 1,231 cM total map distance, with an average distance of 7.2 cM between markers. This is the first microsatellite based map published for *Arachis* and the first map based on sequences that are publicly available. Gobbi et al., (2006) constructed a 'B genome' map. A total of 93 F2s derived from a cross between *A. ipaensis* (KG30076) and *A. magna* (KG30097), both diploid species with B genome were used in the study. A total of 94 polymorphic markers were mapped spanning 11 linkage groups and with a total distance of 754.8 cM. The size of each linkage group ranged from 5.6 to 130.7 cM. Leal-Bertioli et al., (2009) developed a genetic map using 93 F_2 plants derived from a cross between two diploid wild A-genome *Arachis* species, *A. duranensis* × *A. stenosperma*. A total of 369 markers were mapped into 10 linkage groups spanning a total distance of 2532 cM. These 369 markers included 188 SSRs, 80 legume anchor markers, 46 AFLPs, 32 NBS profiling, 17 SNP, four RGA-RFLP and two SCAR markers. Moretzsohn et al., (2009) has constructed a B-genome map, complement to the previously published map of A-genome of *Arachis,* and thereby provided an entire framework for the tetraploid genome. The map was based on a F2 population of 93 individuals obtained from the

cross between the diploid *A. ipaënsis* (K30076) and *A. magna* (K30097). It included 149 loci mapped onto 10 linkage groups and covered a total map distance of 1294 cM. Varshney et al., (2009a) constructed the first SSR based genetic linkage map in cultivated groundnut using 318 RILs obtained from a cross of TAG 24 x ICGV 86031. A total of 135 out of 150 SSR loci were mapped on 22 linkage groups with the total span of 1270.5 cM and with an average intermarker distance of 9.4 cM. As an extension of work by Varshney et al. 2009, Ravi et al., 2011 developed a comprehensive and refined map with 191 SSR loci into 22 linkage groups, spanning a length of 1785.4 cM and with an average of 9.3 cM between loci. Foncek et al., (2009) developed a BC1F1 mapping populations with 88 lines comprising 2 wild diploid accessions (*A. duranensis* V14167 diploid AA and *A. ipaënsis* KG30076 diploid BB), a tetraploid AABB amphidiploid (*A. ipaënsis* × *A. duranensis*) 4X, called AiAd and a cultivated tetraploid AABB variety (Fleur 11). The amphidiploid were developed by crossing *A. ipaënsis* KG30076 (B genome) with *A. duranensis* V14167 (A genome). The resulting F1 was doubled with colchicine to produce a fertile fixed synthetic amphidiploid. Fleur 11, a local peanut variety grown in Senegal, is a Spanish type short cycle variety, high yielding and tolerant to drought. A BC1F1 and a BC2F1 populations deriving from the cross between Fleur 11 used as female recurrent parent and the amphidiploid AiAd were produced. The resulted genetic linkage map has 298 loci on 21 linkage groups spanning

a total map distance of 1843.7 cM with an average distance of 6.1 cM between adjacent markers.

Hong et al., (2010) developed composite linkage maps from three RIL mapping populations that consisted of 22 linkage groups with 175 SSR markers spanning a total composite map length 885.4 cM, with an average marker density of 5.8 cM. Khedikar et al., (2010) constructed a molecular genetic linkage map in cultivated groundnut from a mapping population consisting of 268 recombinant inbred lines obtained from a cross TAG-24 x GPBD-4 using 67 microsatellite markers. A total of 59 markers mapped on 13 linkage groups spanning 909.4 cM with an average marker interval of 15.2 cM. Sarvamangla et al., (2011) constructed a molecular genetic linkage map in cultivated groundnut from in a mapping population consisting of 146 RILs obtained from a cross TG 26 x GPBD 4 using 53 SSRs. A total of 45 markers mapped on 8 linkage groups spanning 657.9 cM with an average marker interval of 14.6 cM.

However, the above mapping studies in groundnut resulted in a lack of a comprehensive/saturated molecular genetic map based on a mapping population derived from the cross of two cultivated (4x) groundnut varieties/cultivars. This may be mainly due to two main reasons: (i) non availability of the mapping population with diverse genetic background that segregates for agronomic traits, and (ii)

unavailability of adequate and appropriate genomics tools to detect existing generic variation in primary gene pool (Varshney et al., 2006).

Greatly improved genetic maps, particularly those derived from SSRs, can contribute immensely to future groundnut improvement by plant breeders. From the review of literature it is evident that mapping of genomes is very advantageous and provides us information about the various genes that are associated with traits of agronomic importance. However, mapping populations derived from wild species showed considerable amount of polymorphism but dissipates in the successive generations. Hence, there is an exigency to explore various new molecular marker technologies like SNPs and DARTs rather than targeting wild species based material, which can track down the molecular variation in groundnut and also need for development of a well saturated and consensus map for the cultivated groundnut.

2.5.4 Marker-trait association

Marker-trait association can be determined by two ways (i) by linkage-based approach, and /or (ii) by linkage disequilibrium (LD) based association mapping. In several crops, genetic mapping based approaches were used to identify the QTLs/genes for a trait (Gupta and Varshney 2004). Recently, LD-based association mapping has been used for trait mapping (Varshney and Tuberosa 2007a).

Table 1: Details of genetic linkage maps constructed in groundnut

Mapping Population/ Population type	Marker system	Features of the maps		References
		LGs	mapped loci/genome coverage (cM)	
A. stenosperma × *A. cardenassi (F2)*	RFLP	11	117/1063	Halward et al.1993
A. kuhlmanni × *A. diogoi (F2)*	AFLP	12	102/1068	Milla 2003
A. stenosperma × *(A. stenosperma* × *A. cardenassi) (BC)*	RAPD	11	167/800	Gracia et al. 2005
A. duranensis × *A. stenosperma (F2)*	SSR	11	204/1231	Moretzsohn et al. 2005
A duranensis × *A. stenosperma (F2)*	SSR, AFLP, SNP, RFLP, SCAR	10	369/2532	Leal-Bertioli et al. 2009
A. ipaensis × *A. magna (F2)*	SSR	10	149/1294	Gobbi et al. 2006; Moretzsohn et al. 2009
ICG 12991 × ICGVSM 93541 (F2)	AFLP	5	12/139	Herselman et al. 2004
TAG 24 × ICGV 86031 (RIL)	SSR	22	191/1785	Varshney et al. 2009b; Ravi et al. 2010
Yueyou 13 × Zhen Zhuhei (RIL)	SSR	19	132/685	Hong et al. 2010
Yueyou 13 × FU 95-5 (RIL)	SSR	21	109/541	
Yueyou 13 × J 11 (RIL)	SSR	13	46/402	
TAG 24 × GPBD 4 (RIL)	SSR	20	188/1922	Khedikar et al. 2010; Sujay et al. 2011
TG 26 × GPBD 4 (RIL)	SSR	21	54/1963	Sarvamangla et al 2011 and Sujay etal. 2011
A. hypogaea × *(A. batizocoi x (A. cardenasii* × *A. diogoi)) (BC)*	RFLP	23	370/2210	Burrow et al. 2001
Aiad × Fleur 11 (BC)	SSR	21	298/1844	Foncéka et al. 2009

2.5.4.1 Linkage map based marker-trait association

Three methods have been widely used for conducting marker–trait association by using linkage maps: (i) single marker analysis (SMA), (ii) simple interval mapping (SIM), and (iii) composite interval mapping (CIM) (Tanksley 1993 and Liu 1998).

2.5.4.1.1Mapping populations used for QTL interval mapping

The construction of genetic linkage map mainly requires a segregating population (i.e. a population derived from sexual reproduction). The parents that are selected for the mapping population should differ for one or more traits of interest. Population sizes used in genetic mapping studies should range from 50 to 250 individuals (Mohan et al., 1997), however for high-resolution mapping, large populations are required. Generally in self-pollinating species, mapping populations originate from parents that are both highly homozygous (inbred), while in cross pollinating species, the situation is more complex because most of these species do not tolerate inbreeding as they are polyploidy (contain several sets of chromosome pairs). However, mapping populations used for mapping cross pollinating species may be derived from a cross between a heterozygous parent and a haploid parent (Wu et al. 1992). Several different populations may be utilized for mapping within a given plant species, with each population type possessing advantages and disadvantages (Paterson 1996). F2 populations (derived by selfing F1 hybrids), and backcross (BC) populations (derived by crossing the F1

hybrid to one of the parents) are the simplest types of mapping populations developed for self-pollinating species as they are easy to construct and require only a short time to produce. Inbreeding of individual F2 plants produces recombinant inbred (RILs) lines, which consist of a series of homozygous lines, each containing a unique combination of chromosomal segments from the original parents. The major disadvantage for producing RIL populations is the length of time required, usually six to eight generations. Doubled haploids (DH) are another type of mapping population that is produced by regenerating plants by the induction of chromosome doubling from pollen grains, however, the production of DH populations is only possible in species that are amenable to tissue culture (e.g. cereal species). The major advantage of RIL and DH populations are: (i) they produce 'true-breeding' or homozygous lines that can be multiplied and reproduced without genetic change occurring there by allowing for conducting the replicated trials across different locations and years; and (ii) seeds may be transferred between different laboratories for further linkage analysis and the addition of markers to existing maps ensuring that all collaborators examine identical material (Paterson 1996 and Young 1996). Therefore, RIL and DH mapping populations serve as 'eternal' resources for QTL mapping.

In the last few decades, different research groups all over the world developed several mapping populations using diverse parents for a combination of traits in groundnut (Table 2). In the initial stages,

mapping populations are developed with the criteria to map a maximum number of loci in a single map by selecting the parents with diverse origin; however, with the increase importance of trait mapping, mapping populations were developed more recently targeting the economically important traits such as biotic and abiotic stresses and agronomic related traits.

Table 2: List of mapping populations using diverse parents for a combination of traits in groundnut research

Population	Segregating lines	Segregating traits
AA genome		
A. stenosperma × *A. cardenassi (F2)*	-	-
A. stenosperma × *(A. stenosperma* × *A. cardenassi) (BC1F1)*	44	-
A. kuhlmanni × *A. diogoi (F2)*	179	Tomato spotted wilt virus (TSWV)
A. duranensis (PI 475887) × *A. duranensis* (Grif 15036) (F3)	98	
A. duranensis × *A. stenosperma (RIL)*	87	Late leaf spot resistance, transpiration response to drought stress, various aspects of plant morphology
BB genome		
A. ipaensis × *A. magna (RIL)*	93	Rust, various aspects of plant morphology
AABB genome		
A. hypogaea cv. IAC-Runner 886 × (*A. ipaensis* × *A. duranensis*) (RIL)	93	Rust and late leaf spot resistance, various morphological and domestication traits
ICG 12991 × ICGVSM 93541 (F2)	200	Aphid vector of groundnut rosette disease
TAG 24 × ICGV 86031 (RIL)	318	Drought related traits

TAG 24 × GPBD 4 (RIL)	266	Late leaf spot and rust resistance
TG 26 × GPBD 4 (RIL)	146	Late leaf spot and rust resistance
Tamrun OL01 × BSS 56 (RIL)	88	Yield parameter and oil content
Yueyou 13 × Zhen Zhuhei (RIL)	142	Protein content
Yueyou 13 × FU 95-5 (RIL)	84	Oil content
Yueyou 13 × J 11 (RIL)	136	Resistance to *Aspergillus flavus* and aflatoxin contamination
CG7 × ICGV-SM 94584 (F5)	111	Groundnut rosette disease
JL24 × ICGV-SM 94584 (F5)	219	Groundnut rosette disease
CG7 × ICGV-SM 90704 (F4)	338	Groundnut rosette disease
Chalimbana × ICGV-SM 90704 (F4)	597	Groundnut rosette disease
JL24 × ICGV-SM 90704 (F4)	151	Groundnut rosette disease
ICGV 93437 × ICGV 94114 (F5)	107	Leaf rust
ICGV 93437 × ICGV 95342 (F5)	466	Leaf rust
ICGV 93437 × ICGV-SM 95714 (F5)	105	Early leaf spot
ROBUT 33-1 × ICGV-SM 95714 (F5)	186	Early leaf spot
Tifrunner × Bailey High O/L (F5)	400	Oleic acid; early and late leaf spot
Tifrunner × C76-16; Florida-07 × C76-16 (F5)	400	Drought tolerance and reduced PAC
Tifrunner × NC 3033; Florida-07 × NC 3033 (F5)	400	*Cylindrocladium* black rot (CBR) disease
Tifrunner × SPT 06-06; Florida-07 × SPT 06-06 (F5)	400	Early and late leaf spot
Florida-07 × Bailey High O/L (F5)	400	White mold disease
Tifrunner × Olin (F3)	550	Oleic acid; maturity
Tifrunner × NM Valencia A (F3)	225	Tomato spotted wilt virus (TSWV); Maturity
Tifrunner × Florunner (F3)	700	Tomato spotted wilt virus (TSWV)
Florida-07 × Olin (F3)	450	Sclerotinia
Florida-07 × NM Valencia A (F3)	270	Oleic acid; tomato spotted wilt virus (TSWV); *Sclerotium rolfsii*

Florida-07 × Florunner (F3)	460	Oleic acid; tomato spotted wilt virus (TSWV); *Sclerotium rolfsii*
Florida-07 × SSD6 and Tifrunner × SSD6 (F3)	66-400	Early and late leaf spot
PI 158839 (554CC) × Tifguard (F5)	400	Nematode resistance; drought tolerance
Gregory × Tifguard (RIL)	78	Nematode resistance; late leaf spot; seed traits
SunOleic 97R × NC94022 (RIL)	354	Tomato spotted wilt virus (TSWV); oil quality
Tifrunner × GT-C20 (RIL)	246	Tomato spotted wilt virus (TSWV); early and late leaf spot; maturity
Yueyou 13 × Zhen Zhuhei and Zhen Zhuhei ×Yueyou 13 (F2)	156	Dark purple testa
A. hypogaea × *(A. batizocoi* × *(A. cardenasii* × *A. diogoi)) (BC1F1)*	78	Wild introgression
A. hypogaea cv. Fleur11 × (*A. ipaensis* × *A. duranensis)* (BC2)	59	Wild introgression

2.5.4.1.2 Approaches for QTL mapping

Two different statistical approaches are mainly used for analyses of linkage mapping based QTL mapping are: (i) the SMA method (single marker analysis), and (ii) the CIM method (composite interval mapping). SMA is the simplest method used for detecting QTLs associated with single markers. The statistical methods used for SMA include (i) *t*- tests, (ii) ANOVA (Analysis of variance), and (iii) linear regression. Linear regression is most commonly used because in this method as the coefficient of determination (R^2) from the marker will explain the phenotypic variations which arose from the QTL linked to the marker. This method is generally used in BSA (bulk segregant analysis) approach for trait mapping, however, this methods has some disadvantages such as: (i) the farther a QTL is from a marker, it is less likely to be detected as

the recombination occurring between the marker and the QTL; and (ii) the magnitude of the effect of a QTL is generally underestimated. The use of a large number of segregating markers covering the entire genome (intervals less than 15 cM), may minimize both problems (Tanksley 1993). Linkage map-based trait mapping approaches employ the SIM method that uses linkage maps and analyses intervals between adjacent pairs of linked markers along chromosomes (Lander and Botstein 1989). Therefore, use of linked markers for analyses under SIM is considered to be statistically more powerful than single-point analysis as the recombination between the markers and the QTL (Liu 1998). The most likely location of QTLs and their genetic effects were initially detected by composite interval mapping (CIM) using the WinQTL Cartographer, version 2.5 (Wang et al., 2007). The CIM approach combines interval mapping with linear regression and includes additional molecular markers in the statistical model in addition to an adjacent pair of linked markers for interval mapping (Zeng et al., 1993 and 1994). This method is more precise and effective at mapping QTLs as compared to single-point analysis (SMA) and SIM, especially when linked QTLs are involved.

QTLs identified are mainly classified into two major types based on the presence or absence of epitasis: (i) main-effect QTLs (M-QTLs), defined as single Mendelian factors at which effects on a given phenotype arise from allelic substitution and are most likely to be detected by marker-trait association using single factor ANOVA or by interval mapping models (Lander and Botstein 1989; Li et al., 1997 and Zeng

1994) and (ii) epistatic QTLs (E-QTLs), defined as loci at which trait values are determined by interactions between alleles at two or more loci and also are detected by the association between (Li et al., 1997). Differentiations of these two types of QTLs are critical to understand the genetic basis of quantitative trait variation (Li 2000).

For identification of candidate QTL regions for traits of interest, two types of trait mapping were used: (i) interval mapping to identify main effect QTLs (M-QTLs) and (ii) epistatic interaction analysis (EIA) to identify interactions between different QTL regions or epistatic QTLs (E-QTLs).

2.5.4.1.3 QTL analysis for economically important traits in groundnut

In groundnut, several mapping populations were developed using diverse parents for a combination of traits. The most important traits included biotic stresses (tomato spotted wilt virus, leaf rust, early leaf spot, late leaf spot, aphid vector of groundnut rosette disease, cylindrocladium black rot disease, sclerotinia and nematode resistance), abiotic stresses (drought related traits such as drought tolerance), nutritional quality (aflatoxin contamination, oil content, oleic acid) and several agronomic traits. The attempts made to map the economically important traits prior to the availability of SSR markers in groundnut were mostly through BSA. However, BSA was used for identification of linked markers especially for nematode resistance (Burow et al., 1996 and Garcia et al., 1996) and aphid vector of groundnut rosette disease

(Herselman et al., 2004) using the markers RAPD and AFLP respectively. The above strategy was also applied in mapping yield and yield related parameters using SSR markers (Liang et al., 2009). Similarly, in cases for resistance to foliar disease such as leaf rust (Khedikar et al., 2010), nutritional quality traits (Sarvamangla et al., 2011), resistance to nematodes (Nagy et al., 2010) and high oleate trait (Chen et al., 2002) were identified. To date, most of the trait mapping studies were conducted for biotic stress related traits, in recent years due to the availability of relatively larger number of molecular markers especially SSRs, and advanced mapping populations such as RILs, linkage mapping based marker analysis has also been undertaken to identify the QTLs related to drought related traits. (Varshney et al., 2009a and Ravi et al., 2011) (Table 3).

Table 3: QTLs identified for economically important traits in groundnut

Traits	No. of QTLs identified	Phenotypic Variance explained (%)	References
Biotic stress			
Late leaf spot (LLS)	39	1.70-67.98	Khedikar et al., 2010; Sujay et al., 2011
Leaf rust	27	1.70-82.96	
Resistance to *Aspergillus flavus* invasion	6	6.2-22.7	Liang et al., 2009
Aphid vector of groundnut rosette disease	8	1.18-76.16	Herselman et al., 2004
Abiotic stress			
Transpiration (T)	7	4.36-18.17	Varshney et al., 2009; Ravi et al., 2011
Transpiration efficiency (TE)	7	4.47-18.12	
Specific leaf area (SLA)	13	3.48-13.29	
Leaf area (LA)	3	7.25-11.51	

SPAD Chlorophyll Meter Reading (SCMR)	29	5.72-19.53	
Biomass	7	4.25-20.32	
Canopy Conductance (ISC)	7	3.28-22.24	
Total dry matter (TDM)	3	4.34-22.39	
Agronomic traits			
Shoot dry weight (ShDW)	10	5.03-22.09	Varshney et al., 2009; Ravi et al., 2011
Pod weight (PW)	7	4.17-8.73	
Seed weight (SW)	5	4.18-8.22	
Haulm weight (HW)	6	3.78-33.66	
Pod mass/plant	3	13.1-18.3	Liang et al., 2009
Mature pod/plant	3	11.9-12.3	
Number of branches	7	8.1-17.3	
Number of fruit branches	1	17.5	
Height of main axis	7	8.2-12.8	
Stem diameter	4	7.8-24.1	

Leaf length	1	12.4	
Leaf width	3	13.2-18.7	
Leaf length / width ratio	3	15.4-18.9	
Yield parameter	5	9.19-17.69	Selvaraj et al., 2009
Oil content	7	1.5-9.5	Liang et al., 2009; Selvaraj et al., 2009; Sarvamangla et al., 2011
Protein content	10	1.4-9.7	Liang et al., 2009; Sarvamangla et al. 2011
oil quality	10	1.5-13.3	Sarvamangla et al., 2011

2.6 Consensus map development

Dense genetic linkage maps are cornerstones for a spectrum of biotechnology and breeding applications such as trait mapping through quantitative trait locus (QTL)/ association mapping, marker-assisted breeding and map-based cloning. It is not possible to map all the available markers in a single mapping population in a given crop, however, several individual genetic linkage maps are constructed for specific traits of interest with fewer mapped loci. Recently, with the increased interest and applications towards comparative genetics, researchers were been gathering data from multiple populations and lines of the same species and trying to integrate into a single map called consensus map which serves as a excellent platform for representing the position and order of markers in whole genome.

Further, as compared to mapping based on a single population, mapping with multiple populations data provides several advantages such as (1) mapping large number of loci onto a single map, (ii) determining relative position of transferable markers, (iii) determining stability of locus position across the genome, (iv) providing evidence for chromosomal rearrangements (Beavis and Grant 1991 and Kianian and Quiros 1992), gene duplication (Kianian and Quiros 1992 and Gentzbittel et al.,1995) and assisting in the assignment of linkage groups to chromosome (Beavis and Grant 1991), (v) providing the basis for comparative genomic studies among related species and subspecies

(Kianian and Quiros 1992; Hauge et al., 1993 and Gentzbittel et al., 1995) and (vi) providing genetic information for greater genomic coverage (Sewell et al., 1999). The one and foremost application of any dense consensus genetic linkage map is to identify chromosomal segments associated with traits of interest through QTL analysis that provides the information about contribution of several loci along with their interactions in a segregating cross (Borevitz and Chory 2004). Consensus genetic maps have been constructed in several crop species such as maize (Beavis and Grant 1991 and Falque et al., 2005), soyabean (Song et al., 2004 and Choi et al., 2007), barley (Wenzl et al., 2006 and Marcel 2007). Groundnut is lagging behind in this area except a recent report on comprehensive genetic map with 175 loci using three mapping populations (Hong et al., 2010).

3. MATERIALS AND METHODS

3.1 Development of Novel Simple Sequence Repeat (SSR) Markers

In order to increase the repertoire of molecular marker resources in groundnut, novel microsatellite or simple sequence repeat (SSR) markers were developed by constructing a SSR-enriched library that was subsequently screened for polymorphic markers between the parental lines of two RIL populations and then mapped by genotyping the two RIL populations.

3.1.1 Construction of microsatellite enriched genomic DNA library

A new SSR-enriched genomic DNA library was developed from the cultivated groundnut genotype ICGV 86031 using the bead capture enrichment protocol of Glenn et al., (2005). ICGV 86031 was derived from a cross between F 334A-B-14 and NC Ac 2214 during 1982 at ICRISAT, Patancheru, India. It is a high-yielding line with many desirable traits and with multiple resistance or tolerance to insect pests and bud necrosis disease. The microsatellite library was enriched using two types of oligo sequences (AAG) 8, (CT) 10, (AG) 8, (TG) 12 and combinations of these.

Young tender leaves of ICGV 86031 15 days old seedlings were collected from the greenhouse at ICRISAT, Patancheru, India. Genomic DNA was extracted using the modified CTAB (Cetyl Trimethyl Ammonium Bromide) method (Sambrook and Russel 2001). The leaf samples were

frozen in liquid nitrogen, 3 g of frozen leaves ground in a precooled mortar and pestle and transferred to a 50 ml falcon tube. To this, 15 ml of DNA extraction buffer (2% CTAB, 100 mM Tris, 20 mM EDTA (Ethylenediaminetetraacetic acid), 1.4 M NaCl, pH 8.0) preheated at 60˚C and 200 mg of PVP (Polyvinylpyrolidone) were added. The contents were mixed gently by swirling and inverting the tube for 3-4 hrs at room temperature and the incubated at 60˚C in a water bath for 45 min with occasional mixing. The tubes were removed, cooled to room temperature, and an equal volume of chloroform-isoamyl alcohol (24:1) was added. The contents were mixed by inversion for 10 minutes and centrifuged at 12000 rpm for 15 minutes. The clear aqueous upper layer was transferred to a new tube and reextracted with an equal volume of chloroform-isoamyl alcohol (24:1). To the final aqueous solution, an equal volume of absolute ethanol was added, mixed by inversion and placed at -20˚C for 15 min. Genomic DNA was pelleted by centrifuging at 14000 rpm for 20 min at 10˚C and the pellet was washed with 70% ethanol. The DNA pellet was dried in a DNA concentrator (Thermo) and suspends the DNA in 1 ml of sterile double distilled water.

When DNA is fully suspended, add 20 µl of RNaseA (5 mg/ml) and incubated at 37˚C for 1 hour. DNA was extracted by adding an equal volume of phenol-chloroform-isoamylalcohol (25:24:1), mixing briefly and spinning at 13000 rpm for 15 min. To the top aqueous layer, 1/10 volume of 3 M sodium acetate and twice the volume of absolute ethanol were added, mixed by inversion and kept at -20˚C for 1 hour. The DNA

was then precipitated by centrifuging at 12000 rpm for 15 min at 10˚C. The pellet was washed thrice with 70% ethanol, dried and dissolved in 800 µl of sterile double distilled water, and stored at -20˚C. The quantity of extracted DNA was estimated based on the intensity of uncut DNA. DNA quantification and purity was checked by measuring the optical density at 260 nm and 280 nm using a UV visible spectrophotometer.

Steps involved in constructing SSR enriched library are as follows:

1. Restriction digestion of genomic DNA;
2. Ligation of double-stranded linkers to the digested DNA;
3. Enrichment of the microsatellite library using streptavidin magnetic beads;
4. Ligation of microsatellite–enriched DNA fragments and plasmid vector (TOPO VECTOR-Invitrogen);
5. Transformation via electroporation; and
6. Selection and amplification of positive colonies.

Genomic DNA (2.0 g) was digested with *Rsa*I and *Xmn*I (New England BioLabs,UK) in a reaction volume of 25 µl having final concentration of 1X NEB buffer, 10 U/µl of *Rsa*I, 10 U/µl of *Xmn*I, 50 mM NaCl, and incubated at 37˚C for 5 h. Digested samples were stored at -20˚C. Complete digestion was confirmed by the presence of a dense smear at 100-1000 bp range following agarose gel electrophoresis. Double stranded (ds) linker was prepared by adding equimolar concentrations

(10 µM) of Super SNX24 (5' GTTTAAGGCCTAGCTAGCAGAATC) and Super SNX24 + 4p (5'pGATTCTGCTAGCTAGGCCTTAAACAAAA) primers which were initially single stranded. The reaction was carried out in 0.5 ml eppendorf tube in a total volume of 200 µl containing 100 mM NaCl, heated to 95°C for 5 min and gradually cooled down to room temperature to favour annealing and formation of the double stranded linker.

Ligation of ds linkers to digested DNA fragments for enrichment of DNA fragments was performed in the molar ratio of 1:10 (DNA fragments: ds linkers). The reaction was carried out in a 0.2 ml microfuge tube containing 6.0 picomoles of DNA fragments, 60 picomoles of double stranded linker in a total reaction volume of 50µl having final concentration of 1X ligase Buffer (NEB), 50 U/µl of high concentration T4 DNA ligase enzyme (NEB) and incubated at 16°C overnight. The samples were stored at -20°C until further use.

Ligation of ds linkers to digested genomic DNA fragments was confirmed by setting up two PCR reactions with super SNX24 primer, one in which 2.0 µl of linker ligated DNA was used as template and the other with 4.0 µl of linker ligated DNA as template in a reaction volume of 20 µl containing, 1X PCR buffer (10 mM Tris HCl, pH 5.3, 50 mM KCl), 1.5 mM $MgCl_2$ (Invitrogen), 0.1 mM dNTPs (Amersham), 0.5 picomoles of super SNX24, 0.2 units of *Taq* polymerase (Invitrogen) along with one negative control where there was no linker ligated DNA. The PCR

programs included an initial denaturation step of 2 min at 95˚C, followed by 30 cycles of 20 sec denaturation at 95˚C, 20 sec annealing at 60˚C, extension was carried out at 72˚C for 15 min and a hold temperature of 15˚C. PCR reactions were done in a DNA thermocycler (Peltier Thermocycler) with heated lid. PCR amplification was checked by agarose gel electrophoresis. Repeat enrichment of genomic DNA fragments was done using biotinylated repeat oligonucleotides. In the current study, four biotin labeled primers were used and hybridization reactions were carried out individually with the five oligonucleotides at their respective hybridization temperatures (Table 4).

Table 4: List of hybridization temperatures for different biotinylated oligonucleotides

Biotinylated repeat oligo	Hybridization temperature
$(CT)_{10}$	42˚C
$(TG)_{12}$	45˚C
$(AG)_{14}$	45˚C
$(AAG)_8$	40˚C
Mixture	45˚C

A homogenized Streptavidin magnetic bead (50 µl of 10 µg/µl, NEB) was aliquoted in a 1.5 ml tube to which 250 µl of TE (10 mM Tris pH 8.0, 2 mM EDTA) was added and washed twice. Beads were captured using a magnetic particle concentrator. The beads were then washed twice in 1X hybridization solution and finally resuspended in 150 µl of 1X hybridization solution. The hybridization reaction was carried out in a

0.2 ml microfuge tube having final concentration of 0.25 ng/µl of linker ligated DNA fragments, 1.0 picomol/µl of biotinylated repeat oligo, 12X SSC, 0.02% SDS (hybridization buffer) in a total reaction volume of 50 µl, incubated at 95˚C for 5 min and quick chilled on ice for 2 min. Samples were incubated at respective hybridization temperature of biotinylated probe for 1 h in the thermal cycler (Eppendorf mastercycler gradient).

For conjugation, the hybridization mix was transferred into a 1.5 ml eppendorf tube to which 50 µl of Streptavidin magnetic beads (10 µg/µl, NEB) were added and incubated at room temperature for 30 min with constant gentle agitation. After conjugation, the eppendorf tube was placed in a magnetic particle concentrator (MPC) and supernatant was removed. The bead-hybridized fragment complex was washed 2 times each for 5 min by adding 400 µl 2X SSC, 0.1% SDS. The complex was further washed with 400 µl of 1X SSC, 0.1% SDS four times each for 5 min at room temperature. Each time the MPC was used to collect the beads and the supernatant was collected with a P200 pipettor, and saved for troubleshooting. The solution was heated within 5-10˚C of the T_m for the oligo used (45-50°C).

After washing, 200 µl of Tris Low EDTA (Ethylenediaminetetraacetic acid) was added, tapped gently and incubated at 95˚C for 5 min. The eppendorf tube was placed immediately in a magnetic stand and the supernatant containing DNA fragments enriched with oligonucleotide

repeats was pipetted out in a microcentrifuge tube. After capturing the beads, 22 µl of NaOAc-EDTA (Sodium Acetate-EDTA) solution was added to the supernatant and mixed by pipetting up and down. To this, 444 µl of 95% ethanol was added, mixed by inverting the tube, kept on ice for 15 min or longer, and centrifuged at full speed for 10 min. The supernatant was removed using a pipette and the "enriched Gold" DNA was air dried to form a pellet which was resuspended in 25 µl of TLE and stored at 4°C.

A total of 5 PCR reactions for each of the 2 eluted samples were carried out in a DNA thermocycler (master cycler gradient) with heated lead. Each 25 µl reaction volume contained about 2.0 µl of eluted DNA, 1X PCR buffer (Invitrogen), 1.5 mM $MgCl_2$ (Invitrogen), 0.1 mM dNTPs (Amersham), 0.5 picomoles of super SNX24 primer, 0.3 units *Taq* polymerase (Invitrogen). The PCR program included an initial denaturation step of 2 min at 95°C, followed by 30 cycles of 20 sec denaturation at 95°C, 20 sec annealing at 60°C, 1.5 min extension at 72°C, final extension at 72°C for 30 min, and a hold temperature of 15°C at the end. The PCR product was run on a 1.5% agarose gel to visualize 500 bp smears. All the 5 PCR reaction products were pooled and stored at 4^0C until further use. Ligation reactions were performed individually for all the repeat enriched DNA fragments obtained by using two biotinylated repeat oligonucleotides. Ligation reactions were performed in a 10 µl reaction volume containing 1 µl of pCR2.1-TOPO vector

(Invitrogen, Carlsbad, USA), 5 µl of PCR enriched product, 1 µl 10X T_4 DNA ligation buffer (NEB), 2.4 U/µl of T_4 DNA ligase (NEB) and incubated at 14˚C overnight. The samples were stored at -20˚C until used and further transformed into *Escherichia coli* competent cells (TOP10, Invitrogen, Carlsbad, USA).The transformation mix was plated on LB Amp^+ (Lysogeny Broth) agar plates coated with IPTG (Isopropyl β-D-1-thiogalactopyranoside) and X-gal (5-bromo-4-chloro-indolyl-β-D-galactopyranoside) and incubated at 37˚C overnight to allow for blue and white colonies selection. Subsequently, white colonies that are consider being the SSR positives were picked with a sterile toothpick and suspended in 20 µl of sterile millipore H_2O, out of which 5 µl was used as template for colony PCR and 15 µl was kept aside for primary culture inoculation in a 96 well plate containing 200 µl each well of LB-amp (Luria Broth-ampicillin). This primary culture was kept in incubator shaker at 37˚C and 200 rpm overnight. Colony PCR were performed using a 10 µl reaction volume with 5 µl of colony suspension as template, 1X PCR buffer (Invitrogen), 1.5 mM $MgCl_2$ (Invitrogen), 0.1 mM dNTPs (Amersham), 0.5 picomoles of M13 forward and reverse primers, 0.3 units *Taq* polymerase (Invitrogen). The PCR program included an initial denaturation step of 2 min at 94˚C, followed by 35 cycles of 30 sec denaturation at 94˚C, 30 sec of annealing at 55˚C, extension at 72˚C for 2 min, final extension at 72˚C for 1.0 min, and a final hold temperature of 15˚C.

3.1.2 Sequencing of microsatellite enriched clones

The colonies having insert sizes more than 300 bp were selected according to colony PCR results visualized on a 1.5% agarose gel. The primary cultures (20 µl) grown in 200 µl of LB Amp^+ in a 96 well plate and derived from colonies that had more than 300 bp inserts were sub-cultured in 5 ml of LB Amp^+ medium and kept at 37˚C in a orbital shaker (Thermo Electron Corporation, California, USA) at 200 rpm overnight for plasmid isolation.

The plasmid DNA from individual clones was isolated using the standard alkaline lysis method (Sambrook and Russell 2001). The overnight culture was transferred into 1.5 ml sterile eppendorf tube and centrifuged at 3200 rpm for 5 min at 4˚C. Supernatant was discarded and the remaining culture was also transferred into an eppendorf tube and centrifuged at 3200 rpm for 5 min at 4˚C. The supernatant was discarded and the tube inverted on a paper towel to remove the entire supernatant. The pellet was resuspended in 200 µl of Lysis I and mixed thoroughly in a vortex mixer. To this, 200 µl of freshly prepared Lysis II was added, the contents mixed by gently inverting the tube 5-6 times and kept in ice for 5 min. To this, 300 µl of potassium acetate was added; the tube inverted 5-6 times gently and kept in ice for 5 min. The tubes were centrifuged at 3200 rpm for 15 min and the supernatant was transferred into a sterile 1.5 ml eppendorf tube to which 10 mg/ml RNase A was added and incubated at 37˚C for one hour. An equal volume

of chloroform: isoamyl alcohol was added, mixed by inversion and centrifuged at 3200 rpm for 20 min. The supernatant was transferred to a sterile 1.5 ml eppendorf tube to which an equal volume of isopropanol was added, mixed by inversion and centrifuged at 3200 rpm for 15 min. The supernatant was discarded and the pellet was washed with 70% ethanol 2 times. The pellet was dried in a DNA concentrator (Thermo) and dissolved in 30 µl of 5 mM Tris. One µl of each plasmid was run in a 1.0% agarose gel along with standard of 1µl *Hind*III (Fermentas, USA) digest marker to check the quality and quantity of plasmid.

A set of 72 positive clones were sequenced only in one direction (5') using M13F-pUC (-40) as sequencing primers, by adopting Sanger's dideoxy sequencing method and BigDye Terminator version 3.1 kit ABI 3700 (Applied biosystems, USA). For sequencing, microsatellite enriched clones, forward and reverses sequencing polymerase chain reactions (PCR) were performed separately in 10 µl reaction volume containing 2 µl of BigDye Terminator version 3.1 (BDT v 3.1) reaction mix (Applied Biosystems, Foster city, USA), 0.5 µl of 5X reaction buffer (Applied Biosystems, USA). 3.2 picomoles of forward and reverse primers for respective reactions, 1 µl of plasmid (100 ng/µl) and 6 µl of sterile water. The cycle sequencing PCR profile used involved 30 sec of intial denaturation followed by 40 cycles of 10 sec at 96 ˚C (denaturation), 5 sec at 50˚C (primer annealing) and 60˚C for 4 min (primer extension) as per instruction manual of BDT v 3.1 sequencing kit (Applied Biosystem, USA). PCR products were given an ExoSAP treatment followed by ethanol

washes in order to remove excessive polyA overhangs and unused dNTPs and then the samples were sequenced using an ABI 3700 Genetic Analyzer (Applied Biosystems, USA).

3.1.3 Primer design and synthesis

A total of 144 sequence reads were obtained from the sequencing of the 72 positive microsatellite enriched clones. The sequences were cured to remove the remnants of vector sequences from both 5' and 3' ends using the vecscreen program at NCBI (http://www.ncbi.nlm.nih.gov/vecscreen/vecscreen.html). Following vector trimming, the 144 sequence reads were formed into contigs (alignment of forward and reverse sequences) using DNABaser v2. The CAP3 programme was used to remove the sequence redundancy. As a result, the contigs and singletons thus obtained from CAP3 assembly (Huang and Madan 1999) were used in the FASTA format in a single file for microsatellite search using *MI*cro*SA*tellite (*MISA*) (Thiel et al., 2003), a tool for identification and localization of both perfect and compound SSRs (two individual microsatellites interrupted by up to 100 bp). The sequences from SSR-enriched library were used for designing primer pairs for SSR using Primer3 programme V 3.0 (Rozen and Skaletsky 2000) (http://frodo.wi.mit.edu/) in a batch file. The SSR markers developed from the microsatellite-enriched library were named as **ICGM** (ICRISAT Groundnut Microsatellite followed by the clone ID).

3.1.4 Optimization and validation of SSR markers

In order to check the amplification of 23 primer pairs, a PCR was performed using two genotypes (ICGV 86031 and TAG 24) in a 5µl reaction volume consisting of 0.5 µl of 10X PCR buffer (Sib Enzymes, Russia), 0.3 µl of 25 mM $MgCl_2$ (Sib Enzymes, Russia), 0.3 µl of 2 mM dNTPs, 0.5 µl of (1 picomole/µl M13 tailed forward primer: 2 picomole/µl reverse primer), 0.1 U (0.2 µl of 5U/µl) of *Taq* DNA polymerase (Sib Enzymes, Russia), and 1 µl of 5 ng DNA template in 96-well microtiter plate. A common touch down PCR profile was performed with 3 min of initial denaturation cycle, followed by first 5 cycles of 94°C for 20 seconds, 65°C for 20 sec and 72°C for 30 sec, with 1°C decrease in temperature per cycle, then 40 cycles of 94°C for 20 sec with constant annealing temperature (59°C) for 20 sec and 72°C for 30 sec, followed by a final extension at 72°C for 20 min. The PCR products together with a 100 base pair ladder were tested for amplification in a 1.2% agarose gel containing 0.5 µl/10ml ethidium bromide (10 mg/ml) by running it at a constant voltage of 80V for 30 min. The amplification was visualized under UV illumination using Uvi Tech gel documentation system (DOL-008.XD, England).

3.2 Construction of Genetic Linkage Map

In the current study, two new mapping populations comprising of 176 and 188 Recombinant Inbred Lines (RILs) segregating for drought tolerance traits viz., transpiration (T), transpiration efficiency (TE), specific leaf area (SLA) and SPAD chlorophyll meter reading (SCMR) were

used. The two mapping populations are derived from the cross ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76. The RILs along with both the parents for were used for phenotyping and genotyping.

3.2.1 Salient features of parents and mapping populations

The salient features of parents of mapping population are as follows:

- **ICGS 76** also know as ICGV 87141, is a high-yielding Virginia botanical type variety, developed at ICRISAT. It matures in 120 days in the rainy season, and has a shelling percentage of 73%. This variety was selected by the bulk pedigree method and derived from a cross between an adapted variety, TMV 10 and an early-maturing source line, Chico. Its pedigree is (TMV 10 x Chico) F2B2-NIB1-B1-B1-B1-B1-B1-B1-B1-B1-B1 and it is mainly adapted to low-input rainfed conditions. It is tolerant to bud necrosis disease and has good recovery from the mid-season drought. It has 30% pod yield and 36% seed yield with oil content of 43% and superiority over the popular varieties such as Kadiri 2, TMV 10, and M 13. Its productivity potential is up to 2.5-3.5 t/ha under good management conditions.

- **CSMG 84-1:** It is a new high-yielding, early maturing, rust-resistant and drought tolerant Virginia runner variety developed at Groundnut Research Station, Manipuri, Uttar Pradesh and ICRISAT. This Virginia variety was selected from MA 10. It appears to be more tolerant to thrips, leaf miners, termites, white grub, foliar disease and pod borer

than the controls (M 13, MA 10, and M 335). It has a very good shelf life and has less insect pest damage during storage. This new hypogaea type has wide range of adaptability when evaluated in agronomic trials for sowing, irrigation and fertilizer schedules.

- **ICGS 44** also known as ICGV 87128 is an improved high yielding bunch variety, bred and developed at ICRISAT. Its pedigree is (Robut 33-l)-l-5-B1-B1-B1-B1-B1 and it is derived from a single plant selection made in a natural hybrid population of an Indian variety Robut 33-1 (Kadiri 3). It has two-seeded small to medium-sized pods and tan colored seeds. This vulgaris type is tolerant to drought and can withstand bud necrosis. It is relatively photoperiod insensitive has good recovery from mid-season drought and is average in its response to end-of-season drought. The shelling turnover is 70%, oil content 48% and protein content 23%. It matures in 110 to 120 days when grown during summer and can yield 3000 to 4000 kg/ha.

The two RIL mapping populations, ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76, at the F_9:F_8 generation were developed at ICRISAT centre, for drought related traits as single seed descendants from the F_3 generation onwards and the remaining F_8 seeds were advanced to F_9/F_{10} for further phenotyping and genotyping (Krishnamurthy et al., 2007; Vadez et al., (unpublished)).

3.2.2 Genotyping of mapping populations

3.2.2.1 DNA isolation of respective parents and RILs

DNA was extracted from fresh furled leaves of the parental genotypes and 176 RILs of ICGS 76 × CSMG 84-1 (F_9 generations) and 188 RILs of ICGS 44 × ICGS 76 (F_8 generation) population using a high-throughput mini-DNA extraction method as per Cuc et al., 2008.

3.2.2.1.1 High-throughput mini-DNA extraction

A. Sample preparation

- Leaves were harvested from 15 days old seedlings.
- Leaf tissue of 70-100 mg was placed in a 12 × 8-well strip tube with strip cap (Marsh Biomarket, USA) in a 96 deep-well plate together with two 4 mm stainless steel grinding balls (Spex CertiPrep, USA).

B. CTAB extraction

- To each sample, 450 µl of preheated (at 65ºC for half an hour) extraction buffer (100 mM Tris-HCl (pH-8, 1.4 M NaCl, 20 mM EDTA, CTAB (2-3% w/v), β-mercaptoethanol) was added and secured with strip caps.
- Samples were homogenized in a Geno Grinder 2000 (Spex CertiPrep, USA), following the manufacturers instructions, at 500 strokes/min for 5 times at 2 min interval.

- Plate was fitted into locking device and incubated at 65°C for 10 min with shaking at periodical intervals.

C. Solvent extraction

- To each sample, 450 µl of chloroform-isoamylalcohol (24:1) was added and mixed thoroughly by inverting twice.
- The plate was centrifuged at 5500 rpm for 10 min, and the aqueous layer (300 µl) transferred to fresh strip tubes (Marsh Biomarket, USA).

D. Initial DNA precipitation

- To each sample, 0.7 vol (210 µl) of isopropanol (stored at –20°C) was added and inverted gently to mix.
- The plate was centrifuged at 5000 rpm for 15 min. The supernatant was decanted from each sample and the pellet air dried for 20 min.

E. RNase treatment

- A volume of 200 µl low salt TE (10 mM Tris EDTA (pH-8)) and 3 µl RNase was added to each sample and incubated at 37°C for 30 min.

F. Solvent extraction

- A volume of 200 µl of phenol-chloroform-isoamylalcohol (25:24:1) was added to each sample and inverted twice to mix.
- The plate was centrifuged at 5000 rpm for 5 min.
- The aqueous layer was transferred to a fresh 96 deep-well plate (Marsh Biomarket, USA).
- A volume of 200 µl chloroform-isoamylalcohol (24:1) was added to each sample and inverted twice to mix.
- The plate was centrifuged at 5000 rpm for 5 min, and the aqueous layer transferred to fresh 96 deep- well plate
- A total of 315 µl ethanol-acetate solution (30 ml ethanol, 1.5 ml 3 M NaOAc (pH-5.2)) was then added to each sample and placed at -20°C for 5 min.
- Plate was again centrifuged at 5000 rpm for 5 min.

- The supernatant was decanted from each sample and the pellet washed with 70% ethanol.
- The plate was centrifuged at 6000 rpm for 10 min.
- The supernatant was again decanted from each sample and samples air dried for 1 hour.
- The pellet was resuspended in 100 µl low-salt TE and stored at 4°C.

3.2.2.1.2 DNA quantification

DNA was quantified by loading the samples in a 0.8% agarose gel containing 0.5 µl/10 ml Ethidium bromide (10mg/ml). The DNA was normalized to 5 ng/µl concentration by comparing visually the diluted DNA samples with standard lambda DNA molecular weight markers (2.5 ng/µl, 5 ng/µl, 10 ng/µl) in an 0.8% agarose gel ran in 0.5X TBE (Tris borate EDTA) buffer at a constant voltage (80 V) for 20 min. The images of gels were documented under UV illumination using Uvi Tech gel documentation system (DOL-008.XD, England).

3.2.3 Screening of SSR markers

In addition to the 23 novel SSRs developed in the present study, a total of 3215 simple sequence repeats (SSR) markers available in the public domain and/or accessed through collaborators (Table 5) were also used to screen the polymorphism across the parents of the two mapping populations and then employed for genotyping the respective mapping populations. The forward primers of all these markers were synthesized with M13 tail for their ease in genotyping on ABI-3700 automatic DNA sequencer (PE- Applied Biosystems, California, USA).

Table 5: Details on markers used for screening the polymorphism in two mapping populations

Source	Marker Series	No. of Markers screened
Ferguson et al., 2004	pPGPseq, pPGSseq	226
Mace et al., 2007	Chaet, Dal, Lup, Stylo, Ades, Amor	51
Cuc et al., 2008	IPAHM	104
Moretzsohn et al., 2004; 2005	Ah, gi, RN, ML, RI, TC, AC	338
Proite et al., 2007	RM,RN	53
He et al., 2003	PM	59
S J Knapp et al., (Unpublished)	GM	2217
Hopkins et al., 1999	Ah	26
Palmieri et al., 2002; 2005	AP	18
Wang et al., 2007	S	123
Total		3215

3.2.3.1 Amplification

For parental screening to assess the polymorphism and further genotyping of the respective mapping populations, a common PCR profile was used for the entire set of SSRs. All PCR reactions were performed in 5 µl reaction volume consisting of 1 µl of 5 ng DNA template, 0.3 µl of 2 mM dNTPs, 0.5 µl of (1 picomole/µl M13 tailed forward primer, 2 picomole/µl reverse primer) and 1 µl of 2 picomole/µl of M13 labeled dye, 0.1 U (0.2 µl of 5 U/µl) of *Taq* DNA polymerase (Sib Enzymes, Russia), 0.5 µl of 10X PCR buffer (Sib Enzymes, Russia), and 0.3 µl of 25 mM $MgCl_2$ (Sib Enzymes, Russia). In addition, the fluorescent dyes 6-FAM, VIC, NED, PET were used in the PCR reaction mixture for detection in the ABI 3700. PCR amplifications were performed in an ABI thermal cycler (PE Applied biosystems, CA). A touch down PCR amplification profile was used with 3 min of initial denaturation cycle, followed by first 5 cycles of 94°C for 20 seconds, 65°C for 20 sec and 72°C for 30 sec, with 1°C decrease in temperature per cycle, then 40 cycles of 94°C for 20 sec with constant annealing temperature (59°C) for 20 sec and 72°C for 30 sec, followed by a final extension at 72°C for 20 min. The PCR products together with a 100 base pair ladder were tested for amplification in a 1.2% agarose gel containing 0.5 µl/10ml ethidium bromide (10 mg/ml) ran at a constant voltage of 80V for 30 min. The amplification was visualized under UV illumination using Uvi Tech gel documentation system (DOL-008.XD, England).

3.2.3.2 SSR Fragment Analysis

After confirming the PCR amplification on 1.2% agarose gel, five post-PCR multiplex sets were constructed based on the allele size range estimates and the type of forward primer label of the markers. Markers that had different labels and allele size ranges were considered for a set. For post PCR multiplexing, 1.5 µl PCR product of each of 6-FAM, VIC, NED and PET-labeled products were pooled (according to above mentioned criteria) and mixed with 7 µl of Hi-Di formamide (Applied Biosystems, USA), 0.25 µl of the LIZ-500 size standard (Applied Biosystems, USA) and 1.5 µl of sterile distilled water. The pooled PCR amplicons were denatured and size fractioned using capillary electrophoresis on an ABI 3700 automatic DNA sequencer (Applied Biosystems, USA). Allele sizing of the electrophoretic data thus obtained was done using Genemapper® software version 4.0 (Applied Biosystems, USA).

3.2.3.3 Data Analysis

After completion of capillary electrophoresis, files generated by ABI machine were processed using Genemapper® software version 4.0 (Applied Biosystems, USA). The GeneScan option assigns the product sizes to all PCR amplicons based on their relative mobility with internal LIZ size standard and the Genotype option assigns the product sizes for

each markers. Raw allele calls derived were processed through the AlleloBin programme (Prasanth et al., 2006) which uses repeat motif as a reference to call the perfect allele. Based on the amplicon sizes in the parents, the segregation data were scored for all the optimized primers. Lines having the allele of "female parent" were scored as "A", the "male parent" as "B", alleles from both the parents as "H", and missing data as "-". Since the present study involved two mapping populations with four different parents, the allele scoring was as follows:

(i) for mapping population ICGS 76 × CSMG 84-1

'A' – allele of female parent (ICGS 76)

'B' – allele of male parent (CSMG 84-1)

(ii) for mapping population ICGS 44 × ICGS 76

'A' – allele of female parent (ICGS 44)

'B' – allele of male parent (ICGS 76)

(iii) and for both mapping populations

'H' – heterozygous (presences of both parent alleles)

'-' – missing data (amplification failed)

3.2.3.4 Linkage map construction

Chi-square (X^2) tests were performed on the genotypic data to test the null hypothesis of expected 1:1 Mendelian segregation on all the scored markers using JoinMap 4.0 (Stam 1993). In mapping population ICGS 76 × CSMG 84-1, 64 out of 128 markers genotyped showed segregation distortion (SD). In the second mapping population ICGS 44 × ICGS 76, 11 markers were distorted out of 87 markers that were genotyped; however, due to low availability of polymorphic markers in both the mapping populations, even the distorted markers were used for linkage map construction and further QTL analysis.

Genotyping data obtained for all the polymorphic marker loci on the respective RIL mapping populations were used for linkage analysis using Mapmaker/EXP v.3.0 (Lander et al., 1987 and Lincoln et al., 1993). The markers were classified into linkage groups (LGs) using linear regression of pairwise analysis using a minimum LOD threshold of 5.0 and maximum recombination fraction (θ) of 0.35 for both the mapping populations. The most likely marker order within each LG was estimated by comparing the log-likelihood of the possible orders of markers using multipoint analysis "Compare" command. The "Try" and "Build" commands were also used to determine the most likely placement of the unlinked markers, and subsequent orders were tested using the "Ripple" command with "Error Detection" and "Use Three Points" options enabled.

The distance between neighboring markers were calculated using the multipoint analysis implemented in the "Map" command. The Kosambi mapping function (Kosambi 1944) was used to estimate the map distances in centimorgans (cM). The inter-marker distances calculated from Mapmaker were used to construct linkage map using MapChart v.2.2 (Voorrips 2002).

3.3 Phenotyping of mapping populations for drought related traits

3.3.1 Phenotypic traits

The mapping population ICGS 76 × CSMG 84-1 (comprising 176 F_9 lines) was phenotyped for several drought related traits. Phenotypic data was collected for traits such as transpiration efficiency (TE), transpiration (T) and biomass dry weight (DW) during post rainy season in 2008 under well watered (WW) and water stress (WS) regimes following the protocols given in Krishnamurthy et al., 2007. For taking the traits observations, the plants were grown in 28 cm diameter pots that were filled with 10 kg of Alfisols soil collected from the ICRISAT research station and suitably fertilized. Three seeds of each genotype were planted and the pots thinned to one healthy seedling per pot at two weeks after sowing. Pots were then saturated with water, and left to drain overnight. To avoid evaporation, plants were bagged around the stem and regular weighing was done and plant transpiration measured. The water stress treatment was applied by allowing plants to loose no more than 100

g/day water, following previous procedures (Krishnamurthy et al., 2007). An extra set of plants was used to calculate the biomass before starting the experiment. In the second season (2009 postrainy season), the population was phenotyped for specific leaf area (SLA), transpiration efficiency (TE), leaf dry weight (LDW), biomass dry weight (DW), transpiration (T), SCMR and leaf area (LA) under well watered (WW) conditions only.

The second mapping population, ICGS 44 × ICGS 76 (comprising 188 F8 lines) was also phenotyped for drought related traits and harvest index (HI). Phenotyping was done in the field under both fully irrigated conditions and intermittent drought stress conditions. In case of intermittent drought stress conditions, stress was applied from 40 days after sowing, by skipping the irrigation once every two times that the fully irrigated control was watered, so that the amount of water received in the water stress treatment was about half of that in the fully irrigated control. This population was mainly phenotyped for vegetative weight at harvest (VW), shoot plus pod dry weight (ShDW) and the harvest index (HI) during post rainy season in year 2008. The methods that were employed for recording the observations of drought related traits for both mapping populations are explained briefly in the following section.

Transpiration efficiency (TE, g biomass/kg water transpired)

Transpiration efficiency was calculated as the ratio of the biomass increased during the experimental period, divided by the amount of water used during that time.

TE was calculated as:

$$TE = (DM2-DM1) / (W2-W1) + WA$$

Where DM1 = the mean shoot biomass in a set of pot harvested at four weeks after sowing; DM2 = shoot biomass at harvest; W1 = weight of the pot at the time of mulching beads; W2 = weight of the pot at time of final harvest; and WA = the water added to individual cylinder after regular weighing.

Transpiration (T, g/plant)

Transpiration is the amount of water evaporated through the leaf stomata and was calculated using the gravity method by regularly (daily) weighing of the pots and the soil surface of the pots bagged with polyethylene bags in order to prevent evaporation from the soil.

Biomass dry weight (DW, g/plant)

Biomass dry weight is the sum of the shoot and pod dry weights.

Specific leaf area (SLA, cm^2/g)

SLA was calculated using the following equation:

SLA = Leaf area (cm^2)/Leaf dry weight (g)

Leaf area (LA)

At the time of harvest, 130 days- after-sowing plant parts were separated into leaf, stem and pods. Specific leaf area was measured using a leaf area meter (Model LI-3100, Lincoln, Nebraska, USA)

Vegetative weight at harvest (VW, g/plant)

Dry weights of stem and leaf were measured after keeping them to 60^{o}C in a hot air oven for 72 hours.

Pod weight (g)

Pod weights were measured after drying under natural sunlight and temperature at 60^{o}C for 72 hours in a hot air oven.

Harvest index (HI)

Harvest index was calculated as the ratio of economical yield (pod yield) and the total biomass (total dry weight) of plant using the following relationship:

Harvest index (HI) = Pod yield/ Pod yield + Shoot and root dry weight

Shoot plus pod dry weight (ShDW, g/plant)

The shoot plus pod dry weight was calculated as the weight of leaves and stems together after drying in a hot air oven at 60^{o}C for 72 hours.

SPAD

Soil plant atmospheric device (Model SPAD 502, Konica Minolto Sensing, Inc.) was used to measured specific leaf chlorophyll content

(SCMR). Most of the measurements were made during sunny days from 10 to 12:30 am India Standard Time.

Leaf dry weight (LDW, g/plant)

The weight of dried leaves (without stems) was measured after drying the leaves in hot air oven for 60^{o}C for 72 hr.

3.3.2 Phenotypic data analysis

3.3.2.1 Analysis of variance (ANOVA)

The analysis of variance for different drought related traits for the years 2008 and 2009 was performed to test the significance of difference between genotypes and pooled analysis of the data to assess the contribution of different sources to the total variation by following Panse and Sukhatme (1961). GenStat (12th Edition) was used to calculate general ANOVA using phenotyping data from 2008 and 2009 years.

3.3.2.2 Correlation coefficient (r)

Correlation coefficient (r) among the different traits was calculated by using GenStat (12th Edition) software.

3.3.2.3 Heritability (h^2)

Broad sense heritability sense was calculated as the ratio of genetic variance to the total phenotypic variance (Hanson et al., 1956) and expressed as a percentage.

$$h^2 = \frac{\sigma_g^2}{\sigma_p^2} \times 100$$

Where, σ_g^2 = Genotypic variance and σ_p^2 = Phenotypic variance

Heritability (broad sense) estimates were categorized into low (5-10%), medium (10-30%), high (30-60%) and very high (>60%).

3.4 Quantitative trait analysis

The candidate QTL regions for drought related traits were identified using two QTL mapping approaches: (i) internal mapping (IM) was used to identify the main-effect QTLs (M-QTLs) while, (ii) epistatic interaction analysis (EIA) was used to identify epistatic interactions between different QTL regions or epistatic QTL (E-QTL). However, the most likely location of QTLs and their genetic effects were detected by composite interval mapping (CIM; Zeng 1993 and 1994) using QTL Cartographer, v.2.5 (Wang et al., 2007). For each trait, the analysis was carried out using data from individual environment and/or from pooled data. CIM analysis is performed using the Model 6 after scanning the genetic map

and estimating the likelihood of a QTL and its corresponding effects at every 1 cM, while using significant marker cofactors to adjust the phenotypic effects associated with other positions in the genetic map. The number of marker cofactors for the background control was set by forward-backward stepwise regression. A window size of 10 cM was used, and therefore cofactors within 10 cM on either side of the QTL test site were not included in the QTL model. When separated by a minimum distance of 20 cM (Ungerer et al., 2002) two peaks on one chromosome were considered as two different QTL. Otherwise, the higher peak was chosen to more closely approximate the position of the QTL. The relative contribution of a genetic component (R^2 and h^2) was calculated as the proportion of the phenotypic variation explained (PVE). The QTLs explaining more than 20% phenotypic variation were considered as major QTLs. The additive effects and R^2 of the detected QTL were estimated by the Zmapqtl procedure inbuilt in QTL Cartographer.

QTLNetwork 2.0 programme based on mixed linear model (Yang et al., 2005) was used to identify epistatic QTLs (E-QTLs) conditioning the drought related traits. EIA analysis was carried out using Genotype Matrix Mapping (GMM) software ver. 2.1 (Isobe et al., 2007; http://www.kajusa.or.jp/GMM) that looks for interactions between different loci. Using GMM two and three loci interactions were tested. The search range was kept default set by the program, based on the input data and minimum number of corresponding samples was set to one. A

QTL Network 2.0 program, based on a mixed linear model (Yang et al., 2005) was also used to determine epistatic QTLs (E-QTL) conditioning drought related traits.

3.5 Construction of consensus map

In the present study, two consensus genetic linkage maps were constructed using the marker segregation data from

(i) three individual RIL mapping populations: TAG 24 × ICGV 86031, a previously studied mapping population by Varshney et al., 2009a and Ravi et al., 2011, ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76, the current mapping populations; and

(ii) ten RILs and one BC population from the international groundnut community.

3.5.1 Marker segregation data for eleven mapping populations

Details regarding the SSR marker segregation data obtained from collaborators on ten recombinant inbred lines (RILs) and one backcross (BC) mapping populations were given in Table 6. The populations, for which marker segregation data were assembled, for the convenience in the present study, have been referred as RIL1 to RIL11 and BC1.

Three mapping populations (RIL1, RIL2, and RIL3), developed at ICRISAT, segregated for drought tolerance related traits (Gautami et al.,

2012a), two mapping populations (RIL4 and RIL5), developed at UAS-Dharwad, segregated for foliar disease resistance (Sujay et al., 2012) and two populations (RIL9 and RIL10), developed at UGA and HAAS, segregated for tomato spotted wilt virus (TSWV). In the case of RIL6, RIL7 and RIL8, developed at GAAS, Yueyou 13 (Y13), a Spanish type with high yield was the common female parent. While the RIL6 segregates for oil content, the RIL7 segregates for protein content and the RIL8 segregates for resistance to *Aspergillus flavus* and aflatoxin contamination (Hong et al., 2010). The remaining BC1F1population was developed using a wild tetraploid AABB amphidiploid (*A. ipaënsis* KG30076 × *A. duranensis* V14167), called AiAd (Fávero et al., 2006) and a cultivated tetraploid AABB variety (Fleur 11).

Table 6: Source of marker data used for constructing the international reference consensus genetic map

Mapping populations/population type	Population size	no. of mapped loci	Trait segregation
TAG 24 × ICGV 86031 (RIL1): (F_8)	318	211	Drought tolerance related traits
ICGS 76 × CSMG 84-1 (RIL2): (F_9)	176	128	Drought tolerance related traits
ICGS 44 × ICGS 76 (RIL3): (F_8)	188	87	Drought tolerance related traits
TAG 24 × GPBD4 (RIL4): (F_7)	266	209	Late leaf spot and rust resistance
TG 26 × GPBD 4 (RIL5): (F_7)	146	209	Late leaf spot and rust resistance
Yueyou 13 × Zhen Zhuhei (RIL6): ($F_{4:6}$)	142	146	Protein content
Yueyou 13 × FU 95-5 (RIL7): ($F_{4:6}$)	84	124	Oil content
Yueyou 13 × J 11 (RIL8): ($F_{4:6}$)	136	64	Resistance to *Aspergillus flavus* and aflatoxin contamination
Tifrunner × GT-C20 (RIL9): ($F_{2:6}$)	248	261	Tomato spotted wilt virus (TSWV) resistance and several agronomic traits
SunOleic 97R × NC94022 (RIL10): ($F_{2:6}$)	352	197	Tomato spotted wilt virus (TSWV) resistance and several agronomic traits
Fleur11 × AiAd (synthetic amphidiploid) (BC1): BC_1F_1	88	339	Several agronomic traits

This population segregated for several agro-morphological and drought related traits (Foncéka et al., 2012).

The segregation data was obtained for 211 marker loci for RIL1 (Varshney et al., 2009a and Ravi et al., 2011), 128 and 87 markers loci for RIL2 and RIL3 (Gautami et al., 2012a), 209 marker loci each for RIL4 and RIL5 populations (Khedikar et al., 2010; Sarvamangla et al., 2011 and Sujay et al., 2012), and 146, 124 and 64 marker loci data for RIL6, RIL7 and RIL8, respectively (Hong et al., 2010). For RIL9 and RIL10, segregation data were obtained for 261 and 193 marker loci, respectively (Qin et al., 2012). The BC1 population contributed segregation data for maximum number (339) of marker loci (Foncéka et al., 2009). Genotyping data as mentioned above have been provided in Table S5 in Gautami et al., 2012b.

3.5.2 Construction of component genetic maps

The ten RIL and one BC1F1 mapping populations were selected based on the robustness, parental diversity and segregation for economically important agronomic traits. The entire data set comprising 1961 segregating markers obtained from all the eleven mapping populations were subjected to chi-square (x^2) tests to examine distortion from the expected 1:1 segregation using “Locus genotype frequency” function of JoinMap V 3.0 (Stam 1993) (Figure S1 in Gautami et al., 2012b).

Individual maps were reconstructed using MAPMAKER/EXP V 3.0 (Lander et al., 1987) and Kosambi mapping function (Kosambi 1944) to assemble linkage groups by maximum-likelihood for respective mapping populations. Marker clusters were identified using a minimum LOD score of 5.0 and a maximum recombination fraction (h) of 0.35. The most likely marker order within each linkage group was estimated by comparing the log-likelihood of the possible orders of markers using multipoint analysis "Compare" command. The "Try" command was also used to determine the most likely placement of the unlinked markers, and subsequent orders were tested using the "Ripple" command with "Error Detection" and "Use Three Points" options enabled. The distance between neighboring markers were calculated using the multipoint analysis implemented in the "Map" command.

3.5.3 Construction of an international reference consensus genetic map

An international reference consensus genetic map was constructed using the markers mapped in 10 RILs and one BC mapping populations. As groundnut is an allotetraploid, the homologous versus homeologous relationships between linkage groups of the different component maps need to be taken into consideration before constructing the consensus map. As a first step, the sub-genome origin of each linkage group of the different component maps were identified by considering a set of 58 single dose SSR markers (Gautami et al., 2012b, Table S1) that

consistently amplify only one locus either on the A or B sub-genomes. Secondly, all linkage groups belonging to the same homology group were then merged with the software MergeMap (Wu et al., 2008).

In MergeMap, a consensus marker order was calculated considering marker order from individual maps by processing the cluster sequentially. For each cluster, the defined orientation was identified by flipping some of the constituent linkage groups. During this process, the software flags the problematic markers and then produces a consensus directed acyclic graphs (DAGs) of clusters by resolving the conflicts. Briefly, the input to MergeMap was a set of DAGs from each individual map, and the output was a set of consensus DAGs that were further simplified and then each consensus DAG linearised using a mean distance approximation to give the final consensus map that was consistent with all or nearly all the markers in the individual input maps. For each cluster, three graphs in the dot format were produced and saved as lgx.dot, lgx_consensus.dot and lgx_linear.dot files respectively, where x is the id of the cluster. The further visualization of these graphs can be viewed with the GraphViz software tool. Among the three output graphs, the lgx.dot graph highlights the conflicts among the individual maps and also shows which marker occurrence is being deleted by the MergeMap. Further, the lgx_consensus.dot gives the view of the simplified consensus DAG while the lgx_linear.dot shows the final linearised consensus map. Therefore, the consensus map coordinates

from MergeMap were normalized to the arithmetic mean cM distance for each linkage group from the three individual maps. Finally, the graphic maps for each LG were generated using Mapchart 2.2 (Voorrips 2002).

For efficient visualization of individual and consensus maps and their comparison, mapping data were put in the comparative mapping programme (CMap version 1.01 http://www.gmod.org/cmap). This mapping programme helps in assessing the congruency of marker positions and order by making a pairwise comparison among different genetic maps. Taking into consideration of the common loci that exist among the various genetic maps, a highly conserved marker order was manifested. Subsequently, all the developed 11 individual genetic maps and the reference consensus map were aligned together using CMap (Gautami et al., 2012b).

4. RESULTS

4.1 Development of SSR markers from enriched genomic-DNA library

With the aim to develop new SSR markers in groundnut, a SSR-enriched library was constructed from the cultivated genotype ICGV 86031 using bead capture enrichment protocol by Glenn et al., (2005). The microsatellite library was enriched using two types of oligo sequences (AAG) 8, (CT) 10, (AG) 8, (TG) 12 and mixtures of these.

The quantity of DNA extracted from the leaf samples of ICGV 86031 was determined by measuring the absorbance at 260 nm and calculated to be ~196 ng/µl with an A260/A280 ratio of 1.72, indicating a high quality DNA. Digestion of genomic DNA was found complete, as indicated by a uniform smear visualized on a 1.5% agarose gel (Figure 1). As 2.0 µg of genomic DNA was for digestion in a final volume of 25µl, the concentration of digested DNA was approximately 80 ng/µl.

Ligation of ds linkers to size selected *Rsa*I digest was confirmed by PCR amplification with linker specific primer SuperSNX24. A thick smear was formed between 300bp-1kbp regions (Figure 1), when 2 µl of linker ligated DNA was taken as template compared to 4 µl of linker ligated DNA, which indicated the successful ligation of ds linkers to all size selected *Rsa*I digested fragments. Hybridization of DNA fragments with biotinylated repeat oligonucleotides was achieved by incubating the

mixture at the respective hybridization temperatures. The reaction was confirmed by PCR using linker SuperSNX24 primer. The smear detected between 300-500 bp regions indicated the successful hybridization of repeat containing DNA fragments (Figure 1).

The presence of both blue and white colonies indicated the presences of inserts in the vector. In the first instance, a total of 150 white colonies were screened for the presence of inserts using colony PCR. Among these, 96 colonies were found to be positive for inserts, as visualized on a 1.5% agarose gel. The amplification profiles of colony PCR results are shown in Figure 2.

A total of 96 SSR positive clones were selected for isolation of plasmid DNA. The insert size in these clones was estimated in the range of 100 to 300 bp. Sequencing of plasmid DNA for these 96 positive clones resulted in good quality sequences in 65 cases. The microsatellite sequence data for these 65 clones were submitted to Genbank under accession numbers FI857100 to FI857164 (Table 7) to make the sequences available to public and make use of this study for further developments of genetic markers.

Table 7: Features of new set of groundnut SSR markers developed

Primer ID	Genbank ID	SSR motif	Primer sequence (5'-3')	PCR profile	Expected size (bp)	Observed size (bp)	Status
ICGM01A03	**FI857101**	(CTAG)2	**F**:CGGAATTAGCCCTTGGTTTA **R**:CCCCCAACATTGGAGTTTAG	60-55	111	150	M
ICGM01A04b	**FI857102**	(TCCA)2	**F**:CATGGCACACTTTCCACATC **R:**GGAACGCTGAGGATCGATAA	60-55	104	118	P
ICGM01A05a	**FI857103**	(CTAG)2	**F**:CGAATTCGCCCTTGTTTAAG **R**:TCTTACAGGAGCAACAGCACA	60-55	179	146	P
ICGM01A05b	**FI857103**	(AGTG)2n(AGTG)2n(AGTG)2	**F:**TGCTTTCCTGGGATACCTTG **R:**GCGGCACTTATTAGCCGAT	60-55	263	240	P
ICGM01A06a	FI857104	(CTAG)2	**F**:GGAATTCGCCCTTGGTTTA **R**:ACATGGATGGAAAGCTCTGG	60-55	162	-	NA
ICGM01A06b	**FI857104**	(TCCA)2	**F**:AAATGGAGCGTTCTCAATGG **R**:GTCTTCTTTGGGCATTTGGA	60-55	183	200	M
ICGM01A06c	**FI857104**	(CACT)2	**F:**CCAGAGCTTTCCATCCATGT **R**:TGCTAAAATCCACTTTTGGG	60-55	248	262	M

ICGM01A08a	FI857105	(CTAG)2n(TTCT)2n(TTTC)2	**F:**CGAATTCGCCCTTGTTTAAG	60-55	219	-	NA
			R:TAGGAATGGCATGCAATGAA				
ICGM01A080b	FI857105	(CAGGG)2	**F:**TTCATTGCATGCCATTCCTA	60-55	133	-	NA
			R:TTTGAGGATGCCAAGGGTTA				
ICGM01A08c	**FI857105**	(AATG)2ctctac(TTCT)2	**F:**CAAGCATTAACCCTTGGCAT	60-55	185	200	M
			R:AGTTGTTGCTGTCGCCTTTT				
ICGM01A09	FI857106	(CTAG)2n(GATG)2	**F:**CGAATTCGCCCTTGTTTAAG	60-55	266	-	NA
			R:TTTGCAGAGCATGCTGTCTT				
ICGM01A10a	FI857107	(CTAG)2	**F:**CGAATTCGCCCTTGTTTAAG	60-55	156	-	NA
			R:CCACATTAGGTGTTCGGCTT				
ICGM01A10b	**FI857107**	(CCA)3	**F:**AAGCCGAACACCTAATGTGG	60-55	247	260	P
			R:CGAAATCATTATGCGTGGTG				
ICGM01A11a	FI857108	(CTAG)2n(CTCC)2n(TCTA)2	**F:**CGAATTCGCCCTTGTTTAAG	60-55	265	-	NA
		(CCT)3n(TTCTC)2n(CCTGAA)2					
			R:GCTTCTCTTAGCTTTCTCTTCA				
ICGM01A11b	**FI857108**	(TTACT)2n(AATT)2	**F**:ACAACAATGCCAAAATGCAG	60-55	242	256	P
			R:GGTTCAACCCATTGATCTTGA				
ICGM01A11c	**FI857108**	(AAAG)2	**F:**GCCATTGGAGCAAACAACTT	60-55	278	280	P

			R:TGCCAAGAAGGCATTGACTA				
ICGM01A12a	FI857109	(CTAG)2n(TGA)3	**F**:CGAATTCGCCCTTGTTTAAG	60-55	116	-	NA
			R:TCCACGGAGAGCTTCTTGTT				
ICGM01A12b	**FI857109**	(ATA)3n(CTAC)2n(TGT)3	**F:**CACTGGTTTCCAGTTTGGGT	60-55	225	110	P
			R:GCTAGTGCATGAACACACGG				
ICGM01A12c	**FI857109**	(CCAG)2n(AAGA)2	**F:**TAAGCTTGGAGCGGCTAGAG	60-55	172	187	P
			R:CCTTCTCGCTTTGGTTTCAC				
ICGM01B02	FI857111	(CTAG)2	**F:**GGAATTCGCCCTTGTTTAAG	60-55	139	-	NA
			R:TTCTAGCCATGCATAACCCC				
ICGM01B03a	**FI857112**	(CTAG)2	**F:**GGAATTCGCCCTTGGTTTA	60-55	140	175-200	M
			R:GCTACTGAGGAACAACGGGA				
ICGM01B03b	FI857112	(ACTG)2	**F:**CCCTTGGTTTAAGGCCTAGC	60-55	196	-	NA
			R:ACCTCCCCAAGTAGGATTCG				
ICGM01B04a	**FI857113**	(CTAG)2	**F:**GAAATTCGCCCTTGGTTTAAG	60-55	243	200-250	M
			R:GAATGAAACTTGGAAGGCCA				

F: Forward primer; R: Reverse primer; M: Monomorphic; P: Polymorphic; NA: Not Amplified

4.1.1 Mining for Simple Sequence Repeats (SSRs)

On mining 65 sequences with the *MISA* (MIcroSAtellite identification tool) perlscript represented 22878 kb, SSRs were found in only 64 sequences. In total, 186 SSR motifs were identified in the 64 sequences with a frequency of one SSR per 2.9 kb (Table 8).

Following the definition of Weber (1990), 61% of SSRs were identified to be perfect, imperfect 4% and compound repeats 35%. While twelve SSRs contained tetra-nucleotide repeats (52%), nine (39%) compound repeats, one tri-nucleotide repeat (4%) and one contained penta-nucleotide repeat (4%). In terms of abundance of a particular SSR, the CTAG repeat motif was found most abundant (30%).

Table 8: Summary of *MISA* search

Total number of sequences examined	65
Total size of examined sequences (kb)	22878
Total number of identified SSRs	186
Number of SSR containing sequences	64
Number of sequences containing more than 1 SSR	54
Number of SSRs present in compound formation	81

For developing the new markers based on the SSRs isolated, design of primer pairs was attempted for all 26 sequences. Primer pairs could be designed for only 23 sequences (35%) (Table 7). In the remaining cases, sequences flanking SSR regions were too short to design primers. The

newly developed SSR markers were designated as **ICGM** (ICRISAT Groundnut Microsatellite followed by clone ID). For testing the amplification of these new SSR markers, two groundnut genotypes ICGV 86031 and TAG 24 were used for PCR amplification. As a result, only 14 (61%) primer pairs amplified scorable amplicons that are bolded and highlighted in Table 7.

4.1.2 Polymorphism assessment of novel SSRs

Screening of 14 functional markers with 2 genotypes (ICGV 86031 and TAG 24) revealed polymorphism with 8 markers (57%) and 6 markers were found to be monomorphic. These 8 polymorphic markers amplified a total of 18 alleles with an average of 2.25 alleles per locus. The PIC (polymorphism information content) values ranged from 0.13 to 0.36 with an average of 0.25. However, highest PIC value was observed with primer pair ICGM01A11c (0.36), followed by ICGM01A05a and ICGM1A12c (0.35) and lowest value was observed with ICGM01A04b (0.13) (Table 9).

Table 9: Polymorphism assessment of novel set of SSR markers

Marker ID	No. of Alleles	PIC value
ICGM01A04b	2	0.13
ICGM01A05a	2	0.35
ICGM01A05b	2	0.26
ICGM01A10b	3	0.22
ICGM01A11b	2	0.20
ICGM01A11c	2	0.36
ICGM01A12b	2	0.16
ICGM01A12c	3	0.35
Mean	2.25	0.25

4.2 Construction of Genetic linkage maps

4.2.1 Screening of SSR markers on the parental genotypes of two mapping populations

The numbers of SSRs isolated in this study are low, and thus results based on this study may not be of much significance to speculate the frequency and abundance of SSRs in the groundnut genome and for further mapping and trait studies.

A total of 3215 SSR markers (both genomic and EST based) available in public domain and/or accessed from other sources/various collaborators were used to screen the polymorphism on the parental

genotypes of the two new mapping populations, ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 (Table 10) and were subsequently used in the present study for construction of two new linkage maps and further for identifying QTLs associated with drought related traits.

In summary, after screening a total of 3238 SSR markers on the parental genotypes of the two mapping populations, 128 polymorphic loci on ICGS 76 × CSMG 84-1 and 87 polymorphic loci on ICGS 44 × ICGS 76 were obtained.

Table 10: Details on markers used for developing two genetic maps

Markers source	No. of Markers screened	ICGS 76 × CSMG 84-1		ICGS 44 × ICGS 76	
		Polymorphic loci	Percent polymorphism	Polymorphic loci	Percent polymorphism
Ferguson et al., 2004	226	10	4.42	10	4.42
Mace et al., 2007	51	0	0	0	0
Cuc et al., 2008	104	13	12.5	10	9.61
Moretzsohn et al., 2004 ; 2005	338	33	9.76	19	5.62
Proite et al., 2007	53	4	7.55	1	1.89
He et al., 2003	59	8	13.56	7	11.86
S J Knapp et al., (Unpublished)	2217	55	2.62	37	1.76
Hopkins et al., 1999	26	1	3.85	0	0
Palmieri et al., 2002; 2005	18	0	0	0	0
Wang et al., 2007	123	4	3.25	3	2.44
Total	3215	128	3.98	87	2.71

4.2.2 Genotyping of polymorphic markers in two mapping populations

Out of 3238 (23 SSR markers from current study and 3215 from other sources), only 128 (3.9%) and 87 (2.7%) markers showed polymorphism between the parental genotypes of ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 respectively. These polymorphic markers were further used in genotyping the sets of 177 (ICGS 76 × CSMG 84-1) and 188 (ICGS 44 × ICGS 76) lines of the respective mapping populations. While genotyping the mapping population ICGS 76 × CSMG 84-1, segregation data were scored at two loci for two markers (GM2724 and GM2233). As a result, segregation data were obtained for a total of 128 loci using 126 polymorphic markers.

The segregation data obtained for the two mapping populations were used to construct the genetic linkage maps and for further trait studies respectively.

4.2.3 Linkage maps construction

The major objective of the present study is to develop two new intra-specific genetic linkage maps ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 for cultivated groundnut since there is no comprehensive linkage maps available. The linkage map was constructed using software

MAPMAKER/EXP v.3.0 (Lander et al., 1987 and Lincoln et al., 1993) multipoint analysis with minimum of LOD score 5 and maximum recombination fraction (h) of 0.35 were set as threshold for linkage group determination.

4.2.3.1 ICGS 76 × CSMG 84-1 linkage map

The chi-square (x^2) tests were conducted to test the Mendelian segregation ratio (expected 1:1) for the genotyping data obtained for 128 polymorphic loci for the mapping population ICGS 76 × CSMG 84-1. A total of 75 (58.6%) of the loci showed the expected 1:1 segregation pattern ($p<0.05$) and were used to establish the linkage groups. Due to paucity of polymorphic markers seen in cultivated groundnut, all the 128 markers were taken into consideration for constructing the linkage map. Using a minimum LOD score of 5.0 and a maximum recombination fraction (h) of 0.35, a total of 119 markers out of 128 polymorphic SSR loci were integrated onto 20 linkage groups (LGs) with a total map distance of 22082 cM (Figure 3A), while 9 markers remained unlinked. The number of markers mapped per linkage group ranged from two (LG3, LG16, and LG18) to ten (LG7). The LG8 of the genetic map spanned the largest genetic map distance of 278.5 cM, followed by LG5 and LG1 with 238.2 cM and 204 cM. The LG18 with 0.3 cM covered the least map distance among all the LGs. The inter-locus gap distance ranged from 0.3 cM (LG18) to 37.1 cM (LG15). The number of marker loci mapped along

with the respective map distance, map density and inter-marker distances are given in Table 11.

4.2.3.2 ICGS 44 × ICGS 76 linkage map

Similarly, in case of second mapping population also, chi-square (x^2) tests were conducted to test the Mendelian segregation ratio (expected 1:1) for the genotyping data obtained for 87 polymorphic loci. A total of 82 (94%) of the markers showed the expected 1:1 segregation pattern ($p<0.05$) and were used to establish the linkage groups. In this mapping population also, all the 87 polymorphic markers were taken into consideration for constructing the linkage map. Using a minimum LOD score of 5.0 and a maximum recombination fraction (h) of 0.35, a total of 82 markers out of 87 polymorphic SSR loci were integrated onto 15 LGs spanning a total map distance of 831.4 cM (Figure 3B), with 5 loci unlinked. The number of markers mapped per linkage group ranged from two (LG6, LG8, LG13 and LG14) to 14 (LG7) and with mean distance between the markers were 5.5 cM. LG7 of the genetic map spanned the highest genetic map distance of 109.4 cM and followed by the LG4 and LG3 with 102.8 cM and 93.1 cM, while LG14 with 6.3 cM covered the least map distance among all the LGs. The inter-locus gap distance ranged from 34.26 cM (LG4) to 4.83 cM (LG10). The number of marker loci mapped along with the respective map distance, map density and inter-locus gap distances are given in Table 11.

In summary, the two new linkage maps were constructed based on ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 mapping populations and were used in the identification and mapping of QTLs for drought related traits.

Table 11: Details on different linkage groups of two genetic linkage maps

Linkage group	Number of mapped loci		Map length (cM)		Map density (cM)		Inter-locus gap distance (cM)	
	ICGS 76 × CSMG 84-1	ICGS 44 × ICGS 76	ICGS 76 × CSMG 84-1	ICGS 44 × ICGS 76	ICGS 76 × CSMG 84-1	ICGS 44 × ICGS 76	ICGS 76 × CSMG 84-1	ICGS 44 × ICGS 76
LG1	7	4	204	19.5	29.14	4.88	34	6.5
LG2	5	5	63.5	52.4	12.7	10.48	15.87	13.1
LG3	2	13	14	93.1	7	7.16	14	7.76
LG4	8	4	188.6	102.8	23.57	25.7	26.94	34.26
LG5	10	11	238.2	57	23.82	5.18	26.46	5.7
LG6	8	2	80	10.8	10	5.4	11.42	10.8
LG7	10	14	138.8	109.4	13.88	7.81	15.42	8.41
LG8	9	2	278.5	9	30.94	4.5	34.81	9
LG9	6	5	125.4	85.8	20.9	17.16	25.08	21.45
LG10	4	4	88	14.5	22	3.63	29.33	4.83
LG11	8	4	147.4	89.7	18.42	22.43	21.05	29.9
LG12	4	5	88	85.9	22	17.18	29.33	21.47
LG13	8	2	96.9	8.4	12.11	4.2	13.84	8.4
LG14	2	2	2.6	6.3	1.3	3.15	2.6	6.3
LG15	3	5	74.2	86.8	24.73	17.36	37.1	21.7
LG16	5	-	47.7	-	9.54	-	11.92	-
LG17	7	-	83.5	-	11.92	-	13.91	-
LG18	2	-	0.3	-	0.15	-	0.3	-
LG19	4	-	58.7	-	14.67	-	19.56	-
LG20	7	-	189.9	-	27.12	-	31.65	-
Total	119	82	2,208	831.4	-	-	-	-
Mean	5.95	5.47	110.41	55.43	16.79	10.41	20.72	13.97

4.3 QTL mapping for drought related traits

4.3.1 Phenotyping data analyses

In order to identify quantitative trait loci (QTLs) for drought related traits in the two intra-specific mapping populations, the phenotyping was done for drought tolerance traits for two years (2008 and 2009) for the mapping population ICGS 76 × CSMG 84-1 and for the second mapping population ICGS 44 × ICGS 76 for the year 2008 with the collaboration from the Crop Physiology Division of ICRISAT, Patancheru. The phenotypic data was obtained for seven traits (Transpiration efficiency, Transpiration, Total dry weight, Shoot dry weight, leaf area and SPAD chlorophyll meter reading (SCMR) for the mapping population ICGS 76 × CSMG 84-1 and three traits (vegetative wt/plant, pod wt/plant and harvest index) for the mapping population ICGS 44 × ICGS 76. The phenotyping data of parents and RILs of the two mapping populations were subjected to analysis of variance (ANOVA). In the present study since the phenotypic data were obtained for one environment hence single environment ANOVA was conducted and showed moderate variations and low heritability for all the traits studied in both the mapping populations (Appendix 1 and Appendix 2). The effects of genotype x environment (GE) interactions, however was not observed to be significant. Similarly, the broad-sense heritability (h2b.s), grand mean, SED and LSD were observed to be moderate to low in both mapping populations (Appendix 1).

The detailed analysis of phenotypic data showed lower incidence of tolerance towards the female parent in both the mapping populations; however, the means of both the RILs were within the parental limits and all traits showed continuous distribution indicating their polygenic nature except SPADWW09 (Appendix 2) which was inclined towards drought tolerant parent (CSMG 84-1).

Therefore, QTL analysis based on genotyping data and phenotyping data on the two mapping populations (ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76) as mentioned above has been further discussed in detailed in the following sections.

4.4 QTL analyses for drought related traits

4.4.1 Identification of main effect QTLs (M-QTLs) using QTL Cartographer and QTL Network

4.4.1.1 ICGS 76 × CSMG 84-1 mapping population

Genotypic and phenotypic data (for two successive years 2008 and 2009) obtained on 176 lines of the mapping population ICGS 76 × CSMG 84-1, were analyzed for identification of the main effect QTLs (M-QTLs) using the software QTL Cartographer version 2.5 (Basten et al., 1994) and QTL Network programmes.

QTL Cartographer V2.5, following using composite interval mapping (CIM) method, detected a total of twenty-four M-QTLs in ICGS 76 × CSMG 84-1 mapping population. Out of twenty-four M-QTLs, six M-QTLs for TE, with phenotypic variance explained (PVE) ranging from 5.63 to 18.12%, nine M-QTLs for T with PVE 4.83 to 18.17%, three M-QTLs for TDW with PVE 6.62 to 22.39%, and five M-QTLs for SDW with PVE 5.03 to 22.09% were identified. However, for SPAD no M-QTL could be detected in the population (Table 12 and Figure 4A).

Similarly by using QTL Network programme, a total of seven drought related M-QTLs were identified in ICGS 76 × CSMG 84-1, mapping populations. Out of seven M-QTLs, three M-QTLs for TE with PVE ranging from 3.31 to 4.25% were detected along with single M-QTL each for T (3.21% PVE), TDW (6.04% PVE), SDW (5.50% PVE) and SPAD (2.51% PVE). The details regarding the position, markers associated and PVE for the drought related QTLs were given in Appendix Table 3A and Appendix Figure 4A.

Therefore, for the mapping population ICGS 76 × CSMG 84-1, a total of thirty one M-QTLs were detected by using QTL Cartographer and QTL Network programme for two successive years 2008 and 2009.

4.4.1.2 ICGS 44 X ICGS 76 mapping population

In the similar manner, genotypic and phenotypic data (for one season) obtained on 188 lines of the mapping population ICGS 44 ×

ICGS 76, were analyzed for identification of the main effect QTLs (M-QTLs) using the software QTL Cartographer version 2.5 (Basten et al., 1994) and QTL Network programmes.

As a result, by using the CIM method, three M-QTLs could be detected for HI measured under well-watered with PVE ranging from 6.39 to 40.10%. Similarly using QTL Network programme, two M-QTLs were identified. Single M-QTL each for HI (3.29% PVE) and Veg wt/pl (2.28% PVE) could be detected (Table 13 and Figure 4B). The details regarding the position, markers associated and PVE for the drought related QTLs are given in Appendix Table 3B and Appendix Figure 4B.

Therefore, for the mapping population ICGS 44 × ICGS 76, a total of 5 M-QTLs were detected by using QTL Cartographer and QTL Network programme for the year 2008.

Table 12: M-QTLs for drought tolerance identified by QTL Cartographer and QTLNetwork in ICGS 76 × CSMG 84-1

Traits	QTL Cartographer		QTLNetwork	
	No. of QTLs identified	**PVE (R^2 %)**	**No. of QTLs identified**	**PVE (R^2 %)**
Transpiration efficiency (TE)	6	5.63-18.12	3	3.31-4.75
Transpiration (T)	9	4.83-18.17	1	3.21
Total dry weight (TDW)	3	6.62-22.39	1	6.04
Shoot dry weight (ShDW)	5	5.03-22.09	1	5.5
SPAD chlorophyll meter readings (SCMR)	-	-	1	2.51

Table 13: M-QTLs for drought tolerance identified by QTL Cartographer and QTLNetwork in ICGS 44 × ICGS 76

Traits	QTL Cartographer		QTLNetwork	
	No. of QTLs identified	**PVE (R^2 %)**	**No. of QTLs identified**	**PVE (R^2 %)**
Harvest index (HI)	3	6.39-40.10	1	3.29
Vegetative weight/plant (Veg wt/pl)	-	-	1	2.28

4.4.2 Identification of epistatic QTLs (E-QTLs) using QTL Network and GMM in two mapping populations

Drought, a polygenic trait and involves complex interactions among several traits that contribute towards drought tolerance. Hence in the present study, a focus was made to identify epistatic QTLs (E-QTLs) that arose due to the interaction between the M-QTLs that are detected for different drought related traits using two programmes namely QTLNetwork and Genotype Matrix Mapping (GMM).

By using the QTLNetwork programme, a total of ten E-QTLs were detected in two mapping populations (ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76). Among these ten E-QTLs, two E-QTLs each were detected for TE with the PVE 2.44-2.91% and T with PVE 7.29-9.01%, while one E-QTL each for ShDW with PVE 7.64%, LA with PVE 11.09%, LD with PVE 7.65%, Total DW has PVE 8.89%, SPAD with PVE 4.77% and Veg wt/pl with PVE 7.66% (Table 14).

Table 14: Epistatic QTLs (E-QTLs) identified for drought component traits using QTLNetwork in two mapping populations

QTLi	Marker interval/ Associated marker	Position (cM)	QTLj	Position (cM)	Marker interval/ Associated marker	Additive effect	PVE (R^2%)
ICGS 76 × CSMG 84-1							
TEWW08_AhII	RNoX681-IPAHM596	12	TEWW08_AhXVIII	0	GM1515-GM2745	-0.1134	2.91
TEWW08_AhII	IPAHM254-PM499	45.6	TEWW08_AhXVIII	0	GM1515-GM2745	0.0764	2.44
TWW08_AhXI	Ah193-GM1992	19.6	TWW08_AhXV	62.8	GM1021-GM1570	-24.2	9.01
TWS08_AhV	pPGS Seq17F06-gi-832	218	TWS08_AhXVII	12.5	GM1135-GM1907	56.18	7.29
ShDWWS08_AhV	TC2D08	148.8	ShDWWS08_AhXVII	18.4	GM2638-GM2637	-0.215	7.64
LA WW09_AhIII	TC4G02-GM745	0	LA WW09_AhXIII	96.7	pPGS Seq19B01-TC3B05	-20.34	11.1
LDWWW09_AhVI	RN34H10-TC7C06	0	LDWWW09_AhIX	32	TC7E04-GM1949	1.132	7.65
Total DWWW09_AhIV	TC1D02- pPGS Seq9G05	158	Total DWWW09_AhXVI	19.3	GM1937-GM1990	1.736	8.89
SPAD WW09_AhXII	TC4G10-GM692	82.8	SCMR WW09_AhXV	57.8	GM1021-GM1570	2.1802	4.77
ICGS 44 × ICGS 76							
Veg Wt/pl Stress08_AhIII	TC11F12-GM2528	49	Veg Wt/pl Stress08_AhIV	21	TC3E05-GM2032	-3.6204	7.66

Epistatic interaction analysis (EIA) involves interaction of QTLs for two and three loci by using the GMM programme. Using this analysis, thirty seven E-QTLs were detected in ICGS 76 × CSMG 84-1 and twenty six E-QTLs in ICGS 44 × ICGS 76 mapping populations respectively (Table 15). For TE, 18 E-QTLs with PVE range 12.67-44.77%, three E-QTLs (15.8-56.56%) for T, six E-QTLs with PVE range 12.69-18.72%, for ShDW, two E-QTLs (29.99-30.87%) for LD, two E-QTLs (34.07-35.32%) for Total DW, three E-QTLs with PVE range 36.33-44.69% for SPAD, four E-QTLs (9.94-13.28%) for Veg wt/pl, ten E-QTLs with PVE range 23.69-36.02% for Pod wt/pl and twelve E-QTLs for HI (8.42-15.11%) were identified. The above E-QTLs involved three loci interactions, while only one E-QTL obtained for ShDW with PVE 14.59% involved two loci interactions (Gautami et al., 2012a; ESM 9). An example for marker-loci interaction for transpiration efficiency (TEWS) in the ICGS 76 × CSMG 84-1 population and pod weight/plant in ICGS 44 × ICGS 76 mapping population detected by using GMM software were represented in the Figure 5A and Figure 5B respectively. This defines the (a) graphical presentation of three-locus interactions and their positions on the genetic map. In this case, the linkage groups are arranged in tandem as a *circle* and *triangles* in the *circle* that represent the interaction of a three-locus combination and (b) graphical presentation of inter-acting loci and allele type by genotype matrices (GMs) and a genotype matrix network (GMN) that shows the significant

locus/allele combinations of three interacting loci. In this case the matrices and the connecting lines indicate GMs and GMNs.

From this study, it is noted that the number of E-QTLs identified and PVE observed by QTLNetwork were found to be very low when compared to the number of E-QTLs identified and PVE observed by GMM.

Table 15: Summary of epistatic interactions at three- and two-loci identified with Genotypic matrix mapping (GMM) in two mapping populations

Traits	Three-loci interactions		Two-loci interactions	
	No. of QTLs identified	**Phenotypic variation efficiency (R^2 %)**	**No. of QTLs identified**	**Phenotypic variation efficiency (R^2 %)**
ICGS 76 × CSMG 84-1 mapping population				
Transpiration efficiency (TE)	18	12.67-44.77	-	-
Transpiration (T)	3	15.8-56.56	-	-
Shoot dry weight (ShDW)	6	12.69-18.72	1	14.59
Leaf area (LA)	2	29.99- 30.87	-	-
Leaf dry weight (LDW)	2	29.99-30.87	-	-
Total dry weight (TDW)	2	34.07-35.32	-	-
SPAD chlorophyll meter readings (SCMR)	3	36.33-44.69	-	-
ICGS 44 × ICGS 76 mapping population				
Vegetative weight /plant (Wt/pl)	4	9.94-13.28	-	-
Pod wt/pl	10	23.69-36.02	-	-
Harvest index (HI)	12	8.42-15.11	-	-

PVE : Phenotypic variance explained

4.5 Construction of consensus genetic map using three ICRISAT RIL populations segregating for drought tolerance traits

Genetic maps developed for three populations (TAG 24 ×ICGV 86031 – the earlier map developed by Varshney et al., 2009a and Ravi et al., 2010 and ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 -two maps from the present study) segregating for drought tolerance traits were used for developing a consensus genetic map. However, all the three maps that were used in the present study were constructed with MAPMAKER/EXP V 3.0 (Lander et al., 1987) using the same mapping functions. Forty nine loci were common between genetic maps based on TAG 24 × ICGV 86031 and ICGS 76 × CSMG 84-1 populations, 33 loci between the genetic maps based on TAG 24 ×ICGV 86031 and ICGS 44 × ICGS 76 populations, 40 loci between genetic maps based on ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 populations, while 13 markers were common among all the three maps. By using these common markers across three maps, a consensus map was developed with MergeMap. In this context, the most-dense genetic map based on TAG 24 ×ICGV 86031 population, with maximum number of mapped loci (191) was taken as a framework map inorder to combine mapped marker loci from the other two maps based on ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 populations.

Integration of different LGs from individual maps to develop the consensus map is given in (Table 16).

Based on the common markers between the individual maps, it was observed that most of the linkage groups were consistent with few exceptions among the individual maps (Gautami et al., 2012a; ESM 11). Details on comparison of different LGs of the consensus map with the three individual maps can be referred from Gautami et al., 2012a; ESM 2.

In brief, the consensus map has 293 SSR loci integrated into 20 linkage groups, and spanning a map distance of 2841 cM (Table 17 and Figure 6). The map length in consensus map ranged from 6.3 cM (LG_AhXX) to 293.4 cM (LG_AhIV) with a mean of 142.0 cM. The number of markers per LG ranged from 2 (LG_AhXX) to 31 (LG_AhVII) (Table 2). The density of markers on the consensus map ranged from 3.15 cM (LG20) to 19.86 cM (LG12) and with an average marker density of 9.96 cM (Table 17). The inter-locus gap distance ranged from 5.68 cM (LG14) to 22.7 cM (LG20), with a mean distance of 11.08 cM per marker (Table 17). Out of 293 mapped loci, 65% (191 loci) marker intervals were less than 10 cM, 27% (79 loci) between 10-30 cM and 8% (23 loci) greater than 30 cM (Table 16).

Table 16: Summary of consensus map based on the three mapping populations (TAG 24 × ICGV 86031, ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76)

Homologous linkage group			Consensus linkage group	Mapped loci	Length (cM)
TAG 24 × ICGV 86031 (frame map)	**ICGS 76 × CSMG 84-1**	**ICGS 44 × ICGS 76**			
LG1	LG1		LG_AhI	17	178.21
LG2,LG20	LG2,LG14	LG2	LG_AhII	16	96.21
LG3,LG21	LG3,LG9,LG20	LG3	LG_AhIII	28	225.93
LG4	LG4	LG4	LG_AhIV	16	293.37
LG5	LG5	LG5,LG15	LG_AhV	28	233
LG6	LG6	LG6	LG_AhVI	16	157.95
LG7,LG19	LG7,LG17	LG7	LG_AhVII	31	198.09
LG8	-	-	LG_AhVIII	19	105.9
LG9	-	-	LG_AhIX	9	59.8
LG10	LG10,LG8	LG10,LG8	LG_AhX	16	256.17
LG11	LG11	-	LG_AhXI	15	135.74
LG12	LG12	-	LG_AhXII	8	158.9
LG13	LG13	LG13,LG12	LG_AhXIII	20	236.19
LG14	-	-	LG_AhXIV	11	110
LG15	-	-	LG_AhXV	5	67.6
LG16	LG16	LG9	LG_AhXVI	10	51.14
LG17	-	LG1	LG_AhXVII	7	44.1
LG18	LG18	-	LG_AhXVIII	11	102.6
LG22	LG19,LG15	LG11	LG_AhXIX	8	123.6
-	-	LG14	LG_AhXX	2	6.3
			Total	293	2840.80

Table 17: Features of consensus genetic map based on three RIL mapping populations

Features	Consensus map
Linkage groups (LGs)	20
Mapped loci	293
Max. markers/group	31
Min markers/group	2
Total map distance (cM)	2840.80
Average map density (cM)	9.96
Average inter-locus distance(cM)	11.08

4.5.1 Mapping M-QTLs and E-QTLs onto the consensus map

In addition to the 36 identified M-QTLs in ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76, a total of 117 M-QTLs detected in TAG 24 ×ICGV 86031 (Varshney et al., 2009a and Ravi et al., 2011) were also taken into consideration in this study inorder to place them onto newly developed consensus map. A total of 153 M-QTLs were identified from the three mapping populations for the drought related traits and were placed onto 16 of the 20 linkage groups, while no M-QTL could be mapped on 4 linkage groups (LG_AhII, LG_AhXV, LG_AhXVIII and LG_AhXX) on the newly developed consensus map. Ten or more than ten M-QTLs were mapped on LG_AhV (21), LG_AhVII (19), LG_AhXI (16), LG_AhX (14), LG_AhIV (12), LG_AhVIII (10), LG_AhXIII (10) and LG_AhXVII (10). While, less than ten M-QTLs were detected on LG_AhIX (8), LG_AhIII (6), LG_AhXIX (5), LG_AhVI (4), LG_AhXII (4), LG_AhI (3)

and LG_AhXVI (3) and single M-QTL was mapped on LG_AhXIV (Figure 6).

A total of 25 E-QTLs identified from the three mapping populations were distributed on 15 LGs of the newly developed consensus map. However, on five LGs no E-QTL could be found (LG_AhVIII, LG_AhX, LG_AhXV, LG_AhXVII and LG_AhXX). Five E-QTLs were detected in LG_AhIII, four in LG_AhVII, three each in LG_AhIX, LG_AhXI, LG_AhXIII and LG_AhXVI. Two E-QTLs each in LG_AhII, LG_AhIV, LG_AhV and LG_AhVI while one E-QTL each in LG_AhI, LG_AhXII, LG_AhXIV, LG_AhXVIII and LG_AhXIX.

A total of 178 QTLs (153 M-QTLs and 25 E-QTLs) that are associated with 25 drought related traits based on three mapping populations (TAG 24 ×ICGV 86031, ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76) were placed onto the newly developed consensus map. Several QTL clusters were found scattered on 14 LGs (LG_AhIII, LG_AhIV, LG_AhV, LG_AhVI, LG_AhVII, LG_AhVIII, LG_AhIX, LG_AhX, LG_AhXI, LG_AhXII, LG_AhXIII, LG_AhXVI, LG_AhXVII and LG_AhXIX) of the newly developed consensus map (Table 18). The region GM1949-TC7E04 (29.3 cM) on LG_AhIII harbours five QTLs for LDW, T, ShDW, TDW and TE traits. TC1D02-TC3E05 (31 cM) region. The pPGSeq19D06-PM418 (37.8 cM) region on LG_AhIV harboured seven

and six QTLs respectively for HaulmWt, SCMR, TDW, VegWt/pl, SLA, ShDW, canopy conductance (ISC) and T. LG_AhV had two clusters i.e., GM630-TC6E01 (39.2 cM) with 18 QTLs for PodWt, SeedWt, TDM, HaulmWt, TE, T and ISC while GM2584-pPGSSeq17F06 (74 cM) with five QTLs for HI, T and TDW. PM375-GM1867 (25.1 cM) on LG_AhVII harboured 16 QTLs for LA, SeedWt, PodWt, TDM, T, SLAHar, Biomass, ShDW, DWInc and TE. On LG_Ah VIII, nine QTLs for the traits SLA, Haulmwt, SCMR, ShDW and TE are harboured in the region pPGPSeq3A06-IPAHM406 (50.4 cM).

Similarly, five QTLs were present in pPGPSeq2B09-GM634 region (17.9 cM) on LG_AhIX for SCMR, ISC and LA traits. LG_AhXI harboured two clusters i.e., genomic region GM2350-TC4H02 (52.2 cM) with sixteen QTLs for the traits initial DW, SLA, T, TDM, HaulmWt, Delta13C, Biomass, SCMR and TEbis while GM1971b-TC4H02 region (48.9 cM) harboured twelve QTLs for T, HaulmWt, Biomass, SLA, SCMR, TE and TDM. Eight QTLs were found on LG_AhXIII in GM1911-PM733b region (28.3 cM) for the traits SLA, SCMR, T and ShDW. Nine QTLs were clustered on LG_AhXVI in GM2050-GM1494 region (39.0 cM) for HI, VegWt/pl, TDW, PodWt/pl and ShDW while seven QTLs were mapped on LG_AhXVII in the region GM1418-S11 (34.3 cM) for the traits HI, SLA, and SCMR. Similarly, genomic region GM1021-

GM1570 (21.3 cM) harboured 3 QTLs on LG_AhXIX for TDW, SCMR and T.

Table 18: QTL clusters identified for biomass, SCMR, yield and drought related traits

Cluster No.	LGs	Marker interval	Position (cM)	No. of QTLs	Traits	PVE (R^2 %)
1	LG_AhIII	GM1949-TC7E04	29.3	5	LDW, TW, ShDW, TDW,TE	3.64-22.39
2	LG_AhIV	pPGSeq19D06-PM418	37.8	6	SLA, ISC04, T, ShDW	3.91-22.24
3	LG_AhIV	TC1D02-TC3E05	31.0	7	Haulmwt, SCMR, TDW, Veg wt	5.06-33.36
4	LG_AhV	GM2584-pPGSSeq17F06	74.0	5	HI, T, Total DW	6.91-7.29
5	LG_AhV	GM630-TC6E01	39.2	18	T,TE, ShDW, Pod Wt, Seed Wt, Haulmwt, TDM, DWInc,	1.7-13.44
6	LG_AhVII	PM375-GM1867	25.1	16	LA, Seed wt, Pod wt, TDM, T, SLAHar, Biomass, ShDW, DWInc, TE	2.93-9.85
7	LG_AhVIII	pPGPSeq3A06-IPAHM406	50.4	9	SLA, Haulmwt, SCMR, ShDW, TE	3.90-9.87
8	LG_AhIX	pPGPSeq2B09-GM634	17.9	5	SCMR , ISC, LA	6.23-10.49
9	LG_AhX	GM2444-IPAHM165	25.5	4	SCMR	7.10-12.15
10	LG_AhX	TC9F04-TC4D09	16.5	7	SCMR, Pod wt, Haulmwt, LA, TE	4.67-7.74
11	LG_AhXI	GM2350-GM2724a	52.2	4	Initial DW, SLA, Delta13C04	4.19-20.32
12	LG_AhXI	GM1971b-TC4H02	48.9	12	T, Haulmwt, ISC, Biomass, SLA, SCMR, TE, TDM	3.44-12.60
13	LG_AhXIII	GM1911-PM733b	28.3	9	SLA, SCMR, T, ShDW	3.11-13.96
14	LG_AhXVI	GM2050-GM1494	39.0	6	HI, Veg wt, TDW, Podwt, ShDW	6.62-40.10
15	LG_AhXVII	GM1418-S11	34.3	9	SCMR,HI, SLA	5.41-19.53
16	LG_AhXIX	GM1021-GM1570	21.3	3	TDW, SCMR, T	2.51-9.87

PVE: Phenotypic variance explained

4.6 An international reference consensus genetic map for tetraploid groundnut

4.6.1 Marker segregation data

The segregation data for a total of 1961 markers were assembled for all the eleven mapping populations with markers ranging from 64 markers (RIL8) to 339 markers (BC1) per population respectively (Gautami et al., 2012b; Table S5). A chi-square test was conducted inorder to test the null hypothesis of segregation ratios of 1:1 for all the ten RIL mapping populations and 3:1 for the BC1F1 mapping population at the threshold of p=0.05. The component genetic maps exhibited variable degrees of segregation distortion ranging from 3.45% (RIL8) to 52.34% (RIL2) and the LG wise segregation pattern of markers in each mapping populations are shown in Gautami et al., 2012b; Figure S1.

4.6.2 Component genetic maps

All the component genetic maps that are used in constructing the reference consensus genetic map in the present study were constructed using MAPMAKER/EXP V 3.0 (Lander et al., 1987) using the Kosambi mapping function (Kosambi 1944) and can be visualized in CMap database at http://cmap.icrisat.ac.in/cmap/sm/gn/gautami/. The numbers of mapped loci ranged from 46 (RIL8) to 332 (BC1) per individual genetic maps. The map distance covered from 357.4 cM (RIL8)

to 2208.2 cM (RIL2) with a range of inter-locus gap distances from 2.5 cM (BC1) to 18.6 cM (RIL2) (Table 19).

Table 19: Features of the component and reference consensus genetic maps

Maps	Linkage groups	Mapped loci	Map length (cM)	Map density (cM)	Inter-locus gap distance (cM)	References
RIL1	22	191	1785.4	9.35	9.39	Varshney et al., 2009b; Ravi et al., 2011
RIL2	20	119	2208.2	18.56	18.71	Gautami et al., 2012
RIL3	15	82	831.4	10.14	10.26	Gautami et al., 2012
RIL4	20	188	1922.4	10.23	10.28	Khedikar et al., 2010; Sujay et al., 2012
RIL5	21	181	1963	10.85	10.91	Sarvamangala et al., 2011; Sujay et al., 2012
RIL6	19	133	793.1	6.01	6.05	Hong et al., 2010
RIL7	21	109	503.1	4.62	4.65	Hong et al., 2010
RIL8	13	46	357.4	7.76	7.94	Hong et al., 2010
RIL9	26	233	1304.9	5.6	5.62	Qin et al., 2012
RIL10	22	193	917.45	5.3	5.35	Qin et al., 2012
BC1	21	332	847.4	2.53	2.56	Foncéka et al., 2009
Reference consensus genetic map	20	897	3863.6	4.42	4.54	-

4.6.3 Construction of an international reference consensus map

Availability of adequate number of common markers and their distribution among eleven genetic maps facilitated integration of all the component genetic maps into one integrated or consensus map using MergeMap Software. While integrating component genetic maps, some discrepancies were observed for names of markers segregating in more than one mapping population i.e. anchor markers. However, to facilitate integration, uniformity in marker naming was maintained for all the markers in all the individual maps and in reference consensus map. For example, 'pPGPseq xxx' and pPGSseqxxx' were represented as 'seqxxx', and 'XIPxxx' as 'IPAHMxxx' to maintain the uniformity. Multiple segregating bands identified with one microsatellite primer pair in a mapping population have been usually indicated with lower case letters; for example two bands (loci) for IPAHM287 SSR marker (primer pair) became IPAHM287a and IPAHM287b. In addition, two CAPS (cleaved amplified polymorphic sequence) markers i.e., ahFAD2A and ahFAD2B were also mapped in the present consensus map (Gautami et al., 2012b; Table S1).

The genotyping data for 1961 markers obtained on eleven mapping populations were used for merging multiple genetic maps (Gautami et al., 2012b; Table S5). Building a consensus map is not possible without common or bridge loci present on each LG (Varshney et al., 2007b). A bridge marker was considered as such when it had an identical name

and should have a similar position in different mapping populations that are underpinned. Markers with the same name that mapped to different positions in different populations were not considered to be common or bridge markers.

However a minimum of three common markers per linkage group should be considered while, in the present study, at least one common marker per LG is also taken into consideration in some LGs because of low polymorphism observed (Gautami et al., 2012b; Table S5). Therefore, one should select appropriate common loci and compile a consensus map using a single pair of linked loci at a time only when they give similar recombination frequencies between individual populations.

Based on the number of common markers between individual genetic maps, most of the LGs were found to be consistent with few exceptions that can be visually assessed from http://cmap.icrisat.ac.in/cmap/sm/gn/gautami/ (Gautami et al., 2012b; Table S5). Out of 897 mapped markers, 542 markers were found to be unique i.e. mapped only in one mapping population, while the remaining 355 markers were common, i.e. they were mapped in at least two mapping populations (187 markers were common between two maps, 72 markers between three maps, 57 markers between four maps, 20 markers between 5 maps, between 6 maps 13 markers are common, 3 markers between 7 maps, 2 markers between 8 maps and one marker is common between 9 maps) and these markers served as anchor points or

bridge markers for the reference map construction (Table 20). The groupings of different LGs from individual genetic maps to develop the reference consensus map were given in Gautami et al., 2012b; Table S2. Therefore, in the consensus genetic map, a total of 355 (39.6%) markers are anchor markers present on all 20 linkage groups. The remaining 542 (60.4%) markers are unique to the individual genetic maps.

In the newly constructed reference consensus map, seventy homeologous loci were identified on “a” and “b” linkage groups (Figure 7), which facilitate the detection of ten homeologous pair and named from a1 to a10 and b1 to b10 based on the same loci detected on BC1 map (the framework map) developed by Foncéka et al., 2009. Out of these seventy homeologous loci, eleven loci were located between the group a1-b1 and a3-b3, eight loci between a2-b2 and a4-b4 and four loci between a9-b9. Except for the groups between a1-b1, a3-b3 and a4-b4 markers order and inter-loci map distance were well conserved between homeologous groups (Figure 7).

Table 20: Summary of common loci between genetic maps for different mapping populations

Mapping population	No. of mapped loci	No. of mapped loci used in consensus map	Number of markers in common with *n* other mapping populations								
			***n* = 0**	***n* =1**	***n* = 2**	***n* = 3**	***n* = 4**	***n* = 5**	***n* = 6**	***n* = 7**	***n* = 8**
RIL1	191	178	55	36	35	27	11	8	3	2	1
RIL2	119	81	39	12	7	11	5	4	2	1	0
RIL3	82	72	18	14	16	8	9	4	1	1	1
RIL4	188	176	19	67	28	32	16	9	2	2	1
RIL5	181	168	17	72	23	31	12	8	3	1	1
RIL6	133	114	27	28	18	17	10	9	3	1	1
RIL7	109	96	12	30	14	20	7	9	2	1	1
RIL8	46	36	10	4	7	6	4	3	0	1	1
RIL9	233	194	85	43	19	23	12	7	2	2	1
RIL10	193	145	51	40	19	18	7	7	1	2	0
BC1	332	324	209	28	30	35	7	10	2	2	1

In some cases, the same marker mapped single locus on different linkage groups in different mapping populations were not considered as the same loci and were considered as unique loci (with the same name) in the reference consensus genetic map. However, twenty nine (26%) primer pairs detected duplicated non-homeologous loci between linkage groups (e.g., seq12F07 detected two loci, one on a1 and one on a10; and IPAHM171 detected three loci on a6, b1 and b8) (Figure 7 and Gautami et al., 2012b; Table S1).

In brief, the reference consensus map is comprised of 895 SSR and 2 CAPS loci distributed over 20 LGs. Nomenclature of LGs in the reference consensus map was followed in the same way as in the framework map (BC1) developed by Foncéka et al., 2009. The map density in the reference consensus map ranged from 2.5 cM (a1) to 6.4 cM (a8) with an average of 4.3 cM per marker. The inter-locus gap distance ranged from 1.5 cM (a1) to 5.4 cM (a8), with a mean value of 4.5 cM per marker (Table 21). Among the 20 LGs, a1 possess maximum marker loci (70) followed by a3 (65), a5 (61) and b3 (60) respectively, while a2 and b9 have only 23 and 21 loci, respectively (Figure 7 and Table 21). The low number of SSR loci mapped on a2 and b10 may be because of the lack of polymorphism on these two LGs. For example, the consensus LG a2 is built with seven LGs of the different component genetic maps, among which four LGs have only two mapped loci. Therefore, for these small LGs additional markers are needed for increasing the map density. However, in the

consensus map, some gaps are observed at the distal ends of the a2, b2, a3, a5, b5, a8, a9 and b9 and a10 linkage groups. Of the 897 mapped loci, 290 loci (32%) of the marker intervals were less than 1 cM, while 369 loci (41%) marker intervals were between 1-5 cM, 143 loci (16%) 5-10 cM, 66 loci (7%) 10-20 cM, and 29 loci (3%) marker intervals were greater than 20 cM.

Table 21: Features of the reference consensus genetic map

LGs	No. of mapped markers	Map distance (cM)	Map density (cM)
a1	70	175.07	2.50
b1	51	300.44	5.89
a2	23	91.59	3.98
b2	30	162.81	5.43
a3	65	272.52	4.19
b3	60	282.02	4.70
a4	56	152.40	2.72
b4	42	177.66	4.23
a5	61	232.63	3.81
b5	33	167.28	5.07
a6	57	275.79	4.84
b6	24	99.03	4.13
a7	43	188.96	4.39
b7	34	114.37	3.36
a8	42	267.23	6.36
b8	47	144.34	3.07
a9	56	267.42	4.78
b9	21	125.86	5.99
a10	47	199.16	4.24
b10	35	166.99	4.77
Total	897	3863.57	-

4.6.4 Features of the reference consensus genetic map

SSR markers are the marker of choice in many breeding applications. Hence in the newly constructed reference map an attempt was made to understand the distribution of different SSR motifs as well as the polymorphism information content (PIC) values for these markers.

Out of 895 SSR loci integrated into the reference consensus map, information on repeat motifs was available for 788 SSR loci. Of the 788 SSRs, 612 SSR loci represent simple repeat motifs and 176 SSR loci contain compound repeat motifs. Among simple repeat motifs contained SSR loci, 47.6% (375 SSR loci) are comprised of di- (NN) repeats followed by 28.7% (226) tri-nucleotides (NNN) repeats. The longer repeat classes, i.e. tetra- (NNNN, 8 loci) and hexa-nucleotide (NNNNNN, 3 loci) represented 1.4% of the SSR loci (Gautami et al., 2012b; Table S3). In the case of the compound repeats containing SSR loci, 93 loci were comprised of NN repeats and the remaining 83 loci comprised with mixed repeats.

Of the 897 mapped marker loci, the information on PIC values was available for 526 SSR marker loci (Gautami et al., 2012b; Table S3). Based on genotypes surveyed in earlier studies, 144 marker loci have PIC value >0.50 while majority of the loci (181) have 0.31-0.40 PIC value (Gautami et al., 2012b; Figure S2). Average PIC values of individual LGs varied from 0.55 (a2) to 0.81 (a1).

In the present study, an attempt has been made to divide the genetic map into 20 cM long BINs for making the consensus map more informative. As a result, a total of 203 BINs were created ranging from 5 (a2 and b6) to 16 (b1) with an average of 4 per linkage group in the reference groundnut genetic map. These BINs carry 1 (a10_02, a10_08 and a10_09) to 20 (a10_04) with an average of 4.41 marker per BIN. While categorizing highly informative SSR markers based on available PIC values, 36 BINs have at least one marker that has >0.70 PIC value and 111 BINs carry at least one marker with >0.50 PIC value. A total of 166 BINs have the marker loci with <0.50 PIC value and 23 BINs do not have the information available on PIC values. A total of 13 BINs do not have any marker.

Finally, a total of 58 genome specific SSR markers were identified for deciphering the relationships between LGs of the different component maps. Therefore, these 58 genome specific SSR markers are of great interest for subgenome assignment of SSR loci in cultivated x cultivated mapping studies. Moreover these markers were also used in diversity analysis studies as they give access to the diversity at the diploid genome level allowing differentiating the structural heterozygosity linked to polyploidy from true heterozygosity.

4.6.5 Relationships between the reference genetic map and component genetic maps

A good congruence was developed between marker orders and positions among component maps and the reference consensus map except for a few exceptions (http://cmap.icrisat.ac.in/cmap/sm/gn/gautami/ and in Gautami et al., 2012b; Table S1). Comparison of a3 and b8 for all the component genetic maps and the reference consensus map, for example, has been shown in Figure 8.

4.6.6 Comparison of reference consensus map with diploid genetic maps of groundnut

The newly constructed international reference consensus genetic map was compared with the diploid genetic maps of groundnut (AA and BB genome maps published by (Moretzsohn et al., 2005 and Moretzsohn et al., 2009). The linkage groups of the reference consensus map in the present study are named similar to the linkage groups named in (Foncéka et al., 2009) (i.e. a1 to a10 and b1 to b10). While in the maps, of AA and BB genome the linkage groups were named as Group 1 to Group 11 and B1 to B10 respectively. Syntenic studies between the newly developed reference consensus genetic map and AA genome map assessed 68 common SSR markers and 43 between BB genome maps (Gautami et al., 2012b; Table S4). Overall, a good collinerity was observed for the corresponding LGs of the two diploid maps, with all the

ten LGs of the newly constructed reference consensus genetic map, with a few exceptions in the marker positions of some markers. The comparison of six LGs of the reference consensus genetic map with AA and BB maps are shown in the Figure 9. The number of common SSR markers per homologous linkage groups varied between 2 and 10 with AA map and with BB map between 1 and 9.

5. DISCUSSION

The present study deals with (i) development of SSR markers from an enriched genomic DNA library, (ii) screening of parental polymorphisms (ICGS 76, CSMG 84-1 and ICGS 44) and genotyping of the respective mapping populations using SSR markers obtained from various sources, (iii) integration of the polymorphic markers in two new genetic linkage maps, (iv) construction of dense consensus genetic map for cultivated groundnut, and (v) identification of QTLs using various linkage mapping approaches. These results have been discussed in detail in the context of available studies.

5.1 Developments of SSR Markers

SSR markers have become a widely used molecular marker for plant genetics and breeding applications in recent years. Despite the fact that hundreds of SSR markers have been isolated in groundnut using SSR-enriched library and BAC end sequence approaches (Hopkins et al., 1999; He et al., 2003; Ferguson et al., 2004; Moretzsohn et al., 2005; Mace et al., 2007; Cuc et al., 2008; Liang et al., 2009; Yuan et al., 2010; Bertioli et al., (unpublished) and Knapp et al., (unpublished)), the narrow genetic background of cultivated groundnut germplasm requires the development of SSR markers in larger numbers so that these can be used in groundnut genetics and breeding. With an objective of increasing the number of SSR markers, a new SSR enriched library for (AAG), (CT),

(AG) and (TG) repeat motifs was constructed from ICGV 86031, a cultivated groundnut genotype.

Construction and screening of partial genomic libraries and sequencing of SSR positive clones have been considered to be an effective method for SSR isolation (Rafalski et al., 1996). Enrichment of genomic DNA libraries for SSRs enhances the SSR isolation efficiency (Edwards et al., 1996). Out of the 65 positive clones, 29 clones had unique sequenced SSRs (44.6%) and primer pairs could be designed for twenty three SSR containing sequences (35.4%). Even though a lesser number of positive clones were used for SSR isolation in the present study, the results obtained are comparable to the earlier SSR isolation studies in groundnut. In the case of Hopkins et al., (1999), 66 (55.0%) out of the 120 sequenced "positive" clones had SSRs, but primer pairs could be designed for only 26 (21.7%). Gao et al., (2003) identified 14 (5.5%) unique SSR-containing sequences in 256 clones. Similarly He et al., (2003) sequenced 401 randomly picked clones resulting from AFLP pre-amplification based protocol, of which 83 (20.7%) were unique SSRs, and primer pairs were designed for 56 (14.0%). Ferguson et al., (2004) identified 348 (21.3%) SSRs by sequencing 1,627 clones, merely 226 (13.9%) primers could be designed. The SSR enrichment efficiency depends on many factors such as the choice of restriction enzyme used for library construction, the SSR probes used for enrichment and optimization of PCR profile and conditions. Therefore, the approach used

in the present study seems to be considerably efficient enrichment strategy for SSR isolation in groundnut. Moreover, in the present study it is observed that all the SSRs identified had different repeat motifs that were not totally complementary to the sequences of oligonucleotide probes used for library enrichment. In fact in earlier studies of Gimenes et al., (2007) also observed that 37% of SSRs isolated had a different repeat motif. The ATT repeat motif which is considered most abundant and highly informative in several legume species like soybean (Akkaya et al., 1992) and chickpea (Huttel et al., 1999) was not observed in the present study. These observations could be explained by the fact that the total number of SSR positive clones used in the present study is far lower than the earlier studies.

5.1.1 Polymorphism assessment of newly developed SSR markers

Polymorphism assessment of the 14 functional markers with two cultivated genotypes revealed polymorphism for eight markers (57%). Therefore, the percentage polymorphism observed in the present study is found to be higher than in other studies (He et al., (2003) (33%) and Ferguson et al., (2004) (28%). The average number of alleles (2.25) and PIC values (0.25) observed in the present study are comparatively lower than the earlier studies. While in case of Moretzsohn et al., (2004), the average number of alleles observed (5.33) and average PIC value (0.56), was observed to be higher and can be explained by the fact that

Moretzsohn et al., (2004) used a higher number of accessions (60) to test polymorphism compared to the present study (23 accessions).

5.2 Marker polymorphism from various sources and genetic maps

As a result of collaborative efforts made in last five years worldwide, nearly 4,000 SSRs were developed by the groundnut community. The parental genotypes of two mapping populations ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 have been screened with a total of 3,215 SSR markers. However, a very low level of polymorphism was observed (3.9% for ICGS 76 × CSMG 84-1 and 2.7% for ICGS 44 × ICGS 76). This may be attributed mainly to two reasons: (i) a narrow genetic diversity in the cultivated groundnut gene pool (Young et al., 1996; Varshney et al., 2009a; Hong et al., 2010; Ravi et al., 2011 and Sarvamangla et al., 2011), and (ii) highly conserved regions (cDNA) as the source of majority (94% EST derived) of SSR markers used (Varshney et al., 2005).

As groundnut is tetraploid crop species, 2 markers (GM2724 and GM2233) in ICGS 76 × CSMG 84-1 mapping population amplified more than one polymorphic locus. Amplification of more than one locus may be due to the polyploidy nature of the crop and has been reported in earlier studies (Hopkins et al., 1999; Krishna et al., 2004; Kottapalli et al., 2007; Varshney et al., 2009a; 2009b; Ravi et al., 2011 and Hong et al., 2010). This also suggests variability between genomes for these loci

and their potential use in comparative mapping between AA and BB genomes.

Recently, a few genetic maps based on RIL populations have been developed in cultivated groundnut (Varshney et al., 2009a; Hong et al., 2010; Khedikar et al., 2010; Ravi et al., 2011 and Sarvamangla et al., 2011) and only one population namely TAG 24 × ICGV 86031 has been used for developing the genetic map and QTL analysis for drought tolerance traits. In the present study, two RIL populations namely ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 segregating for drought tolerance were used to develop two new genetic maps. Together with the genetic map TAG 24 × ICGV 86031 developed from the earlier studies (Ravi et al., 2011) and the two new genetic maps developed in the present study, three genetic maps have become available for constructing a dense consensus map for drought tolerance traits.

LGs for each of the individual linkage maps were resolved without conflicting marker assignments using MAPMAKER/EXP V 3.0 and the parameters as described in earlier materials and methods. The stringent mapping parameters adopted for individual map construction resulted in 20 LGs for ICGS 76 × CSMG 84-1 and 15 LGs for ICGS 44 × ICGS 76 genetic maps. A non-uniform marker distribution was also observed in the maps which may be caused by (i) a non-random sampling of the genome, (ii) uneven distribution of the recombination rate along the LGs (Tanksley et al., 1992), and (iii) clustering tendency of some markers due

to their preferential targeting of some genomic regions (Castiglioni et al., 1999).

5.3 Identification of QTLs for drought related traits

5.3.1 M-QTLs for drought related traits

Drought tolerance is one of the major constraints for low productivity in groundnut and the challenge is to develop drought tolerant varieties. Drought tolerance is a complex trait and controlled by several genes with high environmental influence. Due to above reasons, selection based on phenotypic data is not reliable. To overcome this problem, molecular markers linked with drought tolerance as well as its component traits can be utilized to select drought tolerant breeding lines with higher precision and accuracy. In order to apply marker-assisted selection (MAS), QTLs/genes need to be identified. To identify QTLs for drought tolerance, an extensive study was done in TAG 24 × ICGV 86031. Varshney et al., 2009a and Ravi et al., 2011 identified several M-QTLs and a large number of E-QTLs for drought tolerance related traits in different seasons. Since the QTLs identified in the previous study revealed large number of QTLs with low phenotypic variance, it was imperative to understand complex nature of drought tolerance and its component traits as well as validating the QTLs detected in the previous studies. To validate the results in the previous study or to identify the new QTLs, a QTL analysis for drought tolerance related traits was under

taken on the two mapping populations ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76. Since the trait chosen in the present study is highly complex, more than one software analysis program was used to detect the QTLs. In addition, due to high environmental influence on this trait, two different programs, QTL Network and GMM, were used to study the environmental interactions between different loci. A total of 36 M-QTLs and 10 E-QTLs were identified for drought related traits in both the mapping populations. Interestingly, M-QTLs identified by QTL Cartographer were also identified by QTL Network. Also, the numbers of QTLs identified by QTL Network were comparatively less than those identified by QTL Cartographer. Similar results were also observed in earlier studies in TAG 24 × ICGV 86031 (Ravi et al., 2011). The M-QTLs identified for TE on LG _Ah VI, T on LG_Ah IX, and TDW on LG_Ah V and, ShDW on LG_Ah IX for RIL-2 were identified by both the programmes (Appendix 3A). The same QTLs identified by both the programmes may be considered to be more accurate/dependable QTLs than those detected by only one program. However, the value of such QTLs obtained can be confirmed only by assessing them in multi-location trials or in different genetic backgrounds. On the other hand, a single QTL each was identified by QTL Network for SCMR in ICGS 76 × CSMG 84-1 and Veg wt/pl and HI in ICGS 44 × ICGS 76 which may be considered false positives and hence, need further validation. In general, alleles with moderate to high additive effects were identified for majority

of the traits under study. However, alleles with medium additive effects were detected in the earlier study using TAG 24 × ICGV 86031 (Ravi et al., 2011). The combination of these favorable alleles derived from both the tolerant (positive additive effect) and the susceptible (negative effect) parents may confer more tolerance to drought. Alleles that improve the trait being derived from parents agronomically inferior have also been identified for several plant species (Xiao et al., 1998; Frary et al., 2004 and Yoon et al., 2006). Since QTLs with low to moderate phenotypic variation were detected similar to earlier study (Ravi et al., 2011), appropriate molecular breeding methods such as marker-assisted recurrent selection (MARS) should be deployed.

5.3.2 E-QTLs for drought related traits

Majority of the studies suggested that quantitative variation is determined by few QTLs with a relatively large effect and large number of QTLs with smaller effects. Apart from main effect QTLs (M-QTLs), epistatic QTLs (E-QTLs) which arise due to interactions of different loci in a particular cross, also plays significant role towards controlling a particular trait (Jannink 2007 and Isobe et al., 2007). In the present study, EIA undertaken with GMM and QTLNetwork revealed several epistatic QTLs. GMM could detect a total of 63 interactions among three-loci and only one interaction between two-loci for different drought component traits.

Results obtained in the present study showed several epistatic interactions for TE (18 interactions) followed by HI (12 interactions), pod weight (10 interactions) and ShDW (7 interactions). As expected, the numbers of E-QTLs identified by GMM were more than the M-effect QTLs. Furthermore, the PVE of these QTL interactions was comparatively higher than the M-effect QTLs identified by QTL Cartographer. Similar results were also observed in the earlier studies for TAG 24 × ICGV 86031 in groundnut (Ravi et al., 2011) and for plant persistency in rye (Klimenko et al., 2010). This clearly indicates the importance of these interactions for a complex trait such as drought tolerance that is highly influenced by the environment. Hence, apart from considering Main-effect QTLs (which are less in number), selection of these interacting loci (E-QTLs) while improving drought tolerance is a must. QTLNetwork identified less number of epistatic QTLs for TE (3 QTLs) and ShDW (2 QTLs), while no QTL was detected for Veg wt/pl and HI in both the populations. The variation in detecting QTLs by different programs may be due to different algorithms used by GMM and QTLNetwork. GMM is capable of comparing multiple QTL interactions at the same time, which would make it more advantageous in identifying epistatic interactions as compared to QTLNetwork.

A considerable number of QTLs were identified in the present study for drought related traits with less phenotypic variation for different drought component traits similar to earlier study for TAG 24 × ICGV

86031 (Ravi et al., 2011). Therefore, the results observed from the earlier study for TAG 24 × ICGV 86031 L-1 and the present studies made for ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76; suggest that drought tolerance is governed by a large number of M-QTLs and E-QTLs each with a small phenotypic variation. Stacking of all these minor QTLs is not possible through marker-assisted backcrossing (MABC), as MABC can only be successful in transferring a few QTLs from one genetic background to another (Ribaut et al., 2010). Therefore, alternative and more efficient approaches that allow selection for several QTLs with small effects (Ribaut and Ragot 2007; Bernardo 2008 and Varshney and Dubey 2009) such as MARS or GS will be more useful for the improvement of drought tolerance in groundnut.

5.4 Consensus map for cultivated groundnut developed from three RIL mapping populations

Availability of a high density genetic map in a crop species is must to initiate genetical and molecular breeding activities. The alternate way is to map several marker loci mapped in different partial individual genetic maps through development of consensus map. Development of a consensus map is very useful in such crops like groundnut where a high density genetic map is not available. To achieve this, two individual maps developed from ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 in the current study along with the map developed from TAG 24 × ICGV 86031

populations (Ravi et al., 2011) were used for development of consensus map. The newly developed consensus map consists of 293 SSR loci distributed over 20 linkage groups. Fourteen out of 20 linkage groups possessed more than 10 markers. LG7 was the highest dense linkage group with 31 markers followed by LG3 and LG5 with 28 markers. LG20 and LG14 were very small with only two and five markers respectively (Figure 7 and Gautami et al., 2012a; ESM 11). These small linkage groups could be artificial and additional genetic markers are needed to improve the linkage analysis. The observed total map distance of the newly developed consensus map (2840.8 cM) was almost equal to the expected genome length of groundnut genome (2800 Mb/1C) representing the random distribution of SSR markers across the whole genome.

The markers placed on the consensus map were consistent with respect to order on the LGs with the map developed earlier by Ravi et al., 2011 with few minor differences. This conservative property of the cultivated genome makes the consensus map reliable and successful. The consensus map removes large gaps present in the individual maps except in LGs where the poor coverage might be due to lack of polymorphism for markers screened in those regions. Therefore, this microsatellite dense tetraploid consensus map provides a means to consolidate the information of the marker order and position from three different individual maps and also lays an excellent platform for further

QTL mapping of economically important traits. Moreover the newly developed consensus map shows the position of microsatellites at an average density of 9.96 cM per marker that makes the map useful for several molecular breeding activities and physical mapping.

To the best of our knowledge, this newly developed consensus map is the first SSR rich-dense consensus map for cultivated groundnut. Similar efforts were done by Hong et al. 2010 and they developed a composite map for tetraploid groundnut with 175 loci using three mapping populations with a total map distance of 885.4 cM. For comparable areas, the size of the consensus map developed in the present study was consistently larger than the composite map developed by Hong et al., 2010, which may be due to use of different programs for development of consensus map. Moreover, this consensus map was more dense and accurate because all the maps were developed at the same centre *i.e.,* ICRISAT, India and by using the same set of SSR markers (3,221) for studying marker polymorphism among the parental genotypes. Furthermore, the present consensus map has the merit of being the first SSR-based consensus map for drought related traits as all the three populations were segregating for drought related traits which allowed us to place all the mapped QTLs onto consensus map. The present consensus map possesses a large number of markers spanning the full genome that can be used to genotype individuals for detecting recombinants, fixing loci, restoring a recurrent genetic background,

assembling complex genotypes in complex crosses (Gupta et al., 1999 and Somers et al., 2004), comparative mapping and map-based cloning. Future prospects include adding more microsatellite markers, SNP-based and DArT markers to the consensus map, thus producing a highly saturated map and which helps for a thorough alignment to the physical map of groundnut as well as implementation of the map in several molecular breeding activities in groundnut.

5.5 Candidate genomic regions for drought tolerance on consensus map

All the three mapping populations (TAG 24 × ICGV 86031 from earlier study, ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 from the present study segregated for drought related traits. Hence, it was worthwhile to place all drought related QTLs identified in the individual maps onto the newly developed consensus map. This helped enhance the understanding about the distribution of QTLs related to drought tolerance and a few yield related traits on the cultivated groundnut genome. A total of 178 QTLs (153 M-QTLs and 25 E-QTLs) associated with 25 drought and yield related traits were distributed on 14 LGs.

Interestingly, several of these QTLs were found clustered at 16 specific genomic regions (Table 18). The genomic region bracketed by PM375-GM1867 (23.9 cM) on LG_AhVII possessed 16 QTLs for traits such as LA, Seed wt, PodWt, TDM, T, SLAHar, Biomass, ShDW, DWInc,

ShWt and TE. Similarly the genomic region GM630-TC6E01 (39.2 cM) on LG_Ah V contained 18 QTLs for the traits T, TE, ShDW, Pod Wt/pl, Seed Wt, HaulmWt, TDM and DWInc. These regions have QTLs for yield and yield component from the field experiment under mild stress with co-mapping of seed weight QTLs under WW and WS conditions, and also co-mapping of growth attribute from other phenotyping experiments. An added value of that region was the co-mapping of TE QTLs from earlier experiments, which fits the hypothesis that TE would contribute under situations of mild water stress (Ratnakumar and Vadez 2011). The GM1971b-TC4H02 region on LG_AhXI (48.9 cM) harbored 12 QTLs for T, HaulmWt, carbon discrimination ratio, biomass, SLA, SPAD, TE and TDM, and is interpreted as being a "growth" region. Interestingly, three out of these four biomass clusters also harbored yield and yield component QTL, which is explained by the mild stress in that field experiment. Our interpretation is that under such conditions genotypes favoring plant growth are likely to achieve higher yields. Similarly, four clusters harbored a total of 26 QTLs for SCMR on LG_AhIX (17.9 cM, pPGPSeq2B09-GM634), LG_AhX (25.5 cM, GM2444-IPAHM165), LG_AhXIII (28.3 cM, GM1911-PM733b) and LG_AhXVII (34.3 cM, GM1418-S11). Two clusters were also harboring QTLs related to leaf characteristics, include leaf area, leaf thickness but also leaf conductance and plant transpiration on LG_Ah IV and LG_Ah VII. These clusters are particularly important since leaf conductance and

transpiration condition, the rate at which plant would use a limited water resource and can be important alleles to include in a breeding scheme targeting relatively severe stress conditions or to exclude in a breeding scheme targeting relatively moderate stress conditions. The phenotypic variance for biomass related traits and SCMR ranged from 2.93-22.39 and 3.11-19.53 respectively. In pPGPSeq2B09-GM634 region QTLs are harboring for SCMR trait, but also for canopy conductance (ISC) and leaf area (LA). Our prediction on that QTL is of a region controlling leaf N (nitrogen) status in conjunction with the leaf expansion processes (more leaf expansion leading to less N cm^{-2} and then lower SPAD reading), both being then indirectly involved in setting the level of canopy conductance, itself likely to play an important role for specific drought conditions. The region on LG_AhXIII in GM1911-PM733b with six QTLs for the traits SLA, SPAD, T, and ShDW is interpreted as another region controlling the N status of the plant.

The region on LG_AhXVI at GM2050-GM1494 (39 cM) with six clustered QTLs for HI, Veg wt, TDW, Pod wt and ShDW traits was particularly interesting because it harbored HI QTL from ICGS 44 × ICGS 76, dry weight QTLs from ICGS 76 × CSMG 84-1 and yield and shoot QTL from TAG 24 × ICGV 86031. As mentioned above, a recent finding indicated that lines having lower canopy could be better adapted to intermittent stress conditions (Ratnakumar and Vadez 2011) by limiting

the effect of stress on reproduction, thereby the link with the HI. Seven QTLs were mapped on LG_AhXVII in the region GM1418-S11 (34.3 cM) for the traits HI, SLA, and SPAD and GM1021-GM1570 region (21.3 cM) harboured 3 QTLs on LG_AhXIX for TDW, SPAD and T traits. Apart from above, three clusters harbored a total of 23 QTLs were observed for drought related traits on LG_AhIV (37.8 cM, pPGSeq19D06- PM418), and (31.0 cM, TC1D02-TC3E05) and LG_AhVIII (50.4 cM, pPGPSeq3A06-IPAHM406). The traits mapped under this region showed phenotypic variance of 3.91-33.36%. The TC1D02-TC3E05 region harboured QTL for SPAD reading from TAG 24 × ICGV 86031, which can be taken as a proxy for nitrogen status. It was interesting to find that this same locus also harboured QTL for biomass parameters from ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76. Another region pPGSeq19D06-PM418 on LG_AhIV harbored QTL for SLA, which represents processes of leaf thickening, but also QTL for LA and transpiration rate (ISC04, in g water used cm^{-2} h^{-1}), which represents leaf conductance. Depending on the stress intensity, leaf conductance is important for drought adaptation (Kholova et al., 2010a and Zaman-Allah et al., 2011a; b), as it drives plant transpiration and depends on the degree of leaf expansion (leaf area) and thickening (SLA). Here also, that region appeared to control similar traits, since from the earlier study of Ravi et al., 2011 a QTL for transpiration (T) was also found in the same region.

Two clusters for yield related traits with 25 QTLs on LG_AhV (39.2 cM, GM630-TC6E01) and LG_AhX (16.5 cM, TC9F04-TC4D09) were also observed with phenotypic variance ranging from 1.7-13.44%. The region at GM630-TC6E01 (39.2 cM) on LG_AhV with 18 QTLs for the traits such as pod wt, seed wt, TDM, HaulmWt and T, were identified from different phenotyping experiments. The fact that yield and component QTL co-map with shoot biomass and transpiration QTL from other experiments agrees with the fact that the stress effect in the field experiment of TAG 24 × ICGV 86031 was very mild (200 mm of rain received during the stress period) and therefore traits related to growth were those most related to high yield performance. This was also confirmed by the fact that pod and seed weight QTLs under WW and WS conditions co-mapped. Although the region between GM2584 and pPGSSeq17F06 (74 cM) is relatively large, it was also interesting since it harbors HI QTL from ICGS 44 × ICGS 76 while T and shoot biomass QTL from ICGS 76 × CSMG 84-1. The relationship between the two types of traits is in the recent finding that genotypes with smaller canopy can fare better under intermittent drought stress (Ratnakumar and Vadez 2011). Such clusters can be considered as hotspot genomic regions for further study and utilization in improving crop productivity through introgression of these genomic regions. Further studies are required to dissect these regions to identify tightly linked markers for the QTLs with high phenotypic variation as well as for the introgression either in the same genetic

background for the improvement of crop productivity under water stressed conditions.

Thus, the present study revealed a total of 16 genomic regions with 137 QTLs related to biomass, yield and drought component traits possessing several candidate genes for further exploration and utilization for QTL pyramiding and cloning. For the complex traits such as biomass, yield and drought which are controlled by several genes, many QTLs with low to moderate phenotypic variance are reported and can only be tackled through modern breeding approaches such as marker-assisted recurrent selection (MARS) or genomic selection (GS) (Ribaut and Ragot 2007; Bernardo 2008 and Varshney and Dubey 2009). Since, majority of the components of biomass, yield and drought are correlated, clustering of QTLs controlling different components at specific genomic region has much significance and of practical use for crop improvement for these traits. Therefore, some key genomic regions, containing QTLs for aforementioned traits may be harnessed through marker-assisted selection (MAS) approach to enhance drought tolerance in the elite cultivars/varieties.

5.6 An international reference microsatellite consensus map

SSR markers have already proven to be preferrrable over other molecular markers to undertake basic and applied research because (i)

they are the most co-dominant and easily transferable markers, (ii) display a random distribution across genome and (iii) have high levels of intraspecific and intra population allele polymorphism. Today, several high density microsatellite maps are available in rice (Mc Couch et al., 2002), maize (Sharopova et al., 2002), wheat (Somers et al., 2004) and barley (Varshney et al., 2007b and Marcel et al., 2007).

In recent years, significant progress has been made in developing high throughput genotyping and various linkage mapping technologies have asses in placing a large number of marker loci on genetic maps in several crop species (Somers et al., 2004; Langridge et al., 1995; Mace et al., 2009 and Hyten et al., 2012). Therefore with the advent of these technologies, the number of marker loci placed on genetic maps has increased exponentially.

Until recently, groundnut was suffering from a dearth of molecular markers. Extensive collaborative efforts made in last five years worldwide, resulted in development of around 5000 SSRs including both genomic and EST–SSRs. These large collections of microsatellites have been extensively used for estimation of genetic diversity in the gene pool and mapped in different mapping populations segregating for various traits. However, several factors such as the polyploidy nature, large genome size and limited DNA polymorphism, did not allow all the possible SSR markers to map onto a single genetic map. As an alternative, synthesis of an integrated or consensus genetic map provides

an opportunity to avail the saturated genetic maps by merging all the existing genetic maps by exploiting common bridging markers.

In the case of groundnut, a tetraploid crop, genetic mapping efforts have been initiated recently and few genetic maps with 46 to 332 marker loci have been developed (Pandey et al., 2012). To enhance the marker density, a few consensus maps have also been developed using the mapping data from 2-3 mapping populations although the number of mapped marker loci on these maps is no more than 324 loci. The major objective of the present study was the construction of a highly dense map for cultivated groundnut by using a consensus mapping approach. Rather than developing a high-density map with a fine order of markers, our purpose was to develop a framework consensus map with a general order of markers that could be used as a reference map by the international groundnut community for precise genetic studies.

Availability of dense genetic maps have played an important role in helping many plant geneticists and breeders for (i) identifying the molecular markers closely linked with genes of interest, (ii) genome wide association analysis, (iii) understanding various trait mapping of interest (Varshney et al., 2006), (iv) map-based cloning, and (v) initiating genome sequencing projects. A variety of integrated or consensus genetic maps using segregation data from multiple mapping populations have been reported in several crop species, e.g. barley (Langridge et al., 1995; Qi et

al., 1996 and Karakousis et al., 2003), wheat (Somer et al., 2004), and pearl millet (Qi et al., 2004).

Dense genetic maps can be developed mainly by using two approaches to: (i) map large number of marker loci using highly diverse population, and (b) merge the available genetic maps using common markers that were mapped across the populations. The first approach is however quite challenging and laborious, but is precise. Therefore, the second approach was used in the present study. In this context, segregation data for a total of 1961 marker loci generated for 11 (10 RIL and 1 BC) populations were assembled from different institutes. As a first step, component genetic maps were developed for all 11 populations. While comparing the component genetics maps developed in this study with the ones published by the source laboratory, all mapped marker loci could not be integrated into component genetic maps in this study. One of the main reasons for this may be use of a stringent and common approach to develop all the individual genetic maps.

Constructing a consensus map is not possible without bridge markers present across the individual maps on each LG (Varshney et al., 2007b). Bridge markers are those that have an identical name and have a similar map position in different mapping populations. While markers that have the same name but are mapped at different positions in different populations were not considered to be bridge markers. However, in constructing a consensus map, a minimum of three common markers

per linkage group should be taken into consideration but, in the present study, at least one common marker per linkage group was also taken into consideration in exceptional cases because of lower number of markers integrated in some LGs.

In the present study during the process of merging the individual maps for construction of reference consensus map, a major emphasis was given towards obtaining a general order and distance because of the fact that cultivated groundnut is an allotetraploid with a large genome size (2800 Mb/C) and has a narrow genetic base with very low DNA polymorphism. Therefore, slight discrepancies in marker orders as well as positions were observed in a few LGs. (http://cmap.icrisat.ac.in/cmap/sm/gn/gautami/ and Gautami et al., 2012b; Table S1). These discrepancies among different component genetic maps may be due to (i) different sizes of mapping population used, (ii) different mapping populations types used, and (iii) occurrence of genotyping errors (Feltus et al., 2006). Sometimes, small differences might also be due to mapping imprecision rather than real rearrangements.

Therefore, the newly constructed reference consensus map integrated a total of 897 loci (895 SSR and 2 CAPS) with a mean map density of 4.3 cM. This map is considered to be the densest SSR based map so far developed in groundnut community and therefore is proposed as, “an international reference consensus map”. Despite the dense placing of

markers on various LGs, gaps were observed at the distal ends of some LGs (e.g. a2, b2, a3, a5, b5, a8, a9, b9 and a10). Two main reasons for these are (i) high recombination prone regions and such cases were also observed in other mapping studies (Varshney et al., 2009a; Ravi et al., 2011; Hong et al., 2010; Gautami et al., 2012a; Sujay et al., 2012 and Qin et al., 2012), and (ii) under-representation or deficiency of marker loci from these genomic regions in the dataset used for developing the reference consensus map (Varshney et al., 2007b; Varshney et al., 2009a and Ravi et al., 2011).

In present mapping protocol, both the homologous and homeologous relationships of the LGs were taken into consideration to generate the reference consensus map. Therefore, the marker orders are consistent in most of the linkage groups with few exceptions where the marker orders are in opposite orientation. Moreover, maximum markers were mapped onto the consensus map in their original orders similar to the individual maps, but small number of markers were integrated with slight order changes, which may be caused by the computational variation resulting from (i) occurrence of recombination heterogeneity between different populations, (ii) existence of weak linkages in the various LGs of maps, (iii) missing or poor quality data, (iv) using of different mapping algorithms (programmes) while constructing the individuals and the consensus maps and, (v) using different thresholds statistics for creating the consensus map and the individual maps (Gustafson et al., 2009).

Despite the precautionary measures taken in preparing this consensus map, there still could be some disagreement in order of closely linked markers between the individual maps within some LGs intervals. The disagreement may be due to the quality as well as the quantity and distribution along the LGs of the bridge markers used for preparing the consensus map, or to mapping populations, algorithm and stringency criteria of computer programme (Varshney et al., 2007b; Hong et al., 2010 and Gustafson et al., 2009). For example, the mapping populations from which the consensus map was prepared have different numbers and different types of progeny lines. In smaller populations, the chance that informative recombinant progeny lines are present in the population to accurately position markers is lower than in larger populations (Varshney et al., 2007b and Gustafson et al., 2009). Further, even for a given mapping population, different markers were mapped using different subsets of progeny lines in different laboratories. Therefore, the users of the consensus SSR map must consider that the marker order is conditioned by several factors like the progeny lines used and the position of cross over along chromosome within the progeny lines. The precise fine markers order may differ slightly in other populations and users may need to verify the order of closely linked markers in their mapping and breeding populations.

In the newly constructed reference consensus map dinucleotide microsatellites (48%) and trinucleotide microsatellites (29%), are present in higher proportions than the compound (22%) and other types of SSRs (1%). The reason may be that the majority of SSR loci integrated were derived from the genomic DNA libraries that had been enriched for dinucleotide and trinucleotide SSR probes (Pandey et al., 2011 and 2012). Therefore, the availability of different types of SSR loci in a given region will facilitate selection of the SSR repeat motifs of choice in a particular region of interest. Availability of the primer sequences for a total of 885 SSR loci, approximately 90% of all loci integrated in the consensus map, at one place should accelerate the use of SSR markers in groundnut breeding activities. Moreover, the genotyping data has been made available for all the mapped SSR loci in the present study and this will allow the groundnut community to extend the dataset with their own data set further.

Another most important salient feature of the newly constructed reference consensus map is the defining of the 203 BINs in the groundnut genetic map. The marker loci present in these BINs are associated with the PIC values information. One marker from each of such BIN with higher PIC value has also been identified. Using this criteria, a total of 36 BINs have been identified that have at least one marker with >0.70 PIC value and 111 BINs with at least one marker >0.50 PIC value. This information will provide useful information to

select the genome-wide markers that has higher probability of showing polymorphism in the parental genotypes of the mapping populations or germplasm collections and moreover primer sequence information has also been provided for 885 markers (Gautami et al., 2012b).

6. SUMMARY

The conclusions from the present research work are briefly summarized below.

- **Development of novel SSR markers in groundnut:**

A new SSR enriched library was constructed from the genotype ICGV 86031. Sequencing of 96 SSR positive clones provided good quality sequences for 65 clones. The microsatellite sequence data for these 65 clones were submitted to Genbank under accession numbers FI857100 to FI857164 to make the sequences available to public and make use of this study for further developments of genetic markers. Mining of these sequences with MISA (*MI*cro*SA*tellite) search tool could able to design primer pairs for 23 SSR loci, of which 14 (16%) primer pairs yielded scorable amplicons and eight (57%) primer pairs showed polymorphism among two groundnut genotypes (ICGV 86031 and TAG 24). The polymorphism information content (PIC) for the new polymorphic SSR markers ranged from 0.13 to 0.36, with an average of 0.25.Therefore, the present set of newly developed 14 new novel SSR markers can enriches the existing groundnut SSR repertoire.

- **Screening for parental polymorphisms using SSR markers and genotyping of the respective mapping populations**

The parental genotypes of the two recombinant inbred line (RIL) mapping populations (ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS

76) were screened with 3215 SSR markers available in public domain and from various collaborators. In total 128 polymorphic loci on ICGS 76 × CSMG 84-1 and 87 polymorphic loci on ICGS × ICGS 76 were found polymorphic and genotyping data were generated for these markers.

- **Construction of two genetic linkage maps using polymorphic microsatellite markers**

  Features of the map ICGS 76 × CSMG 84-1:

  - Total number of marker loci mapped: 119
  - Number of marker loci per linkage group: 2 to 10
  - Total map distance: 2208.20 cM
  - Average map distance per linkage group: 16.79 cM

  Features of the map ICGS 44 × ICGS 76:

  - Total number of marker loci mapped: 82
  - Number of marker loci per linkage group: 2 to 14
  - Total map distance: 831.4 cM
  - Average map distance per linkage group: 10.41 cM

- **Phenotyping of two mapping populations for drought related traits**

Phenotyping of parents and RILs in the present study showed moderate variations and low heritability for all the traits in both the mapping populations. The effects of genotype x environment (GE) interactions, however was not observed to be significant. Similarly, the broad-sense heritability (h2b.s), grand mean, SED and LSD were observed to be moderate to low in both mapping populations.

The detailed analysis of phenotypic data showed lower incidence of tolerance towards the female parent in both the mapping populations; however, the means of both the RILs were within the parental limits and all traits showed continuous distribution indicating their polygenic nature.

- **Identification of genes/QTLs associated with tolerance to Drought**

Genotyping data for the two RIL mapping populations were analyzed together with phenotyping data for drought related traits respectively. The QTL analysis detected 31 M-QTLs for the mapping population ICGS 76 × CSMG 84-1 and 5 M-QTLs for the mapping population ICGS 44 × ICGS 76 using QTL Cartographer and QTL Network programme. By using the QTLNetwork programme, a total

of ten E-QTLs were detected in two mapping populations and by using the genotypic matrix mapping programme 37 E-QTLs were detected in ICGS 76 × CSMG 84-1 and 26 E-QTLs in ICGS 44 × ICGS 76 mapping populations respectively.

- **Construction of consensus genetic map using three ICRISAT RIL mapping populations segregating for drought related traits and mapping of several M-QTLs and E-QTLs**

Together with the two genetic maps constructed in the present study, and the reference genetic linkage map with 191 SSR loci based on TAG 24 × ICGV 86031 (Ravi et al., 2010), a consensus map was constructed with 293 SSR loci distributed over 20 linkage groups, spanning 2840.8 cM. As all these three populations segregate for drought tolerance related traits, a comprehensive QTL analysis identified 153 M-QTLs and 25 E-QTLs for drought tolerance related traits. Localization of these QTLs on the consensus map provided 16 genomic regions that contained 137 QTLs.

- **Construction of an international reference consensus genetic map for tetraploid groundnut**

Using marker segregation data for 10 RILs and one BC population from the international groundnut community, an international reference consensus genetic map has been developed. This map comprised of 897 marker loci distributed on 20 LGs (a1-

a10 and b1- b10) spanning a map distance of 3863.6 cM with an average map density of 4.4 cM. Highest numbers of markers (70) were integrated on a1 and the least number of markers (21) on b9. The marker density, however, was lowest (6.4 cM) on a8 and highest (2.5 cM) on a1. The reference consensus map has been divided into 20 cM long 203 BINs. These BINs carry 1 (a10_02, a10_08 and a10_09) to 20 (a10_04) loci with an average of 4 marker loci per BIN. Although the PIC value was available for 526 markers in 190 BINs, 36 and 111 BINs have at least one marker with > 0.70 and > 0.50 PIC values, respectively.

In summary, the newly developed genomic resources such as SSR markers and genetic linkage maps will be useful for groundnut genetics and breeding applications. Moreover, the markers and QTLs for drought tolerance related traits will be useful for molecular breeding for drought tolerance in groundnut improvement. Apart from this, the international reference consensus map developed in the present study provides the marker order for maximum markers available in groundnut community and also helpful in aligning new genetic map as well as anchoring genetic map to the future physical map.

7. REFERENCES

1.Akkaya MS, Bhagwat AA et al. Length polymorphisms of simple sequence repeats DNA in soybean. Genetics, 132, 1992 (1131-1139).

2. Alamillo J, Almogura C, Bartels D, Jordano J. Constitutive expression of small heat shock proteins in vegetative tissues of the resurrection plant Craterostigma plantagineum. Plant Mol. Biol., 29, 1995 (1093-1099).

3. Ashraf M. Inducing drought tolerance in plants: Recent advances. Biotechnol. Adv., 28, 2010 (169-183).

4. Azam Ali SN. Environmental and Physiological Control of Transpiration by Groundnut Crops, *Agricultural and Forest Meteorology*, 33, 1984, (129-140).

5. Azhaguvel P, Vidya Saraswati D, Sharma A and Varshney RK. Methodological advancements in molecular markers to delimit the gene(s) for the crop improvement. In J.Teixera da Silva (ed.) *Floriculture, Ornamental and Plant Biotechnology: Advances and Tropical Issues*, Global Science Books, London, UK, 2006, (460-499).

6. Babu VR and Rao DVM. Water Stress Adaptations in The Ground Nut (*Arachis hypogaea* L.) – Foliar Characteristics and Adaptations to Moisture Stress, *Plant Physiology & Biochemistry*, 10, 1983, (64–80).

7. Barrera-Figueroa B, Pena-Castro J, Acosta-Gallegos JA, Ruiz-Medrano R, Xoconostle-Cazares B. Isolation of dehydratation-responsive genes in a drought tolerant common bean cultivar and expression of a group 3 late embryogenesis abundant mRNA in tolerant and susceptible bean cultivars. Functional Plant Biol., 34, 2007, (368-381).

8.Barkley NA, Dean RE, Pittman RY, Wang ML, Holbrook CC, Pederson GA. Genetic diversity of cultivated and wild-type peanuts evaluated with M13-tailed SSR markers and sequencing. Genetical Research, 89, 2007, (93-106).

9. Bartels D and Sunkar R. Drought and salt tolerance in plants. Crit. Rev. Plant Sci., 24, 2005, (23-58).

10. Bartel D. Targeting detoxification pathways: An efficient approach to obtain plants with multiple stress tolerance. Trends Plant Sci., 6, 2001, (284-286).

11. Basten CJ, Weir BS, Zeng ZB. Zmap – a QTL cartographer. Computing Strategies and Software. Proc 5th Congr on Genetics Applied to Livestock Production: Guelph, Ontario, Canada, 1994.

12. Beavis WD, Grant D. A linkage map based on information from four F_2 populations of maize (*Zea mays* L.). Theor Appl Genet 82, 1991, (636–644).

13. Bernardo R. Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci 48, 2008, (1649–1664).

14. Bhagwat AS, Anantharaman R, Pereira S, Krishna TG. Identification of polymorphism by Random Amplified Polymorphic DNA (RAPD) in groundnut (Arachis hypogaea L.) Plant Varieties and Seeds,14, 2001, (119-124).

15. Bhagwat AS, Krishna TG and Bhatia CR. RAPD analysis of induces mutants of groundnut. Journal of Genetics, 76, 1997, (201-208).

16. Blum A. Drought resistance, water-use efficiency, and yield potential: are they compatible, dissonant, or mutually exclusive? Aust J Agric Res 56, 2005, (1159–1168).

17. Blum A. Crop responses to drought and the interpretation of adaptation. Plant Growth Regul., 20, 1996, (135-148).

18. Bohnert HJ and Sheveleva E. Plant stress adaptations making metabolism move. Curr. Opinion Plant Biol., 1, 1998, (267-274).

19. Boote KJ and Ketring DL. Peanut. In: Stewart B.A. And Nielson D.R. (Eds), Irrigation of Agricultural Crops. Asa- Groundnut - A Global Perspective. International Crops Research CSSA-SSSA, Madison, (1990).

20. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32, 1980, (314–331).

21. Boyer JS. Plant productivity and environment. Science, 218, 1982, (443-448).

22. Bravo JP, Hoshino AA, Angelici CMLCD, Lopes CR and Gimenes MA. Transferability and use of microsatellite markers for the genetic analysis of the germplasm of some *Arachis* section species of the genus *Arachis*. Gen. Mol. Biol., 29(3), 2006, (516-524).

23. Bray E. Abscisic acid regulation of gene expression during water-deficit stress in the era of the Arabidopsis genome. Plant Cell Environ., 25, 2002, (153-161).

24. Burrow MD, Simpson CE, Starr JL, Paterson AH. Transmission genetics of chromatin from a synthetic amphidiploid to cultivated peanut (*Arachis hypogaea* L.): Broadening the gene pool of a monophyletic polyploid species. *Genetics*, 159, 2001, (823–837).

25. Burrow MD, Starr JL, Simpson CE and Paterson AH. Identification of RAPD markers in peanut (*Arachis hypogaea* L.) associated with root-knot nematode resistance derived from *A. cardenasii*. *Mol Breed*, 2: 1996, (307–319).

26. Castiglioni P, Ajmone P, van Wijk R, Motto M. AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor Appl Genet., 99, 1999, (425–431).

27. Chandler J and Bartels D. Plant Dessication. In: Plant Responses to Environmental Stresses: from Phytohormones to Genome Reorganization, Lerner, H.R. (Ed.). Marcel Dekker, New York, USA, 1999, (575-590).

28. Chen THH and Murata N. Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. Curr. Opin. Plant Biol., 5, 2002, (250-257).

29. Christmann A, Weiler EW, Steudle E, Grill E. A hydraulic signal in root-to-shoot signaling of water shortage. Plant J., 52, 2007, (167-174).

30. Choi IY, Hyten DL, Matukumalli LK, Song Q, Chaky JM, et al. A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. Genetics, 176, 2007, (685–696).

31. CIAT. Centro internacional de agricultura tropical. Project IP3, Improved Cassava for the Developing World, Annual Report. Apdo Aereo 6713, Cali, Colombia, 2001.

32. Collins NC, Tardieu F, Tuberosa R. Quantitative trait loci and crop performance under abiotic stress: where do we stand? Plant Physiol., 147, 2008, (469–486).

33. Cone KC, McMullen MD, Bi IV, Davis GL, Yim YS, Gardiner JM, Polacco ML, Sanchez-Villeda H, Fang Z, Schroeder SG, Havermann SA, Bowers JE, Paterson AE, Soderlund CA, Engler FW, Wing RA, Coe EH. Genetic, Physical, and Informatics Resources for Maize. On the Road to an Integrated Map. Plant Physiol, 130, 2002, (1598-1605).

34. Crouch JH. Molecular marker-assisted breeding: a perspective for small to medium –sized plant breeding companies. Asia and Pacific seed Association Technical Report No., 30, 2001, (1-14).

35. Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD and Varshney RK. Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Pl. Biol., 8, 2008, (55).

36. Dwivedi S, Bertioli DJ, Crouch JH, Valls JFM, Upadhyaya HD, Favero AP, Moretzsohn MC, Paterson AH. Peanut Genetics and Genomics: Toward Marker-assisted Genetic Enhancement in Peanut (*Arachis hypogaea* L). In: Oilseeds Series: Genome Mapping and Molecular Breeding in Plants. Berlim, Heidelberg: Springer, 2, 2007, (115-151).

37. Dwivedi SL, Crouch JH, Nigam SN, Ferguson ME, Paterson AH. Molecular breeding of groundnut for enhanced productivity and food security in the semi-arid tropics: Opportunities and challenges. In: Donald L. Sparks, editor. Advances in Agronomy. Academic Press, An Elsevier Science Imprint, 2003, (154–221).

38. Dwivedi SL, Gurtu S, Chandra S, Yuejin W, Nigam SN. Assessment of genetic diversity among selected groundnut germplasm. I. RAPD analysis. *Plnt Breed*, 120, 2001, (345-349).

39. Edwards KJ, Barker JH et al. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20, 1996, (759–760).

40. Encyclopaedia of Agricultural Science. www.worldcat.org/./ encyclopedia- of- agricultural-science/.1994.

41. Falque M, Décousset L, Dervins D, Jacob AM, Joets J, et al. Linkage mapping of 1454 maize candidate gene loci. Genetics, 70, 2005, (1957–1966).

42. FAOSTAT: Agricultural crop production statistics, available at www.fao.org/faostat. Food and Agricultural Organization, Rome, Italy. [Searchable database for crop production statistics of different countries].

43. Fávero AP, Simpson CE, Valls JF, Vello NA. Study of the evolution of cultivated peanut through crossability studies among *Ara chis ipaensis*, A. *duranensis*, and *A. hypogaea*. Crop Sci 46, 2006, (1546–1552).

44. Feltus FA, Hart GE, Schertz KF, Casa AM, Kresovich S, et al. Alignment of genetic maps and QTLs between inter- and intraspecific sorghum populations. Theor Appl Genet, 112, 2006, (1295–1305).

45. Ferguson ME, Burow M, Schulze SR, Bramel PJ, Paterson A, Kresovich S and Mitchell S. Microsatellite identification and characterization in peanut (*A.hypogaea* L.). Theor Appl Genet, 108, 2004, (1064-1070).

46. Foncéka D, Hodo-Abalo T, Rivallan R, Vignes H, Faye I, et al. Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. BMC Plant Biol, 12, 2012, (26).

47. Foncéka D, Hodo-Abalo T, Rivallan R, Faye I, Ndoye Sall M, et al. Genetic mapping of wild introgressions into cultivated peanut: a way toward enlarging the genetic basis of a recent allotetraploid. BMC Plant Biol, 9, 2009, (103).

48. Frary A, Fulton TM, Zamir D, Tanksley SD. Advanced backcross QTL analysis of a Lycopersicon esculentum 9 L. pennellii cross and identification of possible orthologs in the Solanaceae. Theor Appl Genet, 108, 2004, (485–496).

49. Garcia GM, Stalker HT, Schroeder E, Lyerly JH and Kochert G. A RAPD-based linkage map of peanut based on a backcross population between the two diploid species *Arachis stenosperma* and *A. cardenasii*. Peanut Sci, 48, 2005, (145-151).

50. Garcia GM, Stalker HT, Kochert G. Introgression analysis of an interspecific hybrid population in peanuts (*Arachis hypogaea* L.) using RFLP and RAPD markers. *Genome*, 38, 1995, (166–176).

51. Galau GW, Hughes DW, Dure L. Abscisic acid induction of cloned cotton Late Embryogenesis Abundant (LEA) messenger RNAs. Plant Mol. Biol., 7, 1986, (155-170).

52. Gao GO, He GH et al. Microsatellite enrichment from AFLP fragments by magnetic beads. Acta Botanica Sinica, 45, 2003, (1266-1269).

53. Gentzbittel L, Vear F, Zhang YX, Breville A, Nicolas P () Development of a consensus linkage RFLP map of cultivated sunflower (*Helianthus annus* L.). Theor Appl Genet, 90, 1995, (1079–1086).

54. Gautami B, Pandey MK, Vadez V, Nigam SN, Ratnakumar P, et al. Quantitative trait locus analysis and construction of consensus genetic map for drought tolerance traits based on three recombinant inbred line populations in cultivated groundnut (Arachis hypogaea L.). Mol Breed, 30, 2012a, (757–772).

55. Gautami B, Fonce´ka D, Pandey MK, Moretzsohn MC, Sujay V, Qin H, Hong Y, Faye I, Chen X, BhanuPrakash A, Trushar MS, Gowda MVC, Nigam SN, Liang X, Hoisington DA, Guo B, Bertioli DJ, Rami JF, Varshney RK. An international reference consensus genetic map with 897 marker loci based on 11 mapping populations for tetraploid groundnut (Arachis hypogaea L.) PLoS One, 7(7), doi: 10.1371/journal.pone.0041213.

56. Gepts P, Beavis WD, Brummer EC, Shoemaker RC, Stalker HT, Weeden, NF, Young ND. Legumes as a Model Plant Family. Genomics for Food and Feed Report of the Cross-Legume Advances through Genomics Conference, *Plant Physiology*, 137, 2005, (1228–1235).

57. Gimenes MA, Hoshino AA et al. Characterization and transferability of microsatellite markers of cultivated peanut (Arachis hypogaea L.). BMC Plant Biol, 7, 2007, (9).

58. Gimenes MA, Lopes CR, Galgaro ML, Valls JFM and Kochert GD. RFLP analysis of genetic variation in species of the section *Arachis*, genus *Arachis* (*Leguminosae*). Euphytica, 123, 2002, (421–429).

59. Glenn TC, Schable NA. Isolating DNA loci: Methods in Enzymology, 395, 2005, (202-222).

60. Gobbi A, Teixeira C, Moretzsohn M, Guimaraes P, Leal-Bertioli S, Bertioli D, Lopes CR and Gimenes M. Development of a linkage map to species of B genome related to the peanut (*Arachis hypogaea* – AABB). In Plant and Animal Genomes XIV Conference, San Diego, CA, USA, 2006, (679).

61. Gowda A and Hegde BR. Moisture Stress and Hormonal Influence on The Flowering Behavior and Yield Of Groundnut, *Plant Physiology*, 66, 1986, (835–837).

62. Gupta PK, Rustgi S, Mir RR. Array-based high-throughput DNA markers for crop improvement. Heredity, 101, 2008, (5–18).

63. Gupta PK, Rustgi S. Molecular markers from the transcribed/expressed region of the genome in higher plants. Funct Integr Genom, 4, 2004, (139-162).

64. Gupta PK, Varshney RK. Cereal genomics: an overview. In: Gupta PK, Varshney RK (eds) Cereal Genomics. Kluwer Academic Publishers, The Netherlands, 2004, (1-18).

65. Gupta PK, Varshney RK, Prasad M. Molecular markers: principles and methodology. In Jain SM, Ahloowalia BS, Brar DS (eds) Molecular Techniques in Crop Improvement. Kluwer Academic Publishers, The Netherlands, 2002, (9-54).

66. Gupta PK and Varshney RK. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica, 113, 2000, (163-185).

67. Gupta PK, Varshney RK, Sharma PC, Ramesh B. Molecular markers and their applications in wheat breeding. Plant Breed, 118, 1999, (369–390).

68. Gupta PK, Balyan HS, Sharma PC, Ramesh B. Microsatellites in plants: a new class of molecular markers. Current Sci., 70, 1996, (45-53).

69. Guo B, Chen X, Hong Y, Liang X, Dang P, Brenneman T, Holbrook C, and Culbreath A. Analysis of Gene Expression Profiles in Leaf Tissues of Cultivated Peanuts and Development of EST-SSR Markers

and Gene Discovery. Int. J. Pl. Gen., 2009, ID715605 DOI:10.1155/2009/715605.

70. Guo BZ, Xu G, Cao YG, Holbrook CC, Lynch RE. Identification and characterization of phospholipase D and its association with drought susceptibilities in peanut (*Arachis hypogaea*). Planta, 223, 2006, (512-520).

71. Gustafson JP, Ma XF, Korzun V, Snape JW. A consensus map of rye integrating mapping data from five mapping populations. Theor Appl Genet., 118, 2009, (793–800).

72. Halward TM, Stalker HT, Kochert G. Development of an RFLP linkage map in diploid peanut species. Theor and Appl Genet, 87, 1993, (379–384).

73. Halward TM, Stalker HT, Larue EA and Kochert G. Use of single-primer DNA amplifications in genetic studies of peanut (Arachis hypogaea L.). Plant Molecular Biology, 18, 1992, (315-325).

74. Halward TM, Stalker HT, Larue EA, and Kochert G. Genetic variation detectable with molecular markers among unadopted germplasm resources of cultivated peanut and related wild species. Genome, 34, 1991, (1013-1020).

75. Hanson CH, Robinson HG, Comstock RE. Biometrical studies of yield in segregating populations of Korean Lespediza. Agron J., 48, 1956, (268-272).

76. Hartl D, Jones E. Genetics: Analysis of Genes and Genomes, Jones and Bartlett Publishers, Sudbury, MA, USA, 2001.

77. Hauge BM, Hanley SM, Cartinhour S, Cherry JM, Goodman HM. An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant J, 3, 1993, (745–754).

78. He GH, Meng R, Gao H, Guo B, Gao G, Newman M, Pittman RN, Prakash CS. Simple sequence repeat markers for botanical varieties of cultivated peanut (*Arachis* hypogaea L.). *Euphytica*, 142, 2005, (131-136).

79. He GH, Meng R, Newman M, Gao G, Pittman RN, Prakash CS. Microsatellites as DNA markers in cultivated peanut (*Arachis hypogaea* L.).BMC Plant Biol., 3, 2003, (3).

80. He GH and Prakash CS. Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. *Genet. Resources and Crop Evol.*, 48, 2001, (347–352).

81. He GH and Prakash CS. Identification of polymorphic DNA markers in cultivated peanut (*Arachis hypogaea* L.). *Euphytica*, 97, 1997, (143-149).

82. Herselman L, Thwaites R, Kimmins FM, Courtois B, Vander Merwe PJA and Seal SE. Identification and mapping of AFLP markers linked to peanut (*Arachis hypogaea* L.) resistance to the aphid vector of groundnut rosette disease. Theor Appl Genet, 109, 2004, (1426–1433).

83. Hilu, KW and Stalker HT. Genetic relationships between peanut and wild species of *Arachis* sect *Arachis* (Fabaceae): Evidence from RAPDs. Pl. Syst. Evol., 198 (3-4), 1995, (167-178).

84. Holbrook CC and Stalker HT. Peanut breeding and genetic resources. Plant Breed. Rev., 22, 2003, (297-356).

85. Holbrook CC, Stephenson MG, Johnson AW. Level and geographical distribution of resistance to Meloidogyne arenaria in the U.S. peanut germplasm collection. Crop Sci., 40, 2000, (1168-1171).

86. Hong Y, Chen X, Liang X, Liu H, Zhou G, et al. A SSR-based composite genetic linkage map for the cultivated peanut (Arachis hypogaea L.) genome. BMC Plant Biol, 10, 2010, (17).

87. Hong SW, Jon JH, Kwak JM, Nam HG Identification of a receptor-like protein kinase gene rapidly induced by abscisic acid, dehydration, high salt and cold treatments in Arabidopsis thaliana. Plant Physiol., 113, 1997, (1203-1212).

88. Hopkins MS, CasaAM, Wang T, Mitchell SE, Dean RE, Kochert GD and Kresovich S. Discovery and characterization of polymorphic simple sequence repeats (SSRs) in peanut. Crop Sci., 39, 1999, (1243-1247).

89. Huang X, Madan A. CAP3: A DNA Sequence Assembly Program. Genome Res, 9, 1999, (868-877).

90. Hubick KT, Farquhar GD, Shorter R. Correlation between water-use efficiency and carbon isotope discrimination in diverse peanut (Arachis) germplasm. Aust J Plant Physiol, 13, 1986, (803–816).

91. Hüttel B, Winter P et al. Sequence-tagged microsatellite markers for chickpea (Cicer arietinum L.). Genome, 42, 1999, (210-217).

92. Hyten DL, Choi IY, Song Q, Specht JE, Carter TE, et al. A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. Crop Sci., 50, 2010, (960–968).

93. Isobe S, Nakaya A, Tabata S. Genotype matrix mapping: searching for quantitative trait loci interactions in genetic variation in complex traits. DNA Res, 14, 2007, (217–225).

94. Jain AK, Basha SM and Holbrook CC. Identification of Drought-Responsive Transcripts in Peanut (*Arachis hypogaea* L.), *Electronic Journal of Biotechnology*, 4, 2001, (59-67).

95. Jannink J. Identifying quantitative trait locus by genetic background interactions in association studies. Genetics, 176, 2007, (553–561).

96. Janamatti VS, Sashidhar R, Prasad IG, Sastry KSK. Effect Of Cycles Of Moisture Stress On Flowering Pattern, Flower Production, Gynophore Length and Their Relationship To Pod Yield In Bunch Types Of Groundnut, *Journal of Agricultural Research*, 2, 1986, (136–142).

97. Jogloy S, Patanothai A, Toomsan S and Isleib TG. Breeding peanut to fit into Thai cropping systems. Proc. of the Peanut Collaborative Research Support Program- International Research Symposium and Workshop, Two Jima Quality Inn, Arlington, Virginia, USA, 1996, (353-362).

98. Kacperska A. Sensor types in signal transduction pathways in plant cells responding to abiotic stressors: do they depend on stress intensity? Physiologia Plantarum, 122, 2004, (159-168).

99. Karakousis A, Gustafson JP, Chalmers KJ, Barr AR, Langridge P.: A consensus map of barley integrating SSR, RFLP, and AFLP markers. Aust J Agric Res, 54, 2003, (1173–1185).

100. Khedikar YP, Gowda MVC, Sarvamangala C, Patgar KV, Upadhyaya HD, et al. A QTL study on late leaf spot and rust revealed one major QTL for molecular breeding for rust resistance in groundnut (Arachis hypogaea L.). Theor Appl Genet, 121, 2010, (971–984).

101. Kholová J, Hash CT, Kočová M, Vadez V. Constitutive water conserving mechanisms are correlated with the terminal drought tolerance of pearl millet (*Pennisetum americanum* L.). J Exp Bot, 61, 2010a, (369–377).

102. Kholová J, Hash CT, Kumar LK, Yadav RS, Kočová M, Vadez V. Terminal drought tolerant pearl millet [*Pennisetum glaucum* (L.) R. Br.] have high leaf ABA and limit transpiration at high vapor pressure deficit. J Exp Bot., 61, 2010b, (1431-1440).

103. Khush GS. Green revolution: Preparing for the 21st century. Genome, 42, 1999, (646-655).

104. Kianian SF, Quiros CF. Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. Theor Appl Genet, 84, 1992, (544–554).

105. Klimenko I, Razgulayeva N, Gau M, Okumura K, Nakaya A, Tabata S, Kozlov NN, Isobe S. Mapping candidate QTLs related to plant persistency in red clover. Theor Appl Genet, 120, 2010, (1253–1263).

106. Knight H and Knight MR. Abiotic stress signalling pathways: Specificity and cross-talk. Trends Plant Sci., 6, 2001, (262-267).

107. Krapovikas A, Valls JFM and Simpson CE. History of Arachis including evidence of A hypogaea L progenitors p 46 In J R Sholar (ed), American Peanur Research Educatronol Socrety, USA. 2000.

108. Krishnamurthy L, Vadez V, Jyotsna Devi M, Serraj R, Nigam SN, Sheshshayee MS, Chandra S, Aruna R. Variation in transpiration efficiency and its related traits in a groundnut (*Arachis hypogaea* L.) mapping population. Field Crops Res, 103, 2007, (189–197).

109. Krishna GK, Zhang J, Burrow M, Pittman RN, Delikostadinov SG, Lu Y, d Puppala N. Genetic diversity analysis in valencia peanut (*Arachis hypogaea* L.) using microsatellite markers. *Cell. Mol. Bio.*, 9, 2004, (685-697).

110. Kochert G, Halward T, Branch WD, Simpson CE. RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor Appl Genet, 81, 1991, (565-570).

111. Kosambi DD. The estimation of map distance from recombination values. Ann Eugen, 12, 1944, (172–175).

112. Kottapalli KR, Burow MD, Burow G, Burke J, Puppala N. Molecular characterization of the U.S. peanut mini core collection using microsatellite markers. Crop Sci, 47, 2007, (1718–1727).

113. Kulkarni J.H, Ravindra V, Sojitra VK, Bhatt DM. Growth, nodulation and N uptake of groundnut (*Arachis hypogaea* L.) as influenced by water deficits stress at different phenophase, *Oleagineus,* 43, 1988, (415–419).

114. Lambers H, Stuart-Chapin III F, Pons TL. Plant Physiological Ecology. Springer-Verlag, New York, USA, 2000, (540).

115. Lander ES, Green P. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, 1, 1987, (174–181).

116. Lander E, Botstein D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 121, 1989, (185–199).

117. Langridge P, Karakousis A, Collins N, Kretschmer J, Manning S. A consensus linkage map of barley. Mol Breed, 1: 1995, (389–395).

118. Lanham PG, Fennel S, Moss JP, Powell W. Detection of polymorphic loci in *Arachis* germplasm using random amplified polymorphic DNAs. *Genome*, 35, 1992, (885–889).

119. Leal-Bertioli SC, José AC, Alves-Freitas DM, Moretzsohn MC, Guimarães PM, et al. Identification of candidate genome regions controlling disease resistance in *Arachis*. BMC Plant Biol, 9, 2009, (112).

120. Lee GJ, Pokala N, Vierling E. Structure and *in vitro* molecular chaperone activity of cytosolic small heat shock proteins from pea. J. Biol. Chem., 270, 1995, (10432-10438).

121. Levine A. Oxidative Stress as a Regulator of Environmental Responses in Plants. In: Plant Responses to Environmental Stresses: from Phytohormones to Genome Reorganization, Lerner, H.R. (Ed.). Marcel Dekker, New York, USA, 1999, (247-264).

122. Li ZK. QTL mapping in rice: a few critical considerations. Rice Gentics IV, Proceedings of the Fouth Publishers, Inc., and Los Banos (Philippines): International Rice Research Institute, 20004, (88).

123. Li ZK, Pinson SRM, Paterson AH, Park WD, Stansel JW. Epistasis for three grain yield components in rice (Oryza sativa L.).Genetics, 145, 1997, (453-465).

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

125. Lincoln S, Daly M, Lander E. Constructing genetic linkage maps with MAPMAKER/EXP. Version 3.0. Whitehead Institute for Biomedical Research Technical Report, 3rd Edn. (1993).

126. Litt M and Lutty JA. A hyper variable microsatellite revealed by in vitro amplification of a dinucleotide repeats within the cardiac muscle actin gene. Am. J. Hum. Genet, 44, 1989, (397–401).

127. Liu B. Statistical Genomics: Linkage, Mapping and QTL Analysis. CRC Press, Boca Raton, USA, 1998.

128. Mace ES, Rami JF, Bouchet S, Klein PE, Klein RR, et al. A consensus genetic map of sorghum that integrates multiple component maps and high-throughput Diversity Array Technology (DArT) markers. BMC Plant Biol, 9, 2009, (13).

129. Mace ES, Varshney RK, Mahalakshmi V, Seetha K, Gafoor A, Leeladevi Y and Crouch JH. In silico development of simple sequence repeat markers within the aeschynomenoid/ dalbergoid and genistoid clades of the Leguminosae family and their transferability to Arachis hypogaea, groundnut. Pl. Sci., 174, 2007, (51–60).

130. Mace E, Yuejin W, Boshou L, Upadhyaya H, Chandra S and Crouch JH. SSR based diversity analysis of groundnut (*Arachis hypogaea* L.) germplasm resistant to bacterial wilt. *Plant Genet Resou: Charact & Uilz*, in press. 2006.

131. Mace ES, Phong DT, Upadhyaya HD, Chandra S, Crouch JH. SSR analysis of cultivated groundnut (*Arachis hypogaea* L.) germplasm resistant to rust and late leaf spot diseases. *Euphytica*, 152, 2006, (317–330).

132. Manly KF, Cudmore RH, Meer JM. Map Manager QTX, cross-platform software for genetic mapping. Mamm Genome, 12, 2001, (930–932).

133. Marcel TC, Varshney RK, Barbieri M, Jafary H, de Kock MJ, et al. A high-density consensus map of barley to compare the distribution of QTLs for partial resistance to Puccinia hordei and of defence gene homologues. Theor Appl Genet, 114, 2007, (487–500).

134. McCouch S. Diversifying selection in plant breeding. PLoS Biol., 2, 2004, (347).

135. McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D, Stein L. Development and mapping of 2240 new SSR markers for rice (Oryza sativa L.). DNA Res, 9, 2002, (199–207).

136. Milla SR, Isleib TG, Stalker HT. Taxonomic relationships among *Arachis* Sect. *Arachis* species as revealed by AFLP markers. *Genome,* 48, 2005, (1-11).

137. Milla SR. Relationships and utilization of *Arachis* germplasm in peanut improvement. *Ph. D. Thesis*, North Carolina State University, USA, 2003, (1-150).

138. Mittler R and Blumwald E. Genetic engineering for modern agriculture: challenges and perspectives. Annu. Rev. Plant Biol., 61, 2010, (443-462).

139. Meisner CA. Peanut Roots, Shoot and Yield and Water Stress. *Dissertation Abstracts International. Biological Sciences and Engineering,* 52, 1991, (38–48).

140. Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T. Genome mapping, molecular markers and marker assisted selection in crop plants. Mol Breed, 3, 1997, (87-93).

141. Montalvo-Hernandez L, Piedra-Ibarra EE, Gomez-Silva L, Lira-Carmona R, Acosta-Gallegos JA et al., Differential accumulation of mRNAs in drought-tolerant and susceptible common bean cultivars in response to water deficit. New Phytol., 177, 2008, (102-113).

142. Montero-Tavera V, Ruiz-Medrano R, Xoconostle-Cazares B. Systemic nature of drought-tolerance in common bean. Plant Signal. Behav., 3, 2008, (663-666).

143. Monneveux P and Belhassen E. The diversity of drought adaptation in the wide. Plant Growth Regul., 20, 1996, (85-92).

144. Moretzsohn MC, Barbosa AVG, Alves-Freitas DMT, Teixeira C, Leal-Bertioli SCM, et al. A linkage map for the B-genome of *Arachis* (Fabaceae) and its synteny to the A-genome. BMC Plant Biol, 9, 2009, (40).

145. Moretzsohn MC, Leoi L, Proite K, Guimara, PM, Leal-Bertioli SCM, Gimenes MA, Martins WS, Valls JFM, Grattapaglia D, Bertioli DAJ. Microsatellite-based, gene-rich linkage map for the AA genome of *Arachis* (Fabaceae). Theor Appl Genet, 111, 2005, (1060-1071).

146. Moretzsohn MC, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME. Genetic diversity of peanut (Arachis hypogaea L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol, 4, 2004, (11).

147. Nageswara Rao RC, Williams JH, Sivakumar MVK and Wadia KRD. Effect of Water Deficit at Different Growth Phases of Peanut. Ii. Response to Drought during Pre-Flowering Phase, *Agronomy Journal*, 80, 1998, (431-438).

148. Nageswara Rao RC, Wright GC. Stability of the relationship between specific leaf area and carbon isotope discrimination across environments in peanut. Crop Sci, 34, 1994, (98–103).

149. Nagy E, Chu Y, Guo Y, Khanal S, Tang S, Lee Y, Dong W, Timper P, Taylor C, Ozias-Akins P, Holbrook C, Beilinson V, Nelson N, Stalker T, Knapp S. Recombination is suppressed in an alien introgression in peanut harboring RNA, a dominant rootknot nematode resistance gene. Mol Breed, 26, 2010, (357-370).

150. Nelson MN, Phan HTT, Ellwood SR, Moolhuijzen Paula M, Hane J, Williams A, O'Lone CE, Nyarko JF, Scobie M, Cakir M, Jones MGK, Bellgard M, Wolko B, Barker SJ, Oliver RP, Cowling WA. The first gene-based map of *Lupinus angustifolius* L.-location of domestication genes and conserved synteny with *Medicago truncatula*. Theor Appl Genet, 113, 2006, (225– 238).

151. Nigam SN, Chandra S, Rupa Sridevi K, Manohar B, Reddy AGS, Nageswara Rao RC, Wright GC, Reddy PV, Deshmukh MP, Mathur RK, Basu MS, Vasundhara S, Vindhiya Varman P, Nagda AK. Efficiency of physiological trait-based and empirical selection approaches for drought tolerance in groundnut. Ann Appl Biol, 146, 2005, (433–439).

152. Nimmakayala P, Jeong J, Asturi SR, Thomason Y, Tallury S, Reddy UK. Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on AFLP and microsatellite polymorphisms. In: Plant and animal genomes XV conference, San Diego, CA, USA, 2007, (427).

153. Nwokolo E. Bambara groundnut (*Vigna subterranea*). In: Food and Feed from Legumes and Oilseeds (E. Nwokolo and Smartt J, eds), Chapman and Hall, London, 1996, (216-221).

154. Osakabe Y, Maruyama K, Seki M, Satou M, Shinozaki K, Yamaguchi-Shinozaki K. Leucine-rich repeat receptor-like kinase1 is a key membrane-bound regulator of abscisic acid early signaling in Arabidopsis. Plant Cell, 17, 2005, (1105-1119).

155. Paik-Ro OG, Smith RL, Knauft DA. Restriction fragment length polymorphism evaluation of six peanut species within the *Arachis* section. Theor Appl Genet, 84, 1992, (201–208).

156. Palmieri DA, Bechara MD, Curi RA, Gimenes MA, Lopes CR. Novel polymorphic microsatellite markers in section Caulorrhizae (Arachis, Fabaceae). Mol Ecol Notes, 5, 2005, (77–79).

157. Palmieri DA, Hoshino AA, Bravo JP, Lopes CR, Gimenes MA. Isolation and characterization of microsatellite loci from the forage species Arachis pintoi (Genus Arachis). Mol Ecol Notes, 2, 2002, (551–553).

158. Pandey MK, Gautami B, Jayakumar T, Sriswathi M, Upadhyaya HD, et al. Highly informative genic and genomic SSR markers to facilitate molecular breeding in cultivated groundnut (Arachis hypogaea L.). Plant Breed, 131, 2011, (139–147).

159. Pandey MK, Monyo E, Ozias-Akins P, Liang X, Guimaraes P, et al.: Advances in Arachis genomics for peanut improvement. Biotech Adv, 30, 2012, (639–651).

160. Pandey RK, Herrera WAT, Villepas AN and Pendelton JW. Drought Response Grain Legumes Under Irrigation Gradient. *Agronomy Journal*, 76, 1984, (557–560).

161. Panse VG, Sukhatme PV. Statistical methods for agricultural workers. ICAR, New Delhi, India, 1961

162. Paterson AH. Making genetic maps. In: Paterson AH (eds) Genome Mapping in Plants, R. G. Landes Company, Austin, Texas, USA, 1996, (23–39).

163. Phillips JR, Oliver MJ, Bartels D. Molecular Genetics of Dessication and Tolerant Systems. In: CAB International Dessication and Survival in Plants: Drying without Dying, Black, M. and H. Pritchard (Eds.). CAB International, New York, 2002, (319-341).

164. Powell W, Machray GC, Provan J. Polymorphism revealed by simple sequence repeats. Trends Plant Sci., 1, 1996, (215-222).

165. Prasanth VP, Chandra S, Jayashree B and Hoisington D. AlleloBin – A program for allele binning of microsatellite markers based on the alogirithm of Idury and Cardon (1997), ICRISAT, 2006.

166. Proite K, Leal-Bertioli SC, Bertioli DJ, Moretzsohn MC, da Silva FR, Martins NF, Guimaraes PM. ESTs from a wild Arachis species for gene discovery and marker development. BMC Plant Biol, 7, 2007, (7).

167. Qi X, Pittaway TS, Lindup S, Liu H, Waterman E, Padi FK, Hash CT, Zhu J, Gale MD, Devos KM.: An integrated genetic map and a new set of simple sequence repeat markers for pearl millet, Pennisetum glaucum. Theor Appl Genet, 109, 2004, (1485–1493).

168. Qi X, Stam P, Lindhout P.: Comparison and integration of four barley genetic maps. Canadian Journal of Genetics and Cytology, 39, 1996, (379-394).

169. Qin H, Feng S, Chen C, Guo Y, Knapp S, et al. An integrated genetic linkage map of cultivated peanut (Arachis hypogaea L.) constructed from two RIL populations. Theor Appl Genet, 124, 2012, (653–664).

170. Rafalski JA, Vogel IM et al. Generating and using DNA markers in plants. In: Non-mammalian genomic analysis: a practical guide. Academic, San Diego, 1996, (75-135).

171. Raghavendra AS, Gonugunta VK, Christmann A, Grill E. ABA perception and signalling. Trends Plant Sci, 15, 2010, (395-401).

172. Raina SN, Rani V, Kojima T, Ogihara Y, Singh KP, Devarumath RM. RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity , varietal identification, and phylogenetic relationships in peanut (Arachis hypogaea) cultivars and wild species . Genome, 44, 2001, (763-772).

173. Ramanjulu S and Bartels D. Drought- and desiccation-induced modulation of gene expression in plants. Plant Cell Environ, 25, 2002, (141-151).

174. Ratnakumar P, Vadez V. Water use patterns, root characteristics, and yield components under intermittent drought stress in groundnut (Arachis hypogaea L.), (Unpublished), 2011.

175. Ravi K, Vadez V, Isobe S, Mir RR, Guo Y, et al. Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (Arachis hypogaea L.). Theor Appl Genet, 122, 2011, (1119–1132).

176. Reddi GHS and Reddy TY. Efficient Use of Irrigation Water. Kalyani Publishers, New Delhi, 1995.

177. Reddy TY, Reddy VR, Anbumozhi V. Physiological Responses of Groundnut (*Arachis hypogea* L.) To Drought Stress and Its Amelioration: A Critical Review. Plant Growth Regulation, 41, 2003, (75–88).

178. Reddy AJ and Rao IM. Influence of Induced Water Stress On Chlorophyll Components Of Proximal And Distal Leaflets of Groundnut Plants, *Current Science*, 5, 1968, (118–121).

179. Renu KC and Suresh KS. Prospects of success of biotechnological approaches for improving tolerance to drought stress in crop plants. Curr. Sci., 74, 1998, (25-34).

180. Ribaut JM, Ragot M. Marker-assisted selection to improve drought adaptation in maize: the backcross approach, perspectives, limitations, and alternatives. J Exp Bot, 58, 2007, (351–360).

181. Ribaut JM, Jiang C, Gonzalez-De-Leon D, Edmeades GO, Hoisington DA. Identification of Quantitative Trait Loci under Drought Conditions In Tropical Maize. 2. Yield Components and Marker Assisted Selection Strategies. Theor Appl Genet, 94, 1997, (887-896).

182. Ribaut JM, Hoisington DA, Deutsch JA, Jian C, Gonzalez De Leon Identification of QTL under Drought Conditions In Tropical Maize. 1. Flowering Parameters and the ASI, Theor Appl Genet, 92, 1996, (905-914).

183. Risch N. Genetic linkage: Interpreting LOD scores. Science, 255, 1992, (803–804).

184. Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB. The use of microsatellite DNA makers for soybean genotypes identification. Theor Appl Genet, 90, 1995, (43–48).

185. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, M iseners S (Eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology, Humana Press, Totowa, NJ, 2000, (365-386).

186. Sambrook J, Russell DR. Molecular Cloning: A Laboratory Manual (Third Edition), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, (2001).

187. Sarvamangala C, Gowda MVC, Varshney RK. Identification of quantitative trait loci for protein content, oil content and oil quality for groundnut (Arachis hypogaea L.). Field Crops Res, 122, 2011, (49–59).

188. Seki M, Narusaka M and Kamiya A. Functional Annotation of a Full-Length Arabidopsis cDNA Collection. *Science*, 296, 2002, (141–145).

189. Sekhar K, Priyanka B, Reddy VD, Rao KV. Isolation and characterization of a pigeonpea cyclophilin (CcCYP) gene, and its over-expression in *Arabidopsis* confers multiple abiotic stress tolerance. Plant Cell Environ, 33, 2010, (1324-1338).

190. Sewell MM, Sherman BK, Neale DB. A consensus map for loblolly pine (*Pinus taeda* L.). I. construction and integration of individual

linkage maps from two outbreed three-generation pedigrees. Genetics, 151, 1999, (321–330).

191. Sharopova N, McMullen MD, Schultz L, Schroeder S, Sanchez-Villeda H, et al. Development and mapping of SSR markers for maize. Plant Mol Biol, 48, 2002, (463–481).

192. Skelton BJ and Shear GM. Calcium Translocation in the Peanut (*Arachis hypogae* L.). *Agronomy Journal*, 63, 1971, (409–412).

193. Somers DJ, Isaac P, Edwards K. A high-density microsatellite consensus map for bread wheat (Triticum aestivum L.). Theor Appl Genet, 109, 2004, (1105–1114).

194. Song QJ, Marek LF, Shoemaker RC, Lark KG, Concibido VC, et al. A new integrated genetic linkage map of the soybean. Theor Appl Genet, 109, 2004, (122–128).

195. Stalker HT, Mozingo LG. Molecular markers of *Arachis* and marker assisted selection. *Peanut Sci.*, 28, 2001, (117–123).

196. Stalker H T, Kochert GD, Dhesi JS. Variation within the species *A. duranensis*, a possible progenitor of the cultivated peanut. *Genome*, 38, 1995, (1201-1212).

197. Stalker HT and Simpson CE. Germplasm resources in *Arachis*. In: Advances in Peanut Science. Edited by Pattee H E, Stalker H T. Stillwater, OK: American Peanut Research and Education Society, Inc., 1995, (14-53).

198. Stalker HT; Phillips TG, Murphy JP, Jones TM. Diversity of Isozyme Patterns in Arachis Species, Theoretical & Applied Genetics, 87, 1994, (746-755).

199. Stalker HT. A morphological appraisal of wild species in section *Arachis* of peanuts. Peanut Science, 17, 1990, (117-122).

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

201. Stebbins GL. Genetics, Evolution, and Plant Breeding. Indian J. Genet. Pl. Breed., 17, 1957, (129-141).

201. Stirling CM, Black CR and Ong CK. The response of groundnut (*Arachis hypogae* L.) to timing of irrigation. II. 14C partitioning and plant water status. *Journal Experimental Botany*, 40, 1989, (1363–1373).

202. Subramanian V, Gurtu S, Nageswara Rao RC, Nigam SN. Identification of DNA polymorphism in cultivated groundnut using random amplified polymorphic DNA (RAPD) assay. *Genome*, 43, 2000, (656–660).

203. Subramaniam VB and Maheswari M. Physiological Responses of Groundnut to Water Stress. *Indian Journal of Plant Physiology*, 2, 1990, (130–135).

204. Sujay V, Gowda MVC, Pandey MK, Bhat RS, Khedikar YP, et al. Quantitative trait locus analysis and construction of consensus genetic map for foliar disease resistance based on two recombinant inbred line populations in cultivated groundnut (*Arachis hypogaea* L.). Mol Breed, 30, 2012, (773–788).

205. Suther DM and Patel MS. Yield and Nutrient Absorption by Groundnut and Iron Availability in Soil as Influenced By Lime and Soil Water, *Journal of Indian Society of Soil Science*, 40, 1992, (594–596).

206. Tang R, Gao G, He L, Han Z, Shan S, Zhong R, Zhou C, Jiang J, Li Y, Zhuang W. Genetic Diversity in Cultivated Groundnut Based on SSR Markers. *J. Genet Gen.*, 34, 2007, (449-459).

207. Tanksley SD. Mapping polygenes. Annu Rev Genet, 27, 1993, (205–233).

208. Thiel et al., *MicroSatellite (MISA) Perl script for mining the SSRs* (http://pgrc.ipk-gatersleben.de/misa/), 2003.

209. Ungerer MC, Halldorsdottir SS, Modliszewski JL, Mackay TFC, Purugganan MD. Quantitative trait loci for inflorescence development in Arabidopsis thaliana. Genetics, 160, 2002, (1133–1151).

210. Upadhyaya HD, Bhattacharjee R, Varshney RK, Gowda CLL, Hoisington DA. ICRISAT Archival Report. International Crops Research Institute for the Semi Arid Tropics, Patancheru, 502 324, 2007.

211. Urao T, Yakubov B, Satoh R, Yamaguchi-Shinozaki K, Seki M, Hirayama T, Shinozaki K. A transmembrane hybrid-type histidine kinase in Arabidopsis functions as an osmosensor. Plant Cell, 11, 1999, (1743-1754).

212. Van Berloo R, Zhu A, Ursem R, Varbakel H, Gort G, van Eeuwijk FA. Diversity and linkage disequilibrium analysis within a selected set of cultivated tomatoes. Theor Applied Genet, 117, 2008, (89-101).

213. Varshney RK, Bertioli DJ, Moretzsohn MC, Vadez V, Krishnamurthy L, et al. The first SSR- based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). Theor Appl Genet, 118, 2009a, (729–739).

214. Varshney RK, Mahendar T, Aruna R, Nigam SN, Neelima K, Vadez V, Hoisington DA. High level of natural variation in a groundnut (*Arachis hypogaea* L.) germplasm collection assayed by selected informative SSR markers. Plant Breed, 128, 2009b, (486–494).

215. Varshney RK, Dubey A. Novel genomic tools and modern genetic and breeding approaches for crop improvement. J Plant Biochem Biotechnol, 18, 2009, (127–138).

216. Varshney RK and Tuberosa R. Genomics-assisted crop improvement: an overview. In: Varshney RK, Tuberosa R (eds) Genomic Assisted Crop Improvement: Genomics Approaches and Platforms. Volume I, Springer, The Netherlands, 2007a, (1-12).

217. Varshney RK, Marcel TC, Ramsay L, Russell J, Röder MS, et al. A high density barley microsatellite consensus map with 775 SSR loci. Theor Appl Genet, 114, 2007b, (1091–1103).

218. Varshney RK, Hoisington DA, Tyagi AK.: Advances in cereal genomics and applications in crop breeding. Trends Biotech, 24, 2006, (490–499).

219. Varshney RK, Garner A, Sorrells ME. Genomics-assisted breeding for crop improvement. Trends Plant Sci, 10, 2005, (621-630).

220. Voorrips RE: Mapchart. Software for the graphical presentation of linkage maps and QTLs. J Hered, 93, 2002, (77–78).

221. Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hornes, M, Fritjers A, Pot J, Paleman J, Kuiper M, Zabeau M. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res*, 23, 1995, (4407-4414).

222. Wang S, Basten CJ, Zeng ZB. Windows QTL cartographer 2.5. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm. 2007.

223. Wang CT, Yang XD, Chen DX, Yu SL, Liu GZ, Tang YY, Xu JZ. Isolation of simple sequence repeats from groundnut. Electronic J. Biotech., 10, 2007, (473-480).

224. Wang WX, Vinocur B, Shoseyov O, Altman A. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends Plant Sci., 9, 2004, (244-252).

225. Wang DL, Zhu J, LiZ K, Paterson AH. Mapping QTLs with epistatic effects and genotype x environment interactions by mixed linear model approaches. Theor Appl Genet, 99, 1999, (1255-1264).

226. Wenzl P, Li H, Carling J, Zhou M, Raman H, Paul E, Hearnden P, Maier C, Xia L, Caig V, Ovesná J, Cakir M, Poulsen D, Wang J, Raman R, Smith KP, Muehlbauer GJ, Chalmers KJ, Kleinhofs A, Huttner E, Kilian A. A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics, 12, 2006, (7-20).

227. Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*, 18, 1990, (6531-6535).

228. Wu Y, Close TJ, Lonardi S.: On the accurate construction of consensus genetic maps. In *Proceedings of LSS Computational Systems Bioinformatics Conference: 26-29 August2008; Stanford* Edited by Peter Markstein, 2008, (285-296).

229. Wu Y, Bhat PR, Close TJ, Lonardi S. Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. PLoS Genet, 4, 2008, (e1000212).

230. Xiao B, Huang Y, Tang N, Xiong L. Overexpression of LEA gene in rice improves drought resistance under field conditions. Theoretical Applied Genet, 115, 2007, (35-46).

231. Yang J, Hu CC, Ye XZ, Zhu J. QTLNetwork 2.0. Available at http://ibi.zju.edu. cn/software/qtlnetwork. Institute of Bioinformatics, Zhejiang University, Hangzhou, China, 2005.

232. Yang L, Zheng B, Mao C, Qi X, Liu F, Wu P. Analysis of transcripts that are differentially expressed in three sectors of the rice root system under water deficit. Mol Genet Genomics, 272, 2004, (433-442).

233. Young ND, Weeden NF, Kochert G. Genome mapping in legumes (Fam. Fabaceae). In: Paterson AH, Austin (eds) Genome mapping in plants. Landes Company, USA, 1996, (211–227).

234. Yoon DB, Kang KH, Kim HJ, Ju HG, Kwon SJ, Suh JP, Jeong OY,Ahn SN. Mapping quantitative trait loci for yield components and morphological traits in an advanced backcross population between Oryza grandiglumis and the O. sativa japonica cultivar Hwaseongbyeo. Theor Appl Genet, 112, 2006, (1052–1062).

235. Yuan M, Gong L, Meng R, Li S, Dang P, Guo B, He G. Development of trinucleotide $(GGC)_n$ SSR markers in peanut (Arachis hypogaea L.). Electronic J Biotechnol; DOI: 10.2225/vol13-issue6-fulltext-6, 2010.

236. Zaman-Allah M, Jenkinson DM, Vadez V. Chickpea genotypes contrasting for seed yield under terminal drought stress in the field differ for traits related to the control of water use. Funct Plant Biol, 38, 2011a, (270-281).

237. Zaman-Allah M, Jenkinson DM , Vadez V. A conservative pattern of water use, rather than deep or profuse rooting, is critical for the

terminal drought tolerance of chickpea. J Exp Bot (in press), 2011b.

238. Zeng ZB. Precision mapping of quantitative trait loci. Genetics, 136, 1994, (1457–1468).

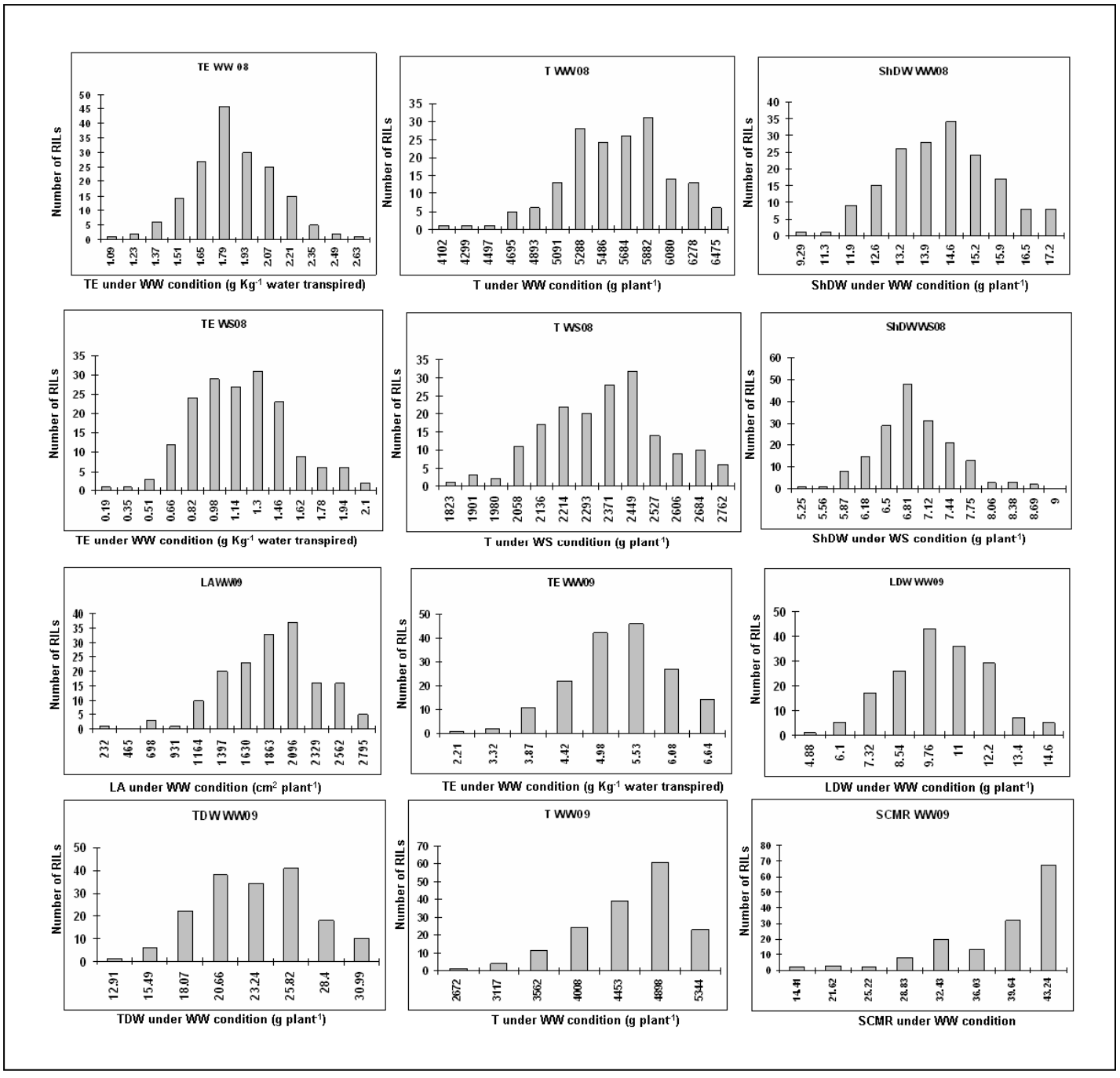
239. Zeng ZB. Theoretical basis of separation of multiple linked gene effects in mapping quantitative trait loci. Proc Natl Acad Sci. 90. 1993, (10972–10976).

Appendix 1: Trait phenotyping data on parents of ICGS 76 × CSMG 84-1 and ICGS 44 × ICGS 76 mapping populations

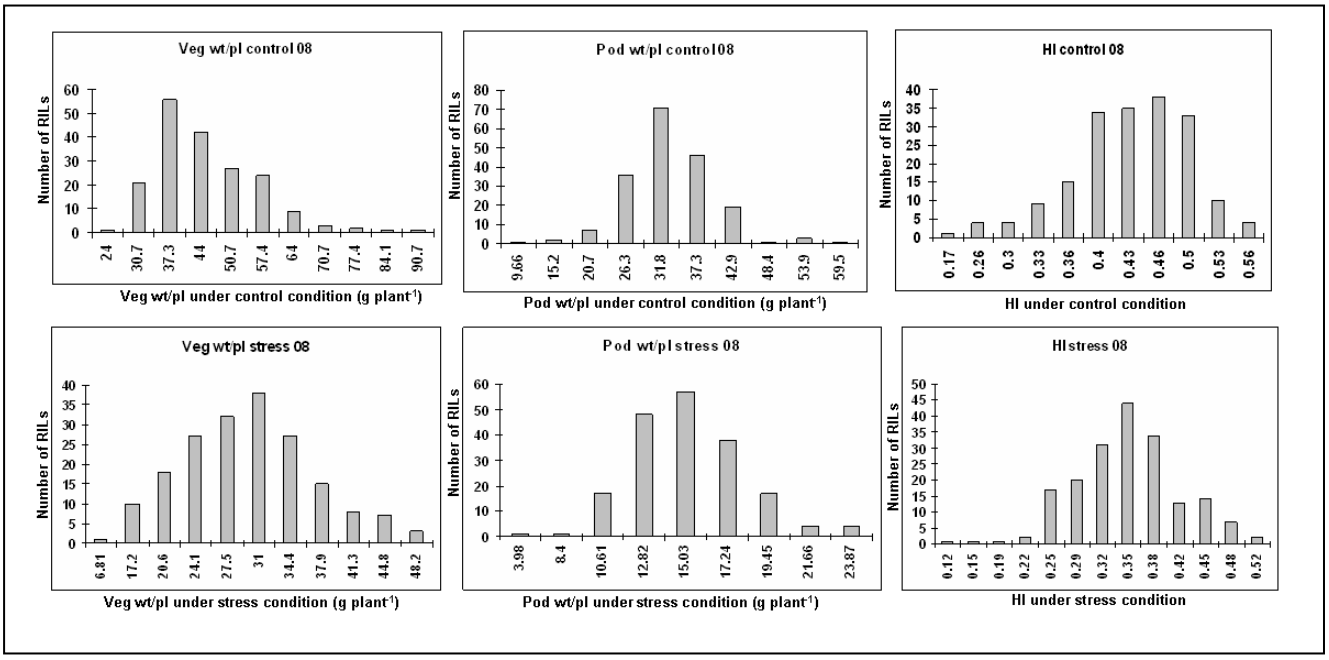
Trait_Year	Parent 1	Parent 2	Variation in RILs	Significance	Heritability (broad sense)	Grand mean	SED	LSD
ICGS 76 × CSMG 84-1								
ShootDW_WW08	14.96	13.77	9.29-17.84	0.5001*	5.61	14.19	2.49	4.89
TotalTR_WW08	4468	5715	4102-6917	0.0128**	19.71	5557.06	707.61	1386.92
TEDW_WW08	2.29	1.48	0.98-3.08	0.021**	2.79	1.82	0.50	0.98
ShootDW_WS08	7.14	5.99	5.11-9.32	0.1493*	3.99	6.75	1.13	2.21
TotalTR_WS08	2327	2169	1824-2841	0.4649*	11.52	2326.87	305.40	598.58
TEDW_WS08	1.10	0.34	0.19-2.27	0.0267**	26.89	1.11	0.48	0.95
LeafArea_WW09	1927	1786	210-2309	0.824*	7.38	1810.12	896.20	1756.55
LeafDW_WW09	11.54	9.44	1.77-8.03	0.3533*	13.38	9.65	4.52	8.85
TotalDW_WW09	23.54	21.34	2.73-15.98	0.6045*	7.49	22.11	8.53	16.72
Transpiration_WW09	5000	3766	12.32-33.57	0.0697**	7.55	4415.53	1361.85	2669.23
TE_WW09	4.79	5.87	2515-5827	0.2097*	6.94	5.08	1.72	3.37
Spad_WW09	40.95	43.10	11.68-47.18	0.6045*	6.26	38.55	14.57	28.55
ICGS 44 × ICGS 76								
VegWt/pl_control08	53.99	47.37	8.41-251.33	0.5284*	10.17	41.38	33.82	66.29
PodWt/pl_control08	36.87	33.98	4.16-103.20	0.5578*	2.25	30.48	18.30	35.88
HI_control08	0.40	0.42	0.08-0.61	0.6643*	27.85	0.43	0.09	0.18
VegWt/pl_stress08	34.96	35.82	7.36-55.25	0.9799*	27.15	27.94	8.88	17.41
PodWt/pl_stress08	16.14	12.38	7.06-39.43	0.2553*	13.62	14.43	4.88	9.57
HI_stress08	0.31	0.26	0.17-0.61	0.1994*	45.01	0.35	0.06	0.13

**significance at 0.05; **significance at 0.01; SED: Standard error of difference; LSD: Least significant difference*

Appendix 2A: Frequency distribution of selected drought tolerance related traits in ICGS 76 × CSMG 84-1 mapping populations



Appendix 2B: Frequency distribution of selected drought tolerance related traits in ICGS 44 × ICGS 76 mapping populations



Appendix 3A: Comparison of M-QTLs identified by QTL Cartographer and QTL Network using CIM method in ICGS 76 × CSMG 84-1 mapping population

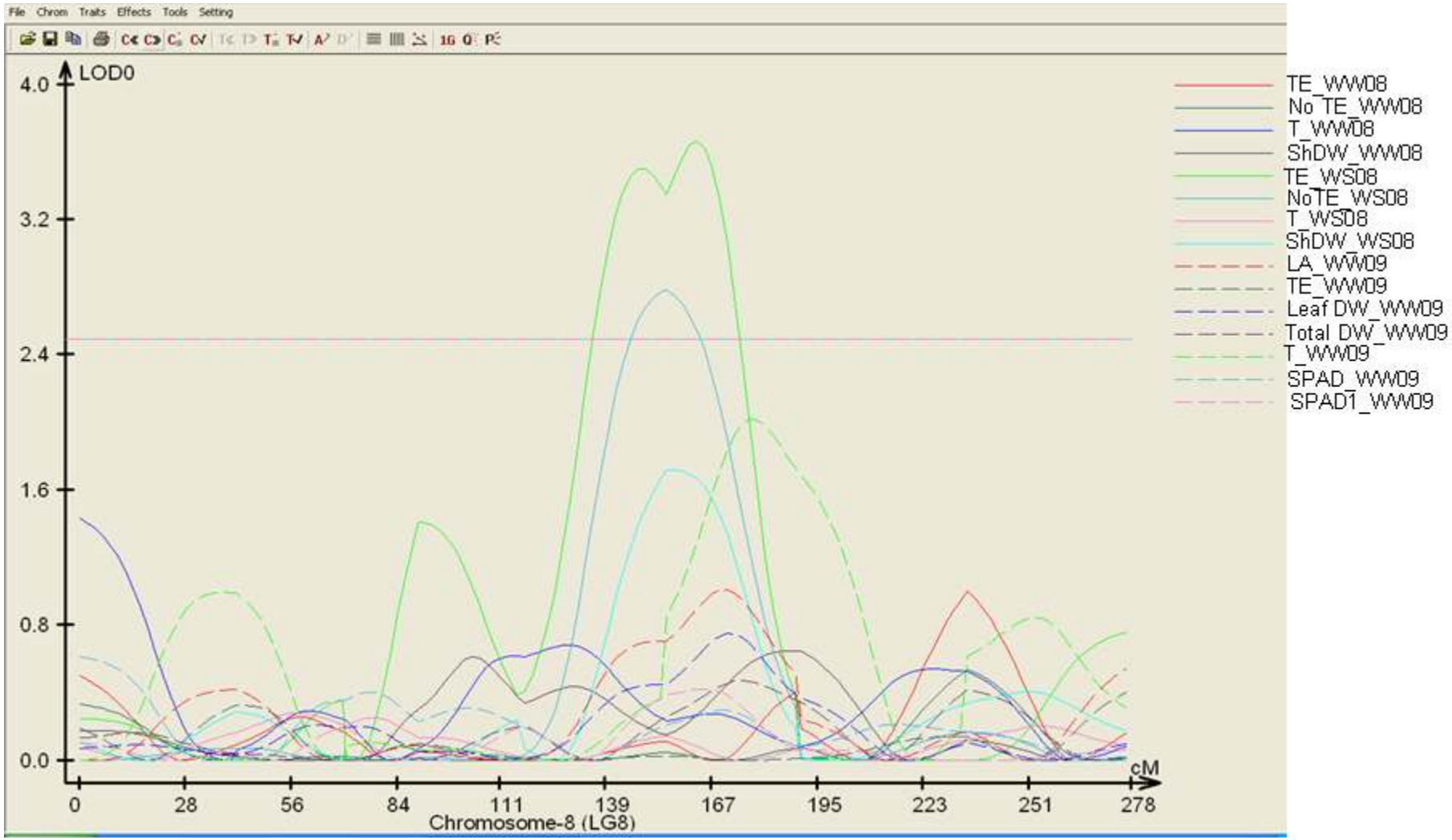
Traits	QTLs identified using QTLCartographer					QTL identified using QTL Network			
	Marker Interval	Position (cM)	Highest LOD score	Range of PVE (R^2 %)	Additive effect	Marker Interval	Position (cM)	Range of PVE (R^2 %)	Additive effect
TEWW08_AhVI	Ah4-101-GM2536	34.71	2.20	7.24	-0.076	Ah4-101-GM2536	36.3	4.75	-0.06
TEWS08_AhVII	PM204	118.01	2.88	6.47	0.101	-	-	-	-
TEWS08_AhVIII	TC9F04-GM2407	147.61	3.51	14.50	-0.144	-	-	-	-
TEWS08_AhVIII	GM2407-TC1B02	165.11	3.63	18.12	-0.161	-	-	-	-
TEWW08_AhV	-	-	-	-	-	GM1571-GM1577	22	3.31	0.09
TEWS08_AhV	TC2D08-Ah4-04	155.81	2.37	13.44	-0.123	-	-	-	-
TEWW09_AhXIII	TC4A02	28.01	2.24	5.63	-0.219	-	-	-	-
TEWW09_AhIX	-	-	-	-	-	TC7E04-GM1949	8	3.64	0.24
					-	-	-	-	-
TWW08_AhI	GM2724b-S54	10.81	3.02	7.52	138.436				
					-	-	-	-	-
TWW08_AhVII	GM1979-GM1919	8.01	2.24	9.64	163.526				
TWW08_AhXI	S1	112.11	2.06	4.82	113.061	-	-	-	-

TWS08_AhIX	TC7E04-GM1949	16.01	2.22	18.17	85.725	TC7E04-GM1949	12	3.21	51.02
TWS08_AhXIX	GM2547	1.41	2.10	4.84	44.592	-	-	-	-
TWW09_AhV	gi-832-TC11H06	237.01	2.87	7.19	166.579	-	-	-	-
					-	-	-	-	-
TWW09_AhVIII	GM2407-TC1B02	179.11	2.02	10.30	196.104				
TWW09_AhXIII	TC4A02	28.01	2.08	4.83	137.689	-	-	-	-
					-	-	-	-	-
TWW09_AhXX	GM2325	8.01	2.37	12.64	268.227				
Total DWWW09_AhV	gi-832-TC11H06	237.01	2.73	6.91	1.081	gi-832-TC11H06	237	6.04	0.76
Total DWWW09_AhIX	TC7E04-GM1949	24.01	2.01	22.39	1.943	-	-	-	-
Total DWWW09_AhXV	GM1021-GM1570	72.81	2.49	6.62	-1.097	-	-	-	-
ShDWWW08_AhIX	TC7E04-GM1949	14.01	2.21	14.40	0.578	TC7E04-GM1949	3	5.5	0.42
ShDWWW08_AhXI	S1	112.11	2.09	5.26	0.357	-	-	-	-
ShDWWW08_AhXV	GM1021-GM1570	64.81	2.45	9.39	-0.467	-	-	-	-
ShDWWS08_AhIV	PM418	24.81	2.20	5.03	0.148	-	-	-	-
ShDWWS08_AhVII	GM1979-GM1919	16.01	3.37	22.09	-0.312	-	-	-	-
SPADWW09_AhXIX	-	-	-	-	-	GM1971-Seq2D12B	34.4	2.51	1.81

Appendix 3B: Comparison of M-QTLs identified by QTL Cartographer and QTL Network using CIM method in ICGS 44 × ICGS76 mapping population

	QTLs identified using QTLCartographer					QTL identified using QTL Network			
	Marker Interval/associated markers	**Position (cM)**	**Highest LOD score**	**Range of PVE (R^2 %)**	**Additive effect**	**Marker Interval/Associated markers**	**Position (cM)**	**Range of PVE (R^2 %)**	**Additive effect**
HI Control 08_AhI	GM1959-IPAHM97	6.01	2.30	6.56	-0.032	-	-	-	-
HI Control 08_AhIX	GM1922-GM2050	12.01	2.42	40.10	-0.050	-	-	-	-
HI Control 08_AhXV	Seq17F06	0.00	2.66	6.39	0.042	-	-	-	-
HI Stress 08_AhXV	-	-	-	-	-	TC11H06-GM2584	64.2	3.29	0.02
Veg wt/pl Control 08_AhIX	-	-	-	-	-	GM1922-GM2050	0	2.28	5.99

Appendix 4A: QTL LOD Peak for all the traits of LG 8 in ICGS 76 × CSMG 84-1 mapping populations



Appendix 4B: QTL LOD Peak for all the traits of LG 9 in ICGS 44 × ICGS 76 mapping populations

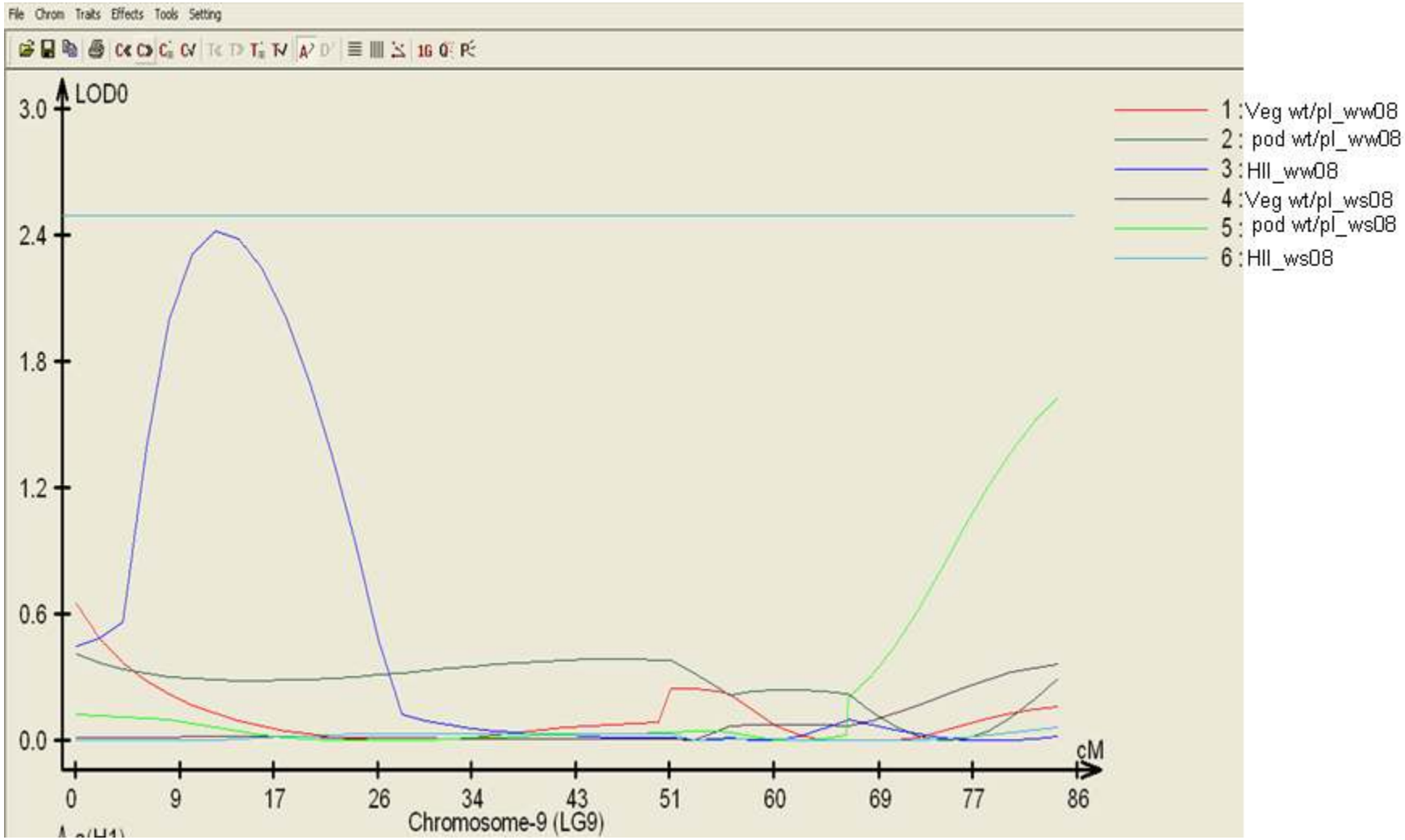
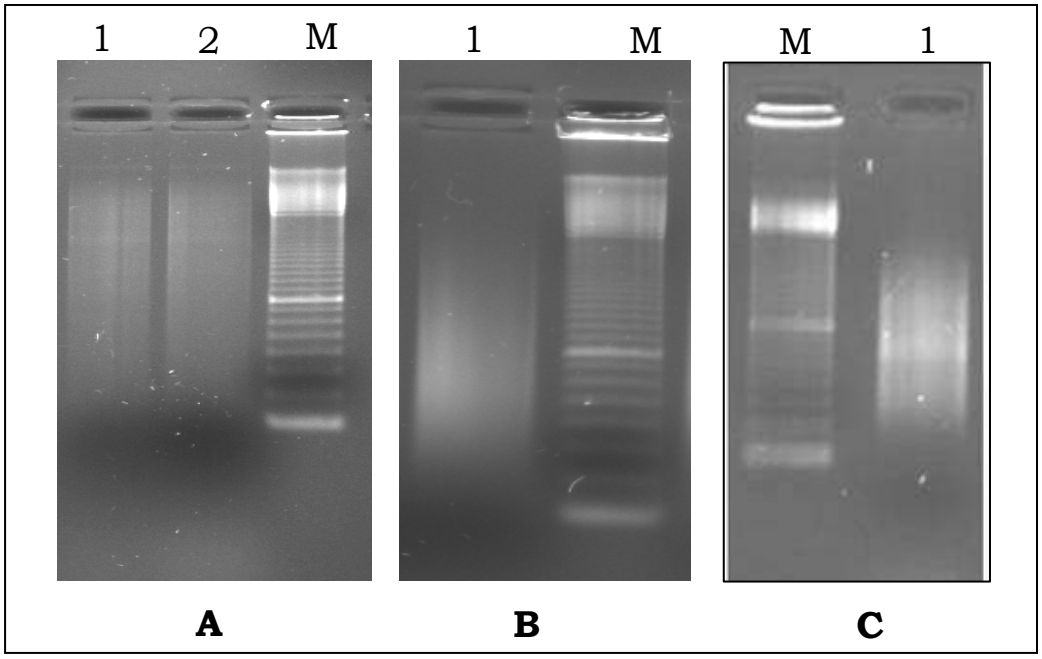


Figure 1: Steps involved in generating SSR enriched libraries



(A) Restriction digestion of genomic DNA, RD= Digested genomic DNA M- 100 bp- DNA ladder, (B) Linker ligation, 1- Linker ligated DNA, M- 100 bp- DNA ladder**,** (C) Enriched SSR genomic DNA fragments, M- 100 bp- DNA ladder, 1- Enriched DNA.

Figure 2: A representative amplification profile of colony PCR screening for the presence of SSR containing inserts.

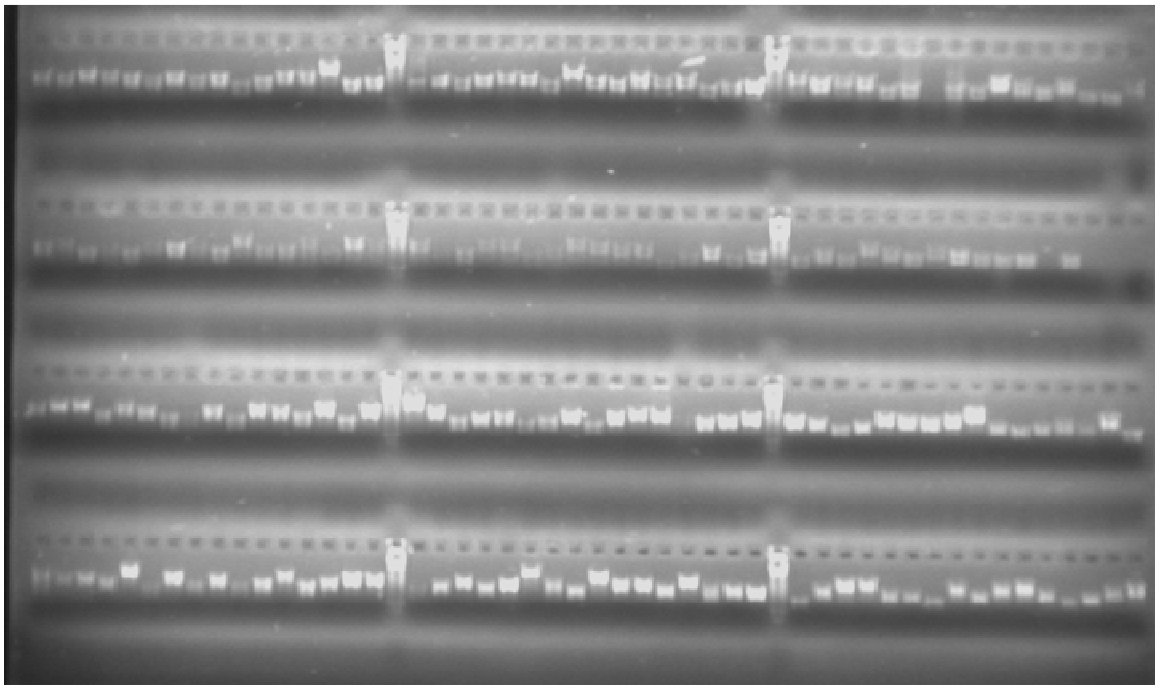


Figure 3A: Genetic linkage map of ICGS 76 × CSMG 84-1

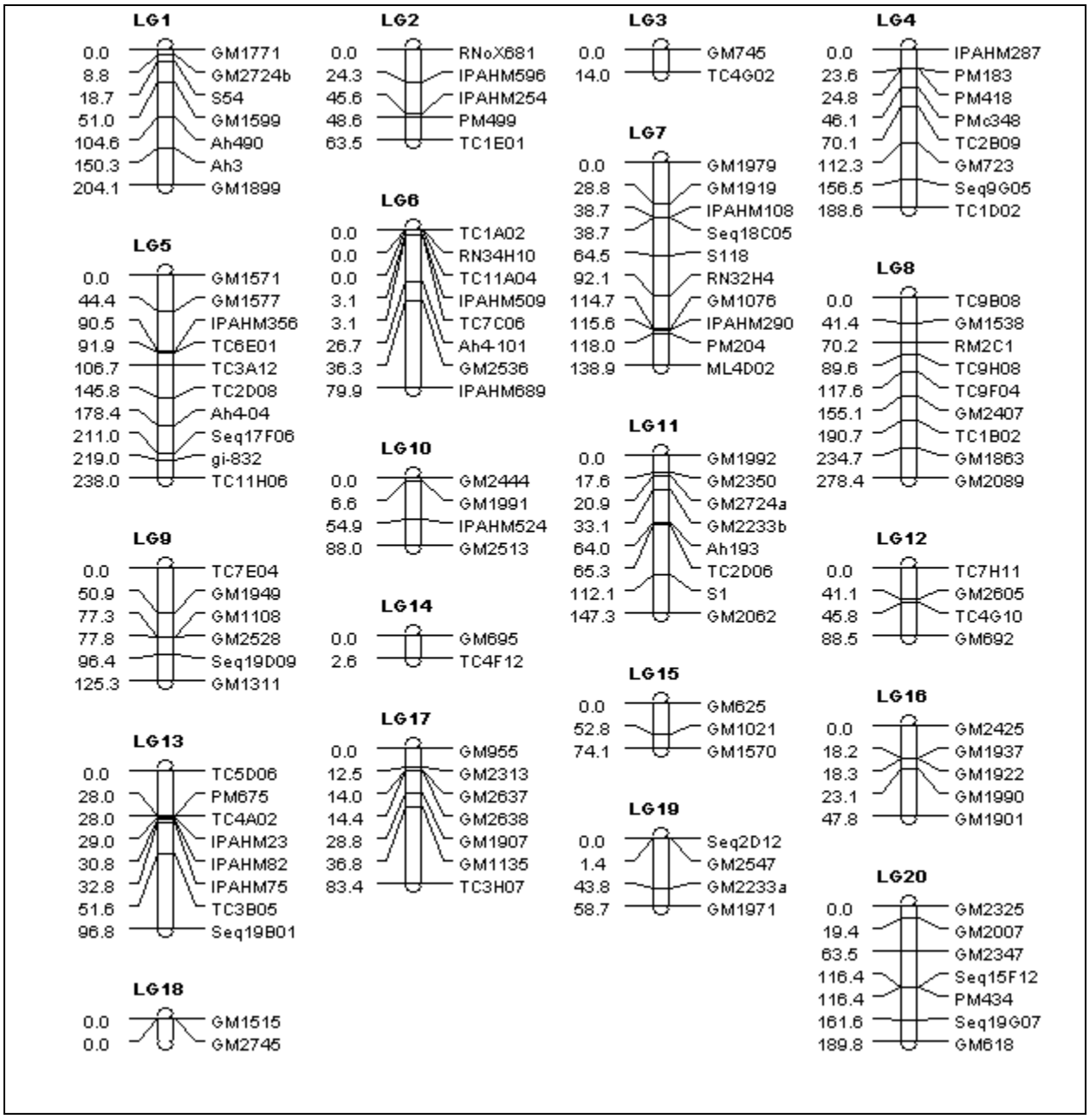


Figure 3B: Genetic linkage map of ICGS 44 × ICGS76

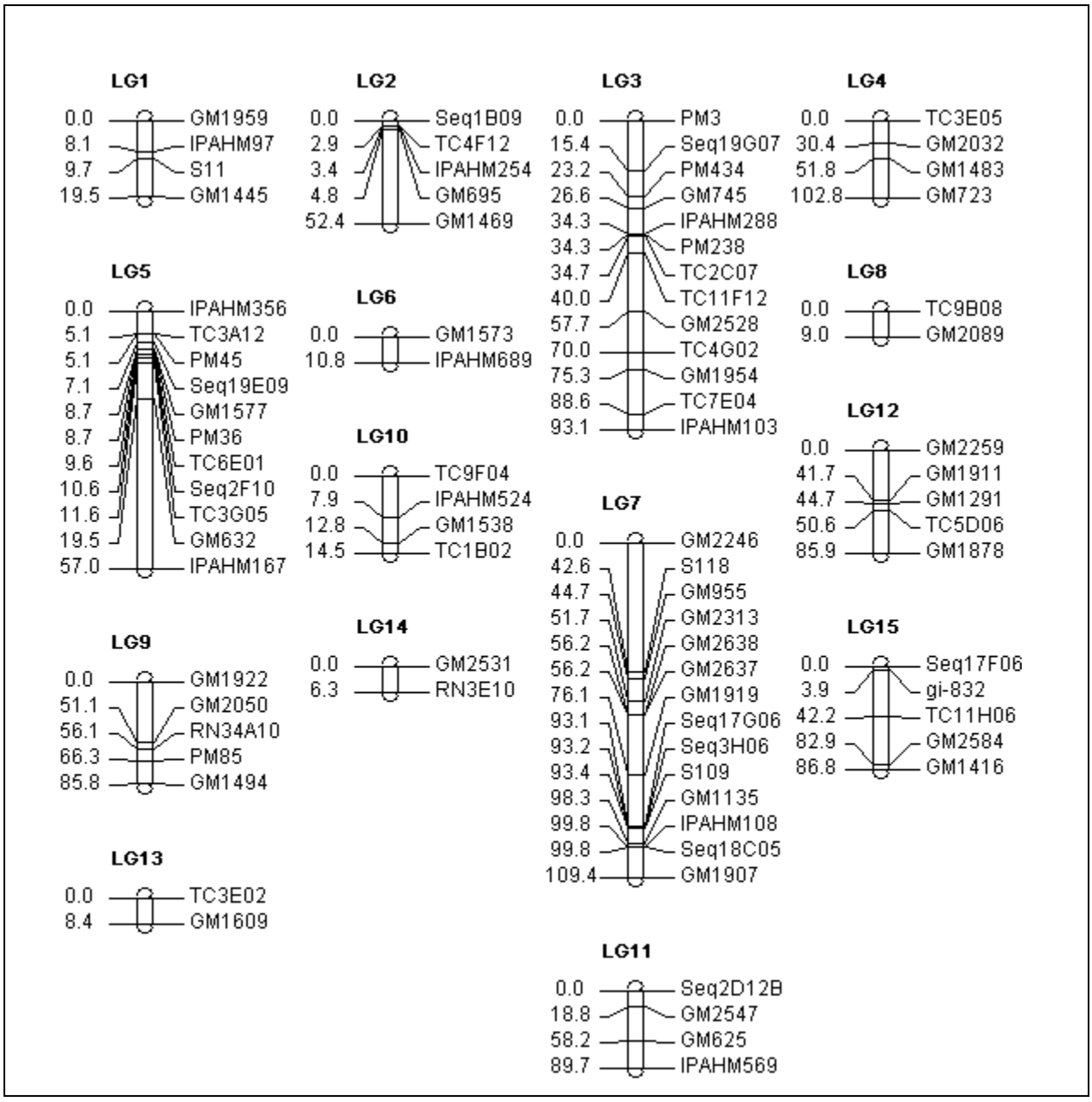
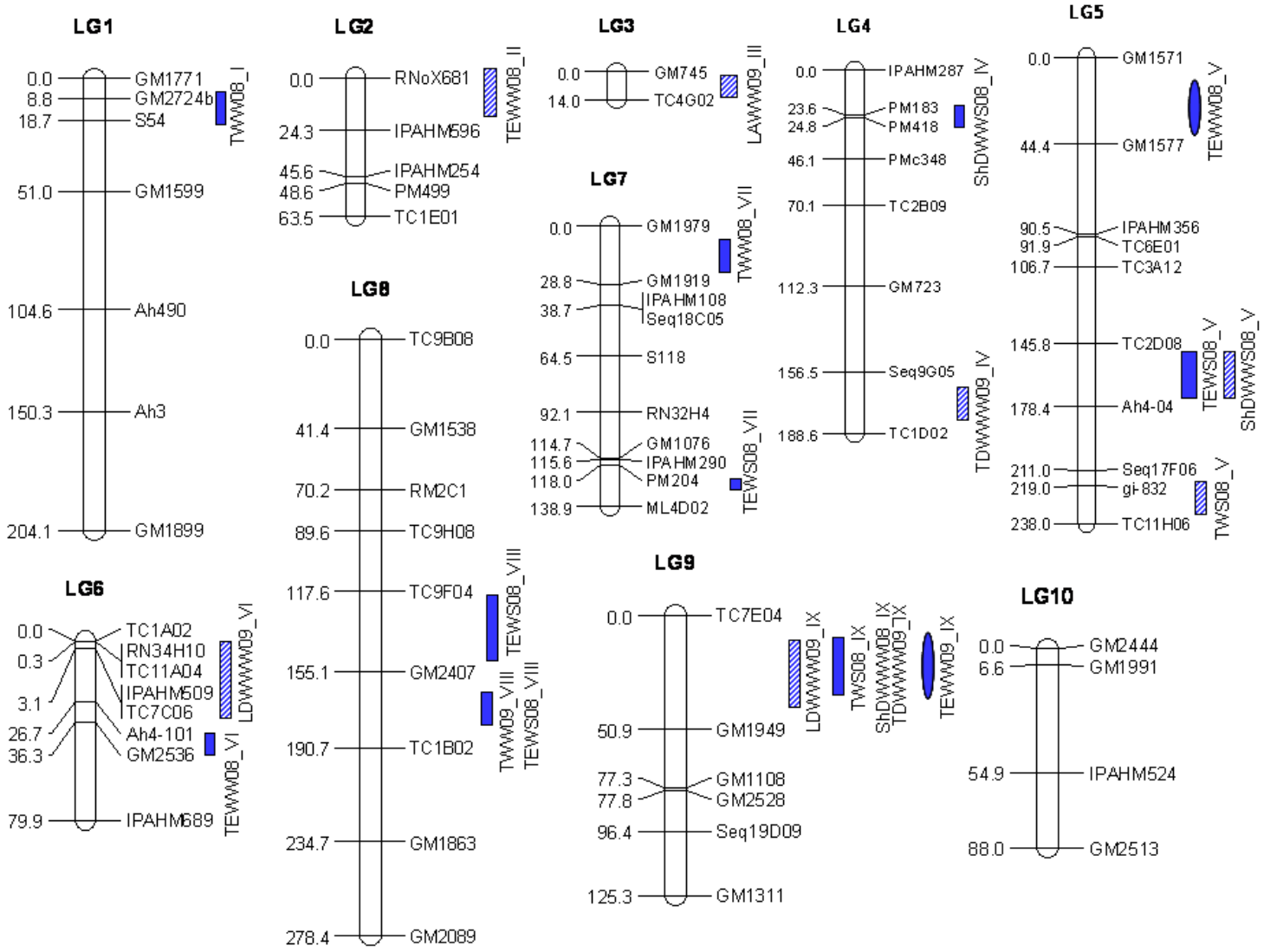


Figure 4A: Genetic linkage map showing QTLs for drought related traits in ICGS 76 × CSMG 84-1



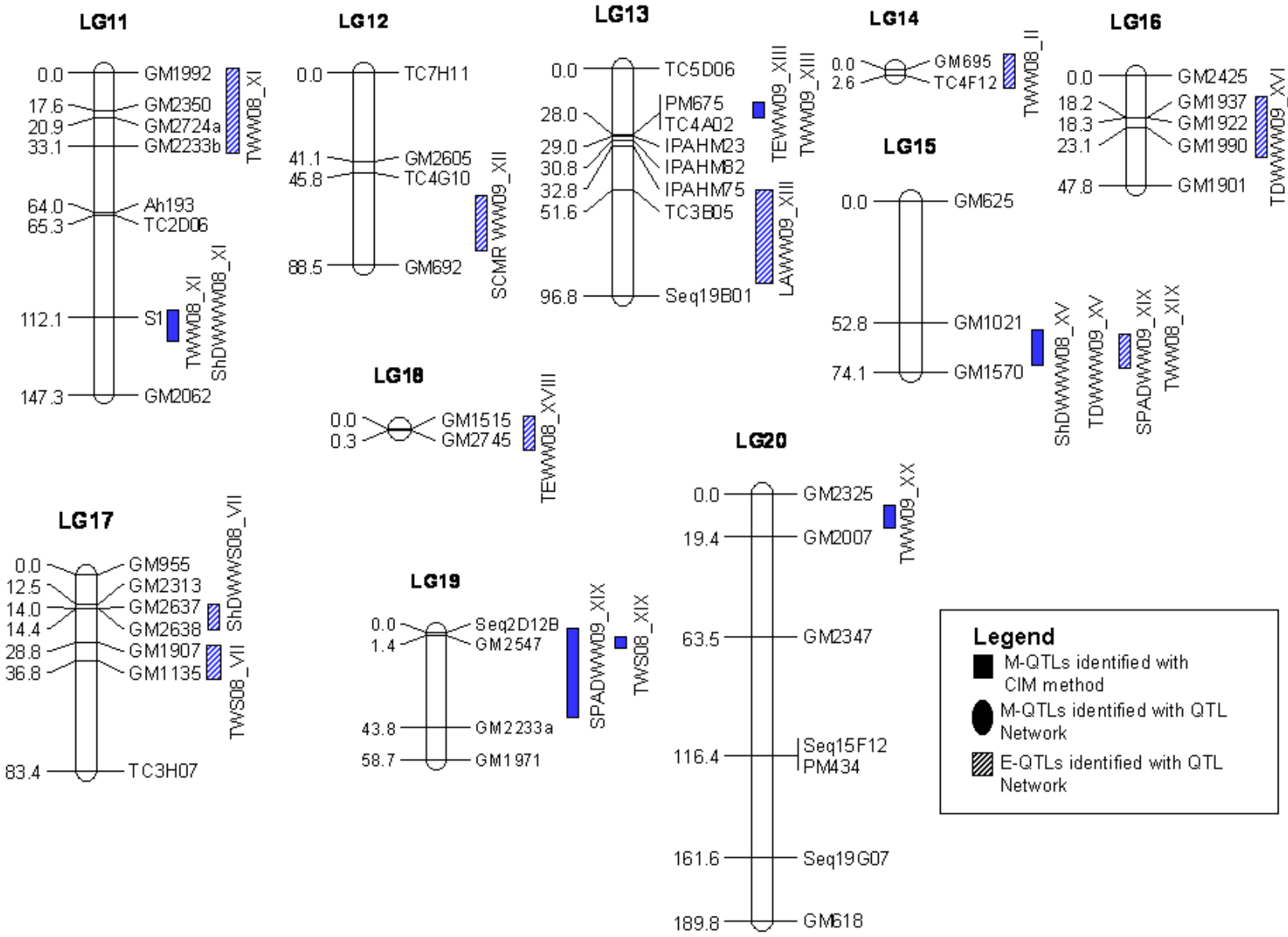
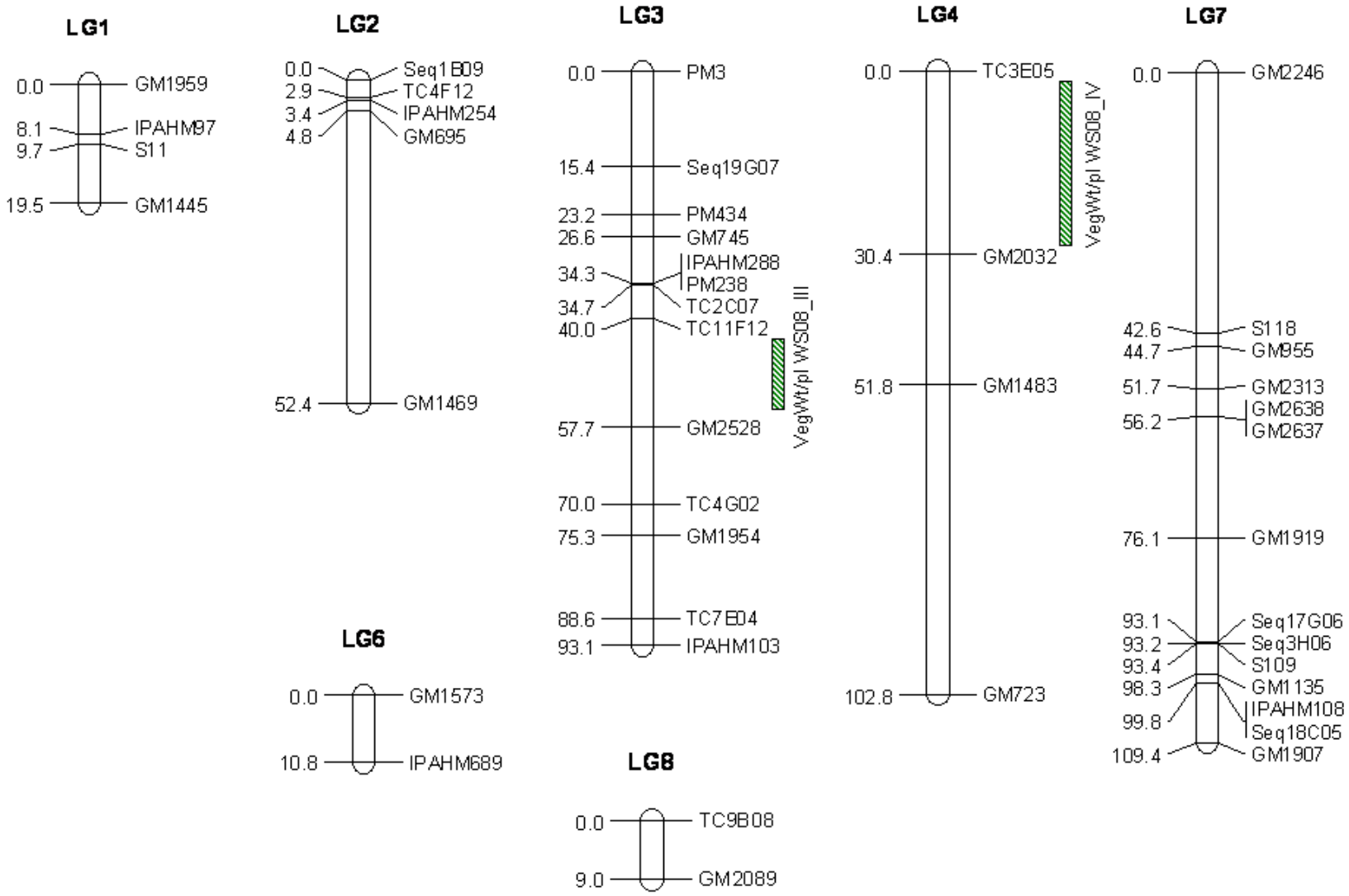

LG11
0.0 GM1992
17.6 GM2350
20.9 GM2724a
33.1 GM2233b
64.0 Ah193
65.3 TC2D06
112.1 S1
147.3 GM2062
TWWW08_XI
ShDWWWW08_XI
TWWW08_XI
LG12
0.0 TC7H11
41.1 GM2605
45.8 TC4G10
88.5 GM692
SCMR WWW09_XII
LG13
0.0 TC5D06
28.0 PM675
29.0 TC4A02
30.8 IPAHM23
32.8 IPAHM82
51.6 IPAHM75
96.8 Seq19B01
TEWWW09_XIII
TWWW09_XIII
LAWWW09_XIII
LG14
0.0 GM695
2.6 TC4F12
TWWW08_II
LG15
0.0 GM625
52.8 GM1021
74.1 GM1570
ShDWWWW08_XV
TDWWWW09_XV
SPADWWW09_XIX
TWWW08_XIX
LG16
0.0 GM2425
18.2 GM1937
18.3 GM1922
23.1 GM1990
47.8 GM1901
TDWWWW09_XVI
LG18
0.0 GM1515
0.3 GM2745
TEWWW08_XVIII
LG17
0.0 GM955
12.5 GM2313
14.0 GM2637
14.4 GM2638
28.8 GM1907
36.8 GM1135
83.4 TC3H07
ShDWWWS08_VII
TWS08_VII
LG19
0.0 Seq2D12B
1.4 GM2547
43.8 GM2233a
58.7 GM1971
SPADWWW09_XIX
TWS08_XIX
LG20
0.0 GM2325
19.4 GM2007
63.5 GM2347
116.4 Seq15F12
PM434
161.6 Seq19G07
189.8 GM618
TWWW09_XX
Legend
M-QTLs identified with CIM method
M-QTLs identified with QTL Network
E-QTLs identified with QTL Network

Figure 4B: Genetic linkage map showing QTLs for drought related traits in ICGS 44 × ICGS 76



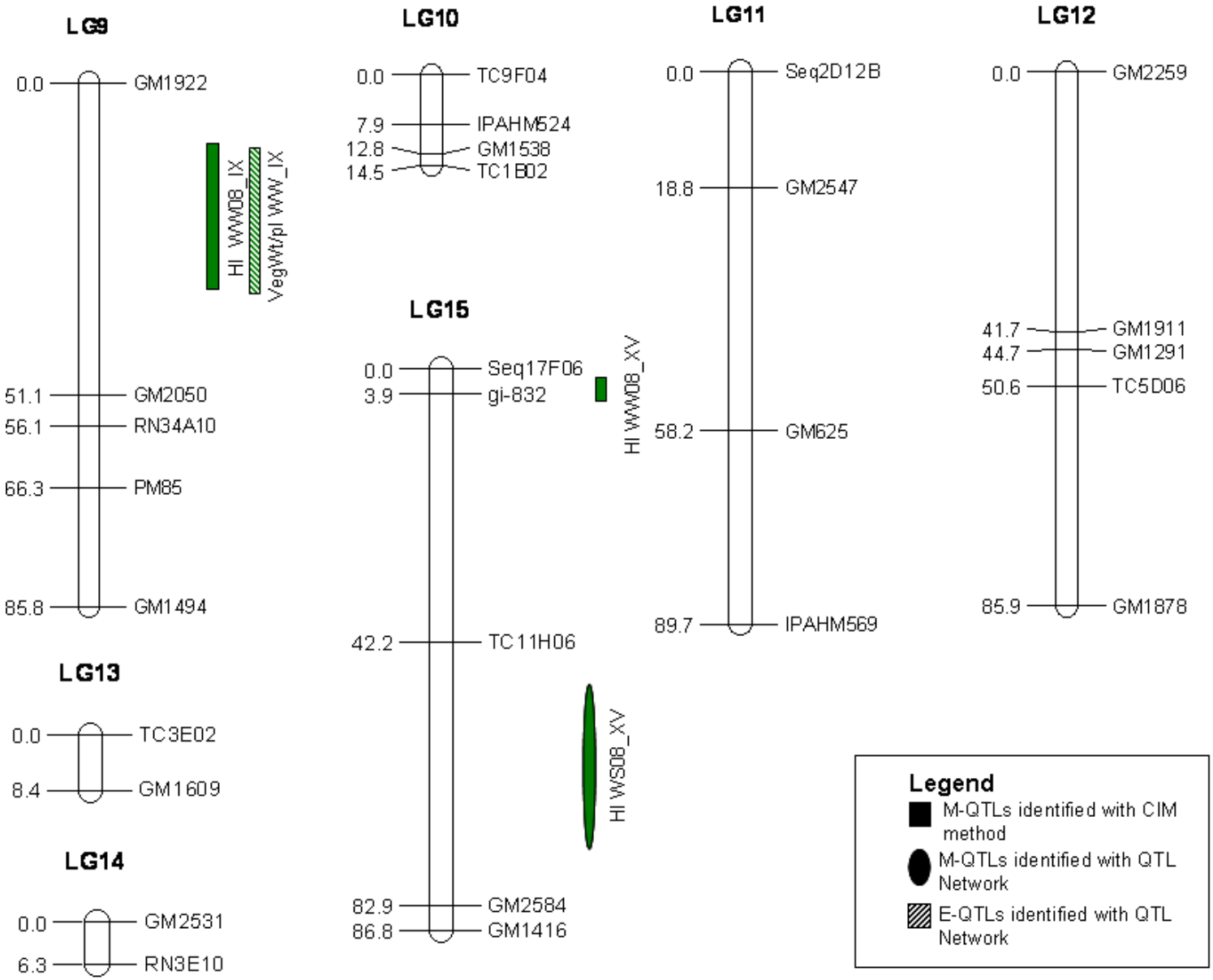

LG9
0.0 GM1922
HI WWW08_IX
VegWt/pl WWW_IX
51.1 GM2050
56.1 RN34A10
66.3 PM85
85.8 GM1494
LG10
0.0 TC9F04
7.9 IPAHM524
12.8 GM1538
14.5 TC1B02
LG11
0.0 Seq2D12B
18.8 GM2547
58.2 GM625
89.7 IPAHM569
LG12
0.0 GM2259
41.7 GM1911
44.7 GM1291
50.6 TC5D06
85.9 GM1878
LG15
0.0 Seq17F06
3.9 gi-832
HI WWW08_XV
42.2 TC11H06
HI WS08_XV
82.9 GM2584
86.8 GM1416
LG13
0.0 TC3E02
8.4 GM1609
LG14
0.0 GM2531
6.3 RN3E10
Legend
M-QTLs identified with CIM method
M-QTLs identified with QTL Network
E-QTLs identified with QTL Network

Figure 5A: The marker loci interaction for transpiration efficiency (TEWS) in ICGS 76 × CSMG 84-1 using GMM software.

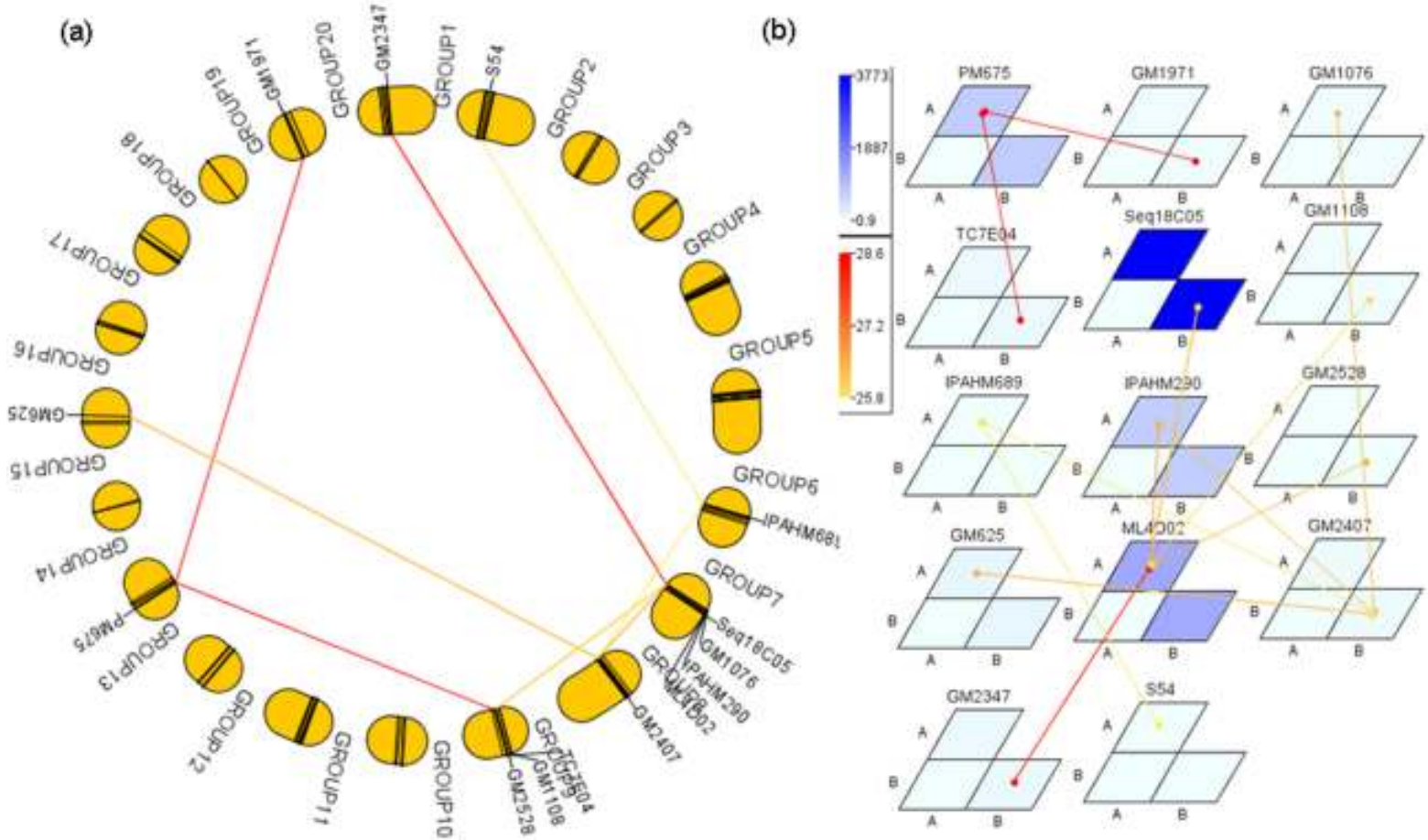


Figure 5B: The marker loci interaction for pod weight/plant in ICGS 44 × ICGS 76 using GMM software.

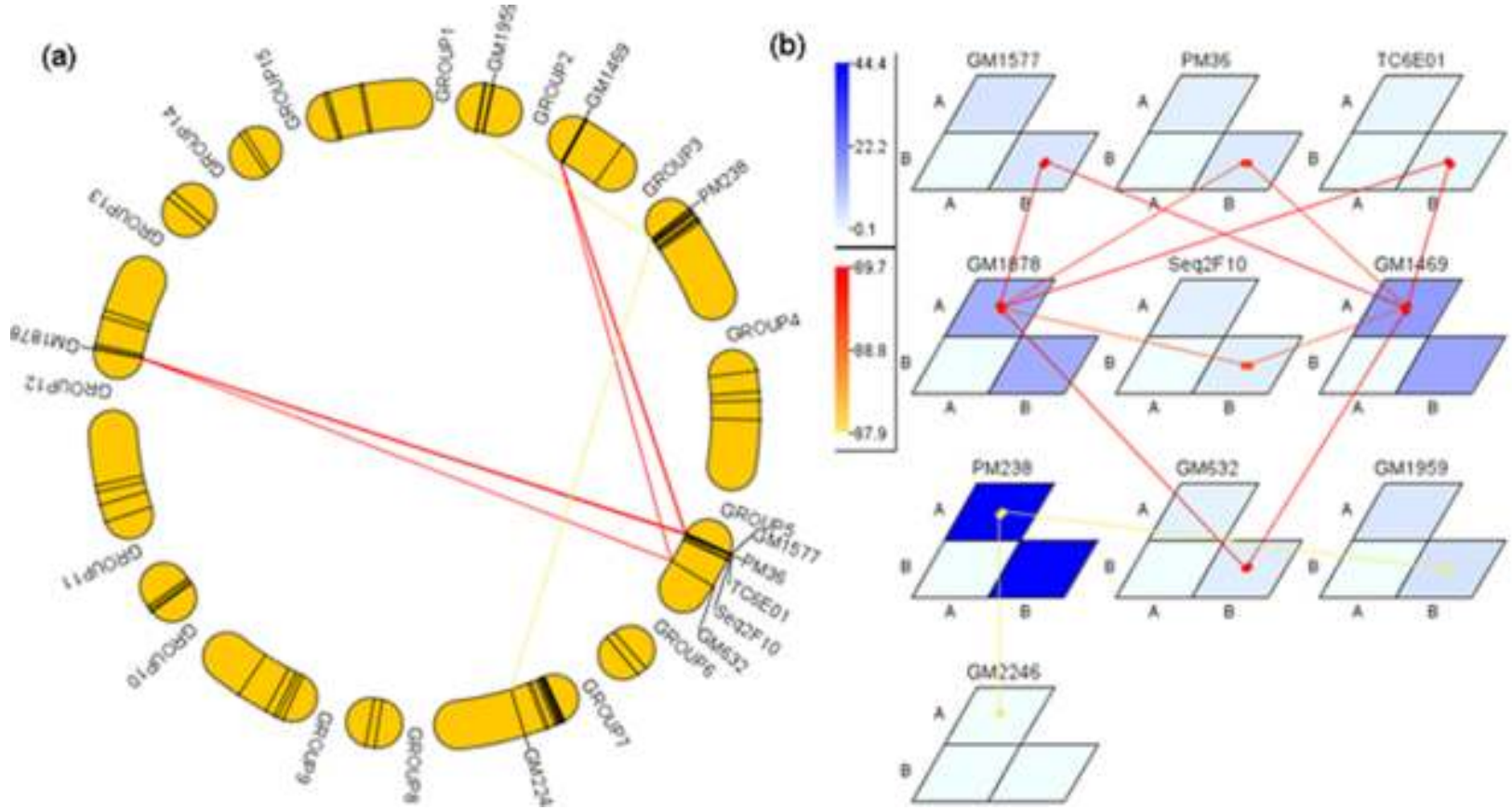
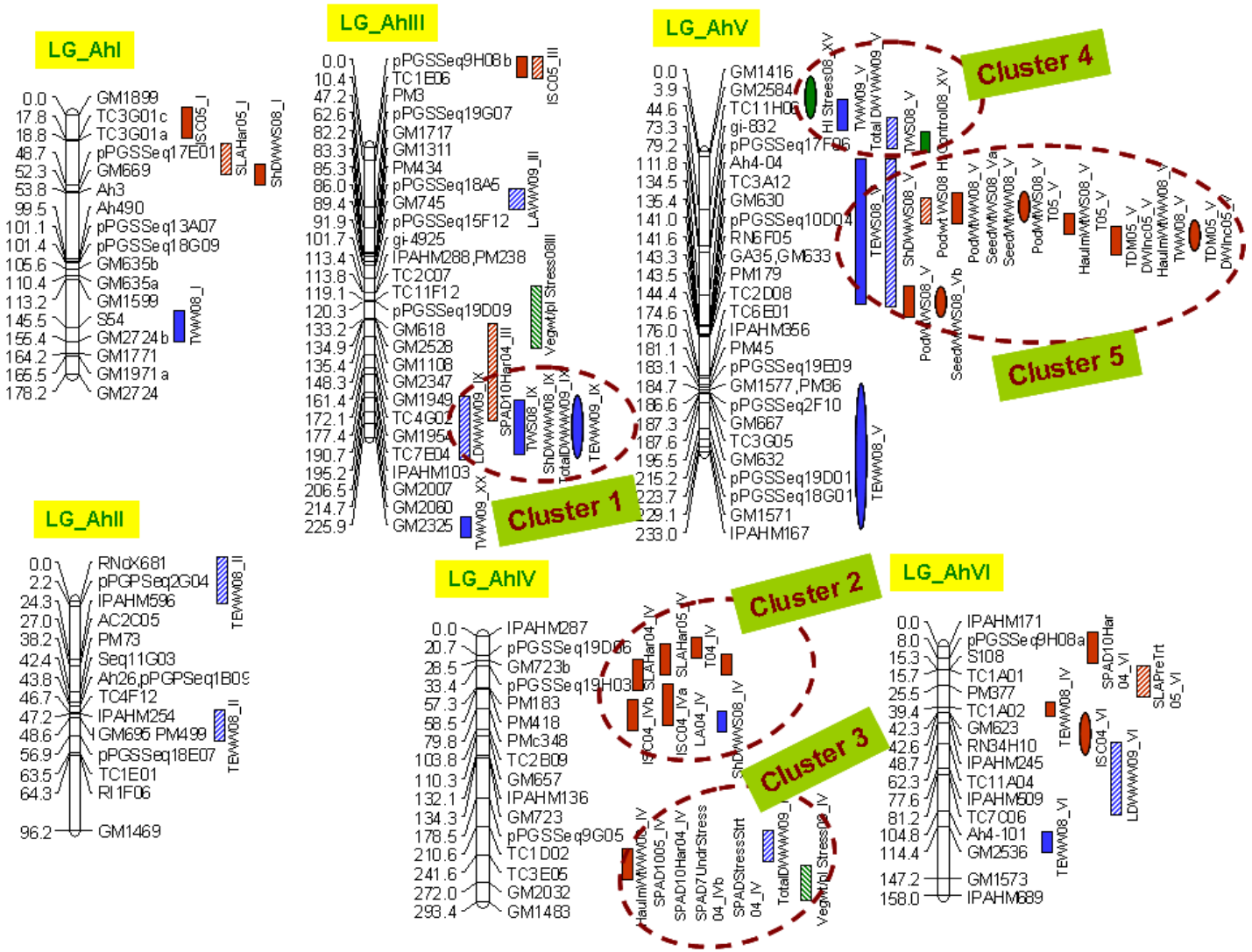
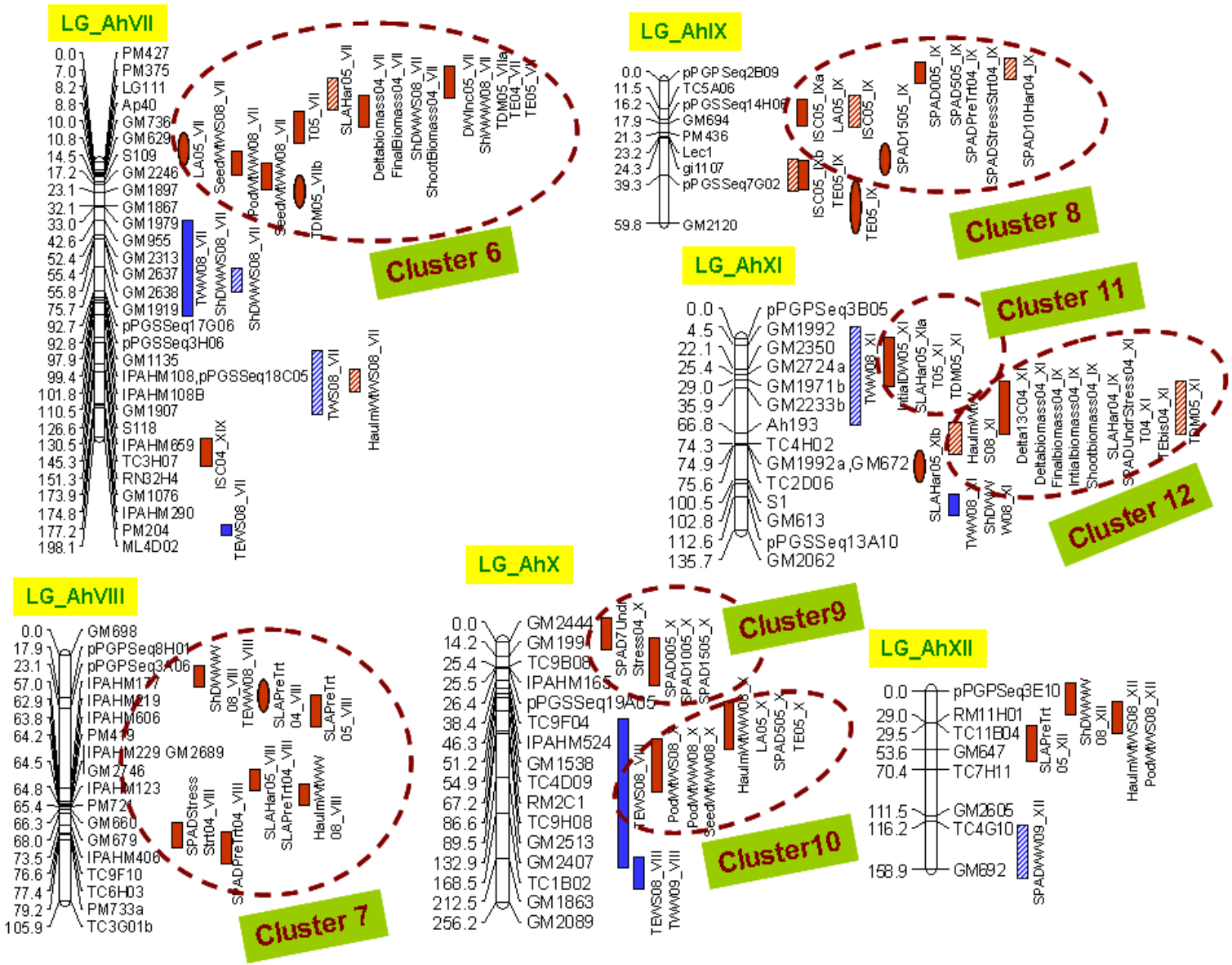


Figure 6: A Consensus map and placing of various M-QTLS and E-QTLs for drought component traits




LG_AhVII
Cluster 6
LG_AhIX
Cluster 8
LG_AhXI
Cluster 11
Cluster 12
LG_AhVIII
Cluster 7
LG_AhX
Cluster9
Cluster10
LG_AhXII

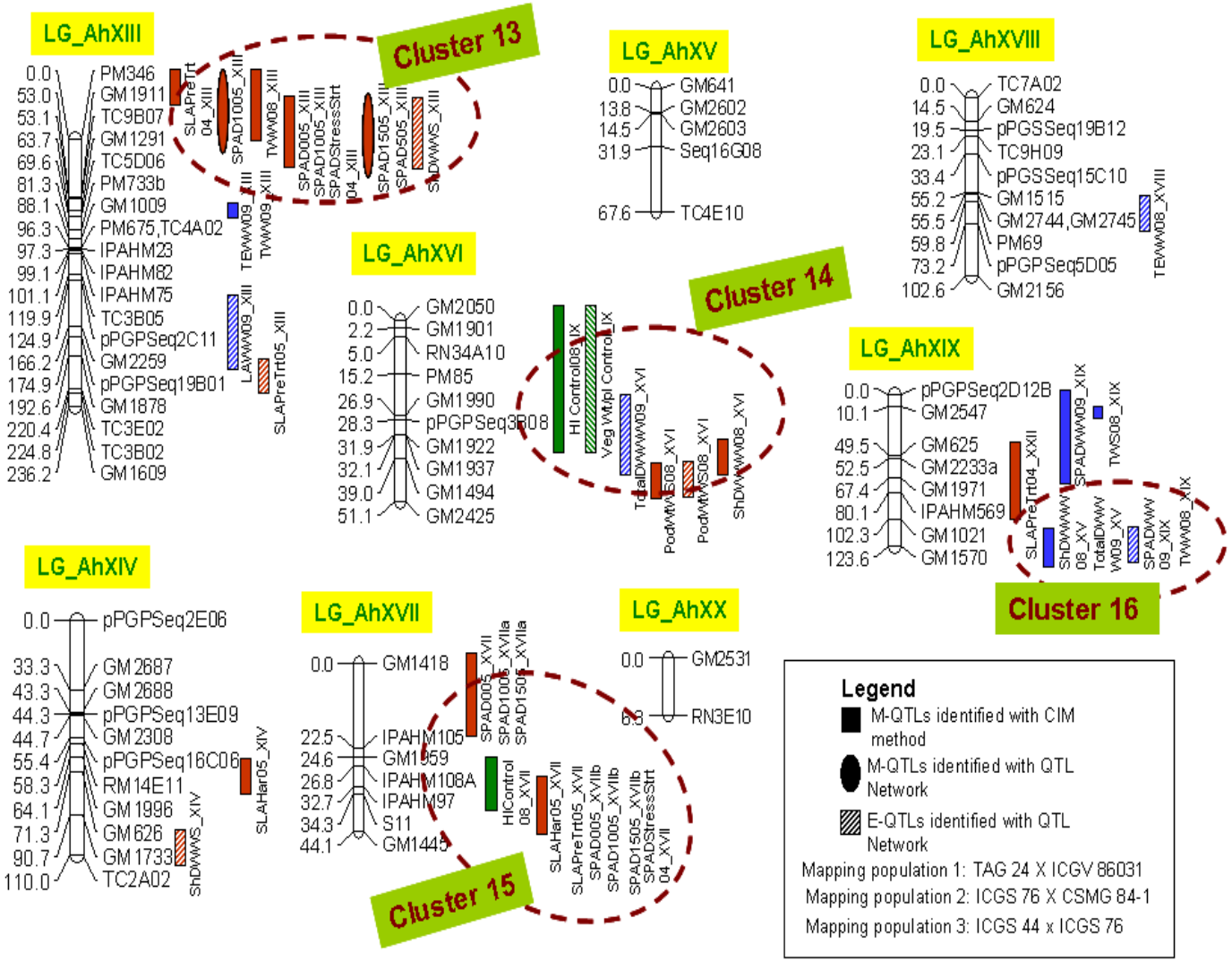
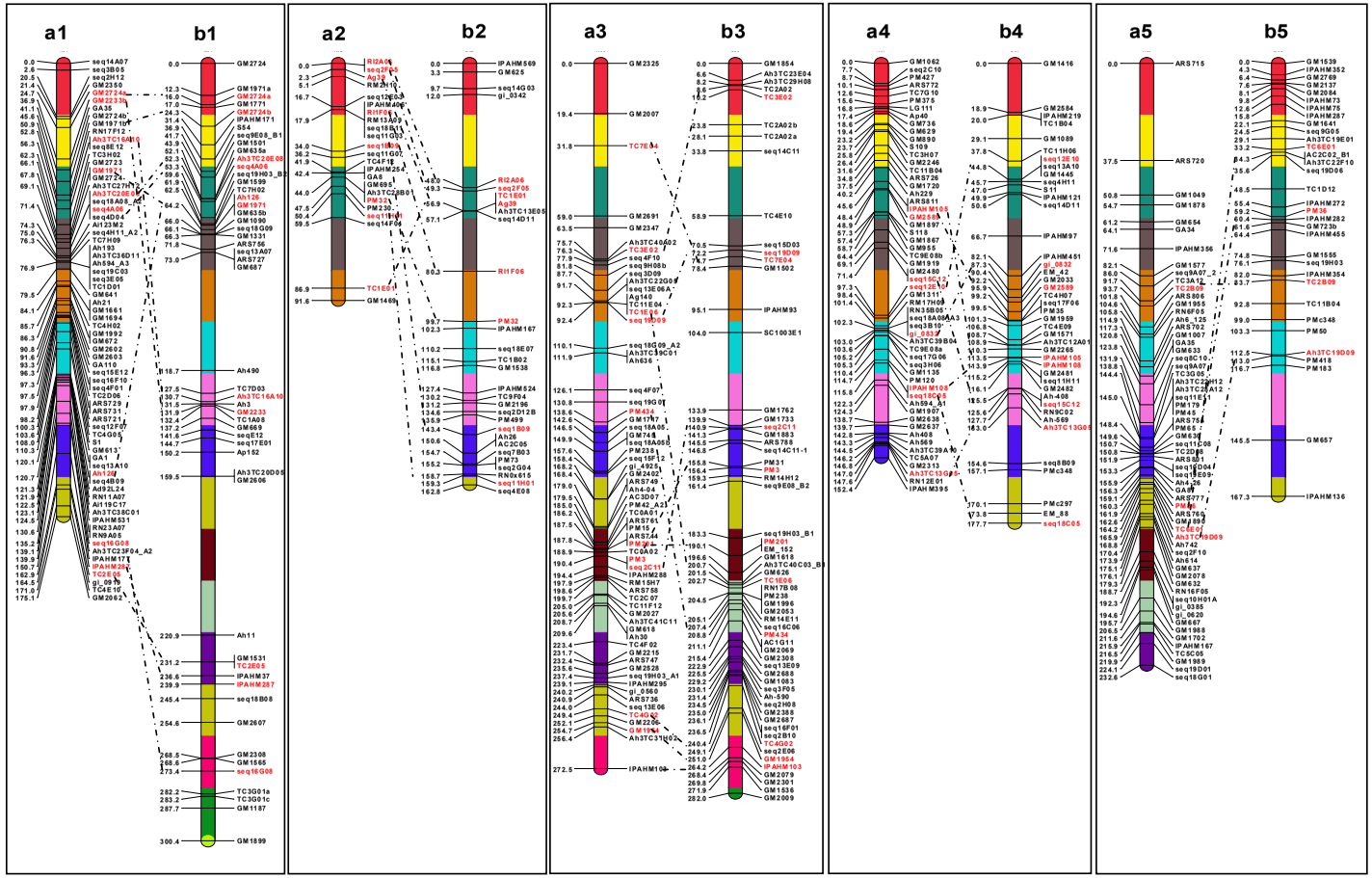

LG_AhXIII
Cluster 13
0.0 PM346
53.0 GM1911
53.1 TC9B07
63.7 GM1291
69.6 TC5D06
81.3 PM733b
88.1 GM1009
96.3 PM675,TC4A02
97.3 IPAHM23
99.1 IPAHM82
101.1 IPAHM75
119.9 TC3B05
124.9 pPGPSeq2C11
166.2 GM2259
174.9 pPGPSeq19B01
192.6 GM1878
220.4 TC3E02
224.8 TC3B02
236.2 GM1609
SLAPreTrt 04_XIII
SPAD1005_XIII
TWW08_XIII
SPAD005_XIII
SPAD1005_XIII
SPADStressStrt 04_XIII
SPAD1505_XIII
SPAD505_XIII
SDWWS_XIII
TEWW09_XIII
TWW09_XIII
LAWW09_XIII
SLAPreTrt05_XIII
LG_AhXV
0.0 GM641
13.8 GM2602
14.5 GM2603
31.9 Seq16G08
67.6 TC4E10
LG_AhXVIII
0.0 TC7A02
14.5 GM624
19.5 pPGSSeq19B12
23.1 TC9H09
33.4 pPGSSeq15C10
55.2 GM1515
55.5 GM2744,GM2745
59.8 PM69
73.2 pPGPSeq5D05
102.6 GM2156
TEWW08_XVIII
LG_AhXVI
Cluster 14
0.0 GM2050
2.2 GM1901
5.0 RN34A10
15.2 PM85
26.9 GM1990
28.3 pPGPSeq3B08
31.9 GM1922
32.1 GM1937
39.0 GM1494
51.1 GM2425
HI Control08_IX
Veg Wt/pl Control_IX
TotalDWWW09_XVI
TotalDWWW08_XVI
PodWtWS08_XVI
PodWtWS08_XVI
ShDWWW08_XVI
LG_AhXIX
Cluster 16
0.0 pPGPSeq2D12B
10.1 GM2547
49.5 GM625
52.5 GM2233a
67.4 GM1971
80.1 IPAHM569
102.3 GM1021
123.6 GM1570
SLAPreTrt04_XXII
SPADWWW09_XIX
TWS08_XIX
ShDWWW 08_XV
TotalDWW W09_XV
SPADWW 09_XIX
TWW08_XIX
LG_AhXIV
0.0 pPGPSeq2E06
33.3 GM2687
43.3 GM2688
44.3 pPGPSeq13E09
44.7 GM2308
55.4 pPGPSeq16C06
58.3 RM14E11
64.1 GM1996
71.3 GM626
90.7 GM1733
110.0 TC2A02
SLAHar05_XIV
ShDWWS_XIV
LG_AhXVII
Cluster 15
0.0 GM1418
22.5 IPAHM105
24.6 GM1959
26.8 IPAHM108A
32.7 IPAHM97
34.3 S11
44.1 GM1445
SPAD005_XVII
SPAD1005_XVIIa
SPAD1505_XVIIa
HIControl 08_XVII
SLAHar05_XVII
SLAPreTrt05_XVII
SPAD005_XVIIb
SPAD1005_XVIIb
SPAD1505_XVIIb
SPADStressStrt 04_XVII
LG_AhXX
0.0 GM2531
6.9 RN3E10
Legend
M-QTLs identified with CIM method
M-QTLs identified with QTL Network
E-QTLs identified with QTL Network
Mapping population 1: TAG 24 X ICGV 86031
Mapping population 2: ICGS 76 X CSMG 84-1
Mapping population 3: ICGS 44 x ICGS 76

Figure 7: An international reference microsatellite consensus genetic map comprising 897 marker loci based on 11 mapping populations



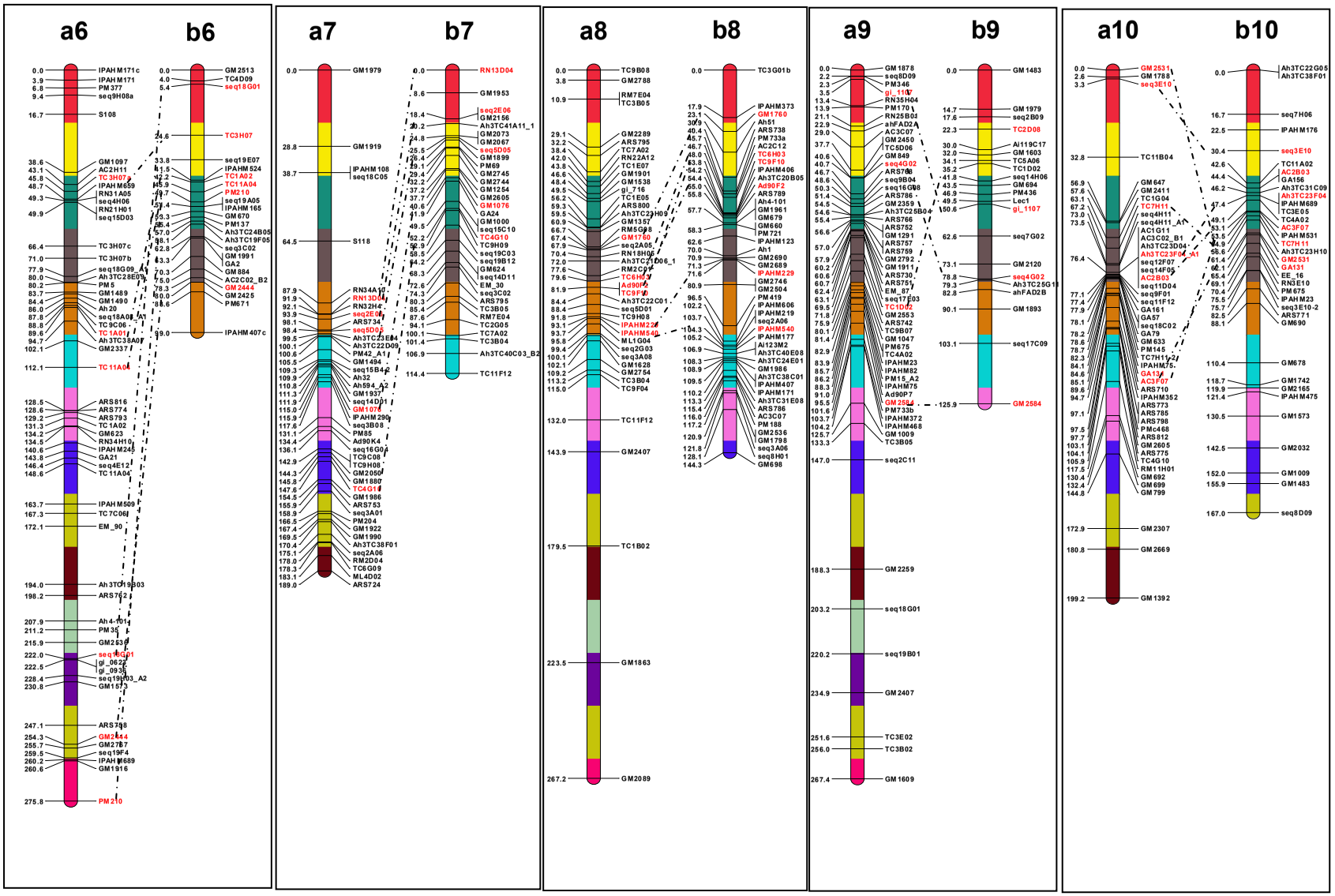

a6
b6
a7
b7
a8
b8
a9
b9
a10
b10

Figure 8: A marker based correspondence for a3 among reference microsatellite consensus and individual genetic maps

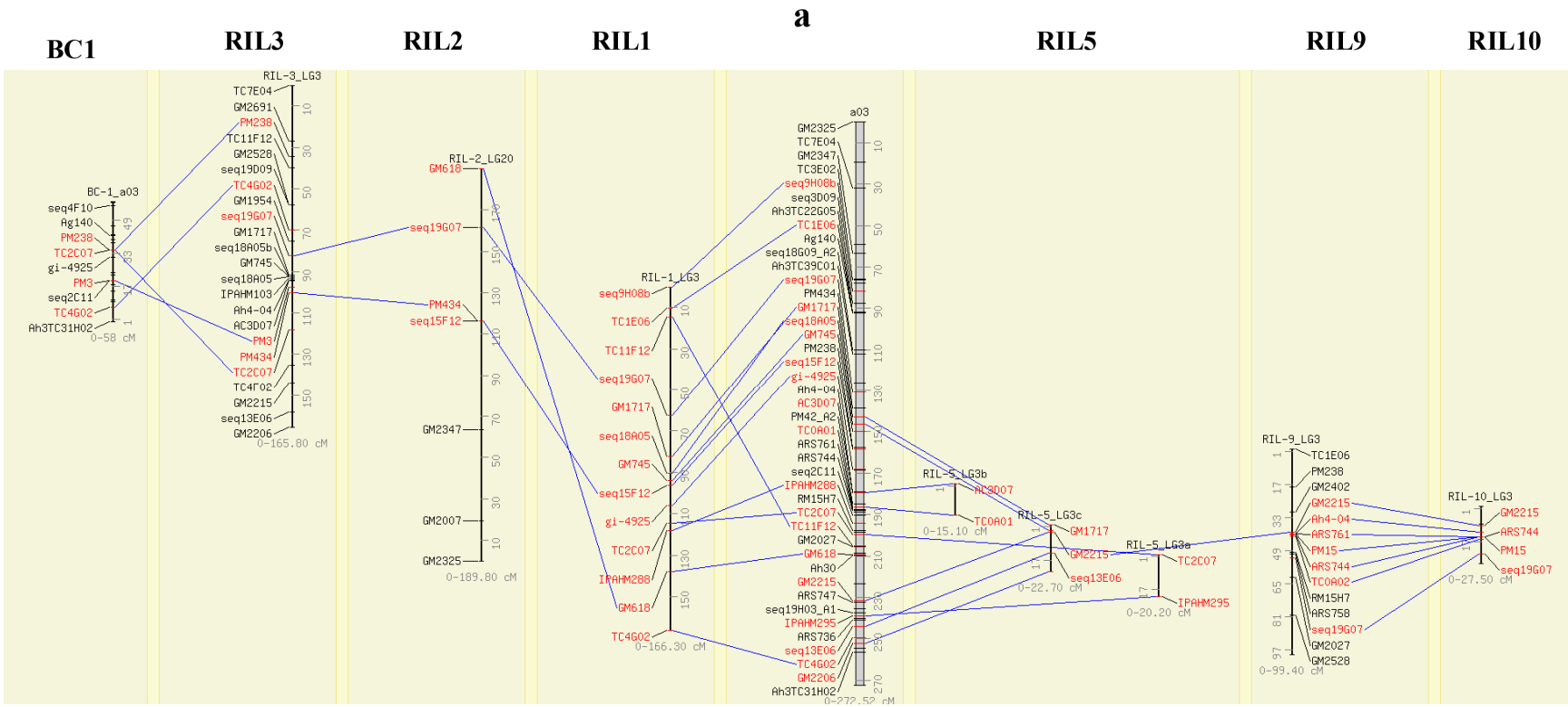
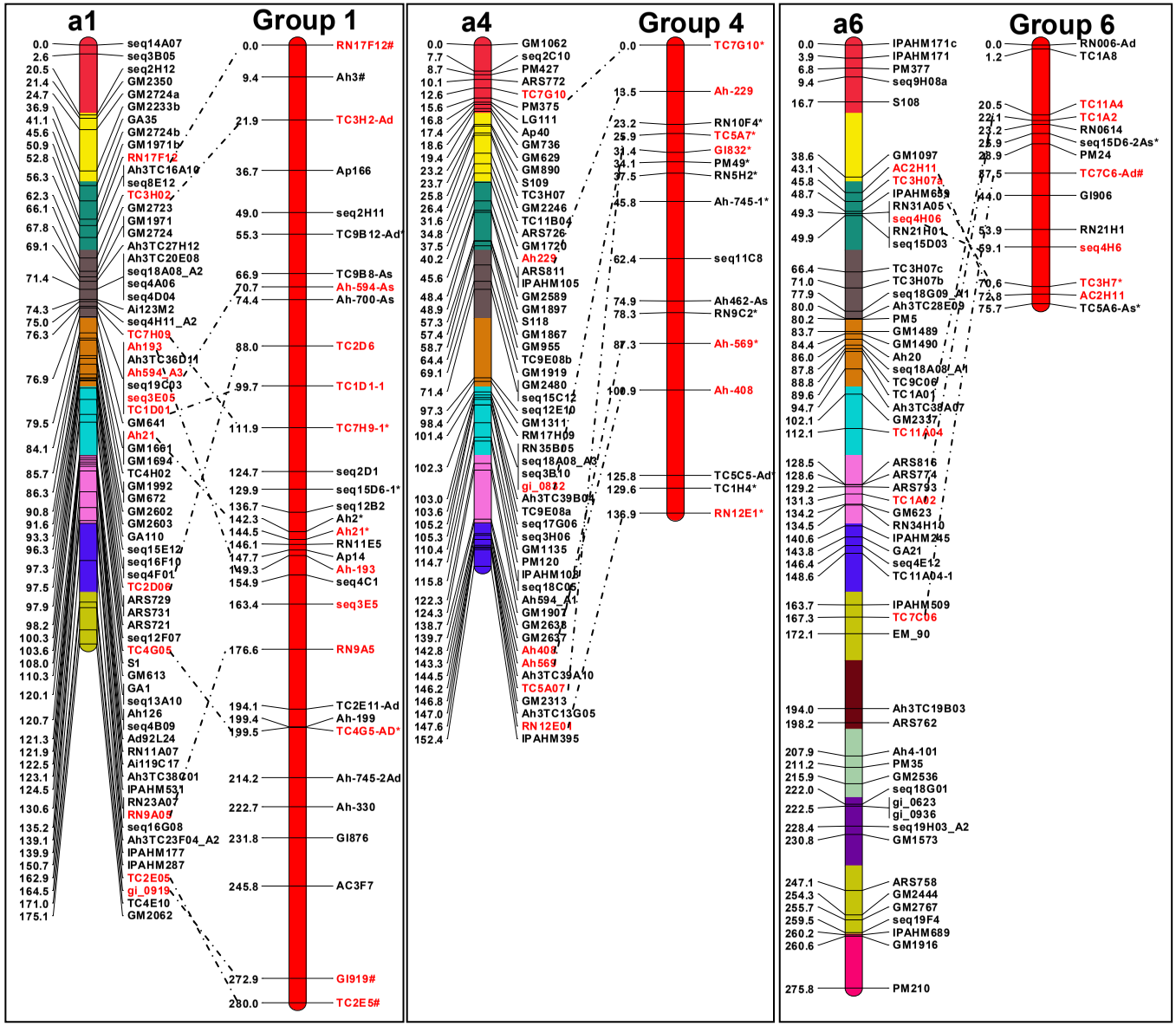
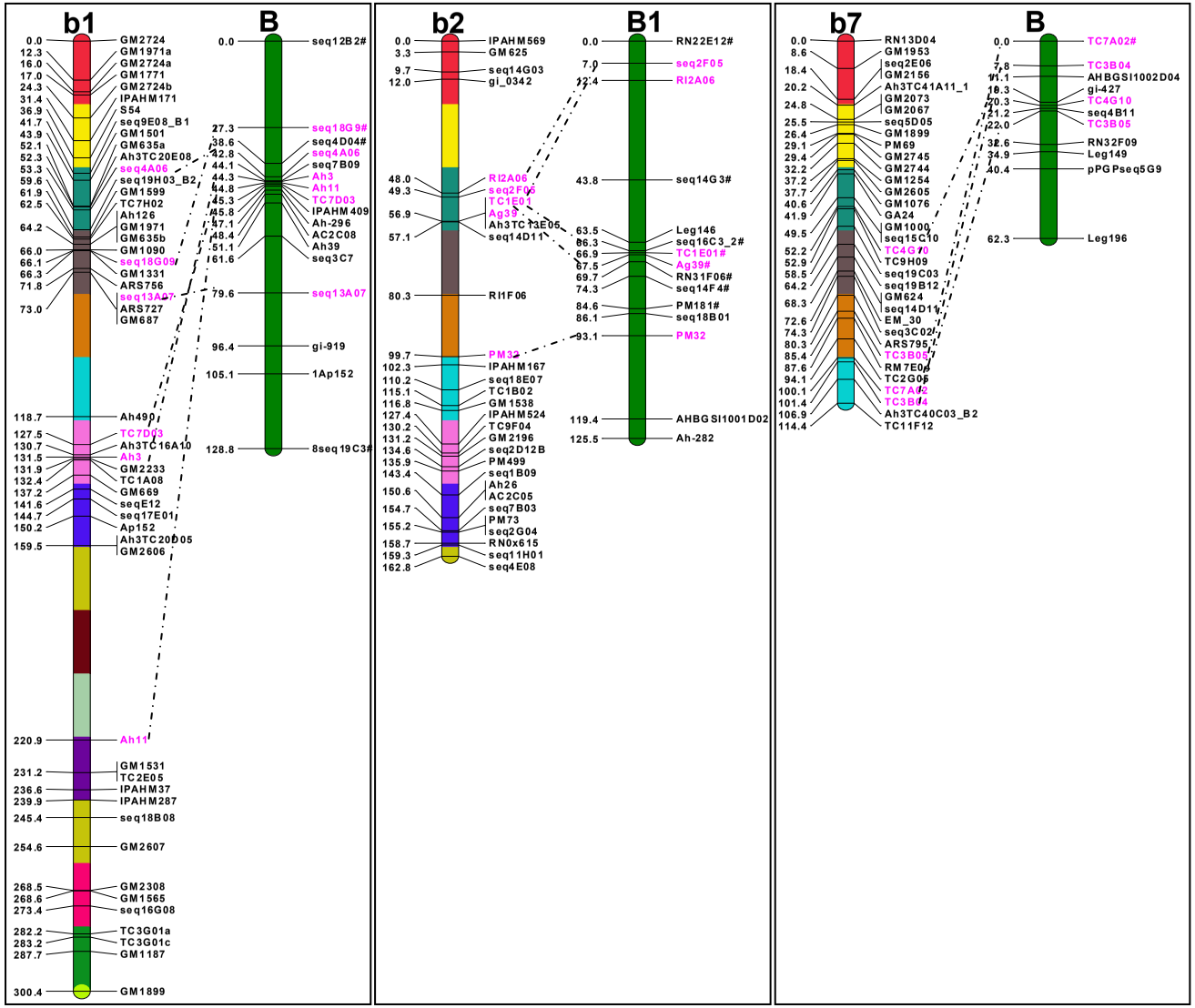


Figure 9: Comparison between the LGs of the international reference microsatellite consensus map and the diploid AA and BB maps




b1
0.0 GM2724
12.3 GM1971a
16.0 GM2724a
17.0 GM1771
24.3 GM2724b
31.4 IPAHM171
36.9 S54
41.7 seq9E08_B1
43.9 GM1501
52.1 GM635a
52.3 Ah3TC20E08
53.3 seq4A06
59.6 seq19H03_B2
61.9 GM1599
62.5 TC7H02
64.2 Ah126
GM1971
GM635b
66.0 GM1090
66.1 seq18G09
66.3 GM1331
71.8 ARS756
73.0 seq13A07
ARS727
GM687
118.7 Ah490
127.5 TC7D03
130.7 Ah3TC16A10
131.5 Ah3
131.9 GM2233
132.4 TC1A08
137.2 GM669
141.6 seqE12
144.7 seq17E01
150.2 Ap152
159.5 Ah3TC20D05
GM2606
220.9 Ah11
231.2 GM1531
TC2E05
236.6 IPAHM37
239.9 IPAHM287
245.4 seq18B08
254.6 GM2607
268.5 GM2308
268.6 GM1565
273.4 seq16G08
282.2 TC3G01a
283.2 TC3G01c
287.7 GM1187
300.4 GM1899
B
0.0 seq12B2#
27.3 seq18G9#
38.6 seq4D04#
42.8 seq4A06
44.1 seq7B09
44.3 Ah3
44.8 Ah11
45.3 TC7D03
45.8 IPAHM409
47.1 Ah-296
48.4 AC2C08
51.1 Ah39
61.6 seq3C7
79.6 seq13A07
96.4 gi-919
105.1 1Ap152
128.8 8seq19C3#
b2
0.0 IPAHM569
3.3 GM625
9.7 seq14G03
12.0 gi_0342
48.0 RI2A06
49.3 seq2F05
TC1E01
Ag39
56.9 Ah3TC13E05
57.1 seq14D11
80.3 RI1F06
99.7 PM32
102.3 IPAHM167
110.2 seq18E07
115.1 TC1B02
116.8 GM1538
127.4 IPAHM524
130.2 TC9F04
131.2 GM2196
134.6 seq2D12B
135.9 PM499
143.4 seq1B09
150.6 Ah26
AC2C05
154.7 seq7B03
155.2 PM73
seq2G04
158.7 RN0x615
159.3 seq11H01
162.8 seq4E08
B1
0.0 RN22E12#
7.0 seq2F05
12.4 RI2A06
43.8 seq14G3#
63.5 Leg146
86.3 seq16C3_2#
66.9 TC1E01#
67.5 Ag39#
69.7 RN31F06#
74.3 seq14F4#
84.6 PM181#
86.1 seq18B01
93.1 PM32
119.4 AHBGSI1001D02
125.5 Ah-282
b7
0.0 RN13D04
8.6 GM1953
seq2E06
18.4 GM2156
Ah3TC41A11_1
20.2 GM2073
24.8 GM2067
25.5 seq5D05
26.4 GM1899
29.1 PM69
29.4 GM2745
32.2 GM2744
37.2 GM1254
37.7 GM2605
40.6 GM1076
41.9 GA24
49.5 GM1000
seq15C10
52.2 TC4G10
52.9 TC9H09
58.5 seq19C03
64.2 seq19B12
68.3 GM624
seq14D11
72.6 EM_30
74.3 seq3C02
80.3 ARS795
85.4 TC3B05
87.6 RM7E06
94.1 TC2G05
100.1 TC7A02
101.4 TC3B04
106.9 Ah3TC40C03_B2
114.4 TC11F12
B
0.0 TC7A02#
7.8 TC3B04
11.1 AHBGSI1002D04
18.3 gi-427
20.3 TC4G10
21.2 seq4B11
22.0 TC3B05
32.6 RN32F09
34.9 Leg149
40.4 pPGPseq5G9
62.3 Leg196