GENETIC DIVERSITY AND LINKAGE ANALYSIS IN CHICKPEA USING DNA MARKERS

A Thesis Submitted to the

University of Pune

For the Degree of

DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY

BY

VRINDA J. SANT

PLANT MOLECULAR BIOLOGY UNIT

DIVISION OF BIOCHEMICAL SCIENCES

NATIONAL CHEMICAL LABORATORY

PUNE 411008 (INDIA)

MAY 2001

DECLARATION

Certified that the work in this thesis entitled **"Genetic diversity and linkage analysis in chickpea using DNA markers",** submitted by **Ms. Vrinda J. Sant** was carried out by the candidate under my supervision. The material obtained from other sources has been duly acknowledged in this thesis.

(Vidya S. Gupta)
Research Guide

TO MY BELOVED PARENTS

INTER-INSTITUTIONAL COLLABORATIVE RESEARCH EFFORT

RESEARCH EMBODIED IN THIS THESIS WAS CARRIED OUT AT

NATIONAL CHEMICAL LABORATORY (N.C.L.), PUNE, INDIA

IN COLLABORATION WITH

MAHATMA PHULE AGRICULTURAL UNIVERSITY
(M.P.K.V.), RAHURI, INDIA

AND

WASHINGTON STATE UNIVERSITY (W.S.U.), PULLMAN, USA

SCIENTIST(S) INVOLVED FROM EACH INSTITUTE

N.C.L. **DR. VIDYA S. GUPTA** (Research Guide) and **DR. P.K. RANJEKAR**

M.P.K.V **DR. R.B. DESHMUKH**

W.S.U. **DR. F.J. MUEHLBAUER**

CONTENTS

Acknowledgements	**i**
List of abbreviations	**ii**
List of figures	**iv**
List of tables	**vi**
Thesis abstract	**vii**
CHAPTER 1	**1**
Conventional and biotechnological approaches towards sustainable chickpea agriculture	
CHAPTER 2	**28**
Potential of DNA markers in detecting divergence and in analyzing heterosis in Indian elite chickpea cultivars	
CHAPTER 3	**52**
Ty1-*copia* retrotransposon like elements in chickpea genome: their identification, distribution and use for diversity analysis	
CHAPTER 4	**72**
Molecular mapping of the RAPD and ISSR markers on the integrated map of chickpea genome and its exploitation for Fusarium wilt resistance tagging in chickpea	
CHAPTER 5	**101**
General discussion	
REFERENCES	**107**
CURRICULUM VITAE	**118**

ACKNOWLEDGEMENTS

After having spent a significant portion of my life in the exhilarating, stimulating and worthwhile atmosphere of NCL in the pursuit of a doctorate in Biotechnology, I would like to acknowledge the large number of people whose support, encouragement and inspiration backed me up constantly.

First and foremost I express my heartfelt gratitude to my advisor Dr Vidya S. Gupta for giving me the opportunity and the means to "do science". Her friendliness, openness and positive approach to solve difficulties indeed boosted me during the trying times in research.

I am deeply indebted to Dr. P. K. Ranjekar, who provided me with many insights into new ideas and suggestions. The unforgettable discussions with him would enthuse me to work with a renewed zest.

I thank Dr R. B. Deshmukh, Director of Research, Mahatma Phule Agricultural University, Rahuri for the collaborative work which involved providing me all the field facilities and also help from his team at the Pulse Reseach Station. I am also thankful to Dr F. J. Muehlbauer, Washington State Univeristy, Pullman, USA for the informative discussions with him during his visits to India.

I would like to make a special mention of my seniors Milind and Dipak whom I admire for their distinctive approach towards science. Their tenacious commitment to work and their ability to ask the hard questions was indeed contagious. I was graced with the friendship of Rajesh who rendered valuable help and made me laugh during the difficult times.

I thank Abhay, Ashok, Aparna and Archana for their constant help and also for making the long journeys to Rahuri extremely memorable and cheerful. I feel lucky to have friends like Mukund, Armaity, Aditi, Swati, Bhushan, and Rahul, whom I thank for their untiring help.

My thanks are due to Medhavinee, Jessica and Sanjay who assisted me in my work. I enjoyed the company of Ajit, Venkat, Sastry, Sami, RK, Anjali, Bimba, Manisha, Renu, Suresh, Shashi, Aparna D, Meena, Renuka, Maneesha, Arundhati, Sadhana and Suvarna.

In the hostel I have had the good fortune of experiencing many joyous moments with Aparna T, K.N.Rao, Rethi, Neelam and Sreedevi. I thank them for their friendship during these past years and for providing me with many laughs and excellent excuses to grab some fun and shoot the breeze.

Many thanks are due to Drs. Mohini, Meena, Nirmala, Lalitha, Shubhada and Mr. B. G. Patil. I appreciate the help rendered by Usha, Indira and Satyali. Thanks are due to Karunakaran, Jagtap, Modak and Kamthe for their technical help in instrumentation.

I am greatly indebted to the Director, NCL, for providing the facilities during the tenure of my work. I acknowledge the Council of Scientific and Industrial Research for providing financial assistance. I express my gratitude to the McKnight Foundation, USA, for supporting my research.

I wish to express my thanks to the unending love, support and sacrifice of my family. I am at loss of words to describe the love showered on me by my brothers, Satish and Shrikant and sisters-in-law, Leena and Jyoti. I express my deepest gratitude to my parents-in-law for the care, support and encouragement.

My husband, Jeetendra, brought joy to my life, a sense of perspective to this work and a sense of possibility to the future. Though he was far away from me throughout my research tenure his regular encouraging emails kept me going. I realize how fortunate I am.

None of my accomplishments would have been possible without the unquestioning love, moral support and blessings of my parents. They always stressed education and made their children's education a priority of their lives.

I thank all my well wishers I may not have mentioned here for helping me carry out this research work successfully.

(Vrinda J.Sant)

LIST OF ABBREVIATIONS

AFLP	Amplified Fragment Length Polymorphism
ASAP	Allele Specific Associated Primer
BAC	Bacterial Artificial Chromosome
bp	Base pair
$CHCl_3$	Chloroform
cM	Centimorgan
CTAB	Hexadecyl Trimethyl Ammonium Bromide
DAF	DNA Amplification Fingerprinting
dATP	Deoxyadenosine 5' triphosphate
dCTP	Deoxycytidine 5' triphosphate
dGTP	Deoxyguanosine 5' triphosphate
DAS	Days after sowing
dTTP	Deoxythymidine 5' triphosphate
DNA	Deoxyribose Nucleic Acid
DOP-PCR	Degenerate Oligonucleotide Primer-Polymerase Chain Reaction
EDTA	Ethylene Diamine Tetra Acetic acid
EST	Expressed Sequence Tag
GCA	General Combining Ability
HCl	Hydrochloric Acid
IAA	Iso-amylalcohol
ICARDA	International Center of Agricultural Research for the Dry Areas
ICRISAT	International Crop Research Institute for Semi-Arid Tropics
ISSR	Inter Simple Sequence Repeat
kb	Kilobase
KCl	Potassium Chloride
LOD	Log of Odds
LTR	Long Terminal Repeat
Mb	Megabase
$MgCl_2$	Magnesium Chloride
µg	Microgram
µl	Microlitre
ml	Millilitre
mM	Millimolar
NaCl	Sodium Chloride
ng	Nanogram
NIL	Near Isogenic Lines
ORF	Open Reading Frame
PAGE	Polyacrylamide Gel Electrophoresis

pmoles	pico-moles
PCR	Polymerase Chain Reaction
QRL	Quantitative Resistance Loci
QTL	Quantitative Trait Loci
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RGA	Resistance Gene Analog
RNA	Ribonucleic Acid
rpm	Revolution per Minute
rt	Reverse Transcriptase
SCA	Specific Combining Ability
SDS	Sodium Dodecyl Sulphate
sec	Second
SCAR	Sequence Characterised Amplified Region
SSAP	Sequence Specific Amplified Polymorphism
SSPE	Sodium Chloride Sodium Dihydrogen Phosphate EDTA
SSR	Simple Sequence Repeat
STMS	Sequence Tagged Microsatellite Site
STS	Sequence Tagged Site
Tris	Tris Hydroxymethyl Amino Methane
TAE	Tris Acetate EDTA
TAPS	3-Tri(hydroxylmethyl) methyl Aminopropane Sulfonic Acid
TBE	Tris Borate EDTA
Tm	Temperature of melting
U	Unit of enzyme
UPGMA	Unweighted Pair Group Method with Arithmatic Mean
V	Volt

LIST OF FIGURES

Figure No	Description	Page No.
1.1	Different forms of chickpea used for consumption	4
1.2 a, b	Typical chickpea field of `Desi' cultivar with pink flowers	6
1.3	Taxonomic classification of chickpea	8
1.4	Relative importance of abiotic and biotic stresses affecting chickpea crop	14
1.5	Increase in the number of different pathogens of chickpea over a period of 17 years.	16
1.6a	Chickpea field affected by *Fusarium oxysporum*	17
1.6b	Chickpea pod affected due to *Ascochyta rabiei* pathogen	17
1.7a, b	Chickpea pod infestated by *Helicoverpa armigera* larvae and *H. armigera* larvae feeding on chickpea plant	18
2.1	RAPD fingerprint of the elite chickpea cultivars with OPF09	36
2.2a	Oligonucleotide fingerprint with *Taq*I/$(GATA)_4$	39
2.2b	Oligonucleotide fingerprint with *Taq*I/$(GATA)_4$	40
2.3	Dendrogram based on the similarity index values of 29 elite chickpea genotypes using microsatellite markers	42
2.4	Dendrogram displaying the similarity index values of the eight genotypes used for diallel analysis	44
3.1a, b	Multiple sequence alignment of the Ty1-*copia* like rt regions obtained from *C. arietinum* (ICC4958) and *C. reticulatum* (PI489777)	59
3.2	Hybridization of CA1 with genomic DNA of *C. arietinum* (Vijay) and *C. reticulatum* (PI489777)	61
3.3a, b	Genomic digest with *Bam*HI and *Dra*I hybridized with CA1 and CR10 inserts	63
3.4	Amplification of DNA from *Cicer* species using the	64

	degenerate primers for rt region of Ty1-*copia* element	
3.5	Dendrogram depicting the similarity indices between *Cicer* species	65
4.1a	Representative segregation of $ISSR850_{0.55}$ in the RILs developed from interspecific cross	84
4.1b	Representative segregation of $OPC18_{0.6}$ and $OPC18_{0.55}$ in the RILs developed from interspecific cross	85
4.2	*C. arietinum* integrated genome map represented as 16 linkage groups	87, 88
4.3	Field trial for testing of RILs generated from intraspecific cross at MPKV, Rahuri	90
4.4 a, b	Fusarium wilt reaction of the control plants in the wilt sick plot during the field trials of the RILs at MPKV, Rahuri	91
4.5	Frequency distribution of disease score for the RILs along with the segregation of the TA110 allele in the RILs	93
4.6	Segregation of the alleles $TA110_{0.220}$ and $TA110_{0.234}$ in the RILs developed from intraspecific cross.	96

LIST OF TABLES

Table No.	Description	Page No.
1.1	A list of *Cicer* species	9
1.2	Nuclear DNA content in *Cicer*	10
1.3	Crossability group of chickpea on the basis of hybridization and fertility of hybrids	12
1.4	Important diseases affecting chickpea and the causative agents	15
2.1	*Cicer arietinum* L. germplasm used for diversity analysis	32
2.2	Primers amplifying polymorphic bands among the chickpea cultivars	37
2.3	Average number of bands produced by enzyme-oligonucleotide probe used in fingerprinting the elite chickpea cultivars	38
2.4	Probability of identical match by chance	38
2.5	Genetic distance, percent midparent heterosis for different agronomic characters, SCA effects in crosses of the diallel set	46
2.6	General combining ability effects of the parents for different characters	47
4.1	Details of the PCR amplifications used for generating molecular markers	77
4.2	Primer sequences for amplifying RGA, STMS and ASAP markers	78
4.3	Chi-square values of the RAPD and ISSR markers	86

THESIS ABSTRACT

THESIS ABSTRACT

Introduction

Chickpea is the second most important cool season legume crop in the world grown in at least 33 countries in Central and West Asia, South Europe, Ethiopia, North Africa, North and South America and Australia (Ladizinsky 1976, Singh 1997). It serves as a main source of dietary protein for more than 80% of the Indian vegetarian population and is also used as feed for livestock. However, this crop is susceptible to various biotic and abiotic stresses and hence several strategies have been attempted to produce varieties resistant to these stresses and also to harvest high yield capacity. For example, interspecific hybridization has been performed in this crop for incorporating traits from phylogenetically related wild species. Among intraspecific hybridizations, diallel mating designs have been used to obtain segregants giving high yields. In recent years, marker assisted selection and pyramiding of agronomically important genes using DNA markers have shown a great potential in producing varieties resistant to biotic and abiotic stress with better yield potential.

In this thesis, I have attempted to throw light on some aspects of chickpea genomics which can be useful for future research perspectives in this crop.

A brief summary of the important findings of my work is given below

1. Potential of DNA markers in detecting divergence and in analysing heterosis in Indian elite chickpea cultivars

Molecular markers such as RAPD and microsatellites were used to study genetic diversity in 29 elite Indian chickpea cultivars with important agronomic characters. It was found that microsatellites were more efficient than the RAPD markers in detecting polymorphism in these genotypes. Among the microsatellites, tri- and tetra- nucleotide probes were used for hybridization with different restriction enzyme combinations. To generate the RAPD markers, 10-mer random primers were used for developing the DNA fingerprints of the 29 elite chickpea cultivars. On the basis of number of bands shared and unshared, similarity index values were calculated and were used to generate the dendrogram which depicted the genetic distance of the chickpea cultivars with respect to each other.

In order to investigate the usefulness of DNA markers in predicting F_1 performance and heterosis in chickpea, 8 genotypes having important agronomic characters were crossed in a diallel set. The F_1`s and their parents in the diallel set

were analyzed for agronomic traits for better parent and midparent heterosis. Although molecular marker based genetic distance did not linearly correlate to heterosis, two heterotic groups could be identified on the basis of the general marker heterozygosity.

2. Ty1-*copia* retrotransposon like elements in chickpea genome; their identification, distribution and use for diversity analysis

Retrotransposons are abundant in the plant genome due to proliferation by reverse transcription. They have been reported to be capable of generating large populations in a relatively short evolutionary time as they have a replicative mode of transposition. The resulting progeny is 'seeded' to new genomic sites creating insertional polymorphism (Voytas *et al*. 1998). Plant genomes can accumulate large amounts of these sequences and do not appear to remove them rapidly. Chickpea is known to be a less diverse genome and there are many attempts to reveal polymorphism in the germplasm In the present study, I have identified potential of Ty1-*copia* retrotransposon like element for diversity analysis among wild and cultivated species of *Cicer.* Further sequence analysis revealed heterogeneity among Ty1-*copia* elements in chickpea.

Ty1-*copia* retrotransposon like elements were amplified from *Cicer* species using primers derived from conserved region of the reverse transcriptase gene. Two sized fragments of ~280 bp and ~650 bp were obtained, which on sequencing showed homology for the Ty1-*copia* reverse transcriptase region. Interestingly, the ~650 bp fragment showed two reverse transcriptase regions, one from Ty1-*copia* and the other from *Tto*1 element fused with each other. The copy number was high in the cultivated *C. arietinum* genome as compared to the wild *C. reticulatum*. Genetic diversity among the *Cicer* species was investigated using the conserved primers, which grouped the wild species and the cultivated *C. arietinum* separately.

3. Construction of linkage map of chickpea and identification of a marker linked to Fusarium wilt resistance

Development of a detailed linkage map of any crop enhances the understanding of genetics and improves the efficiency of crop improvement programs, especially those involving quantitative traits. During last decade, linkage maps based on morphological and biochemical traits are rapidly being integrated with DNA based marker maps.

Till date a detailed linkage map of the chickpea genome is not available. In an international collaborative effort, a linkage map of chickpea genome based on RILs

from *C. arietinum* x *C. reticulatum* cross was developed using STMS, RAPD, ISSR and AFLP markers. In my contribution, the two parents were screened with random primers and primers amplifying regions between simple sequence repeats to generate polymorphic RAPD and ISSR markers, respectively. These identified polymorphic markers were scored on the RIL population and mapped on the linkage groups using the software Mapmaker/Exp 3.0b. It was observed that the RAPD markers were distributed on the various linkage groups unlike the other types of markers that showed clustering.

An intraspecific cross was utilized for identifying a putative marker linked to Fusarium wilt caused by *Fusarium oxysporum* f.sp. *ciceri*. The two parents, Vijay (resistant) and JG-62 (susceptible) were crossed and 166 RILs were developed by single seed descent. The two parents were screened for polymorphism using different types of primers such as RAPD, ISSR and STMS. Primers for resistant gene analogs (RGAs) and allele specific associated primers (ASAP) were also used for amplifying respective loci in the two parents. The identified polymorphic markers were scored on the RILs and linkage analysis was performed using the software Q-Gene. Among the STMS primer, TA110 was analyzed on the 166 RILs for segregation and the marker segregated in an expected 1:1 ratio (X^2=0.433, P<0.05). The segregating marker and the phenotype data were analyzed by simple linear regression using the software Q-gene. The R^2 value indicated an estimated 20.62% contribution to phenotypic variation for Fusarium wilt resistance at P<0.01. Thus, TA110 represents a putative STMS marker linked to the quantitative resistance locus contributing resistance to race 1 of *F. oxysporum* f.sp. *ciceri*.

CHAPTER 1

Conventional and Biotechnological Approaches Towards Sustainable Chickpea Agriculture

1.1. Chickpea: The Most Important Legume Crop in Indian Subcontinent

1.2. Origin, Taxonomy, Occurrence, Cytology and Phylogeny of Chickpea

1.2.1 Chickpea originated in Turkey and adjoining areas of Syria

1.2.2 Taxonomy and occurrence of chickpea

1.2.3 DNA content and karyotype analysis in chickpea

1.2.4 Phylogeny on the basis of crossability and heterochromatin studies

1.3 Biotic and Abiotic Stress: Responsible for Poor Average Yield in Chickpea

1.4 Conventional Breeding Practices for Chickpea Improvement

1.4.1 Interspecific hybridization

1.4.2 Intraspecific hybridization

1.5 Molecular Markers in Chickpea Genomics

1.5.1 Diversity and phylogeny of chickpea as revealed by molecular marker studies

1.5.2 Genetic linkage map of *Cicer* genome

1.5.3 Initial efforts of tagging Fusarium wilt and Ascochyta blight resistance genes in chickpea

1.5.4 Future prospects

1.6 Objectives

1.7 Organization of Thesis.

1.1. Chickpea: The Most Important Legume Crop in Indian Subcontinent

Chickpea (*Cicer arietinum* L.) is the second most important cool season pulse crop in the world and is grown in at least 33 countries including Central and West Asia, South Europe, Ethiopia, North Africa, North and South America and Australia (Ladizinsky and Adler 1976, Singh 1997). It covers 15% (10.2 million hectares) of the area and accounts for 14% (7.9 million ton) of the production of pulses in the world (FAO 1994). India is the largest producer of chickpea accounting for 75% and 73%, respectively, of the world`s share in terms of the area under cultivation and production (Jodha and SubbaRao 1987). Although India produces a large variety of pulses, chickpea alone accounts for 43.2% of the total annual pulse production of 11.79 million tons and is an important component of food in the diets of all income groups (Pushpamma and Geervani 1987). In India, `Desi' type of chickpea with small and brown seeds accounts for nearly 90% area under chickpea cultivation and `Kabuli' type with bold and cream coloured seeds is grown in around 10% area.

Chickpea acquires importance as it provides food for humans as well as for livestock. Furthermore, chickpea pod covers and seed coats can also be used as fodder. It is consumed as a fresh immature green seed, whole seed, dhal and flour. In the Barind region of Bangladesh, where extremely low rains result in lesser choice of crops for farmers, the top twig from each chickpea plant is consumed as a green vegetable. It is also consumed as a delicacy in India during the chickpea-growing season. Figure 1.1 shows the different ways of consuming chickpea all over the world.

In grain legumes, proteins are an important seed component and are responsible for their relevant nutritional and socio-economic impact. The chickpea seed is a good source of carbohydrates and proteins, which together constitute 80% of the total dry seed weight. The crude protein content of chickpea varies from 17% to 24% containing the essential amino acids like tryptophan, methionine and cysteine (Williams and Singh 1987). Thus chickpea serves as a main source of dietary protein for more than 80% of the Indian population which is vegetarian in nature.

Figure 1.1 Different forms of chickpea used for consumption.

Since chickpea plays the pivotal role of supplying protein source in the vegetarian diet, it is also called as the 'poor man's meat'. Supplementation of cereals with high protein legume is potentially one of the best solutions to protein-calorie malnutrition, particularly in the developing countries.

Chickpea is a hardy deep-rooted dryland crop and can grow to full maturity despite conditions that would prove fatal for most crops. It is grown on marginal land and rarely receives fertilizers or protection from diseases and insect pests (Singh and Reddy 1991). Nearly 90% of the crop is cultivated under rainfed conditions mostly on receding soil moisture. Chickpea being a rabi crop is normally sown in the month of October and harvested in January, whereas in Northeastern Australia it is sown in May/June and harvested from October to December (Knights 1993). Figure 1.2 shows a typical chickpea field sown with the desi cultivars. Chickpea, being a legume, produces nitrogen fixing root nodules and can enrich the soil with at least 50 kg of nitrogen per hectare every season. In addition, chickpea can mobilize phosphorus from sources that are not available to other crop plants. This is due to its acidic root exudates rich in citric acid which help to dissolve calcium phosphates. However, chickpea crop is affected by a number of bacterial and fungal pathogens whereas, it has relatively less number of insect pest as compared to other legumes. This could be due to secretion of malic acid and oxalic acid. It was observed that cultivars secreting these acids above a threshold level were relatively resistant to *Helicoverpa armigera* (Khanna-Chopra and Sinha 1987).

In summary, chickpea is a number one legume crop in the developing world and it not only serves as a good source of nutrition to the people but also improves the soil.

a

b

Figure 1.2 (a) and **(b)** Typical chickpea field of Desi cultivar with `pink' flowers

1.2. Origin, Taxonomy, Occurrence, Cytology and Phylogeny of Chickpea

1.2.1 Chickpea originated in Turkey and adjoining areas of Syria

The cultivated chickpea is a crop of ancient origin and is one of the first grain legumes domesticated in the Old World (van der Maesen 1972). Chickpea most probably originated in a region of present day southeastern Turkey and adjoining areas of Syria (Singh 1997). Vavilov (1926) identified southwest Asia and the Mediterranean as the two primary centers of origin, while Ethiopia as a secondary center of origin. Further the author noted the abundance of large seeded chickpea lines around the Mediterranean and small seeded lines in the East. There is a linguistic evidence that large ram seeded cream coloured chickpea reached India via the Afghan capital Kabul about two centuries ago and acquired a name as kabuli chana in Hindi (van der Maesan 1972). Regarding the origin of kabuli and desi, it is almost certain that desi originated first followed by kabuli type developed by selection and mutation (Singh 1997).

1.2.2 Taxonomy and occurrence of chickpea

Chickpea is a self-pollinated diploid (2n=16) grain legume, taxonomically classified to a separate tribe of Cicereae Alef in subfamily Papilionoideae of Leguminoseae family (Figure 1.3) (Kupicha 1981). The genus *Cicer* contains two subgenera - Pseudononis and Viciastrum which include 43 species (9 annual, 33 perennial and 1 unspecified) (van der Maesen 1987) as enlisted in Table 1.1.

The *Cicer* species occur from sea level (e.g *C. arietinum, C. montbretii*) to over 5000m (*C. microphyllum*) near glaciers in the Himalayas. The cultivated species, *C. arietinum* is found only in cultivation and cannot colonise successfully without human intervention. The wild species (e.g *C. reticulatum*, *C. bijugum*) occur in weedy habitats and fallow or disturbed habitats, roadsides, cultivated fields of wheat and other places not touched by man or cattle. The species like *C. pungens* and *C. yamashitae* occur on mountain slopes among rubble while *C. montbretti* and *C. floribundum* can be found in forest soils, broad leaf or pine forests. (www.grep-icrisat.org/cgiar/chickpea).

Figure 1.3 Taxonomic classification of chickpea (Cronquist 1968, Stace 1980)

Kingdom	**Plantae**
Subkingdom	**Embryobionta**
Division	**Tracheophyta**
Subdivision	**Spermatophytina**
Class	**Angiospermopsida**
Subclass	**Dicotyledonidae**
Order	**Rosales**
Family	**Leguminoseae**
Subfamily	**Papilionoideae**
Tribe	**Cicereae Alef.**
Genus	***Cicer***

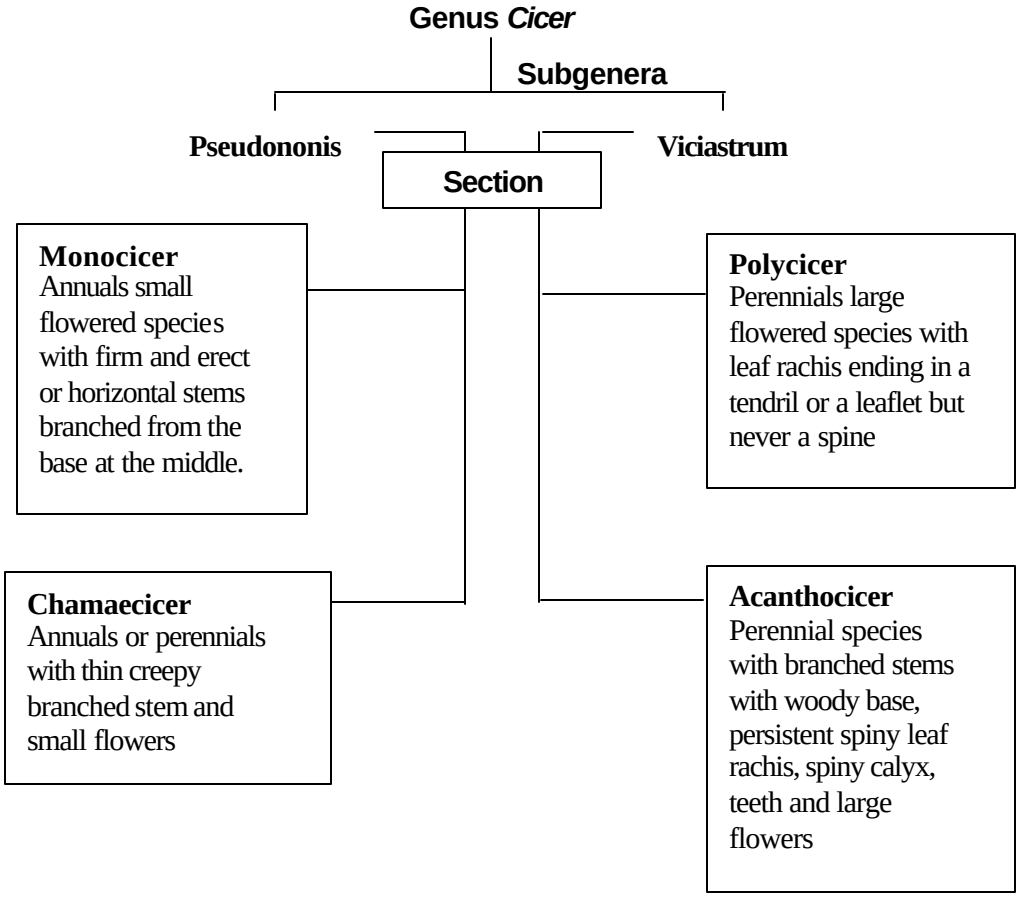


Table 1.1 A list of *Cicer* species (van der Maesen 1987)

	Species	Section	Cycle [a]
1	*C. acanthophyllum* Boriss.	Acanthocicer	**P**
2	*C. anatolicum* Alef.	Polycicer	**P**
3	*C. arietinum* L.	Monocicer	**A**
4	*C. atlanticum* Coss. ex. Maire	Polycicer	**P**
5	*C. balcaricum* Galushiko	Polycicer	**P**
6	*C. baldshuanicum* (M.Pop.) Lincz.	Polycicer	**P**
7	*C. bijugum* K.H.Rech	Monocicer	**A**
8	*C. canariense* Santos Guerra & Lewis	Polycicer	**P**
9	*C. chorassanicum* (Bge) M. Pop	Chamaecicer	**A**
10	*C. cuneatum* Hochst. ex Rich	Monocicer	**A**
11	*C echinospermum* P.H. Davis	Monocicer	**A**
12	*C. fedtschenkoi* Lincz	Polycicer	**P**
13	*C. flexuosum* Lipsky	Polycicer	**P**
14	*C floribundum* Fenzl.	Polycicer	**P**
15	*C. graecum* Orph.	Polycicer	**P**
16	*C. grande* (M.Pop) Korotk.	Polycicer	**P**
17	*C. heterophyllum* Contandr. *et al.*	Polycicer	**P**
18	*C. incanum* Korotk.	Acanthocicer	**P**
19	*C. incisum (*Willd.) K.Maly	Chamaecicer	**P**
20	*C. isauricum* P. H. Davis	Polycicer	**P**
21	*C. judaicum* Boiss.	Monocicer	**A**
22	*C. kermanense* Bornm.	Polycicer	**P**
23	*C. korshinskyi* Lincz.	Polycicer	**P**
24	*C. laetum* Rassulova & Sharipova	?	**?**
25	*C. macracanthum* M. Pop	Acanthocicer	**P**
26	*C. microphyllum* Benth.	Polycicer	**P**
27	*C. mogoltavicum* (M.Pop.) Koroleva	Polycicer	**P**
28	*C. montbretti* Jaub. & Sp.	Polycicer	**P**
29	*C. multijugum* van der Maesen	Polycicer	**P**
30	*C. nuristanicum* Kitamura	Polycicer	**P**
31	*C. oxyodon* Boiss. & Hoh.	Polycicer	**P**
32	*C. paucijugum* (M. Pop.) Nevski	Polycicer	**P**
33	*C. pinnatifidum* Jaub & Sp.	Monocicer	**A**
34	*C. pungens* Boiss	Acanthocicer	**P**
35	*C. rassuloviae* Lincz.	Polycicer	**P**
36	*C. rechingeri* Podlech	Acanthocicer	**P**
37	*C. reticulatum* Ladiz.	Monocicer	**A**
38	*C. songaricum* Steph.ex. DC.	Polycicer	**P**
39	*C. spiroceras* Jaub. & Sp.	Polycicer	**P**
40	*C. stapfianum* K.H. Rech	Acanthocicer	**P**
41	*C. subaphyllum* Boiss.	Polycicer	**P**
42	*C. tragacanthoides* Jaub. & Sp.	Acanthocicer	**P**
43	*C. yamashitae* Kitamura	Monocicer	**A**

[a] P=perennial, A=annual.

1.2.3 DNA content and karyotype analysis in chickpea

In *Cicer*, cytological studies have been carried out on a limited number of species due to nonavailability of living material (van der Maesen 1987). Studies by Ohri and Pal (1991) have clearly shown that seven annual species of *Cicer* differ from each other in definite karyotypic features. The nuclear DNA content shows a 1.95 fold variation among the species. Table 1.2 gives the DNA content values in 7 annual species which form 3 groups. The mean values of DNA content in each group are separated by regular intervals of 0.8 pg while within group variation in 2C DNA amount is insignificant. Cultivars of *C. arietinum* show consistently similar karyotypes (excepting minor variation) and constant DNA amounts with minor differences. More significantly, no differences in 2C DNA content are detected in `Desi' and `Kabuli' cultivars of chickpea despite their ancient divergence and disparate cultivation associated with clear cut phenotypic differences.

Table 1.2 Nuclear DNA content in *Cicer* (Ohri and Pal 1991)

Species	2n=	2C Nuclear DNA (pg)	Group
C. judaicum	16	1.83	I
C. cuneatum	16	2.50	
C. bijugum	16	2.54	
C. pinnatifidum	16	2.56	II
C. reticulatum	16	2.65	
C. echinospermum	16	2.70	
C. arietinum `Desi'	16	3.30	
C. arietinum `Kabuli'	16	3.34	
C. arietinum ICC4918	16	3.39	III
C. arietinum ICC4973	16	3.47	
C. arietinum ICC5003	16	3.57	

In the karyotypic studies, many of the chromosomes of each *Cicer* species have distinctive features which facilitate the determination of their pairing. Remarkably, the maximum karyotypic asymmetry is associated with least DNA amount and successively more symmetrical karyotypes depict greater DNA amounts. The most significant finding is the presence of two satellite pairs (1^{st} and 2^{nd}) in *C. reticulatum* while there is only one such pair in

the remaining annual species. In *C. arietinum* and *C. cuneatum* it is the first pair, in *C. judaicum*, *C. bijugum* and *C. echinospermum* it is the second and in *C. pinnatifidum* it is the eighth indicating that the position of the satellite pair varies in the species.

1.2.4 Phylogeny on the basis of crossability and heterochromatin studies

Harlan and De Wet (1971) have categorized wild species on the basis of their usefulness for improving the cultigen. Species within the primary gene pool are readily intercrossed and produce progenies that are fully fertile or nearly so. Consequently, gene flow between the species of the primary gene pool can be accomplished by conventional breeding methods. Any partial fertility that appears is easily overcome by selection among the progenies. The secondary gene pool contains species which are somewhat distant from the cultigen. In such cases, hybridization to obtain gene flow is more difficult and the progenies have substantial degrees of sterility, usually because of chromosomal rearrangements. The tertiary gene pool contains those species that are related to the cultigen but their hybridization with the cultigen is not possible or the hybrids are completely sterile.

Based on crossability and fertility of hybrids in interspecific crosses, Ladizinsky and Adler (1976) have divided nine annual species of *Cicer* into four crossability groups. The first group includes three species namely, *C. arietinum*, *C. reticulatum* and *C. echinospermum* where *C. reticulatum* has been recognized as a subspecies (Moreno and Cubero 1978) and has been proposed as the putative progenitor of the cultivated *C. arietinum* (Kabir and Singh 1988, Ladizinsky 1975). *C. bijugum*, *C. pinnatifidum, C. yamashitae* and *C. judaicum* represent the second crossability group while *C. chorassanicum* and *C. cuneatum* which can neither be crossed with any other annual species nor with each other, form the third and fourth groups, respectively (Table 1.3) (Muehlbauer *et al.* 1994a).

Later on, Ahmad *et al.* (1988) have examined the barrier to interspecific hybridization within and between the crossability groups of chickpea. In their study, normal pollination and fertilization have been observed in many interspecific combinations, but factors resulting from the genetic disharmony

between the maternal and paternal genomes have been believed to be the cause of sterility. On the basis of the above studies, species in the primary, secondary and tertiary gene pools of chickpea are determined as follows wherein most of the species defined in the second, third and fourth crossability groups, except *C. cuneatum* and *C. yamashitae* defined by Ladizinsky and Adler (1976) formed the tertiary gene pool.

Primary gene pool : *C. arietinum*, *C. reticulatum* , *C. echinospermum*

Secondary gene pool : No species

Tertiary gene pool : *C. bijugum*, *C. pinnatifidum*, *C. judaicum*, *C. chorassanicum* , *C. montbretii*

Table 1.3 Crossability groups of chickpea on the basis of hybridization and fertility of hybrids (Muehlbauer *et al.* 1994a)

Crossability Group	Species
Group I	*C. arietinum, C. reticulatum, C. echinospermum*
Group II	*C. bijugum, C. pinnatifidum, C. judaicum, C. yamashitae*
Group III	*C. chorassanicum*
Group IV	*C. cuneatum*

Tayyar *et al.* (1994) applied the C-banding technique that would allow the identification of chromosomes within the species and also study relationships among the species. The measurements of the proportion of the total genome length occupied by the C-band heterochromatin indicated that the nine annual species of *Cicer* can be divided into two groups: Group I consisting of *C. bijugum* and *C. cuneatum* and Group II including *C. arietinum*, *C. chorassanicum*, *C. echinospermum* , *C. judaicum* , *C. pinnatifidum* , *C. reticulatum* and *C. yamashitae*. The significance of variation in heterochromatin content has been discussed in terms of phylogenetic advancement where a trend is observed for reduction in C-heterochromatin content with evolutionary advancement in the genus *Cicer*. Extensive heterochromatization or diminution of heterochromatin has been observed in several species as one of the major evolutionary pathways (Lavania and Sharma 1983, Fernandez and Davina 1991).

1.3 Biotic and Abiotic Stress: Responsible for Poor Average Yield in Chickpea

Chickpea has a world average yield of 700 kg/ha which is much below its potential of 4 tons/ha due to its susceptibility to various biotic and abiotic stresses as depicted in figure 1.4 (Singh *et al.* 1994a). The most important and destructive stress is due to various pathogens as indicated in Table 1.4 followed by the most important abiotic stress, drought (Singh *et al.* 1994a) resulting in a slow annual growth rate of chickpea production at 1.9% and yield at only 0.6%.

In chickpea, 49 pathogens were reported from 35 countries until 1978. By 1995, the number of pathogens increased to 172, which was reported from 55 countries. Figure 1.5 denotes the increase in different pathogens like fungi, bacteria, viruses over a period of 17 years. Among these, maximum number of pathogens have been reported from India alone with 35 pathogens in 1978 and 89 in 1995 (Nene *et al.* 1996). Among fungal diseases, wilt and blight are the most devastating diseases affecting chickpea in tropical and temperate regions, respectively. Figure 1.6 shows the chickpea field with wilted plants and a pod infected with Ascochyta blight. About 10-100% losses in yield due to Fusarium wilt is a regular feature in chickpea growing states of India. *F. oxysporum* is seed and soil borne and survives in plant debris in soil for more than 5 years. Losses due to Ascochyta blight have been upto 100% in Pakistan and India during epidemic years (Nene 1987). Besides pathogens, the insect pest *Helicoverpa armigera* represents the most important pest of chickpea. It feeds on foliage, flowers and developing seeds with a single larva damaging several pods of chickpea per day leading to severe losses in the crop yield (Figure 1.7).

Among the abiotic stresses, drought is the most important stress in chickpea, since it is mostly grown as a rainfed crop during the winter season in the Indian subcontinent (Singh *et al.* 1994a). Because of poor irrigation facilities in the major chickpea growing areas, about 85% of the crop is unirrigated and grown on soils with limited water storage capacity (Chaudhary *et al.* 1989, Sheldrake *et al.* 1979). Due to this situation, the crop invariably suffers from moisture stress at one or the other stages of development depending on the water availability in the soil. Terminal drought stress, which

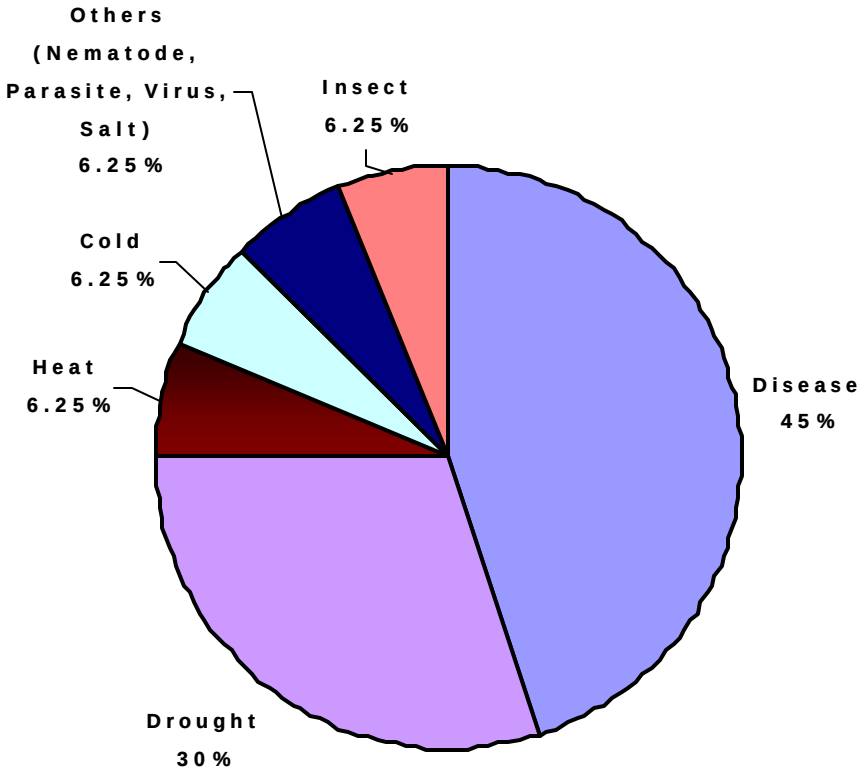


Figure 1.4 Relative importance of abiotic and biotic stresses affecting chickpea crop (Singh *et al.* 1994a).

Table 1.4 Important diseases affecting chickpea and the causative agents (Singh and Reddy 1991)

Disease	Causative agent
Soil borne fungal diseases	
Fusarium wilt	*Fusarium oxysporum* Schlecht.emend Snyd. & Hans. f.sp. *ciceri* (Padwick) Snyd. & Hans.
Verticillium wilt	*Verticillium dahliae* Reinke & Berth
Dry root rot	*Rhizoctonia bataticola* (Taub.) Butler [*Macrophomina phaseolina* (Tassi) Goid]
Collar rot	*Sclerotium rolfsii* Sacc.
Wet root rot	*Rhizoctonia solani* Khun
Black root rot	*Fusarium solani* (Mart.) sacc.
Phytophthora root rot	*Phytophthora megasperma* Drechs
Pythium root and seed rot	*Pythium ultimum* Trow.
Foot rot	*Operculella padwickii* Kheshwalla
Stem rot	*Sclerotinia sclerotiorum* (Lib) de Bary
Foliar fungal diseases	
Ascochyta blight	*Ascochyta rabiei* (Pass.) Lab. (*Mycosphaerella rabiei* Kovachevski)
Botrytis gray mold	*Botrytis cinerea* Pers.ex Fr
Alternaria blight	*Alternaria alternata* (Fr) Kiessler
Stemphylium blight	*Stemphylium sarciniforme* (Cav.) Wills
Rust	*Uromyces ciceris-arietini* (Grogn.) Jacz & Beyer
Viral diseases	
Stunt	Bean (pea) leaf roll virus
Nematode diseases	
Root knot	*Meloidogyne incognita* (Kofoid and White) Chitw., *M.javanica* (Treub) Chitw.
Cyst	*Heterodera ciceri* Vovlas, greco, and Di Vito
Root Lesion	*Pratylenchus thornei* Sher and Allen

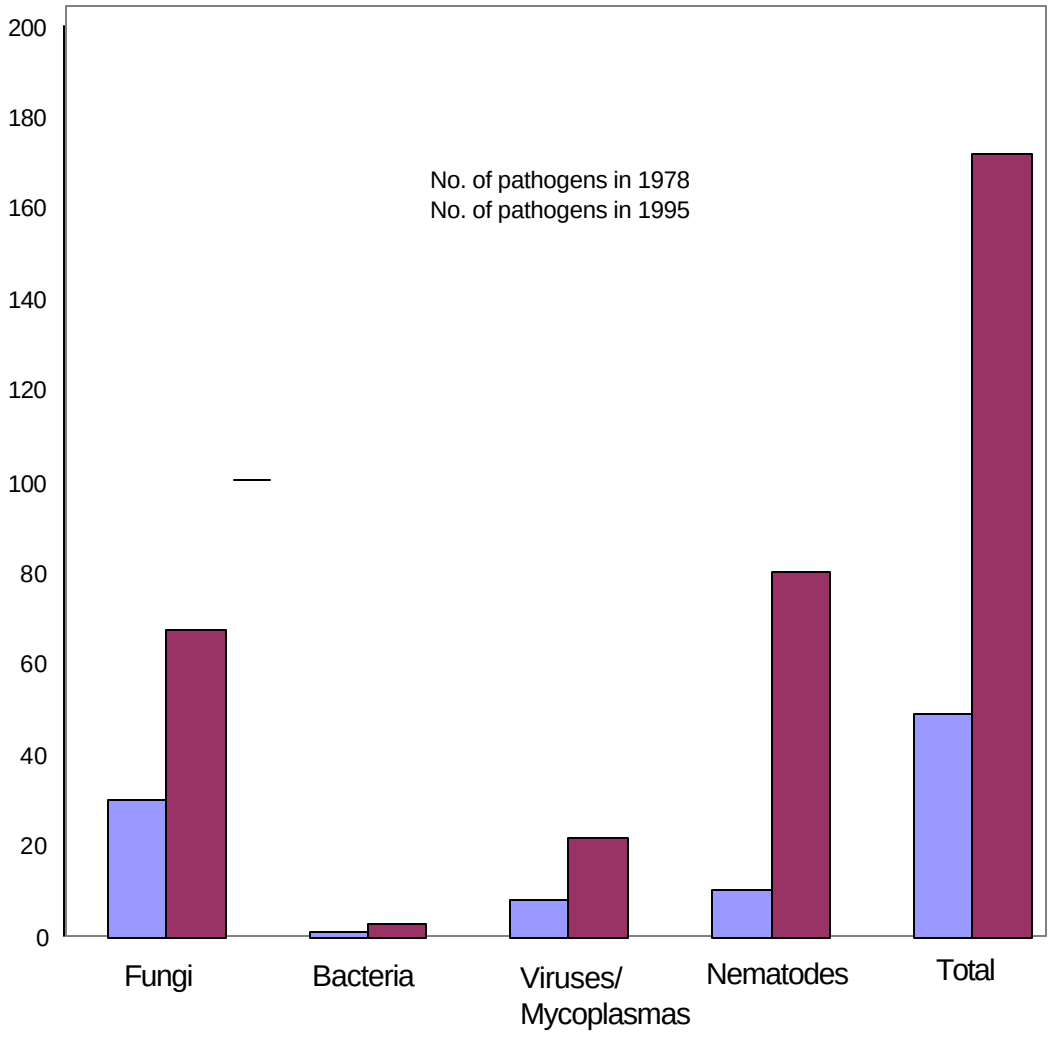


Figure 1.5 Increase in the number of different pathogens of chickpea over a period of 17 years (Nene *et al.* 1996)

a

b

Figure 1.6 (a) Chickpea field affected by *Fusarium oxysporum*. **(b)** Chickpea pod affected due to *Ascochyta rabiei* pathogen

a

b

Figure 1.7 (a) Chickpea pod infested by *Helicoverpa armigera* larvea and **(b)** *H. armigera* larvae feeding on chickpea plant

occurs during the pod-filling phase is a common yield reducer in chickpea (Nageshwara Rao *et al.* 1985 a, b). Pod filling is highly dependant on weather and varies from 8.97 to 56.73% (Pundir *et al.* 1992).

Cold is the second most important abiotic stress wherein the necessity of cold tolerant chickpea cultivars arises due to the yield advantage of winter sown chickpea over traditionally spring sown chickpea (Singh *et al.* 1994a). No correlation has been observed between the tolerance rating at the preflowering and seedling stages but susceptibility to cold is greater at the late vegetative stage than at the seedling stage (Singh *et al.* 1984). Cold resistance in chickpea plant tends to decrease from germination to flowering (Wery 1990).

1.4 Conventional Breeding Practices for Chickpea Improvement

Breeders have always made efforts to release varieties that can sustain the biotic and environmental stress coupled with a good yield. In chickpea, there have been a few efforts to improve quality and quantity through conventional breeding approaches. A brief description of these efforts is as follows.

1.4.1 Interspecific hybridization

Wild relatives that are closely related to the cultivated species, have the potential to provide the needed genetic variation for the improvement of the specific crop under study.

There have been many efforts to identify wild sources of *Cicer* resistant to biotic stresses such as *C. judaicum*, *C. montbrettii*, and *C. pinnatifidum* resistant to Ascochyta blight; *C. judaicum* to Fusarium wilt; *C. pinnatifidum* and *C. judaicum* to grey mold; and *C. bijugum* to cyst nematode (Singh *et al.* 1990). ICARDA has identified important wild species like *C. bijugum*, *C. judaicum* and *C. pinnatifidum* possessing multiple stress resistance (Singh *et al.* 1994a). *C. echinospermum* has seeds of a similar size to those of cultivated chickpea and is resistant to bruchids, leaf miner and Ascochyta blight (Singh *et al.* 1991a). However, to the best of our knowledge there are no reports available so far for the successful introgression of these genes in *C. arietinum*. Among the wild species of *Cicer*, *C. reticulatum* is cross compatible with chickpea (*C. arietinum*) whereas crossability of *C.*

echinospermum with chickpea is low resulting in sterile F_1 hybrids (Ladizinsky and Adler 1976, Pundir and van der Maesen 1983).

The low and unstable yields of chickpea can be ascribed to the narrow genetic base of *C. arietinum*. In 1984, Harlan compiled information on the improvement in chickpea seed yield through hybridization between cultivated species and their wild relatives. In previous studies by Singh and Ocampo (1993), 28-153% hybrid vigour was obtained in the F_1s of crosses between *C. arietinum*, *C. echinospermum* and *C. reticulatum* compared to the 75% reported in intraspecific crosses (Singh *et al.* 1984a). Singh and Ocampo (1993) also found numerous transgressive segregants for high yield in F_2 populations. The performance of these lines suggested that genetic reshuffling, originating from interspecific hybridization, could produce favourable combinations of genes expressing high yield. It was also observed that both the wild species contributed towards the recovery of lines superior in yield and free of any known undesirable traits from the wild species. When the lines were given an organoleptic test, the quality was as good as the cultigen *C. arietinum* suggesting that high yielding lines along with desired traits could be developed from interspecific hybridization in chickpea.

1.4.2 Intraspecific hybridization

Among the 50 diseases affecting chickpea, Ascochyta blight and Fusarium wilt are the most destructive and widespread. ICRISAT has evaluated more than 12,000 accessions and identified several hundreds of resistant sources which have been shared with national programs of which few have been released. Mexico was the first to breed wilt-resistant cultivars (e.g Surutato 77 and Sonora 80), while in India Punjab Agricultural University developed lines GL87078 and GL87079 through hybridization which were resistant to the four predominant races, namely race 1, 2, 3, and 4 of *Fusarium oxysporum* f.sp. *ciceri*. Likewise, ICARDA bred more than 1600 lines resistant to Ascochyta blight and shared with national programs releasing ultimately, 39 cultivars in 12 countries. In 1992, the first Australian cultivar bred specifically for increased Phytophthora resistance, 'Barwon' was released which was developed from a cross between moderately resistant

Russian accession CPI56564 and the slightly resistant Iranian accession ICC2903 (Singh *et al.* 1994a).

Diallel mating designs have been attempted in chickpea to obtain segregants giving high yields and resistance to diseases. Singh (1974) also suggested the possibility of deriving pure lines performing better than or as good as F_1 hybrids in chickpea through diallel mating. Further Singh *et al.* (1992) studied the nature of genetic control of important agronomic traits of chickpea by diallel analysis. They concluded that additive genetic effects were greater than non-additive effects for plant height, days to flowering, and 100 seed mass. Seed mass is a highly heritable and important yield component in chickpea which can be used effectively as an indirect selection criterion for improving seed yield (Singh and Paroda 1986).

Double poddedness in chickpea has also been reported to contribute positively towards higher productivity in chickpea, probably through rapid increase in the sink capacity of the plant and additional photosynthetic activity that takes place in pods (Singh *et al.* 1994). In 1998, Rubio *et al.* developed NILs by crossing a kabuli single podded cultivar grown in Spain and a double podded Indian local desi cultivar, JG62. Their study indicated that there could be differences in yield stability over locations between double and single podded lines with double podded lines being more stable than single podded lines. Further the results indicated that the `double pod' character was fully compatible with the `large seeds' characteristic of the high quality chickpea used for human consumption in Western Mediterranean countries. Singh (1987) had earlier suggested that `double pod' could decrease seed size in chickpea. However, studies by Rubio *et al.* (1998) suggested that single/double podded gene was not linked to any other gene related to seed size and hence, this character could be introduced in most breeding programs.

1.5 Molecular Markers in Chickpea Genomics

1.5.1 Diversity and phylogeny of chickpea as revealed by molecular marker studies

Genetic diversity studies in a crop are important for various aspects such as management of genetic resources, identification of duplicate

accessions in the germplasm and in applied breeding programs. In initial studies low copy sequence RFLP markers revealed very low levels of polymorphism (Udupa *et al.* 1993) while, microsatellite based RFLP markers later demonstrated the potential of detecting higher polymorphism in chickpea (Weising *et al.* 1992, Sharma *et al.* 1995). In their extensive studies, Serret *et al.* (1997) analyzed the genetic diversity among different accessions of cultivated chickpea of the 'Kabuli' type from different countries of the Eastern and Central Asia; and Mediterranean regions using the $(GATA)_n$ probe. The genetic distance data revealed that the genetic diversity in chickpea was the greatest in Pakistan, Iraq, Afghanistan, South-East Russia, Turkey and Lebanon and lower in Iran, India, Syria, Jordan and Palestine. It was further concluded that there are 3 centres of diversity for chickpea namely Pakistan-Afghanistan, Iraq-Turkey and Lebanon.

To throw light on the phylogenetic relationships in *Cicer* species, extensive studies were carried out using various parameters such as allozymes (Kazan and Muehlbauer 1991, Tayyar and Waines 1996), seed storage proteins (Ahmad and Slinkard 1992), RFLP (Patil *et al.* 1995) and RAPD (Ahmad 1999). In these studies, *C. reticulatum* was shown as the progenitor of *C. arietinum* and its implications could be considered at the secondary level, since the annual state represents the primary level in the evolution of chickpea (Ahmad *et al.* 1992).

In their further work, Udupa *et al.* (1999) studied allelic variation for the locus specific $(TAA)_n$ microsatellite loci in a world wide sample comprising 72 landraces, 4 improved cultivars and 2 wild species of the primary gene pool (*C. reticulatum* and *C. echinospermum*). A positive correlation was observed between the average number of repeats (size of the locus) and the extent of variation. All the 76 accessions of cultivated chickpea could be readily distinguished with these markers. The STMS primer pairs flanking $(TAA)_n$ repeats could also amplify polymorphic fragments from *C. reticulatum* and *C. echinospermum*. The microsatellite polymorphism in *C. arietinum* did not reveal any relationship between accession and geographical origin, giving no clue to the mode of geographical dispersal of the species. Based on these results and the previous isozyme and RAPD analysis in chickpea (Ahmad *et al.* 1992, Anonymous 1994), it was concluded that the present world wide

distribution of cultivated chickpea is due to recent dispersal from its site of origin which is the region of present day Turkey, Afghanistan, Pakistan, India and the former USSR (Ladizinsky and Adler 1976).

In continuation with the work on STMS markers, Choumane *et al.* (2000) investigated the potential of these markers for improving chickpea using the primary and secondary gene pool. They examined whether and to what extent STMS primers could be applied to genome analysis of wild *Cicer* species. Their phylogenetic studies on the wild *Cicer* species revealed that the species in the first crossability group were closely related to the perennial *C. anatolicum* and distantly related to *C. cuneatum*. For synteny studies within the genus *Cicer* the suitability of STMS markers derived from chickpea was assessed and two conclusions were arrived at: (i). STMS from chickpea were of limited use as syntenic markers beyond the first crossability group and (ii). Sequence of amplicons derived from species of other crossability groups differed so much from chickpea sequence that they most probably represented different loci.

1.5.2 Genetic linkage map of *Cicer* genome

A complete genetic map representing regions of all chromosomes is important in locating genes of unknown map position. The first genome map in plants was reported in maize followed by the map in rice and *Arabidopsis* using RFLP markers (Joshi *et al*. 1999). In *Cicer*, Gaur and Slinkard (1990) reported the first rudimentary gene map, which included 26 isozyme and 3 morphological loci. Kazan *et al.* (1993) examined the linkage relationships of the genes for several isozymes and morphological traits and extended the linkage map of chickpea on the basis of F_2 population derived from the crosses between two cultivated lines and between cultivated line and the wild species *C. reticulatum* and *C. echinospermum*.

The first molecular map of chickpea using STMS markers was reported by Winter *et al.* (1999). This map was constructed with the objective of obtaining a core map of chickpea based on co-dominant markers to which more dominant markers can be added. In this study, initially locus specific primer pairs were designed for 22 microsatellite containing regions from *C. arietinum* which also generated amplification products in *C. reticulatum*, the

closest relative of *C. arietinum*. Characterization of the simple sequence repeats from *C. arietinum* genome and amplification of the alleles in the wild progenitor made them a favorite choice for linkage map construction of the *Cicer* genome. The map generated using these STMS markers covered 613 cM, whereas the previous chickpea map covered 550cM (Simon and Muehlbauer 1997). Also the fact that these repeats were transcriptionally active implied their importance in map based cloning.

Comparison of linkage groups of pea, lentil and chickpea revealed the existence of common linkage groups consisting of homologous loci (Kazan *et al.* 1993, Simon and Muehlbauer 1997). It was observed that six linkage groups of chickpea containing 17 loci were conserved in pea and lentil. However, there were some portions of the linkage groups that were not conserved in all the three genera probably due to chromosomal repatterning that changed the location of the genes after their divergence from a common ancestor. In the most recent finding by Grant *et al.* (2000), significant synteny between soybean and *Arabidopsis* was observed. Using the data from only 3 linkage groups of soybean and the information currently available from the *Arabidopsis* Genome Initiative, they were able to demonstrate synteny between the two genera and show the commonality of duplicated segments spanning 10-20 cM in soybean with *Arabidopsis*. The existence of synteny can help target the genes among the related crops, facilitate cross utilization of genetic resources and shed light on evolutionary events associated with the divergence of the genomes.

1.5.3 Initial efforts of tagging Fusarium wilt and Ascochyta blight resistance genes in chickpea

Fusarium wilt and Ascochyta blight are the two most important fungal diseases of chickpea common in tropical and temperate regions respectively, of chickpea agriculture. There have been a number of efforts to develop resistance against these diseases. For example ICRISAT in its decade long screening program identified wilt resistance in 160 out of more than 13,500 chickpea accessions tested (Haware *et al.* 1992).

Marker assisted selection can be applied effectively using DNA markers linked to resistance genes for accelerating disease resistance

breeding programs. For example, genes for resistance to *Foc* 4 and *Foc* 5 races of *Fusarium oxysporum* f.sp. *ciceri* causing wilt in chickpea have been mapped to only one linkage group 4 by Simon and Muehlbauer (1997). In a study by Mayer *et al.* (1997), two allele specific associated primer (ASAP) based markers linked to the gene conferring resistance to *Foc* 1 have been mapped on the same linkage group implying that the gene conferring resistance to *Foc*1 is also located on linkage group 4. Thus genes for resistance to Fusarium wilt races 1, 4 and 5 are clustered in chickpea. A similar situation of clustering of resistance genes for different races of pathogen and for different pathogens has also been demonstrated in several other plants like rice and soybean (Kanazin *et al.* 1996, Yu *et al.* 1996). Such a phenomenon may be the result of gene duplication, exon shuffling and recombination processes that are thought to have generated different resistance genes from one or a few ancestral progenitor genes (Michelmore 1996, Hammond-Kosack and Jones 1997). Compared to Fusarium wilt resistance work, less is known about Ascochyta blight resistance which is considered as a quantitative trait. In a study for identifying molecular markers for Ascochyta blight resistance, recombinant in-bred lines from an interspecific cross between *C. arietinum* and *C. reticulatum* were subjected to marker analysis and molecular markers linked to two major QTLs, QTL1 and QTL2 were identified which together accounted for 50.3% of the estimated phenotypic variation, (Santra *et al.* 2000).

1.5.4 Future prospects

Legume breeders have taken interest in chickpea and have contributed many varieties performing well in the fields during last several decades. Progress has been done in chickpea research using several biotechnological tools in the last 10 years. However, research in cereals like wheat and rice has been accelerated due to the integration of conventional and modern biotechnology. Similar progress needs to be achieved in a legume like chickpea, where several genetic resources among the wild germplasm are unexploited. It is necessary to have a high-density chickpea genome map to identify molecular markers linked to agronomically important traits. These markers can assist in breeding programs for generating tailor made varieties

complementary to the region-wise requirements. Further, large insert BAC libraries of any crop are useful for positional cloning and identification of important genes. Soybean is a good example in this context where Danesh *et al.* (1998) have used the BAC library of soybean for identifying clones near a major cyst nematode resistance gene. Meksem *et al*. 2000 have utilized the BAC libraries of soybean for chromosome walking and genome wide physical mapping. Very recently, cDNA library has been used to construct microarrays of 4089 unique genes to ascertain baseline expression data for certain tissues and stages of soybean development as well as various treatments like pathogen challenge or stress response (Vodkin *et al*. 2001). The physical maps have also been used to develop new genetic markers like micro-satellites-in-dels and SNPs in soybean, *Arabidopsis* and *Fusarium solani* in regions of the genome which lack conventional genetic markers (Iqbal *et al.* 2001). These approaches can be used in chickpea improvement since we have a limited understanding of the complex genetic inheritance and functional mechanism of the agronomic traits in chickpea. Concerted efforts of genomic technologies and effective utilization of the genetic space (Cooper 2001) need to be undertaken to obtain new genotypes that are superior to those in hand.

1.6 Objectives

Chickpea being an important legume crop, I started my research work with the following objectives:

1. Genomic diversity analysis using RAPD and microsatellite markers to understand the prevalent genetic variation in the chickpea germplasm. Further, to study the utility of these markers in heterosis analysis in this crop.

2 To investigate the presence, distribution and potential of transposable elements as molecular markers in chickpea since they represent an important part of the plant genomes.

3. Application of DNA markers by mapping them on the genome map of chickpea and also for identifying putative markers linked to Fusarium wilt resistance gene in chickpea.

1.7 Organization of Thesis

The research results obtained by me during the last five years on chickpea genomics have been detailed in following four chapters apart from this first chapter of review of literature.

Chapter 2: Potential of DNA markers in detecting divergence and in analysing heterosis in Indian elite chickpea cultivars.

In this chapter, genetic diversity studies in various Indian elite chickpea cultivars possessing important agronomic traits were studied and genetic distance between 8 cultivars which were crossed in a diallel manner was correlated with heterosis using DNA markers.

Chapter 3: Ty1-*copia* retrotransposon like elements in chickpea genome; their identification, distribution and use for genetic diversity.

This chapter deals with identification and distribution of Ty1-*copia* retrotransposon elements in chickpea genome and exploring the potential of these elements for use as molecular markers.

Chapter 4: Molecular mapping of the RAPD and ISSR markers on the integrated map of chickpea genome and its exploitation for Fusarium wilt resistance tagging in chickpea.

This chapter involves mapping of RAPD and ISSR markers on the integrated map of *Cicer* using recombinant inbred lines developed from the interspecific cross between *C. arietinum* and *C. reticulatum*. Further, segregation of additional markers of different types was studied for using recombinant inbred lines from an intraspecific cross of *C. arietinum* developed for tagging the Fusarium wilt resistance.

Chapter 5: General discussion

Studies on chickpea so far are still in its infancy as compared to the crops like wheat and rice among the cereals and soybean among the legumes. In this chapter, an attempt is made to predict the future line of research in chickpea genome considering the data available in well studied crops like rice and *Arabidopsis*.

CHAPTER 2

Potential of DNA Markers in Detecting Divergence and in Analysing Heterosis in Indian Elite Chickpea Cultivars

The contents of this chapter have been published as a full-

length paper in: Theoretical and Applied Genetics (1999) 98: 1217-1225

2.1 Introduction

2.2 Materials and Methods

2.2.1 Seed material and DNA isolation

2.2.2 RAPD assay

2.2.3 Hybridization based microsatellite analysis

2.2.4 Field evaluation and data collection

2.2.5 Statistical analysis

2.3 Results

2.3.1 Assessment of genetic diversity in elite chickpea germplasm using RAPD and microsatellite DNA markers

2.3.1.1 RAPD-PCR analysis

2.3.1.2 Microsatellite based diversity analysis

2.3.1.3 Estimation of genetic relatedness

2.3.2 Potential of DNA markers in predicting F_1 performance and heterosis in chickpea

2.3.2.1 Polymorphism using molecular markers

2.3.2.2 Genetic distance among parents and clustering of parental lines

2.3.2.3 Hybrid performance and heterosis

2.3.2.4 Correlation of hybrid performance with molecular marker heterozygosity

2.4 Discussion

2.4.1 Usefulness of microsatellites in fingerprinting chickpea genotypes

2.4.2 Divergence among chickpea cultivars based on microsatellite markers

2.4.3 Diallel mating: A molecular approach

2.1 Introduction

Knowledge of genetic diversity and relatedness in the germplasm is a prerequisite for crop improvement programs. Morphological traits (Staub *et al.* 1995) and isozyme markers (Kazan *et al.* 1991) which have been routinely used earlier for this purpose have several limitations such as their limited number, environmental dependence, and temporal and spatial expression. In the last two decades, DNA markers have shown a great promise for cultivar identification (Moser and Lee 1994), diversity analysis (Vasconcelos *et al.* 1996), construction of genetic maps (Song *et al.* 1991) and tagging agronomically important genes (Kelly 1995). Another important application of DNA markers is prediction of heterosis in hybrids. DNA-based markers have also been used extensively to correlate genetic diversity and heterosis in several crops such as maize (Ajmone Marsan *et al.* 1998), oat (Moser *et al.* 1994, O`Donoughue *et al.* 1994), rice (Xiao *et al.* 1996, Zhang *et al.* 1994) and wheat (Zhong *et al.* 1991). It has been reported in maize that measures of similarity based on restriction fragment length polymorphism (RFLP) and pedigree knowledge could be used to predict superior hybrid combinations (Smith *et al.* 1990). However, both low and high correlations between DNA based genetic distance and heterosis are reported in various crops (Barbosa-Neto *et al.* 1996, Melchinger *et al.* 1990, Zhang *et al.* 1994).

Comprehensive investigation towards genetic diversity in chickpea has not been carried out so far, except a report by Sharma *et al.* (1995) who studied the abundance and polymorphism of 38 different microsatellites in four chickpea accessions. In my work I have employed RAPD and microsatellite markers to assess the genetic diversity in 29 Indian elite chickpea genotypes. This data could aid in designing crosses for pyramiding desired traits to produce superior chickpea genotypes. I have also examined the relationship of genetic diversity with hybrid performance and heterosis to investigate the utility of these markers for the same purpose.

2.2 Materials and Methods

2.2.1 Seed material and DNA isolation

Seeds of all the accessions of chickpea (*C. arietinum* L.) were obtained from the Pulses Research Centre of the Mahatma Phule Agricultural University (Mahatma Phule Krishi Vidyapeeth, MPKV), Rahuri, India. Table 2.1 lists the cultivars used, wherein the eight parental lines which were crossed to obtain a diallel set of twenty-eight crosses (excluding reciprocals) during the rabi season (October to January) of 1994-95 are marked separately.

The plants were grown in greenhouse and leaf tissue was collected and frozen at -80^{o}C. Total DNA was extracted from the frozen tissue by the CTAB method with a slight modification (Rogers and Bendich 1988). Young leaf tissue was ground in liquid nitrogen and mixed with 15 ml of CTAB extraction buffer (2% hexadecyltrimethyl ammonium bromide, 100 mM Tris-HCl pH 8.0, 20 mM EDTA pH 8.0, 1.4 mM NaCl, 1% polyvinyl pyrrolidone). The homogenate was then incubated at 60°C for 15 min, emulsified with an equal volume of $CHCl_3$:IAA (24:1) and centrifuged at 10,000 rpm for 10 min. To the supernatant an equal volume of CTAB precipitation buffer (1% CTAB, 50 mM Tris-HCl pH 8.0, 10 mM EDTA pH 8.0) was added, mixed gently and centrifuged at 10,000 rpm for 10 min. The pellet was dissolved in high salt TE buffer (1 M NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA) and was precipitated with twice the volume of chilled ethanol. The precipitate was washed with 70% ethanol, dried and redissolved in TE buffer for further purification. DNA was quantified in a fluorometer according to the supplier`s instructions (Hoefer instrument model TKO-100).

Table 2.1 Details of *Cicer arietinum* L.. germplasm used for diversity analysis.

Genotype	Agronomic characters
*Vijay, Phule G-8505-7	High yield, wilt resistant, drought tolerant temperature tolerant, late sown
Phule G-89224, GCP-102, Phule G-92014 Phule G-92005, *Phule G-12, Phule G-8501-1, Phule G-91025, *ICCV-10, *Phule G-89219, *Phule G-91028	High yield, wilt resistant
JG-74, Phule G-93044, Phule G-92028	Temperature tolerant/late sown, high yield, wilt resistant,
Vishwas (Phule G-5), *Vishal (Phule G-87207)	High yield, bold seeded, wilt tolerant
Phule G-8505-10	High yield, drought tolerant, wilt resistant
Jaki-9324	High yield, bold seeded
Phule G-92007	High yield, bold seeded, wilt resistant, root rot resistant
ICC-410	High yield
ICC-31	Drought tolerant
*ICC-4958	Bold seeded, drought tolerant
JG-315	Temperature tolerant /late sown, wilt resistant.
BG-372	Temperature tolerant/late sown, high yield
*Bheema	Bold seeded
C-235, BG-390 JG-62	High yield, wider adaptability, wilt susceptible Wilt susceptible, twin podded

* indicates the cultivars used for diallel analysis

2.2.2 RAPD assay

RAPD assays were performed using random 10-mer oligonucleotide primers from Operon Technologies Inc., USA. Amplification reaction was carried out in 25 µl volume containing 10 ng of genomic DNA, 1.5 mM $MgCl_2$, 50 mM KCl, 10 mM TAPS [3-tri (hydroxylmethyl) methyl aminopropane sulfonic acid], 0.01% gelatin, 100 µM each of dATP, dCTP, dGTP and dTTP (Amersham Pharmacia biotech, USA), 5 pmoles primer and 0.6 U Taq DNA polymerase (Bangalore Genei Pvt. Ltd., India). Amplification was performed in

a PTC-100 (Perkin Elmer Inc., U.S.A) thermocycler programmed for 5 min at 94°C followed by 45 cycles of 1 min at 94°C, 1 min at 36°C and 2 min at 72°C. This was followed by a final extension at 72°C for 5 min. Amplification products were analyzed by gel electrophoresis in 1.9% agarose in 0.5X TAE buffer pH 8.0 and visualized by ethidium bromide staining. All the reactions were repeated at least 3 times and only reproducible bands were used in further analyses.

2.2.3 Hybridization based microsatellite analysis

Genomic DNA was digested with various restriction enzymes such as *Alu*I, *Dra*I, *Hae*III, *Hinf*I and *Taq*I according to the manufacturer`s instructions (Promega Life Science, USA). The digested DNA was size-fractionated on agarose gels (0.8-1.2 %) in batches with a standard molecular weight marker while keeping all the other conditions constant and later the gels were dried in-vacuo. The dry gels were denatured, neutralized and then equilibrated in 6X SSPE (sodium chloride, sodium dihydrogen phosphate, EDTA, pH 7.4). Oligonucleotides such as $(AAC)_5$, $(GATA)_4$ etc. were labelled with γ-^{32}P-ATP and purified by ion exchange chromatography on DE-52 column (Ali *et al.* 1986). The gels were hybridized with oligonucleotide probe at Tm-5°C (Miyada *et al.* 1987) overnight in hybridization buffer containing 5X SSPE, 5X Denhardt`s reagent, 0.1% milk powder and 0.1% SDS. Stringent washes were given and the hybridized gels were autoradiographed with intensifying screens at -70°C.

2.2.4 Field evaluation and data collection

The F_1 hybrids obtained from 28 crosses between the eight parents along with the parental lines were grown in rabi season of 1995-96 for phenotypic evaluation in a randomized complete block design with three

replications, at MPKV, Rahuri, India. Each plot consisted of 3 m long rows spaced 30 cm apart. Distance between two plants in a row was maintained as 15 cm. Five randomly selected competitive plants in each plot were used for recording the observations. Various agronomic traits were examined for the experimental material. Means over replications were recorded for each trait and used in data analysis. The data obtained was used for combining ability analysis using Griffing`s model I, method 2 (Griffing *et al.* 1956). The analysis of variance for combining abilities and estimation of variance components were carried out. Yield potential was calculated as a function of its yield contributing components with the formula, Yield potential=1/100 x pods/plant x grains/pod x 100 seed weight.

2.2.5 Statistical analysis

Pairwise comparisons of degree of band sharing were made, and similarity index (S.I.) values were calculated by Nei`s method as S.I. = $2 N_{ab} / N_a + N_b$ where N_a = total no. of bands present in lane a, N_b = total no. of bands present in lane b, N_{ab} = no. of bands common to lanes a and b (Nei *et al.* 1979). A dendrogram was constructed using TAXAN version 4.0 software based on the degree of band sharing. Probability of identical match by chance by which two genotypes would show identical band pattern was calculated as $(X_D)^n$ where X_D was the average similarity index value and n was the average number of total bands shared per probe per primer (Wetton *et al.* 1987). Midparent and better-parent heterosis were estimated for seed yield/plant (g), pods/plant, number of seeds/pod, and 100 seed-weight (g) and the significance of the percentage heterosis over better-parent and midparent was tested by the least significant difference method. The relationships between genetic distance and heterosis/hybrid performance were evaluated by regressing heterosis or trait values on the genetic distance in the F_1 hybrids.

2.3 Results

2.3.1 Assessment of genetic diversity in elite chickpea germplasm using RAPD and microsatellite DNA markers.

2.3.1.1. RAPD-PCR analysis

To obtain stable and reproducible DNA fingerprints which could discriminate chickpea cultivars, it was necessary to optimize the PCR protocol with regard to concentrations of template DNA, primer and magnesium ions. Although, chickpea DNAs amplified with different concentrations of magnesium (1.5, 2.0, 2.5 mM), strong and reproducible bands were obtained at 1.5 mM magnesium concentration which was later used in all the experiments. Thirty-five random primers were used to amplify the chickpea DNAs. Each primer produced about 2-12 amplification products with various cultivar DNAs with average number of bands as 8.75 per primer, per accession. The lengths of the amplification products varied from 0.26 kb to 3.0 kb. Out of 35 primers used, 10 primers generated polymorphic patterns as listed in Table 2.2. As shown in figure 2.1, primer OPF09 generates a monomorphic pattern with exception of one band of 600 bp which is present only in cultivars Vijay (lane1), ICC-4958 (lane 3), PG-89224 (lane 6), GCP-102 (lane 7), ICC-410 (lane 9), ICCV-10 (lane 15), PG-5 (lane 17), PG-87207 (lane 18), Jaki-9324 (lane 19), PG-93044 (lane 20), BG-390 (lane 25), PG-12 (lane 27) and Bheema (lane 29). Absence of amplification products using this primer, is observed in lane 28 which could be due to low concentration of template DNA resulting in undetectable PCR products on the agarose gel. Based on the RAPD patterns, pairwise comparison was made between all the genotypes for each primer. A total of 254 amplification products were scored out of which 14.56% were polymorphic in nature. The genetic distance values varied from 0.02-0.22 with an average value of 0.13.

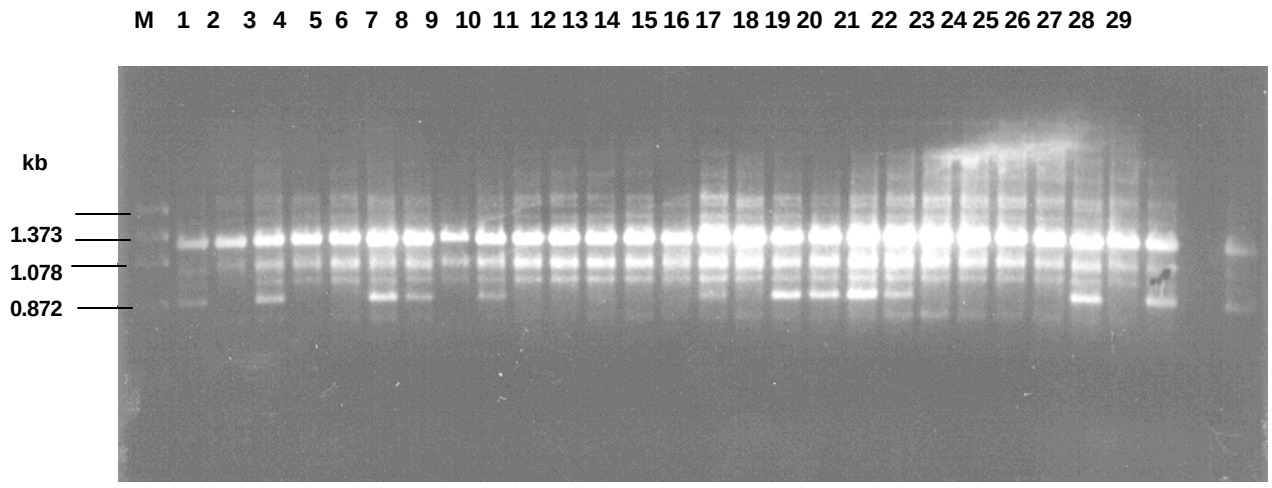


Figure 2.1 **RAPD fingerprint of the elite chickpea cultivars with primer OPF09. M: ØX-174/*Hae*III digest marker. Lane nos. 1 to 29 are: 1. (Vijay), 2. (PG-8505-7), 3. (ICC-4958), 4. (PG-8505-10), 5. (ICC-31), 6. (PG-89224), 7. (GCP-102), 8. (PG-92014), 9. (ICC-410), 10. (PG-92005), 11. (PG-8501-1), 12. (PG-92007), 13. (PG-91025), 14. (JG-74), 15. (ICCV-10), 16. (JG-315), 17. (PG-12), 18. (PG-87207), 19. (Jaki-9324), 20. (PG-93044), 21. (PG-92028), 22. (BG-372), 23. (JG-62), 24. (C-235), 25. (BG-390), 26. (PG-89219), 27. (PG-12), 28. (PG-91028), 29. (Bheema).**

Table 2.2 Primers amplifying polymorphic bands among the chickpea cultivars.

Primer	Sequence	No.of polymorphic bands	Average number of bands
OPA 05	AGGGGTCTTG	5	10
OPA07	GAAACGGGTG	4	9
OPA08	GTGACGTAGG	3	5
OPA10	GTGATCGCAG	5	12
OPA11	CAATCGCCGT	2	6
OPA12	TCGGCGATAG	4	12
OPA 19	CAAACGTCGG	4	8
OPF 09	CCAAGCTTCC	1	5
OPJ 06	TCGTTCCGCA	4	11
OPJ 07	CCTCTCGACA	3	5

2.3.1.2 Microsatellite based diversity analysis

Genomic DNAs of 29 cultivars of chickpea digested with restriction endonucleases *Alu*I, *Hinf*I, *Dra*I, *Taq*I and *Hae*III, individually, separated on agarose gel which were dried and hybridized to various microsatellite probes. Table 2.3 lists the average number of bands per genotype obtained with each enzyme-probe combination. Most of the bands generated by the probes were polymorphic in the chickpea genotypes. Five microsatellites namely $(AAC)_5$, $(AAG)_5$, $(ACG)_5$, $(ACT)_5$ and $(GATA)_4$ yielded clear polymorphic patterns with 12 to 23 average bands per genotype which could distinguish all the genotypes under the present study. Probe $(ACG)_5$, however, yielded less than 6 bands on hybridization with *Taq*I digest. Figures 2.2 a and 2.2 b show the representative pattern of polymorphic bands obtained with *Taq*I-$(GATA)_4$ combination which can fingerprint all the 29 cultivars with bands ranging from 9 kb to 2 kb. As seen in the figure, majority of the lanes show clear hybridization profiles except lane 9 which shows poor hybridization signals.

Table 2.3 Average number of bands produced by the enzyme-oligonucleotide probe used in fingerprinting the elite chickpea cultivars.

Oligonucleotide probe	Restriction Endonuclease	Average no. of bands
$(AAC)_5$	*Dra*I	12.5
	*Taq*I	14.2
$(AAG)_5$	*Taq*I	12.3
$(ACG)_5$	*Taq*I	5.6
	*Hae*III	15.0
$(ACT)_5$	*Dra*I	23.0
	*Hae*III	18.0
$(GATA)_4$	*Alu*I	13.7
	*Hinf*I	14.6
	*Dra*I	13.7
	*Taq*I	14.2
	*Hae*III	17.0

A total of 1916 loci were scored out of which 632 were similar on pairwise comparison of the 29 cultivars. The average genetic distance value for all the oligonucleotide-enzyme combinations listed in Table 2.3 was 0.560 with the lowest and the highest genetic distance values being 0.39 and 0.82, respectively. The probability of identical match by chance $(X_D)^n$ denoting the probability of two genotypes having identical band profiles was calculated to be 7.8×10^{-7}, 2.92×10^{-6}, 5.75×10^{-8}, and 1.77×10^{-6} for probes, $(AAC)_5$, $(ACT)_5$, $(AAG)_5$ and $(GATA)_4$, respectively while 2.32×10^{-25} using all the 4 enzyme-probe combinations (Table 2.4).

Table 2.4 Probability of identical match by chance

Probe	Highest S.I	Lowest S.I	X_D	n	$(X_D)^n$
$(AAC)_5$	0.71	0.078	0.337	13.36	$7.8X10^{-7}$
$(ACT)_5$	0.780	0.352	0.575	23.08	$2.92X10^{-6}$
$(AAG)_5$	0.577	0.000	0.260	12.37	$5.75X10^{-8}$
$(GATA)_4$	0.72	0.123	0.396	15.01	$1.77X10^{-6}$

X_D = Average similarity index for all pairwise comparisons, n=Average no. of bands shared per primer

per probe, $(X_D)^n$ =Probability of identical match by chance, S.I=Similarity index

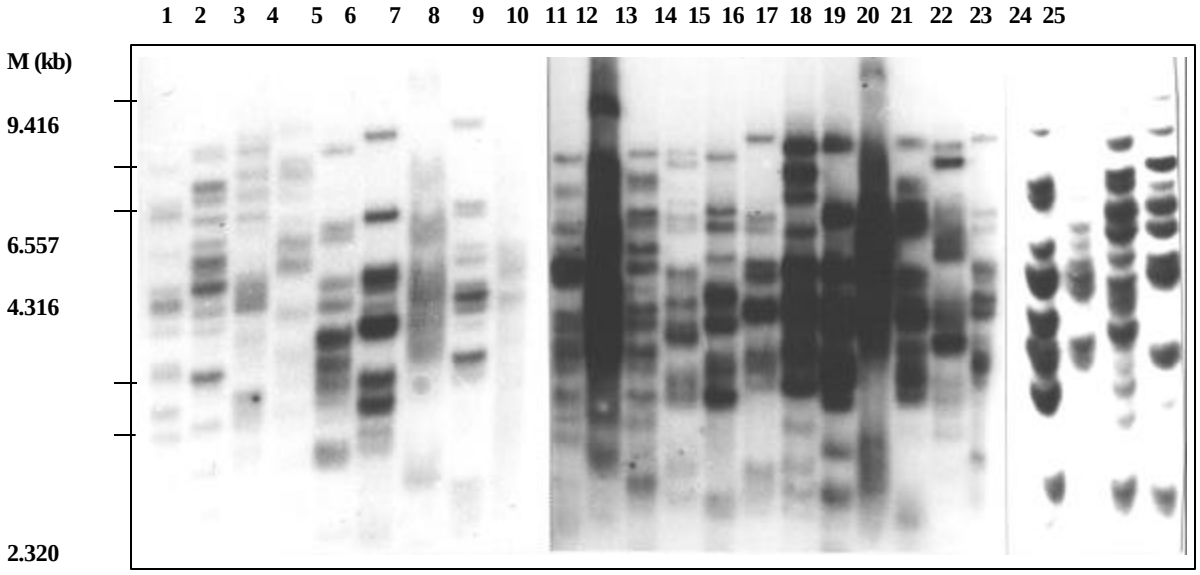


Figure 2.2a **Oligonucleotide fingerprint with *Taq*I-(GATA)$_4$. M: Lambda *Hind*III digest marker. Lane nos. 1 to 25 are 1. (Vijay), 2. (PG-8505-7), 3. (ICC-4958), 4. (PG-8505-10), 5. (ICC-31), 6. (PG-89224), 7. (GCP-102), 8. (PG-92014), 9. (ICC-410), 10. (JG-74), 11. (ICCV-10), 12. (JG-315), 13. (PG-12), 14. (PG-87207), 15. (Jaki-9324), 16. (PG-93044), 17. (PG-92028), 18. (BG-372), 19. (JG-62), 20. (C-235), 21. (BG-390), 22. (PG-92005), 23. (PG-8501-1), 24. (PG-92007), 25. (PG-91025).**

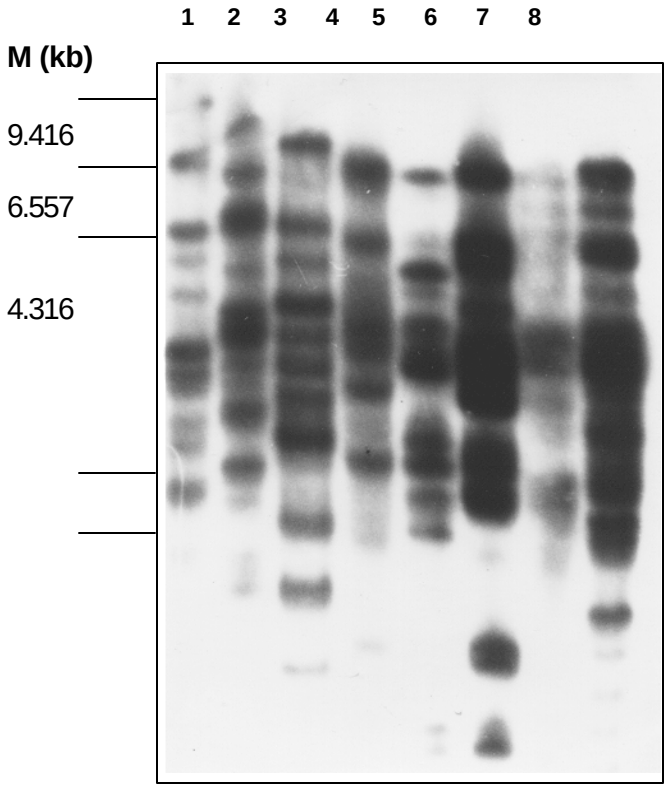


Figure 2.2b **Oligonucleotide fingerprint with *Taq*I-(GATA)$_4$. M: Lambda *Hind*III digest marker. Lane nos.1 to 8 are 1. (PG-89219), 2. (Vijay), 3. (ICCV-10), 4. (PG-12), 5. (PG-91028), 6. (PG-87207), 7. (ICC-4958), 8. (Bheema).**

2.3.1.3 Estimation of genetic relatedness

Computer software TAXAN version 4.0 was used to construct a dendrogram on the basis of similarity index values obtained from microsatellite marker data of the 29 chickpea genotypes (Figure 2.3). Genetic distance calculated from the dendrogram showed a high level of genetic diversity in the range of 0.39-0.82 in chickpea elite germplasm. The chickpea genotypes formed five main clusters with eight cultivars genetically distant and outgrouped from the main clusters. Cultivars Vijay, PG-92014, PG-92007 and PG-91025 formed first cluster (cluster I) at a genetic distance of 0.47 while cultivars GCP-102, JG-74, and Vishal together formed cluster II at a genetic distance of 0.49. Cluster III contained PG-8505-10, ICC-410, PG-8501-1 and ICCV-10 with the cultivar PG-8505-7 loosely clustering in this group at a genetic distance of 0.58. Cluster IV consisted of ICC-4958, JG-315, Vishwas and ICC-31 and Cluster V included cultivars PG-93044, JG-62, BG-390 and PG-92028 at a genetic distance 0.48, while cultivar BG-372 was equidistant (0.49) from both of these clusters. Cultivars PG-89224, Jaki-9324, C-235, PG-91028, PG-92005, PG-89219, PG-12 and Bheema were genetically distinct from the above clustered cultivars as well as from each other with the exceptions of (Jaki-9324 and C-235) and (PG-12 and Bheema) as they grouped within themselves.

2.3.2 Potential of DNA markers in predicting F_1 performance and heterosis in chickpea

2.3.2.1 Polymorphism using molecular markers

Eight parents used in the diallel mating set were selected on the basis of their phenotypic characters and were studied for analyzing DNA polymorphism. Twelve microsatellite-restriction enzyme combinations and 35 RAPD primers were used to generate polymorphic patterns revealing 31.62% and 7.3% polymorphism among the parents, respectively. For increased genome coverage, 388 polymorphic loci including 85 RAPD and 303 microsatellite were considered to study genetic

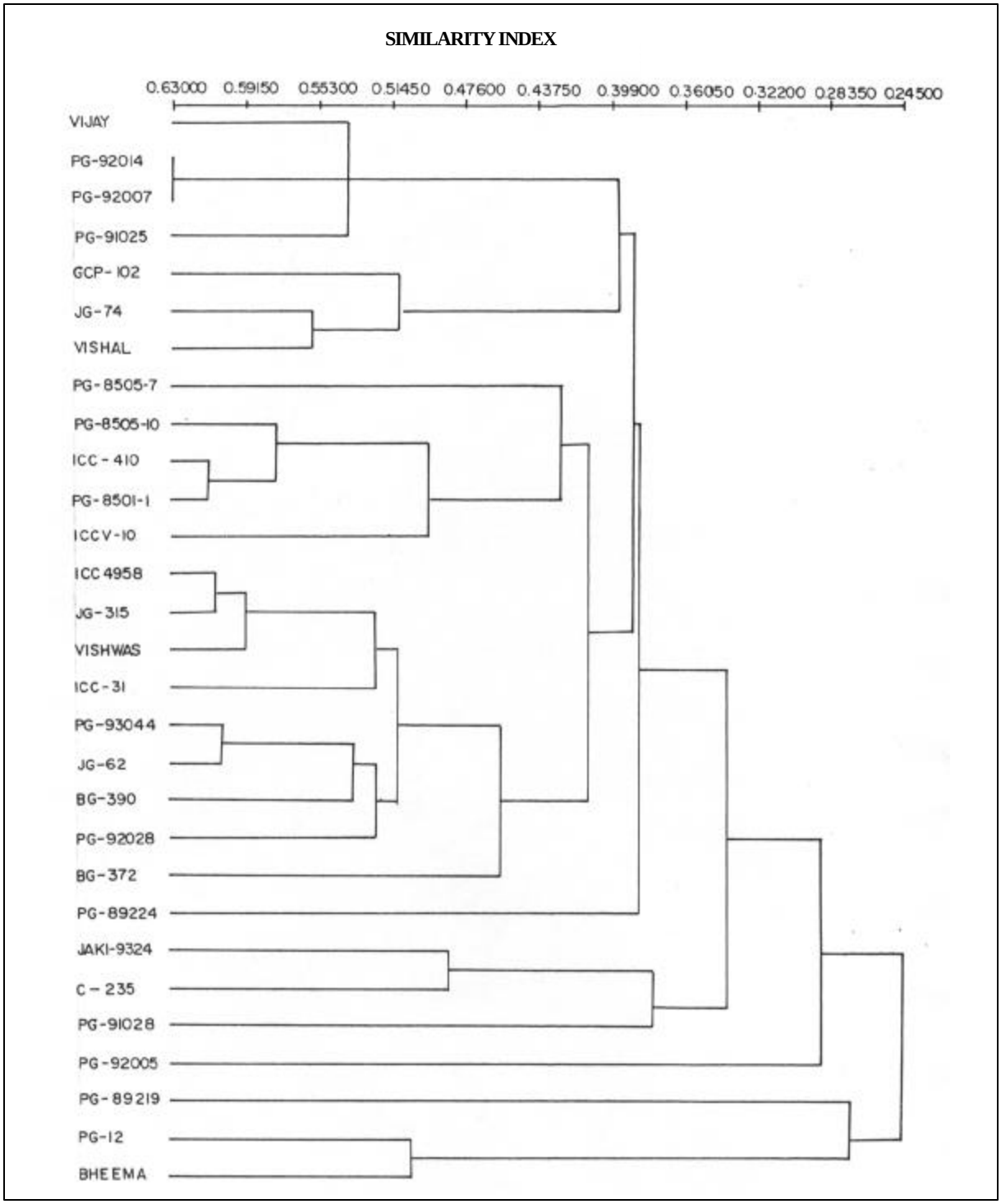


Figure 2.3 Dendrogram based on the similarity index values of 29 elite chickpea genotypes using microsatellite markers.

distance among eight parents under consideration.

2.3.2.2. Genetic distance among parents and clustering of parental lines

Genetic distance was calculated from the Nei's similarity index values for all the 28 combinations of the eight parents considering RAPD and microsatellite approaches individually as well as together. Based on the RAPD markers alone, the genetic distance ranged from 0.09 to 0.27, while that on the basis of microsatellite markers ranged from 0.42 to 0.61. However, genetic distance based on both the molecular markers together ranged from 0.26 to 0.40 which was used to generate dendrogram as in figure 2.4. From the dendrogram, it is seen that the eight parental lines cluster into one major group containing six cultivars, while the two cultivars Vijay and ICCV-10 group out from the others. The major group consists of 2 subgroups, PG-89219 and PG-12 forming one subgroup and PG-91028, PG-87207, Bheema and ICC-4958 forming the other. PG-91028 and PG-87207 from the second subgroup share a common parent, K-850. Bheema and PG-87207, both boldseeded, are the closest at a distance of 0.26, however, pedigree of Bheema was not available to comment on the clustering based on pedigree. When the same eight parents were analyzed for their morphological and yield component traits using D^2 statistics, they clustered into 2 groups (Mahalanobis 1936). Here, Vijay, PG-91028, and ICC-4958 clustered together, ICCV-10, PG-87207 and Bheema formed a second cluster while PG-12 and PG-89219 were outgrouped (data not shown) .

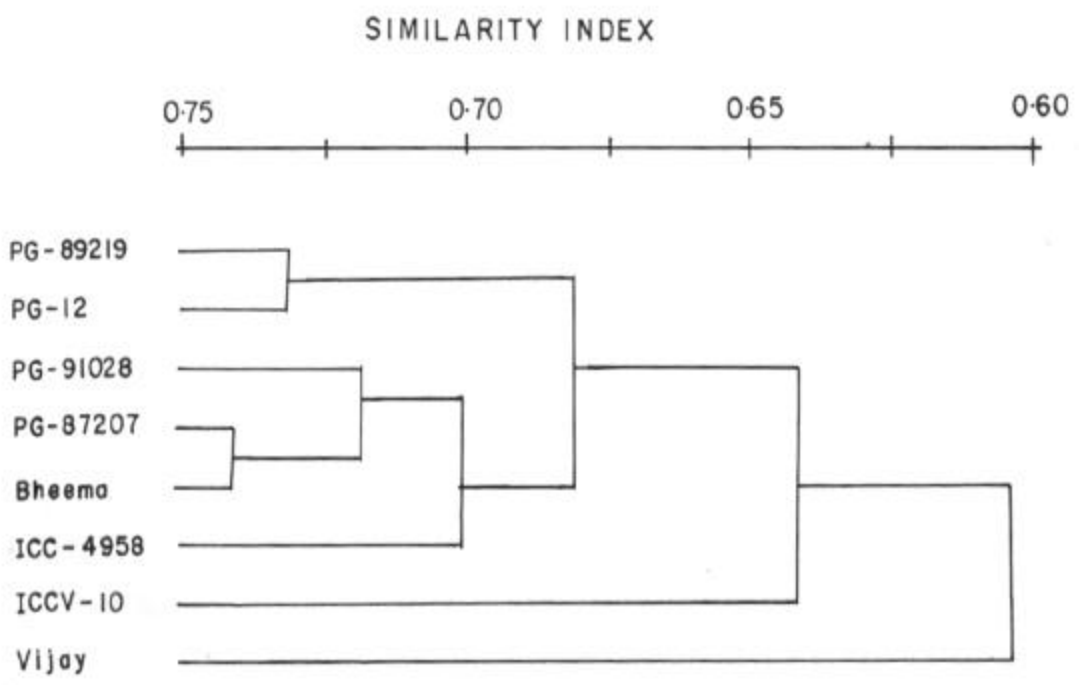


Figure 2.4 Dendrogram displaying the similarity index values of the eight genotypes used for diallel analysis.

2.3.2.3 Hybrid performance and heterosis

The midparent heterosis values as well as specific combining ability effects for the agronomic traits are given in Table 2.5. As observed in this table, the degree of heterosis and hybrid performance vary significantly for each trait. Highest midparent heterosis is obtained for seed yield (76.16%, P=0.01) followed by number of pods/plant (57.64%, P=0.01%), number of secondary branches/plant (30.64%, P=0.01) and 100-seed weight (9.91%, P=0.05). Comparatively number of basal branches/plant and 100-seed weight exhibit low magnitude of heterosis. Among the other traits examined such as plant height, days to 50% flowering and days to maturity, no significant correlation was seen with grain yield. However, a correlation at 1% significance was obtained between plant spread and heterosis for grain yield (r=0.507, P=0.01).

Midparent heterosis in seed yield/plant ranges from 21.76% to 76.16% in the crosses with PG-89219 as one of the parents and is maximum compared to all other crosses. Also heterosis over midparent in number of secondary branches, number of pods/plant and 100-seed weight is more with PG-89219 as one of the parents compared to all other crosses (Table 2.5). As mentioned in Table 2.5, crosses of PG-12, PG-87207, ICC-4958 with PG-89219, Vijay x PG-91028 and ICCV-10 x ICC-4958 exihibit significant specific combining ability effects for seed yield/plant. General combining ability effects were calculated for each trait as presented in Table 2.6. PG-89219 is a good combiner for seed weight/plant, number of seeds/pod, number of pods/plant and number of secondary branches. PG-87207 and ICC-4958 are good combiners for 100 seed-weight, whereas Vijay shows a significant combining ability for number of pods/plant and seed yield/plant. Bheema is a good combiner for 100 seed weight and seed yield/plant and ICCV-10 shows significant combining ability for number of seeds/pod. PG-91028 shows good combining ability for number of basal branches/plant, number of secondary branches/plant and number of pods/plant.

2.3.2.4 Correlation of hybrid performance with molecular marker heterozygosity

Significant positive heterosis resulted when crosses were performed between two subgroups for example PG-89219 and PG-12 from one subgroup with other parents namely, PG-91028, PG-87207, Bheema, ICC-4958, ICCV-10 and Vijay for the yield/plant as shown in Table 2.5. However, the correlation between genetic distance and heterosis for any of the traits was not linear. For example, the highest midparent heterosis of 76.16% was obtained on crossing PG-89219 and PG-12 when the genetic distance between them was 0.2646, whereas midparent heterosis of -13.24% was obtained on crossing PG-87207 and Bheema having a genetic distance of 0.2587. It is evident from the data in Table 2.5 that there is no correlation between genetic distance and the heterosis in different traits.

Table 2.5 Genetic distance, percent midparent heterosis for different agronomic characters, SCA effects in crosses of the diallel set (SCA=specific combining ability).

Cross	Genetic distance	% Mid parent heterosis for						SCA(seed yield/plant)
		No.of sec-branches	No of pods/ plant	No.of basal branches	Plant-spread	100 seed weight	Yield / plant	
PG89219xVijay	0.3996	5.1	0.52	-11.7	19.9	8.09	23.83**	0.329
PG89219xICCV10	0.3549	29.82**	17.88**	1.52	30.50	-3.96	23.11**	0.118
PG89219xPG12	0.2646	1.97	57.64**	8.60	32.27	9.91**	76.16**	13.78**
PG89219xPG9102	0.3129	14.69**	15.40**	-17.88	29.2	0.258	21.76**	0.449
PG89219xPG8720	0.2919	9.21*	10.40**	-13.45	-6.85	-12.46	22.23	5.38**
PG89219xICC4958	0.3540	1.00	46.86**	15.69**	11.98	-20.98	36.74**	6.89**
PG89219xBheema	0.3299	21.88**	13.63**	7.19	14.34	-7.19	10.57	-0.723
VijayxICCV10	0.3556	-13.44	7.26	-6.99	-16.59	1.74	11.31	0.348
VijayxPG12	0.4051	16.20**	21.50**	0.00	4.37	3.10	31.96**	3.982*
VijayxPG91028	0.3596	-21.93	26.23**	-8.33	-13.40	6.54	37.72**	10.753**
VijayxPG87207	0.3729	-1.98	-18.37	-1.48	2.77	2.33	-7.79	-2.083
VijayxICC4958	0.3711	4.95	1.77	8.6	-21.94	-18.17	2.22	-5.246
VijayxBheema	0.3631	-24.11	-0.48	1.23	-29.15	-14.12	-11.27	1.334
ICCV10xPG12	0.3471	18.11**	23.66**	-9.46	-5.96	-5.16	23.08	-2.904
ICCV10xPG91028	0.3180	2.75	5.65	1.23	17.75	7.39	1.24	-1.334
ICCV10xPG87207	0.3131	-12.76	-8.41	-8.14	-3.01	-18.32	-4.94	-1.325
ICCV10xICC4958	0.3467	9.00	56.94**	7.19	-6.10	-18.32	28.48**	8.217**
ICCV10xBheema	0.3146	- 4.58	10.22	-2.43	-23.63	-30.03	8.98	3.022
PG12xPG91028	0.3047	-30.58	17.03	-2.91	1.79	2.20	22.99	2.145
PG12xPG87207	0.3044	-6.62	-7.17	-4.44	-15.69	-6.50	-12.30	-15.544
PG12xICC4958	0.3153	30.64**	39.98**	-0.22	-4.20	-11.71	25.92	3.605
PG12xBheema	0.3060	-4.48	-3.86	-4.11	-14.33	-11.19	0.52	-2.781
PG91028xPG8720	0.2774	-21.46	-0.05	-12.89	-7.53	-10.92	0.33	1.880
PG91028xICC4958	0.2952	-14.28	3.27	-0.22	21.11	-21.99	-10.92	-6.97
PG91028xBheema	0.3031	-11.37	10.84**	-12.33	-1.21	-23.62	-1.31	0.623
PG87207xICC4958	0.2923	-14.68	-24.07	1.34	-13.24	-5.86	22.54	-8.211
PG87207xBheema	0.2587	7.97	-19.50	-16.49	-15.43	-5.22	-13.24	-2.213
ICC4958xBheema	0.2748	4.51	9.24	1.52	-8.07	6.91	8.85	3.98*

* indicates significant at P=0.05, ** indicates significant at P=0.01

Table 2.6 General combining ability effects of the parents for different characters.

	Character	PG-89219	Vijay	ICCV-10	PG-12	PG-91028	PG-87207	ICC-4958	Bheema
1.	Days to 50% flowering	2.91**	-1.783**	0.283	-0.483	-0.483	-0.317	-0.683*	-0.58
2.	Days to maturity	1.620**	-1.583**	-1.45**	1.217**	1.217**	-0.25	-0.517	-1.35**
3.	Plant height	0.816*	-2.012**	-0.727**	0.782*	0.782**	-2.511**	-0.181**	1.98**
4.	Plant spread	0.148	0.148	-1.655**	-0.648**	-0.648**	-1.082**	0.695**	3.78**
5.	No. of basal branches/plant	-0.68**	0.039	0.006	0.056*	0.058**	-0.094**	-0.038	0.062
6.	No. of secondary branches/plant	1.312**	-1.222**	-1.412**	0.968**	0.968**	-0.872**	-1.282**	0.36
7.	No. of pods/plant	19.763**	3.629**	2.50*	-1.95	7.542**	-20.58**	-12.80**	2.067
8.	No. of seeds/pod	0.124**	-0.073**	0.193**	-0.003	-0.003	0.029	-0.132**	-0.1**
9	100 seed weight	-2.707**	-1.068**	-1.843*	-1.787**	-1.787**	1.583**	5.485**	3.40**
10	Seed yield/plant	3.978**	2.715**	0.72	-3.618	-3.618**	-2.325**	1.725**	3.30**

* indicates significant at P=0.05, ** indicates significant at P=0.01

2.4 Discussion

2.4.1 Usefulness of microsatellites in fingerprinting chickpea genotypes

From our data low polymorphism in chickpea germplasm with RAPD markers was observed and hence these markers may not be suitable in revealing genetic diversity in `Desi' chickpea genotypes. Extensive DNA polymorphism, however, has been reported using RAPD markers in several other crop plants (Hilu and Stalker 1995, Morell *et al.* 1994, Ranade and Sane

1996). In the earlier work on pigeonpea from our laboratory, it has been shown that RAPDs could reveal a high degree of polymorphism in wild species of pigeon pea (Ratnaparkhe *et al.* 1995).

Oligonucleotides representing microsatellites have detected high levels of polymorphism in all the chickpea cultivars under present study. When 25 microsatellite-enzyme combinations were attempted on 29 elite chickpea germplasm, 4 microsatellites namely, $(AAC)_5$, $(ACT)_5$, $(AAG)_5$, and $(GATA)_4$ gave distinct fingerprints for the 29 cultivars. From the $(X_D)^n$ values it is clear that $(AAG)_5$ which is the most polymorphic microsatellite could distinguish maximum number of cultivars though it gives less number of bands when compared with other microsatellites (Table 2.3). The microsatellite hybridization in our study has revealed the genetic diversity in the range of 39% to 82% which indicates the overwhelming potential of these oligonucleotide probes in cultivar identification, genetic characterization and relatedness in the chickpea germplasm.

2.4.2 Divergence among chickpea cultivars based on microsatellite markers

Cultivars (PG-92014 and PG-92007) and (GCP-102 and Vishal) clustered together as they shared a common parent, KPG-36 and K-850, respectively, which probably resulted in similar fingerprinting profiles. Cultivars PG-8505-7 and PG-8505-10 selected from the same population of a cross between WR-315 x Sel-436, clustered together at a genetic distance of 0.59. PG-8505-7 is temperature tolerant/late sown as against PG-8505-10, thus differing in the phenotypic characters. PG-92014 and PG-92005 selected from the population [(KPG-36 x P-326) x ICC12271], and ICCV10 and BG372 selected from the population (P-1231 x P-1265) separated by genetic distance of 0.76 and 0.60, thus indicating a divergence of the cultivars from each other during the succeeding cycles of selfing. The diversity thus observed with

microsatellite markers in the chickpea germplasm is probably due to use of landraces in most of the Indian subcontinent and even today these landraces are being used for the development of elite cultivars (Malhotra *et al.* 1987). However, the genetic diversity between the various landraces still remains to be studied and molecular markers such as microsatellites will be greatly useful in quantifying this diversity.

2.4.3 Diallel mating : A molecular approach

Predicting hybrid performance has always been a primary objective in all hybrid crop breeding programs (Hallauer and Miranda 1988). Evaluation of hybrids for their performance in the field is expensive and time consuming. As a result, many parameters such as pedigree information, qualitative and quantitative traits (Smith *et al.* 1990, Wang *et al.* 1992) and biochemical data (Leonardi *et al.* 1991) are being used to study parental diversity leading to prediction of heterosis. In this study a diallel set was attempted and the hybrid performance was correlated with molecular marker heterozygosity which is the first report of its kind in elite chickpea lines.

It was observed that the hybrids were more heterotic for seed yield than for yield component traits. This is obvious because yield (Y) is generally a multiplicative function of three component traits viz, number of pods/plant (P), number of grains per pod (G) and 100-seed weight (W); (Y = PxGxW). Heterosis of component traits would multiplicatively amplify each other to produce a much larger heterosis in the ultimate trait. On applying t-test for the observed and predicted yield in the hybrids a 't' value of 2.329 at $P<0.05$ was obtained. Also a significant positive correlation between plant spread and grain yield indicated the potential avenue for increasing grain yield in chickpea. Secondly, it was observed that one of the parents, PG-89219 when crossed with others, consistently gave significant heterosis. The analysis of variance for combining ability and the estimation of variance components

indicated that additive x dominance gene interaction played important role in the inheritance of seed yield and other characters except 100-seed weight. Hence PG-89219 can be used in the crossing program for developing a new line with high performance.

Although no linear correlation was obtained between genetic diversity and heterosis, it was evident that significant heterosis resulted for hybrids by crossing parents from two different subgroups (Table 2.5 and figure 2.4). Hybrids obtained after crossing parents from the same group gave poor heterosis for yield. Thus our results suggest that the concept of genetic divergence for maximum expression of heterosis has certain limitations in chickpea. It was earlier suggested that hybrids showing heterosis were usually developed from parental lines diverse in relatedness, ecotype, geographic origin etc. (Yuan *et al.* 1985). However, it was observed in maize that heterosis manifested by hybrids developed from genetically diverse varieties was less than that between varieties which were genetically less diverse (Moll *et al.* 1965). In wheat, a low correlation was reported between heterosis and DNA based genetic distance (Barbosa-Neto *et al.* 1996). With different sets of maize cultivars, low correlation was detected between combining ability and RFLP-based genetic distance. Apparently, crosses between extremely divergent parents create a situation where the harmonious functioning of alleles is disrupted. Consequently the physiological functions are not efficient resulting in low heterosis. In fact, doubts were expressed about the usefulness of increased genome coverage for calculating marker distance and correlating it with hybrid performance to improve the efficiency of the prediction (Melchinger *et al.* 1990, Boppenmaier *et al.* 1993). Alternatively, identification of marker loci and genotypes significantly associated with traits of interest was suggested. Thus correlations calculated using specific heterozygosity based on the positive markers would be more significant than those based on general heterozygosity.

In summary, the diversity analysis based on microsatellites in my study has revealed the usefulness of these markers in identification of polymorphism in chickpea genome. Such markers will be highly efficient to identify specific markers linked to trait of interest. Secondly, although the PCR based RAPD markers and microsatellite markers could cluster the eight genotypes under the diallel studies into heterotic groups, the genetic distance and heterosis were not in linear correlation. It is, therefore, essential that specific DNA markers need to be developed in this system for efficient and reliable estimation of genetic distance for predicting heterosis.

CHAPTER 3

Ty1-*copia* Retrotransposon like Elements in Chickpea Genome; their Identification, Distribution and Use for Diversity Analysis

The contents of this chapter have been published as a full-length paper in: GENE (2000) 257: 157-166.

3.1 Introduction

3.2. Materials and Methods

3.2.1 Isolation and analysis of retrotransposon-like sequences from *Cicer* genome

3.2.2 Cloning and sequence analysis of PCR products

3.2.3 Southern blotting and slot blot analysis

3.2.4 PCR based genetic diversity in *Cicer* species

3.3 Results

3.3.1 Presence of Ty1-*copia* retrotransposon like sequences in chickpea

3.3.2. Sequence alignment studies of Ty1-*copia* retrotransposon like elements from chickpea

3.3.3. Copy number estimation of Ty1-*copia* retrotransposon like sequences in chickpea

3.3.4 Ty1-*copia* retrotransposon element detects genetic diversity in *Cicer* species

3.4 Discussion

3.4.1 Reverse transcriptase sequence variation of Ty1-*copia* group retrotransposons

3.4.2 Genomic distribution of Ty1-*copia* retrotransposon like elements in chickpea

3.4.3 Genetic divergence among the *Cicer* species based on the Ty1-*copia* element

3.1 Introduction

Genome size varies remarkably in different plant species, for example it is 145 Mbp in *Arabidopsis thaliana* and 10^5 Mbp in lily (Arumuganathan and Earle 1991). This variation in the DNA content mostly depends on difference in the amount of repeated DNA sequences, since the number of structural genes does not vary much (Bennetzen *et al.* 1996). A large part of repetitive DNA in plants is interspersed throughout the genome. In many cases, the interspersed sequences have been characterized and partly found to be retrotransposon like in nature (Flavell *et al.* 1986). Retrotransposons differ from other transposons by their ability to transpose *via* RNA intermediate that is converted into extrachromosomal DNA by reverse transcription, prior to reinsertion into the genome.

Transposable elements have been separated into two major groups (class I and II) depending on their mode of transposition (Finnegan 1989, Flavell *et al*. 1994). Class I transposable elements include retrotransposons which are distinguished into two major groups: LTR and non-LTR retrotransposons. The LTR retrotransposons have been further subdivided into Ty1-*copia* like and Ty3-*gypsy* like retroelements (Schmidt 1999). The *gypsy* and the *copia* elements were first found in *Drosophila* and they differed in the arrangement of three domains namely, protease, integrase and reverse transcriptase. Retrotransposons of LTR-containing superfamily also include yeast elements Ty (1, 2, 3) wherein Ty1 and Ty2 are *copia*-like and Ty3 is *gypsy*-like retrotransposon (Ivanov and Il'in 1995). The non-LTR retrotransposons terminate with a poly A or A rich sequences called as LINES (long interspersed repetitive elements) and SINES (short interspersed repetitive elements) and are found to be in high copy numbers in plant species (Kumar and Bennetzen 1999). It is proposed that retrotransposons may possibly be involved in gene duplication as well as in regulation of expression of flanking genes (White *et al.* 1994). Insertion of retrotransposon and change in copy number of a gene flanked by a retrotransposon have been seen under stress conditions like pathogen infection, wounding and *in vitro* cell culture or tissue culture (Grandbastien 1998). This transpositional activity contributes to genetic diversity in plant genomes and DNA

polymorphism caused by such activity has been used for linkage analysis (Ellis *et al.* 1998) and for DNA fingerprinting (Fukuchi *et al.* 1993).

In the present study, I have identified the potential of another DNA-based molecular system such as Ty1-*copia* retrotransposon element for diversity analysis among wild and cultivated species of *Cicer.* Sequence analysis has suggested that there is sequence heterogeneity among Ty1-*copia* elements in chickpea while Southern analysis has further indicated that the copy number is high in the cultivated *C. arietinum* as compared to *C. reticulatum*.

3.2. Materials and Methods

3.2.1 Isolation and analysis of retrotransposon-like sequences from *Cicer* genome

C. arietinum seeds were obtained from Mahatma Phule Agricultural University, Rahuri, India and DNAs of other *Cicer* species were a kind gift from Dr. Fred Muehlbauer, USDA (United States Department of Agriculture), Pullman, USA. DNA was extracted from leaf tissue by CTAB method as described in 2.2.1. Primers 5' CAN GCN TTY YTN CAY GG 3' and 3' ATR CAN CTR CTR TAC RA 5' (where R=puRine A & G; Y =pYrimidine C & T; N = aNy base) were obtained from Prof. Andrea Brandes, John Innes Centre, Norwich, UK. They correspond to peptide sequences TAFLHG and YVDDML which flank the internal domain of reverse transcriptase gene of Ty1-*copia* transposon. PCR was performed using DNAs from 3 accessions of *Cicer arietinum*: Vijay, ICC-4958, and JG-62 and one accession of *Cicer reticulatum* (PI489777) with the following cycling conditions (Flavell *et al.* 1992): 94°C for 1 min followed by 35 cycles each comprising 1 min at 94°C, 1 min at 45°C and 1 min at 72°C with a final elongation of 7 min at 72°C. The reaction volume of 20 µl contained all the constituents as in 2.2.2 except that the concentration of *Taq* DNA polymerase used was 2.5 U while that for the primer was 40 pmoles. The PCR products were resolved on 1% agarose and viewed under UV.

3.2.2 Cloning and sequence analysis of PCR products

Two PCR products of size ~280bp and ~650bp were obtained on PCR

amplification in all the accessions which showed a background smear after electrophoresis on 1% agarose gel. These PCR products were cloned in pMos*Blue* T-vector as per the manufacturer`s instructions and transformed into *E.coli* strain XL-1B cells (Amersham Pharmacia biotech, USA). Plasmid was isolated by the alkaline lysis method (Sambrook *et al.* 1989) and purified using the Wizard Plus Miniprep DNA purification system (Promega Corp, USA). Sequencing of the two clones namely CA1 (~280bp) from *C. arietinum* and CR10 (~650bp) from *C. reticulatum* was done by the dideoxy method using the Sequenase Version 2.0 DNA sequencing kit (Amersham Pharmacia biotech, USA.).

The BLAST 2.0 program was used for amino acid and nucleotide sequence homology search against the sequences in the GenBank and EMBL database. Multiple sequences were aligned using the Clustal W (1.8) software (Thompson *et al.* 1994) and amino acid sequences were deduced using the package Sequid.

3.2.3 Southern blotting and slot blot analysis

About 10 µg of genomic DNAs of four *C. arietinum* accessions were digested with *Bam*HI and *Dra*I individually and electrophoresed on 1% agarose gel in TAE buffer, pH 8.0. The inserts of CA1 (~280bp) and CR10 (~650bp) were released by *Hind*III and *Bam*HI digestion, the latter was also digested with *Eco*RI, and resolved on 1.5% agarose gel. These gels were southern blotted on Hybond N membrane as described by Sambrook *et al.* (1989). Specific probes as detailed in section 3.3.4 were radiolabelled by random primed method and were hybridized to blots for 16 hr in buffer containing 5X Denhardt's solution, 5X SSPE, 0.1% SDS and 1% milk powder at 52°C. The blots were washed at room temperature with 0.5X SSPE, 1% SDS twice for 15 min each and then at 60°C for 10 min. The hybridization signals were visualised by autoradiography.

Slot blot analysis of *C. arietinum* (cv Vijay; 1C=1.6 pg) and *C. reticulatum* (cv PI489777; 1C=1.3 pg) DNA (0.25 µg) was performed as described by Pearce *et al.* (1996). CA1 (~280 bp) was quantified spectrophotometrically, diluted serially and transferred on to Hybond N

membrane to use as a control. The hybridization was carried out as detailed above using CA1 as probe to determine its copy number.

3.2.4 PCR based genetic diversity in *Cicer* species

PCR was performed using DNAs of 9 wild *Cicer* species and 10 accessions of *C. arietinum* using the primer pair as described in 3.2.1. Radiolabelled $\propto^{32}$P-dATP was incorporated in the PCR reaction with the annealing temperature at 50°C to avoid non-specific binding of primers. The samples were denatured and loaded on 6% denaturing polyacrylamide gel and amplified bands were viewed by autoradiography.

3.3 Results

3.3.1 Presence of Ty1-*copia* retrotransposon like sequences in chickpea

Degenerate primers flanking Ty1-*copia* rt sequence amplified products of ~280 bp and ~650 bp in both *C. arietinum* and C. *reticulatum*. One of the clones CR10 (~650 bp) from *C. reticulatum* was digested with *Eco*RI giving two bands of size 450 bp and 200 bp which was probed with CA1 (~280 bp) from *C. arietinum*. The latter hybridized to both the fragments of size 450 bp and 200 bp bands as well as to itself (positive control of CA1 insert). Amplification of PCR products of expected molecular weight indicated the possible presence of Ty1-*copia* elements in chickpea.

3.3.2. Sequence alignment studies of Ty1-*copia* retrotransposon like elements from chickpea

CA1 and CR10 clones were sequenced completely and both the sequences were submitted to the DDJB, EMBL and GenBank nucleotide databases under the accession nos. AF264002 and AF264003. The derived amino acid sequences from nucleotide sequence data for both the clones were subjected to BLAST search and related sequences were aligned using CLUSTAL W program.

The CA1 clone consisted of 87 amino acids and 5 stop codons at positions 28, 35, 43, 48 and 62 in the 3rd ORF of this clone. BLASTX search

with predicted amino acid sequence of CA1 showed homology to 86 rt sequences of Ty1-*copia* retrotransposons. Figure 3.1a shows the alignment of the predicted peptide sequence of CA1 with Ty1-*copia* rt from *Solanum tuberosum*, *Nicotiana tabaccum, Lycopersicon chilense, Lycopersicon esculentum, Petunia, Allium cepa* and soybean. A similarity score ranging from 26% to 30% was obtained between CA1 and these sequences.

The predicted peptide sequence of CR10 obtained from *C. reticulatum* consisted of 207 amino acids. There were 18 stop codons in the proposed peptide of rt gene positioned at 15, 16, 33, 34, 49, 54, 71, 98, 110, 113, 115, 121, 128, 141, 147, 153, 159, and 171 in the 3rd ORF. Length of CR10 was greater than the expected length of the Ty1-*copia* rt region. BLASTX search of predicted peptide sequence of CR10 showed homology to Ty1-*copia* rt at the 3rd ORF from amino acid residues 154 to 200 at the *3'* end. The peptide from amino acid residues 154 to 225 was taken for sequence alignment with previously characterized Ty1-*copia* like retrotransposon sequences. Figure 3.1b shows the alignment of this region with *Lycopersicon esculentum, Brassica napus, Solanum tuberosum, Alstroemeria ligtu,* and soybean with a similarity score of 26% to 37%.

As mentioned above, the size of the clone CR10 was longer than the expected ~280 bp for rt region. The 71 amino acids at the 3' end giving homology to Ty1-*copia* like element of CR10 were, therefore, eliminated and the remaining 138 amino acids towards the 5' end were searched for homology using BLASTX programme. It was interesting to find that 83 amino acids (residue no 30-116) out of 138 showed 32% homology in the 2nd ORF to *Tto*1 retrotransposon element from *Nicotiana tabaccum* (Hirochika *et al.* 1996). Total length of the ORF of *Tto*1 is 1338 amino acids, while the region showing sequence homology to CR10 was from residues 825 to 888. In the *Tto*1 element of *N. tabaccum*, the rt domain extended from 810 to 1328 amino acids which showed 38% homology to the rt of *copia* elements. Thus we observed two rt regions, namely, Ty1-*copia* like 71 amino acids and *Tto*1 like 83 amino acids, separated by 159 nucleotides from each other in two different ORFs.

The clones CA1 from *C. arietinum* and CR10 from *C. reticulatum* were compared for nucleotide sequence identity using CLUSTAL W software.

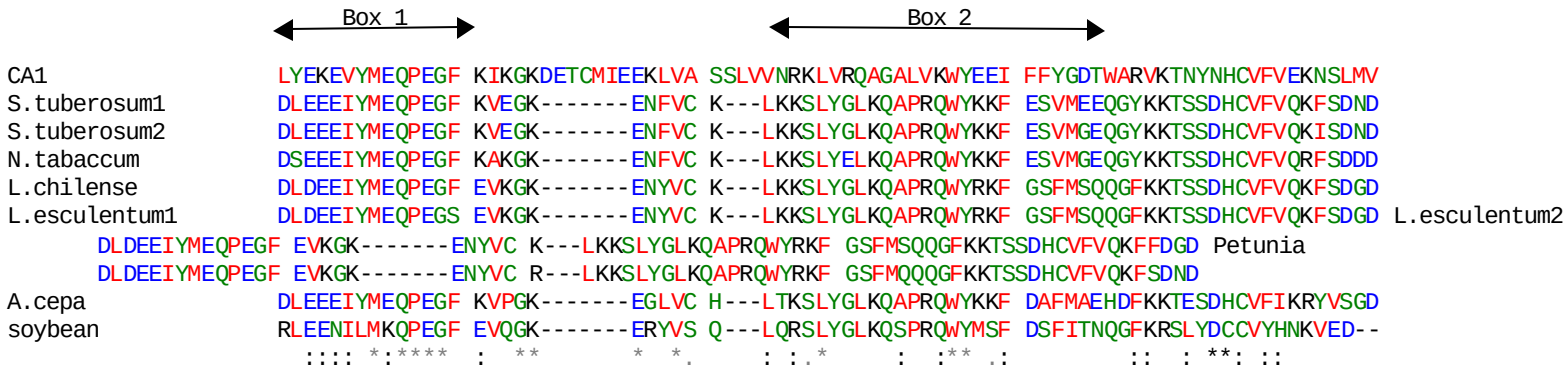


Figure 3.1a

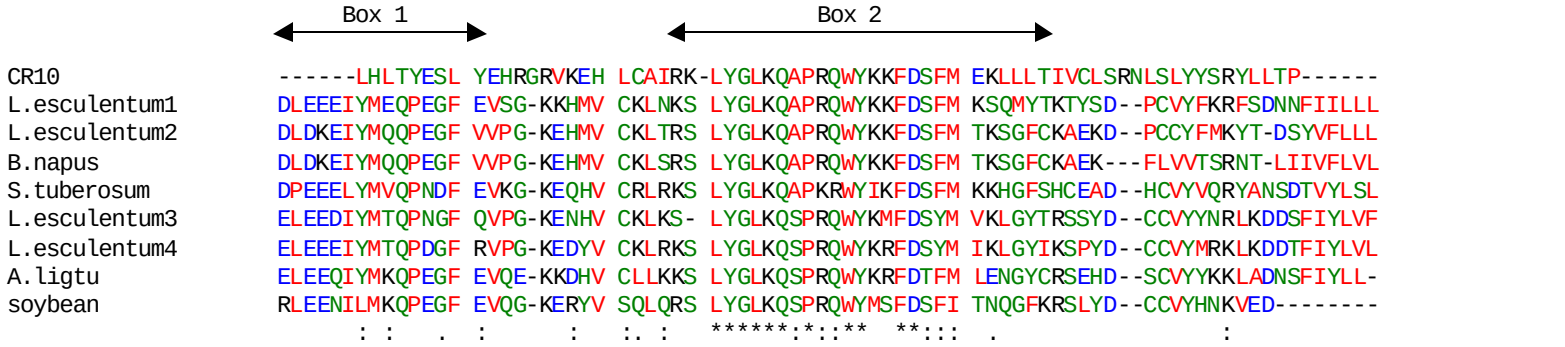


Figure 3.1b

Figure 3.1: Multiple sequence alignment of the Ty1-*copia* like rt regions obtained from *C. arietinum* (ICC-4958) (**3.1a**) and *C. reticulatum* (PI489777) (**3.1b**) using Clustal W program. EMBL accession no. of the sequences used for comparison are as follows; Fig 3.1a:AJ228810, AJ228808, D12829, AJ001211, AF072648, AJ228804, G47758, AJ223065 and soybean; Fig 3.1b: AF072655, I47758, A47759, AAA03500, AF072641, AF072638 and AJ2 59 and soybean (Voytas *et al* 1992). "*" =identical residues, ":" & "." =conserved and semi-conserved substitutions, res ely.

Both showed 41% sequence identity which explained the hybridization of CA1 with CR10 clone as mentioned earlier. However, at the predicted peptide level, the similarity between the two clones was only 11%. This indicated the sequence heterogeneity between the two Ty1-*copia* like retrotransposon elements in chickpea. Further sequence analysis of CA1 and CR10 revealed two regions as in Box 1 & 2 of Figure 3.1 (a and b), with identical residues, conserved residues with respect to hydrophobicity and semiconserved substitutions of residues with respect to charges.

3.3.3. Copy number estimation of Ty1-*copia* retrotransposon like sequences in chickpea

We were interested to know the copy number of at least CA1 element in the *C. arietinum* genomes which also showed homology to CR10 element. The copy number of CA1 was estimated as described in section 3.2.3. Densitometric scanning was used to quantitate hybridizing signals based on which the copy number was estimated per diploid genome. As seen in Figure 3.2, the intensity of hybridization signal of CA1 to $6.25X10^3$ *C. arietinum* genomes was equivalent to control CA1 which contained $2X10^6$ copies while $1.25X10^5$ *C. reticulatum* genomes matched the intensity of control DNA containing $1X10^6$ copies. Based on this, the total number of elements of CA1 estimated was approximately 600 copies and 10 copies per diploid genomes of *C. arietinum* and *C. reticulatum*, respectively. The difference in the hybridization signals of CA1 with that of *C. arietinum* and *C. reticulatum* genome could be due to less copies of exact homologues of CA1 in the *C. reticulatum* genome. Furthermore, this could also be due to sequence mismatch since CA1 was cloned from *C. arietinum* and there might not be close homologues of CA1 in *C. reticulatum*. A possibility of large number of weakly homologous sequences in *C. reticulatum* genome could also be suggested. Six hundred copies of the ~280bp rt sequences in the genome of *C. arietinum* correspond to 1.68×10^5 bp, which is equivalent to about 0.11% of that genome. If all these fragments are a part of complete Ty1-*copia* group retrotransposons (which are at least 5 kb in length), then they would comprise 0.2% of the genome of *C. arietinum*.

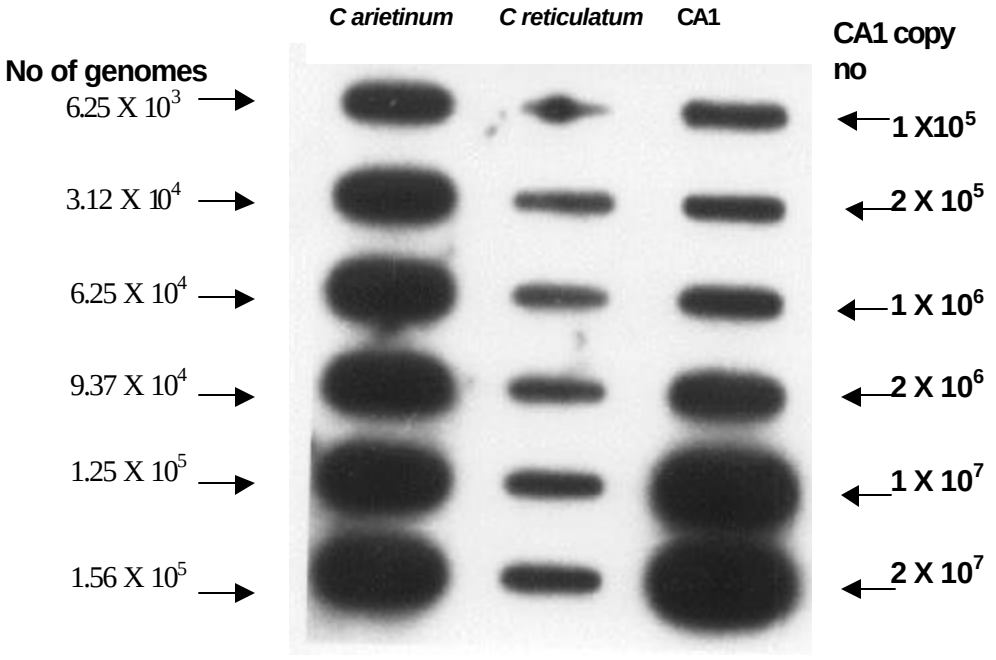


Figure 3.2 Hybridization of CA1 with genomic DNA of *C. arietinum* (Vijay) and *C. reticulatum* (PI489777) corresponding to $6.25x10^3$, $3.12x10^4$, $6.25x10^4$, $9.37x10^4$, $1.25x10^5$ and $1.56x10^5$ genomes and along with clone CA1 corresponding to $1x\ 10^5$, $2x10^5$, $1x10^6$, $2x10^6$, $1x10^7$, and $2x10^7$ molecules.

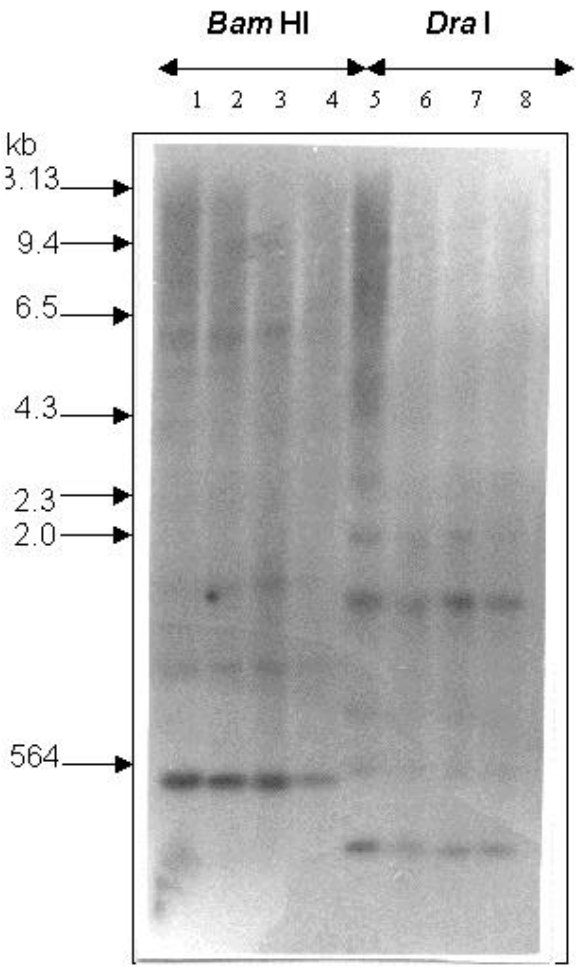
3.3.4 Ty1-*copia* retrotransposon element detects genetic diversity in *Cicer* species

The genomic distribution and variability of the Ty1-*copia* retrotransposon like sequences were studied by DNA blot hybridization using four accessions of *C. arietinum* namely PG-89219, ICC-4958, PG-87207 and Bheema. As seen in Figure 3.3a, hybridization of CA1 to *Bam*HI digested genomic DNAs revealed 3 bands, whereas *Dra*I digested DNAs showed 5 bands in addition to few faint bands in both the cases, with background smear indicating the presence of many copies of this element in *C. arietinum* genome. Bheema cultivar digested with *Dra*I, however, showed absence of two bands with sizes 0.7 kb and 2 kb.

Hybridization of CR10 to *Bam*HI and *Dra*I digested genomic DNAs showed few intense bands wherein genomic DNAs digested with *Bam*HI revealed two bands, and *Dra*I digested DNAs yielded three bands (Figure 3.3b). However, no variation between the four accessions was detected in the profile using this probe. The number of bands detected using CR10 was less than CA1. In a study on genomic organization of the genes encoding floral binding protein (fbp1) in *Petunia*, the probe containing the MADs box having high copy number showed a smear, whereas fbp1 specific probe showed a single band on hybridization (Angenent *et al.* 1992). This might support my observation that CR10 was represented less in the genome as compared to CA1.

In present analysis, the 0.55 kb and 0.9 kb bands in *Bam*HI digested DNAs, and 1 kb band in *Dra*I digested DNA were detected when probed with both CA1 and CR10. Hybridization of both these clones indicated few bands that were common which might be due to the sequence similarity of CA1 with CR10.

I further used the degenerate primers of rt region of Ty1-*copia* for diversity analysis as detailed in section 3.2.4. As seen in figure 3.4, total number of fragments generated were 21 ranging from 0.650 kb to 0.180 kb. More number of bands were detected on PAGE as compared to agarose gel. This is due to high resolving power of a polyacrylamide gel in a specific size range and increased sensitivity of detection due to radiolabel incorporation. This also indicates probability of occurrence of many more Ty1-*copia* retrotransposon like sequences in the *Cicer* genome. Five fragments were common among all the *Cicer* species, while fragment of 0.240kb was specific to *C. arietinum* cultivars which could be regarded as a species specific band. Similarity index values were calculated using Dice coefficient (Sokal and Sneath 1963) on the basis of which a dendrogram (figure 3.5) was constructed using the software TAXAN based on UPGMA algorithm (Sneath and Sokal 1973).



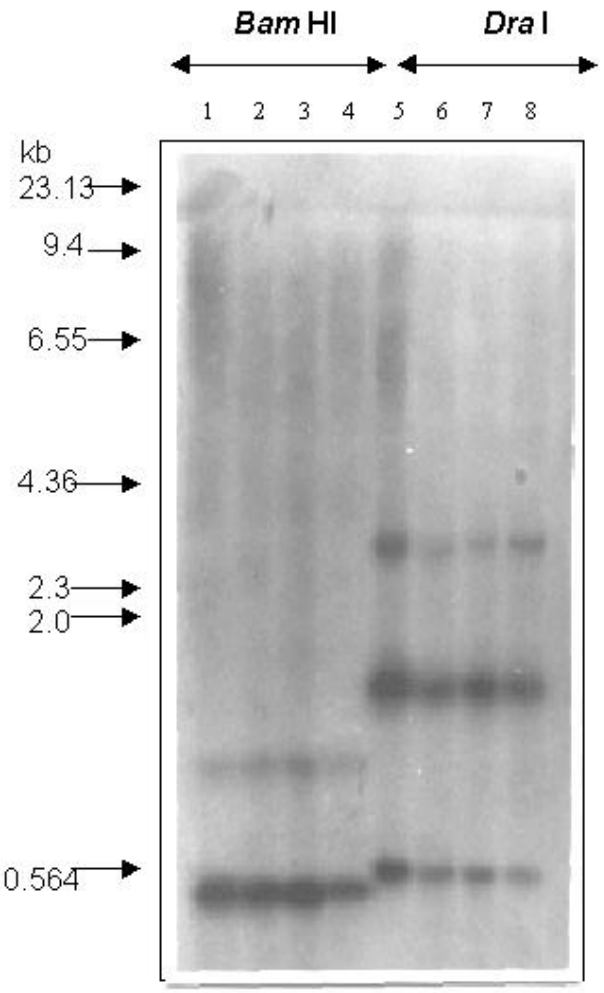


Figure 3.3b

igure 3.3 **Genomic digest with *Bam*HI and *Dra*I hybridized with CA1 (3.3a) and CR10 (3.3b) inserts. o.1&5=PG-89219; 2&6=ICC-4958; 3&7=PG-87207; 4&8=Bheema. Lambda *Hind*III digest was used as a mol *v*eight marker.**

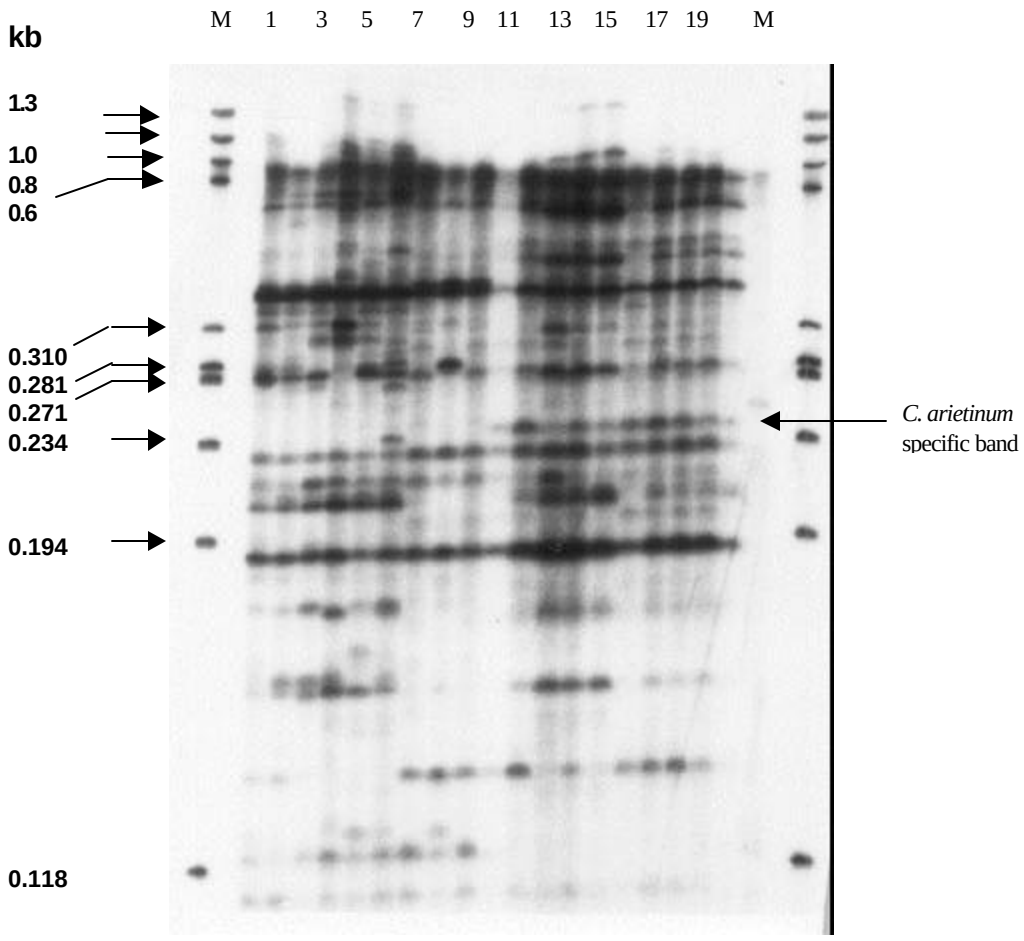


Figure 3.4: Amplification of DNA from *Cicer* species using the degenerate primers for rt region of Ty1-*copia* element with the incorporation of $\propto^{32}$PdATP, resolved on 6% denaturing polyacrylamide gel. M= ϕX174-*Hae*III digest. **1**. *C. acanthophyllum*, **2**. *C. anatolicum*, **3**. *C. nuristanicum*, **4**. *C. microphyllum,* **5**. *C. macracanthum,* **6**. *C. bijugum,* **7**. *C. yamashitae,* **8.** *C. echinospermum,* **9**. *C. reticulatum,* (**10 to 19)** *C. arietinum*, **10.** V-65-R, **11**. ABCP12, **12**. ABCP13, **13**. ABCP14, **14**. USSR, **15**. Vijay, **16**. ICC-4958, **17**. JG-62, **18**. C-104, **19**. WR-315.

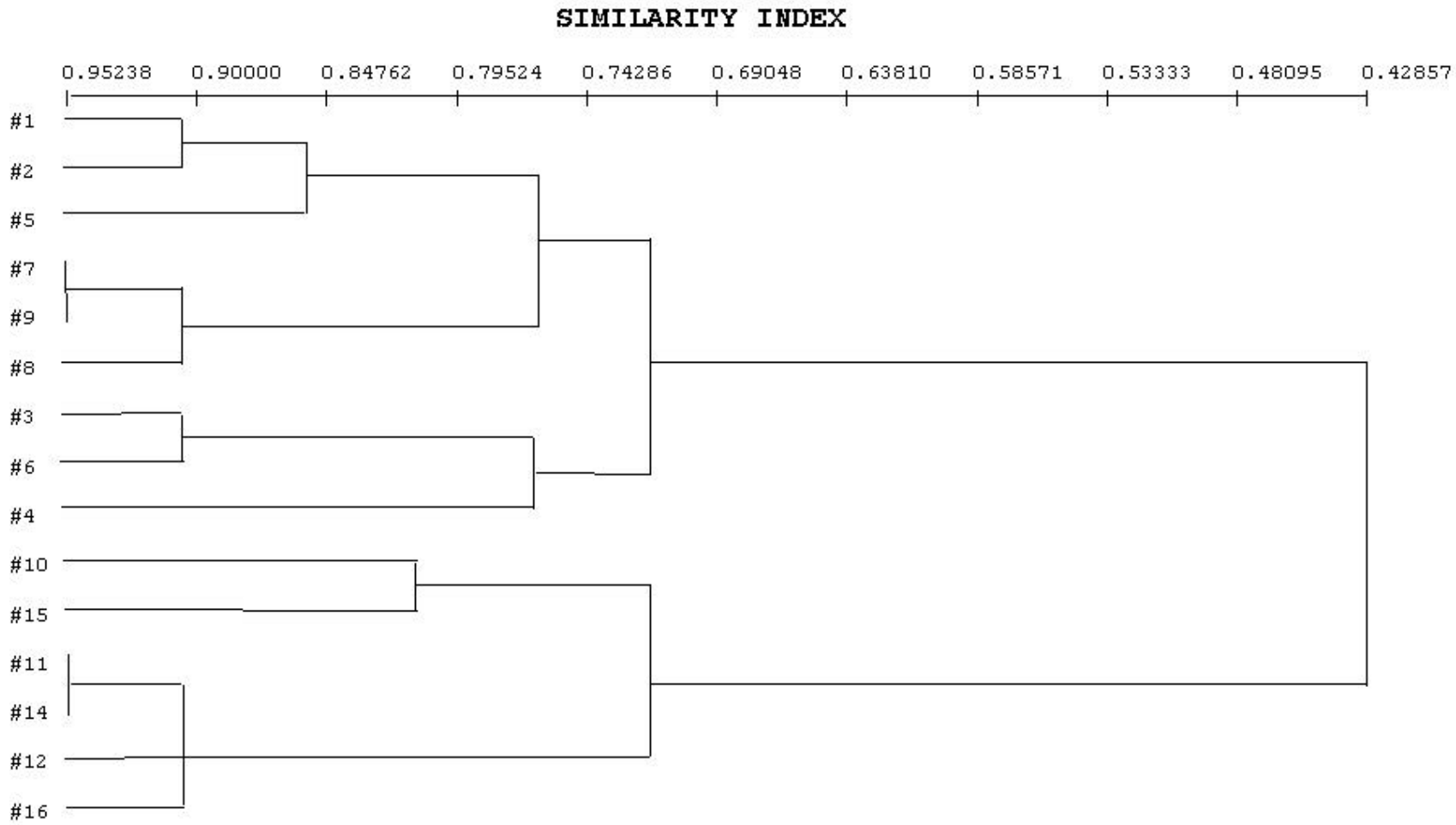


Figure 3.5 Dendrogram depicting the similarity indices between *Cicer* species. The numbers denote various *Cicer* species as described in figure 3.4.

The similarity index values among the wild *Cicer* species ranged from 0.95 to 0.71, whereas among the *C. arietinum* accessions it ranged from 1 to 0.71. Dendrogram in figure 3.5, shows two major clusters, cluster I consisted of all the wild *Cicer* species while cluster II consisted of all the *C. arietinum* accessions. The average similarity index value was 0.85 among the *C. arietinum* accessions and 0.83 among the wild *Cicer* species. The two major clusters showed a similarity index value of 0.42, indicating the divergence of *C. arietinum* from all the wild species of *Cicer.* Cluster II contained two subgroups in which four accessions namely, ABCP14, JG-62, C-104, and WR-315 were not represented as they had a similarity index value of 1.0. This indicated that they were identical to each other at least with respect to the distribution of this element detected on the basis of the degenerate primers.

3.4 Discussion

Reverse transcriptase is the most conserved protein encoded by retroviruses and retrotransposons with five distinct domains in the enzyme (Doolittle *et al.* 1989). The connecting residues of these five domains, however, show significant sequence heterogeneity (Xiong and Eichbush 1990). Most of the Ty1-*copia* elements isolated from plants exhibiting the divergence in this rt region have been identified by a DOP-PCR approach, using primers complementary to the conserved domains (Flavell *et al.* 1992). This technique allows the specific identification of elements of the Ty1-*copia* group, excluding other families of retrotransposons (Voytas *et al.* 1992). I have used the same strategy to detect the presence of Ty1-*copia* group of retroelements in the *Cicer* species.

3.4.1 Reverse transcriptase sequence variation of Ty1-*copia* group retrotransposon

The PCR fragments of the rt gene of Ty1-*copia* elements in my analysis have shown homology to the rt sequences in the gene bank as

mentioned in section 3.3.2. In earlier studies (Xiong and Eichbush 1990) by aligning 82 rt containing elements including: retroviruses, LTR-and non LTR retrotransposons, and group II mitochondrial introns, five distinct domains have been identified and found to be conserved in over 50% of the sequences. The 3rd and 5th domain consisting of the peptide sequence TAFLHG and YVDDML respectively, which are not shown in figure 3.1a and b, have been used by me to amplify the rt region of Ty1-*copia* retrotransposon like elements in chickpea.

In the present sequence alignment studies, Box 1 and 2 of CA1 and CR10 sequences in chickpea (Figure 3.1a and b) denote the 3rd and 4th common domains in the related rt sequences while the region flanking these domains shows sequence heterogeneity. Such sequence heterogeneity has been previously reported in many higher plant genomes. For example, when 31 fragments of the rt genes of the Ty1-*copia* group retrotransposons from potato were sequenced, each one was found to be different with predicted amino acid similarities between individual fragments varying from 5% to 75% (Flavell *et al.* 1992). These formed 13 different subgroups on the basis of sequence heterogeneity of Ty1-*copia* elements. In *A. thaliana*, four randomly selected clones were sequenced and databank searches of the sequences showed homology ranging from 31% to 64% to the rt gene of Ty1-*copia*-like elements (Brandes *et al.* 1997). The relatively low replication accuracy of rt provides an indication that internal variation is programmed into the system, however, many host "defense" activities (e.g cytosine methylation) may also increase variability. The host genome needs to select such altered elements and to maintain them within the genome. Natural selection acting on the host may utilize this variability whenever possible (Kumar and Bennetzen 1999).

The Ty1-*copia* like elements CA1 and CR10 from *Cicer* species in this study have shown presence of stop codons which indicates that they are defective. Such defective retrotransposons have been reported in several plant genomes (Kumar and Bennetzen 1999). Unlike *Drosophila* and yeast, stop codons, deletions or frameshifts have commonly been observed in Ty1-*copia* like retrotransposons from higher plants (Flavell *et al.* 1992). There could be existence of a significant population of active retrotransposons superimposed upon a background of multiple copies of defective elements

which can retrotranspose by the use of trans-acting factors (Flavell *et al.* 1992). These internal variations within retrotransposons can provide raw material for either new transposon properties or new capabilities acquired by the host genome.

Another interesting observation in my work is that the CR10 element from *C. reticulatum* has two rt regions, one Ty1-*copia* like at the 3' end and the other element at 5' end showing homology to the *copia* like rt region from *Nicotiana tabaccum* (*Tto*1). The *Tto*1 like element present at the 5' end may be an insertion into the Ty1-*copia* like element. Such insertions of one element into the other is one of the mechanisms adopted by the host in order to control their copy number in the host genome or to get eliminated from them (Kumar *et al.* 1997). In maize, over 50% of the nuclear DNA is occupied by densely packed retrotransposons (and several unidentified retrotransposons) belonging to Ty1-*copia*, Ty3-*gypsy* and LINEs group of retrotransposons with variable copy number. Insertion of these elements into each other as well as into LTRs has been earlier reported (San Miguel *et al.* 1996). Other larger genomes like barley and lily also have both highly repetitive retrotransposons and many different families of retrotransposons with variable number of copies (Bennetzen *et al.* 1996).

3.4.2 Genomic distribution of Ty1-*copia* retrotransposon like elements in chickpea

Retrotransposon is a class of dispersed middle repetitive sequences which have contributed to the genetic diversity of their host species. Three mechanisms have been reported to be involved generating variability due to retrotransposons among genomes: transposition, homologous recombination between retrotransposons or LTRs and frequent mutation in the form of point mutation, insertion, deletion or methylation of cytosine to thymine in retrotransposon sequences (Asins *et al.* 1999). Due to this feature, retrotransposons have been used as efficient DNA fingerprinting probes in some plant species (Fukuchi *et al.* 1993). The diversity generated by these elements has been studied in maize, pea and barley (Ellis *et al.* 1998, Kalendar *et al.* 1999, Purugganan and Wesseler 1995) and these elements have further been exploited as DNA markers to generate genetic linkage

maps in barley and pea (Ellis *et al.* 1998, Kumar *et al.* 1997, Wang *et al.* 1999). So far, there are no reports of use of retrotransposons as RFLP probes in chickpea. In our studies, we have found genetic variability among *C. arietinum* accessions in *Dra*I digests of genomic DNAs using rt clones of Ty1-*copia* like retrotransposon as probes (Figure 3.3a). The Ty1-*copia* rt sequence from *C. arietinum* (CA1) showed a high copy number in the genome. Copy number of Ty1-*copia* element in *A. thaliana* (Ta elements) and potato (*Tst*1) ranges from 1-10 only. In tobacco, *Tnt*1 shows a copy number of 100, whereas barley genome contains 70,000 BARE I copies. Total Ty1-*copia* elements which consist of heterogenous rt sequences are present in copy number 1,000,000 in *V. faba*, 50,000 in *V. sativa* and 1000 in *V. melanops* (Flavell *et al.* 1997). Studies have shown that genome size variation is correlated with both, the total mass of retrotransposons that are present and the number of different retrotransposon families. In *Arabidopsis*, both LTR and non-LTR retrotransposons account for 4 to 10% of the total nuclear DNA, whereas in maize there are thousands of different families of LTR and non-LTR retrotransposons (Kumar and Bennetzen 1999). Copy number variation for individual Ty1-*copia* retrotransposon subgroups has been observed in *Secale cereale*, where 3 individual rt clones, R9, R24 and R25, representing 3 different retrotransposon subgroups are present in a variable copy number ranging from 300 to 50,000 (Pearce *et al.* 1996). In the present study, we report the copy number of Ty1-*copia* like element, CA1 from *C. arietinum*. However, Ty3-*gypsy* element also exists in *C. arietinum* genome which indicates the possibility of presence of more than one family of retrotransposon elements in *Cicer* genome (Staginnus et al, 1999).

Retrotransposons have been reported to be capable of generating large populations in a relatively short evolutionary time as they have a replicative mode of transposition. Plant genomes can accumulate large amounts of DNA and do not appear to remove these sequences rapidly. It is possible that during evolution retrotransposons have proliferated into larger populations from a few active elements in the host genome (Kumar and Bennetzen 1999). The *C. arietinum* genome as compared to *C. reticulatum* has been shown to possess high copy number of CA1 Ty1-*copia* like element in this study. I can therefore, postulate that the high copy number in *C.*

arietinum could be due to its evolution from its progenitor *C. reticulatum* during which the retrotransposon element could increase its copy number.

To study whether this element is an active retrotransposon, CA1 was hybridized to Northern dot blots of RNA obtained from *C. arietinum* leaf tissue at 15 days after sowing, along with genomic dots. There was no hybridization signal on the RNA dots, except the control DNA, indicating that the element is transcriptionally inactive (data not shown). Most Ty1-*copia* group retrotransposons are transcriptionally inactive and most of them are thought to be fixed in their genomic locations. Active Ty1-*copia* elements have been found in tobacco (*Tnt*1), barley, wheat and rye (BARE 1), the latter showing great deal of insertional polymorphism (Flavell *et al.* 1997). *In-situ* hybridization studies have indicated that *Tnt*1 and BARE1 element are located in the euchromatic region, while in *A. thaliana* and *C. arietinum*, the Ty1-*copia* elements are clustered in the genetically inactive centromeric heterochromatin region (Brandes *et al.* 1997). This supports our observation that Ty1-*copia* is probably inactive in *C. arietinum*. However it has also been reported that the transcriptional activity of such elements could be induced under several biotic and abiotic stresses (Grandbastien *et al.* 1998).

3.4.3 Genetic divergence among the *Cicer* species based on the Ty1-*copia* element

Retrotransposons are abundant in the plant genome due to proliferation by reverse transcription and the resulting progeny is 'seeded' to new genomic sites creating insertional polymorphism (Voytas and Naylor 1998). This could be the reason of obtaining many fragments which have been detected using the sensitive technique described in section 3.2.4. Likewise, more number of fragments has also been detected in a study where PCR was performed with ISSR primers with radioactive label incorporation (Godwin *et al.* 1997). In my previous studies on genetic diversity of elite chickpea cultivars using RAPD and hybridization based microsatellite markers, the similarity indices ranged from 0.78 to 0.98 and 0.24 to 0.63, respectively (2.3.1.2). In this study, the similarity index (0.71 to 1.0) among the *C. arietinum* cultivars based on the retrotransposon element is moderate. We have also shown a *C. arietinum* specific band of molecular weight 0.240kb

which is absent in all the other *Cicer* species (Figure 3.4). Retrotransposons have been used to estimate phylogenetic relationships in cereal plants because each element has an unique transpositional history. For example, the evolutionary relationships between wheat, rice and maize have been studied on the basis of Ty1-*copia* rt domain. The study has shown the presence of elements obtained from distantly related species, indicative of origins prior to the radiation of grass species (Matsuoka *et al.* 1999). Similarly detail phylogenetic studies based on the variation in the retrotransposon sequence may reveal the evolutionary relationships among the legumes.

In summary, this is the first report showing the presence of Ty1-*copia* like elements in *Cicer*. There might be a possibility of presence of many subgroups of Ty1-*copia* elements in *Cicer* genome which can be differentiated on the basis of sequence heterogeneity as revealed by CA1 and CR10 elements in this study. The genetic diversity data show that these elements can be exploited as molecular markers by studying the distribution of LTR sequences. Using these sequences, retrotransposon based SSAP technique can be developed in *Cicer* species. Phylogenetic relationships between the *Cicer* species can also be studied on the basis of sequence heterogeneity in these elements.

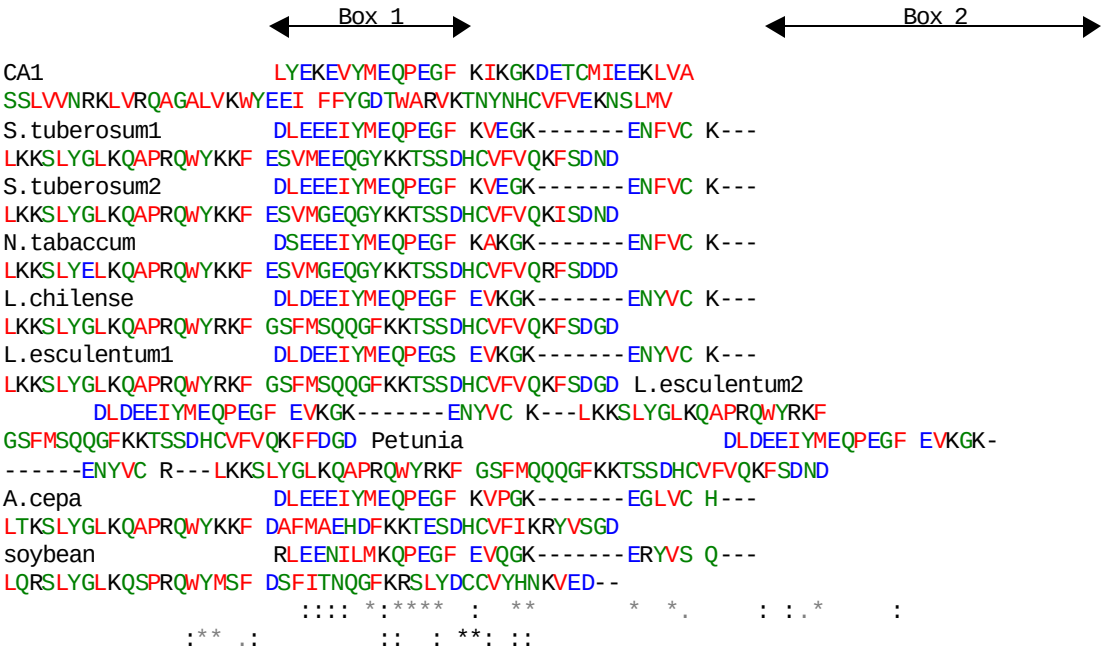


Figure 3.1a

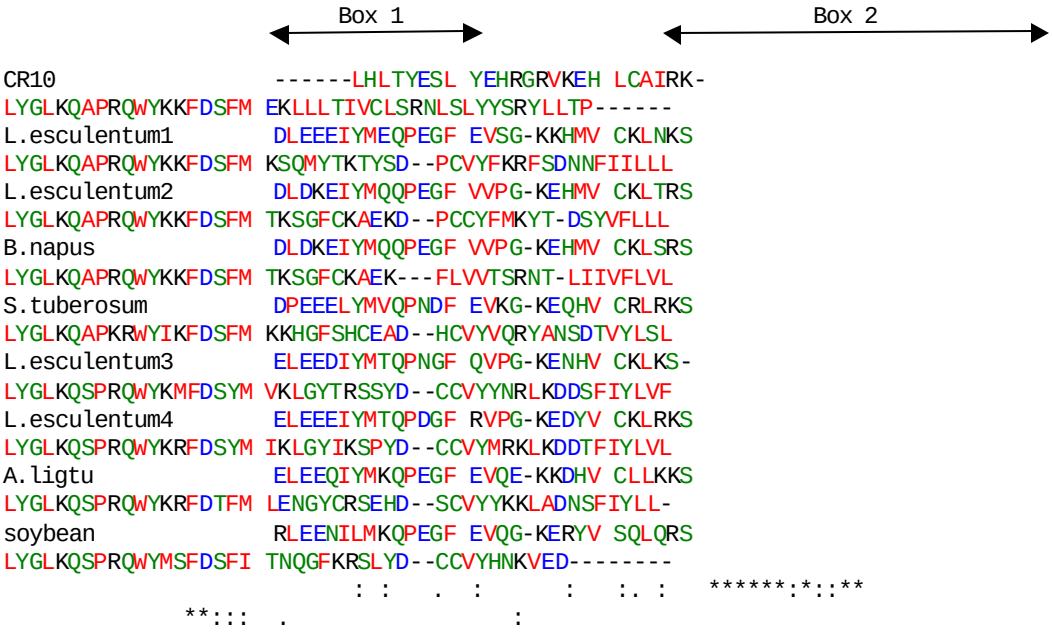


Figure 3.1b

Figure 3.1: Multiple sequence alignment of the Ty1-*copia* like rt regions obtained from *C. arietinum* (ICC-4958) (**3.1a**) and *C. reticulatum* (PI489777) (**3.1b**) using Clustal W program. EMBL accession no. of the sequences used for comparison are as follows; Fig 3.1a:AJ228810, AJ228808, D12829, AJ001211, AF072648, AJ228804, G47758, AJ223065 and soybean; Fig 3.1b: AF072655, I47758, A47759, AAA03500, AF072641, AF072638 and

AJ223557.and soybean (Voytas *et a*l 1992). "*" =identical residues, ":" & "." =conserved and semi-conserved substitutions, respectively.

CHAPTER 4

Molecular Mapping of the RAPD and ISSR Markers on the Integrated Map of Chickpea Genome and its Exploitation for Fusarium Wilt Resistance Tagging in Chickpea

Part of the contents of this chapter have been published as a full-length paper in: Theoretical and Applied Genetics (2000) 101: 1155-1163.

4.1 Introduction

4.2 Material and Methods

4.2.1 Plant material

4.2.2 Chickpea DNA extraction

4.2.3 Primers and PCR amplification

4.2.4 Evaluation for resistance of RILs in the field for *Fusarium oxysporum* f.sp. *ciceri* race 1

4.2.5 Agarose and polyacrylamide gel electrophoresis

4.2.6 Statistical analysis

4.2.7 Linkage analysis

4.3 Results

4.3.1 Linkage mapping of the *C. arietinum* genome using interspecific cross *C. arietinum* X *C. reticulatum*

4.3.1.1 Parental screening for RAPD and ISSR polymorphism

4.3.1.2 Segregation, linkage and distribution of markers

4.3.2 Identification of a putative marker linked to Fusarium wilt resistance race 1 in chickpea

4.3.2.1 Primer selection for detecting parental polymorphism

4.3.2.2 Analysis of phenotypic data

4.3.2.3 Selective genotype analysis

4.3.2.4 Identification of a putative STMS marker linked to Fusarium wilt resistance race 1

4.4 Discussion

4.4.1 Segregation distortion in RILs of the interspecific cross

4.4.2 Status of RAPD markers in linkage mapping

4.4.3 Differential wilting in chickpea genotypes

4.4.4 Genes for resistance to Fusarium wilt are clustered in chickpea

4.1 Introduction

Comprehensive genetic linkage maps have proved to be a powerful tool in genetic studies which permit the localization of genes controlling both simple and complex traits. This in turn is the basis for marker-assisted breeding where flanking markers can be used to introgress specific genes into a desired genetic background. The mapped markers not only make it possible to select for the gene(s), but also enable to select against linkage drag. Further, comparative genome mapping addresses questions related to the mode and rate of chromosomal evolution (Hallden *et al.* 1996). Use of RFLP markers to construct human genetic maps was first suggested by Botstein *et al.* (1980). The development of PCR based markers such as RAPD, ISSR, AFLP and STMS has increased the scope of markers suitable for genetic mapping.

In any gene mapping program, the most important factors are the cultivars used as parents for generating the population, the population type, its size and the generation used for genotypic and phenotypic analysis. The F_2 and backcross populations are the most commonly used populations for genome mapping, as they are simple and fast to develop. Populations such as recombinant inbreds (RI), double haploid (DH) and near isogenic lines (NILs) are immortal lines and can be shared by different laboratories for generation of high density map. In plants, RILs were first used by Jinks *et al* (1981) to estimate the components of variance and subsequently used for mapping storage proteins and fertility restoration genes relative to the nucleolar organizer in wheat and legumin genes in peas (Burr and Burr 1991).

A direct application of genetic linkage maps has been in tagging genes of economic and agronomic importance. There are a vast number of examples of gene tagging in plants. Many biotic resistance genes in various crop plants such as bacterial blight and blast resistance in rice (Yoshimura *et al*. 1995, Blair and McCouch 1997, Hittalmani *et al.* 1995), Fusarium wilt in tomato (Sarfatti *et al*. 1991) and cyst nematode resistance in soybean (Concibido *et al.*1997) have been tagged using molecular markers. Recently, efforts for pyramiding of resistance genes in a single genetic background are ongoing. Several agronomically important traits such as yield, maturity and resistance to abiotic stresses are controlled by relatively large number of loci, each of which makes a small positive or negative contribution to the final phenotypic value of the trait and such loci are termed as “quantitative trait loci” (QTL).

The first extensive integrated *C. arietinum* genome map has been constructed in an international effort among four laboratories placing different types of DNA markers, isozyme markers and genes conferring resistance to Fusarium wilt caused by races 4 and 5 on the linkage maps. The study described in this chapter includes my contribution of mapping RAPD and ISSR markers on the integrated *C. arietinum* genome map. Further, few of these mapped markers along with additional random markers have been used to identify a putative STMS marker linked to the locus conferring resistance to *F. oxysporum* f.sp. *ciceri* race 1 using an intraspecific cross.

4.2 Materials and methods

4.2.1 Plant material

An interspecific cross was made between *C. arietinum* (ICC4958) resistant to Fusarium wilt race 4 and 5 and *C. reticulatum* (PI489777), to generate 118 RILs by single seed descent (SSD) from the F_2 population by Dr. Fred Muehlbauer at the USDA-ARS, Pullman, USA. These RILs were shared between the laboratories involved in this work.

Similarly an intraspecific cross between JG62 (susceptible) and Vijay (resistant) was made in the rabi season of 1995-96 at Mahatma Phule Agricultural University, Rahuri, India and RILs of $F_{4:5}$ generation were developed by SSD method, for the purpose of identifying markers linked to gene conferring resistance to wilt caused by *F. oxysporum* f.sp *ciceri* race 1. Additional features of JG62 are double poddedness, early maturity and small seeds, while Vijay is, drought tolerant, high yielding, having wider adaptability, and high pod number.

4.2.2 Chickpea DNA extraction

DNAs were isolated from the leaf tissue of the parents and RILs using the miniprep method of Doyle and Doyle (1987). One gram of each leaf sample from individual line was submerged in liquid nitrogen and then ground to a fine powder. The powder was quickly transferred to a tube containing 7.5 ml of ice cold extraction buffer (0.35 M Sorbitol, 0.1 M Tris, 5 mM EDTA, pH 7.5). The tube was briefly shaken and 7.5 ml of nuclei lysis buffer (2 M NaCl, 0.2 M Tris, 50 mM EDTA, 2% CTAB, pH 7.5) was quickly added, followed by 3 ml of 5% Sarkosyl solution. Sample sets were incubated in a 65°C water bath for 20 min.

After incubation the tubes were cooled for few minutes and 18 ml of $CHCl_3$/IAA (24:1) was added to each tube. The tubes were then centrifuged at 5000rpm for 15mins. The aqueous layer was removed and extracted again with 15 ml $CHCl_3$/IAA mixture. Finally, DNA was precipitated with two volumes of chilled ethanol and suspended in 1 ml of TE buffer.

4.2.3 Primers and PCR amplification

Various primers used in the present analysis amplified RAPD, ISSR, STMS, RGA and ASAP markers in the respective inter as well as intraspecific populations. The template and primer concentrations and the cycling conditions used are mentioned in Table 4.1.

In all the PCRs mentioned below the reaction mixture contained 0.8 U *Taq* DNA polymerase (Perkin Elmer Int., USA) except, when RGA primers were used the *Taq* DNA polymerase concentration was 1.0 U. The reaction was carried out in 100 mM Tris pH 8.3, 1.5 mM $MgCl_2$, 500 mM KCl, 0.01% (w/v) gelatin, 0.2 mM dNTPs and 0.5 mM spermidine. PCR was performed in Perkin Elmer 9700 thermocycler with heated lid that eliminated addition of mineral oil.

The primer sequences used in this work for generating RGA, STMS and ASAP markers are mentioned in Table 4.2. The RGA primers used for heterologous PCR were deduced from the peptides conserved between resistance genes RPS2 of *Arabidopsis thaliana* and N of *Nicotiana tabaccum* (Leister *et al.* 1996).

Table 4.1 Details of the PCR amplifications used for generating molecular markers

Marker amplified	Template conc.	Primer conc.	Denaturation	Cycling conditions	Elongation
RAPD (Operon Technologies, Inc., USA)	2 ng	5 pm/20 µl	94°C-4 min	5 cycles of:92°C-30 s, 36°C-2 min, 72°C-90 s, followed by 35 cycles of 92°C-5 s, 40°C-20 s, 72°C-90 s	72°C-5 min
ISSR (University of British Columbia, Canada)	20 ng	35 pm/20 µl	94°C-5 min	40 cycles of 94°C-1 min, 50°C-45 s, 72°C-2 min	72°C-5 min
STMS (Winter *et al.* 1999)	50 ng	2 µM/20 µl	94°C-5 min	35 cycles of 94°C-2 min, 55°C-50 s, 60°C-50 s	No elongation step
RGA (Leister *et al.* 1996)	50 ng	0.25µM/50µl	93°C-2 min	35 cycles of 93°C-45 s, 51°/45°/49°C-45 s, 72°C-80 s	72°C-10 min
ASAP (Mayer *et al.* 1997)	40 ng	2 µM/20 µl	94°C-5 min	40 cycles of 94°C-20 s, 62°C-1 min, 3 min ramp to 72°C, 72°C-1 min	72°C-8 min

Table 4.2 Primer sequences for amplifying RGA, STMS and ASAP markers

Primer	Sequence
RGA[1]	
Peptide sequence N-terminal **G G V G K T T**	
s1	5- GGT GGG GTT GGG AAG ACA ACG-3
s2	5- GGI GGI GTI GGI AAI ACI AC -3
Peptide sequence	C-terminal **G L P L A L**
As1[#]	5 CAA CGC TAG TGG CAA TCC 3
As2[#]	5 IAA IGC IAG IGG IAA ICC 3
As3[#]	5 IAG IGC IAG IGG IAG ICC 3
STMS	
TA110(F[2])	5TTCTTTATAAATATCAGACCGGAAAGA3
(R[2])	5ACACTATAGGTATAGGCATTTAGGCAA3
TA37(F)	5ACTTACATGAATTATCTTTCTTGGTCC3
(R)	5CGTATTCAAATAATCTTTCATCAGTCA3
TA27(F)	5GATAAAATCATTATTGGGTGTCCTTT3
(R)	5TTCAAATAATCTTTCATCAGTCAAATG3
TA34(F)	5AAGAGTTGTTCCCTTTCTTTT3
(R)	5CCATTATCATTCTTGTTTTCAA3
TA42(F)	5ATATCGAAATAAATAACAACAGGATGG3
(R)	5TAGTTGATACTTGGATGATAACCAAAA3
TA53(F)	5GGAGAAAATGGTAGTTTAAAGAGTACTAA3
(R)	5AAAAATATGAAGACTAACTTTGCATTTA3
TA106(F)	5CGGATGGACTCAACTTTATC3
(R)	5TGTCTGCATGTTGATCTGTT3
TA113(F)	5TCTGCAAAAACTATTACGTTAATACCA3
(F)	5TTGTGTGTAATGGATTGAGTATCTCTT3
TA114(F)	5TCCATN*TAGAGTAGGATNTTNTTGGA3
(R)	5TGATACATGAGTTATTCAAGACCCTAA3
ASAP	
Cs-27(F)	5AGCTGGTCGCGGGTCAGAGGAAGA3
Cs-27(R)	5AGTGGTCGCGATGGGGCCATGGTG3
UBC-170(F)	5ATCTCTCCTGTGTGTGTG3
UBC-170(R)	5ATCTCTCCTGCATCACAAG3

#antisense primers are written in opposite orientation to the peptide sequence. The first triplet in each of these primers corresponds to the last amino acid in the peptide.

1 The PCR amplification was performed using the primer combination as given by Leister *et al.* (1996) which is as follows: (a): primers s1 plus as1; (b) s2 plus as2 (c) s2 plus as3. Annealing temperatures were 51°C for (a), 45°C for (b) and 49°C for (c).

2=F means forward primer and **R** means reverse primer, *=aNy base

4.2.4 Evaluation for resistance of RILs in the field for *Fusarium oxysporum* f.sp.*ciceri* race 1

One hundred and sixty-six RILs were evaluated for Fusarium wilt resistance in a uniformly wilt sick plot at MPKV, Rahuri which has been used and maintained for screening chickpea genotypes for wilt caused by race 1 of *Fusarium oxysporum* f.sp. *ciceri* (Pawar *et al.* 1992) and was developed according to the method described by Nene *et al.* (1981). Resistance to Fusarium wilt race 1 was checked in rabi season of 1998-99 using the F_4 and in 1999-2000 using the F_5 generation RILs. In 1998-99 ten individual seeds of each F_4 generation RILs, and in 1999-2000 five individual seeds of each F_5 generation RILs, were sown with a spacing of 30X10 cm^2 in the wilt sick plot. After every five rows of test entries one row of susceptible check, JG62, was planted in order to assess the uniformity of inoculum in wilt sick plot. Twenty plants of the genotype Vijay (resistant), ICC4958 (late wilter) and JG62 (susceptible) were sown as checks in the wilt sick plot. Each RIL was scored for wilting symptoms using a scale of 1 to 9 as follows: **1**= >91 Days After Sowing (DAS) completely resistant; **2**= wilted between 81-90 DAS; **3**=71-80 DAS; **4**=56-70 DAS; **5**=41-55 DAS; **6**=31-40 DAS; **7**=21-30 DAS; **8**=11-20 DAS; **9**=0-10 DAS. The first scoring was made when the susceptible check showed symptoms of the disease between 17 to 20 DAS, since no plants wilted in the range (0 to 10 DAS). Throughout the disease scoring period, the individual plants of each row of the RIL were noted independently.

Selective genotyping was used to identify markers in the near vicinity of the resistant locus. Seven resistant and eight susceptible RILs were noted from the 1998-99 and 1999-2000 wilt scoring data and were used for DNA extraction individually. These samples from resistant and susceptible RILs were further analyzed using polymorphic primers identified from the parental survey.

4.2.5 Agarose and polyacrylamide gel electrophoresis

RAPD and ISSR amplified PCR products were resolved on 2% agarose gel in 0.5X TAE electrophoresis buffer and viewed under UV transilluminator. The amplified products obtained using STMS primers were radioactively labelled by incorporating α-^{32}PdATP in each PCR reaction. The samples were denatured and loaded on a 6% denaturing polyacrylamide gel and amplified bands were viewed by autoradiography.

4.2.6 Statistical analysis

Segregation of loci was tested for goodness of fit to the expected Mendelian segregation ratio using X^2 test and the significance was tested at P=0.05. X^2 value was calculated as follows:

$$X^2 = \sum_{i=1}^{p} (n_i - E_i)^2 / E_i$$

where p = no. of classes

n_i = observed number of units falling in class i

E_i = number of units expected to fall into class i, assuming that the hypothesized ratio holds.

4.2.7 Linkage analysis

For linkage analysis Mapmaker version 3.0b was used (Lander *et al*. 1987). Loci were first divided into linkage groups at a LOD score of 4.0 by two point analysis using the 'group' command. Marker order in linkage groups was determined using the 'try' command of the program and the order generated by data of other laboratories was considered as the starting order. Marker order obtained was scrutinized by multipoint analysis applying the 'ripple' function. Final map distances were calculated by applying the 'Kosambi' function (Kosambi 1944) provided by the program.

Computer software program for QTL analysis QGENE (Nelson 1997) was used for simple linear regression analysis. The R^2 value of the marker associated with the Fusarium wilt resistance QTL was used to calculate the proportion of phenotypic variation explained by the QTL.

4.3 Results

4.3.1 Linkage mapping of the *C. arietinum* genome using interspecific cross *C. arietinum* X *C. reticulatum*

4.3.1.1 Parental screening for RAPD and ISSR polymorphism

The two parents *C. reticulatum* (PI489777) and *C. arietinum* (ICC4958) were screened with two hundred and forty random primers for RAPD polymorphism. Among the polymorphic bands, few (5) showed monomorphic pattern in the RILs although polymorphic in parents indicating skewing towards either of the parents and hence were not included in the analysis. Devey *et al.* (1996) have also reported a similar finding in RAPDs while constructing genetic linkage map for *Pinus radiata* based on RAPD, RFLP and microsatellite markers. Finally a total of 12 primers yielded seventeen polymorphic bands which segregated among the RILs.

During ISSR analysis of the two parents, 18 ISSR primers out of 22 generated amplified bands and 8 primers out of these 18 produced 12 polymorphic bands averaging upto 1.5 polymorphic band per primer. Out of 12 polymorphic bands, 8 were found to be segregating in the RIL population and were scored while 4 showed monomorphism as in case of RAPD markers.

Figures 4.1a and b are representative pictures of the ISSR and RAPD loci segregating in the RILs, respectively. The molecular weights of the amplified fragments using both, RAPD and ISSR, primers ranged from 1.5 to 0.5 kb. Figure 4.1a shows the segregation of ISSR fragment of 0.55 kb amplified from *C. reticulatum* using UBC 850 which is present in 12 out of 30 RILs and absent in remaining 16 RILs. Figure 4.1b shows the segregation of two RAPD markers amplified by OPC18 of molecular weights 0.6 kb and 0.55 kb, originating from *C. arietinum* (ICC4958) and *C. reticulatum* (PI489777), respectively. The markers segregated in a co-dominant manner showing no condition of heterozygosity in any of the RILs. The two markers also mapped to the same locus on the linkage group and hence only one locus was considered for further analysis. In a similar work in cowpea, Menendez *et al.* (1997) scored RAPD markers in an intraspecific population for linkage map construction where most RAPDs segregated in a dominant manner and only, a small proportion (5.3%) showed co-dominance. In the same study, polymorphic loci identified by one

primer mapped to different linkage groups however, 4 co-dominant pairs of markers were also observed to map at the same locus.

4.3.1.2 Segregation, linkage and distribution of markers

The 118 RILs obtained from the interspecific cross were evaluated for the segregation of 17 RAPD and 8 ISSR polymorphic loci. Table 4.3 gives the origin, ratio of segregation in the population and the chi-square value for each of these markers. From the 17 RAPD markers, 9 segregated in the expected 1:1 Mendelian segregation ratio and 8 showed distorted segregation. Out of 8 ISSR markers, only one segregated in the 1:1 Mendelian ratio. It was observed that the markers segregating in the RILs were in greater favour of *C. reticulatum* than of *C. arietinum*. For example, out of 17 RAPD markers, 12 originated from the wild parent, *C. reticulatum* and only 5 from *C. arietinum*. Further 7 RAPDs out of 12 originating from *C. reticulatum* and one RAPD out of 5 from *C. arietinum* showed distorted segregation. In case of ISSR markers, four markers out of 8 originated from *C. reticulatum* and the remaining from *C. arietinum*. Only one of the ISSR markers originating from *C. reticulatum* showed a 1:1 Mendelian segregation and the remaining segregated distortedly. All the four markers from *C. arietinum* segregated in a distorted manner.

Linkage analysis using Mapmaker 3.0b mapped four RAPD and ISSR markers each at LOD score 4 on different linkage groups. Among the RAPDs out of 17 markers 76.4% were unlinked. Figure 4.2 shows the chickpea linkage groups with the markers mapped on it. The markers mapped on the linkage groups are indicated by red arrow. There are 16 linkage groups that span a total of 2077.9 cM. The four RAPD markers namely OPI06-3, OPL4-2, OPC18-3 and OPE3-2 are located on linkage group 1, 3, 9 and 14, respectively. In comparison with other types of markers the RAPDs were the only markers that showed no tendency of clustering on the linkage groups.

Among the ISSR markers, (ISSR 8481 and ISSR 888) and (ISSR 8262 and ISSR 889) are located on linkage group 2 and linkage group 5, respectively. Interestingly, the marker ISSR 8262 was located on linkage group 2 on which the Fusarium wilt resistance genes for *Foc* 4 and *Foc* 5 are mapped. This marker was located at a distance of 50 cM from the *Foc* 4 resistance gene. Considering all the 37 ISSR markers mapped on the chickpea genome map, 21.6% were

found to be unlinked and a tendency for clustering was observed similar to other types of markers such as STMS.

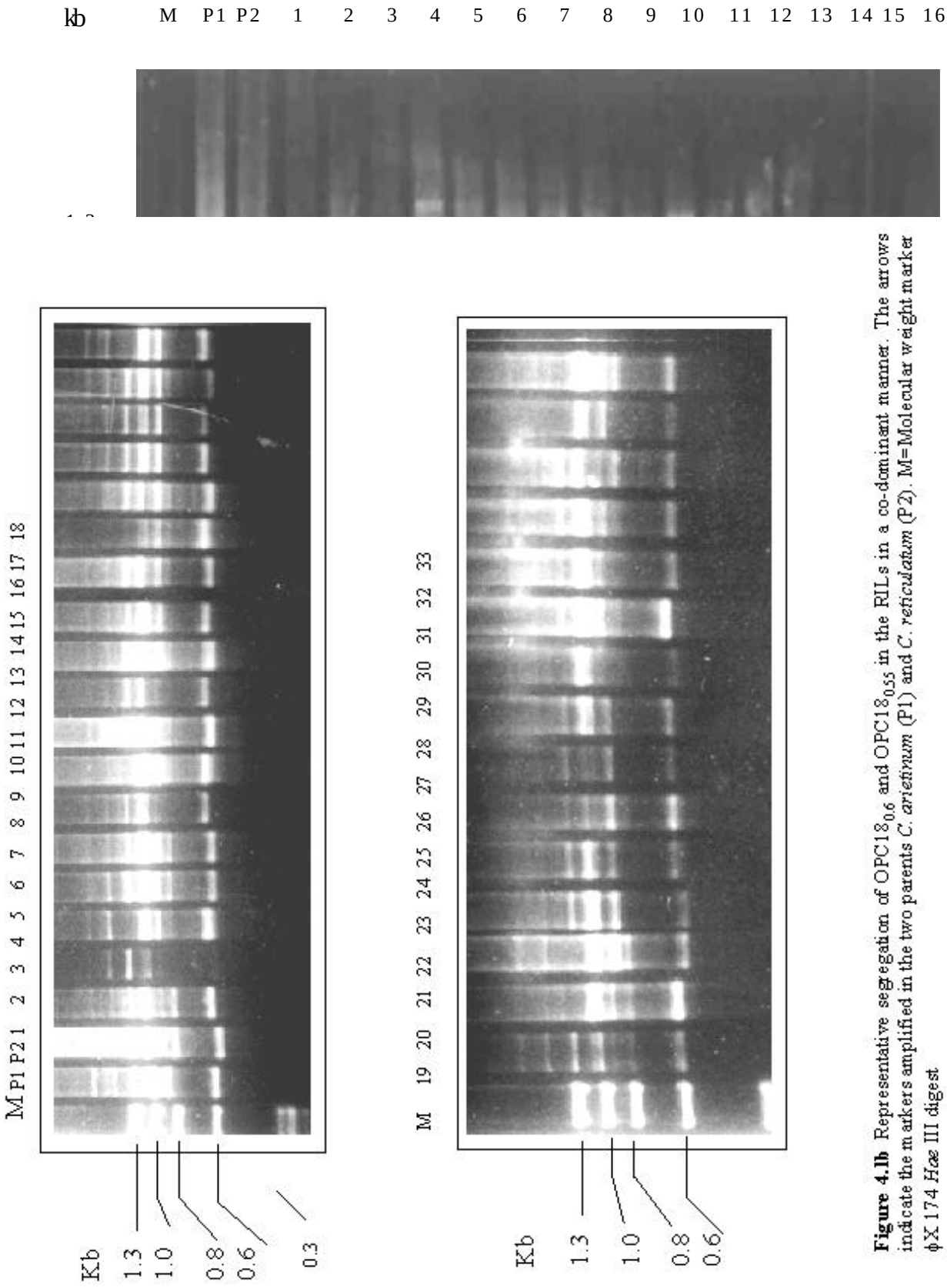


Figure 4.1b Representative segregation of $OPC18_{0.6}$ and $OPC18_{0.55}$ in the RILs in a co-dominant manner. The arrows indicate the markers amplified in the two parents *C. arietinum* (P1) and *C. reticulatum* (P2). M=Molecular weight marker ϕX 174 *Hae* III digest

Table 4.3 Chi-square values of the RAPD and ISSR markers

Marker	Size in kb	Origin	a:b	X^2(P=0.05)*
RAPD				
OPI06-3	0.7	*C. arietinum*	48:53	2.66

OPL4-2	0.6	*C. reticulatum*	52:48	2.88
OPF4-1	1.0	*C. reticulatum*	4:98	77.05
OPL11-1	1.7	*C. arietinum*	55:29	15.52
OPL11-2	0.5	*C. reticulatum*	81:17	38.10
OPC18-3	0.55	*C. reticulatum*	39:35	16.54
OPB17-1	1.5	*C. arietinum*	48:52	2.88
OPB17-2	1.4	*C. reticulatum*	23:64	22.38
OPL14	0.55	*C. reticulatum*	68:47	3.81
OPE3-2	0.9	*C. reticulatum*	26:85	29.91
OPE3-3	0.55	*C. arietinum*	65:47	3.05
OPE4-4	2.0	*C. reticulatum*	65:49	2.30
OPE4-5	1.0	*C. arietinum*	54:51	1.50
OPE7-4	0.9	*C. reticulatum*	48:50	3.42
OPE9-1	1.4	*C. reticulatum*	32:21	36.83
OPE9-2	0.7	*C. reticulatum*	24:26	39.22
OPAC4-1	1.3	*C. reticulatum*	50:60	1.38
ISSR				
ISSR 8272	0.9	*C. arietinum*	78:26	24.57
ISSR 8481	1.1	*C. reticulatum*	58:44	3.83
ISSR 8482	0.75	*C. arietinum*	55:35	10.03
ISSR 8262	0.7	*C. arietinum*	34:77	16.08
ISSR 850	0.55	*C. reticulatum*	38:52	8.30
ISSR 889	0.5	*C. reticulatum*	32:56	12.50
ISSR 888	0.6	*C. arietinum*	53:22	23.81

ISSR 8271	2.0	*C. reticulatum*	39:64	7.20

***Chi-square value at P=0.05 is 3.84**

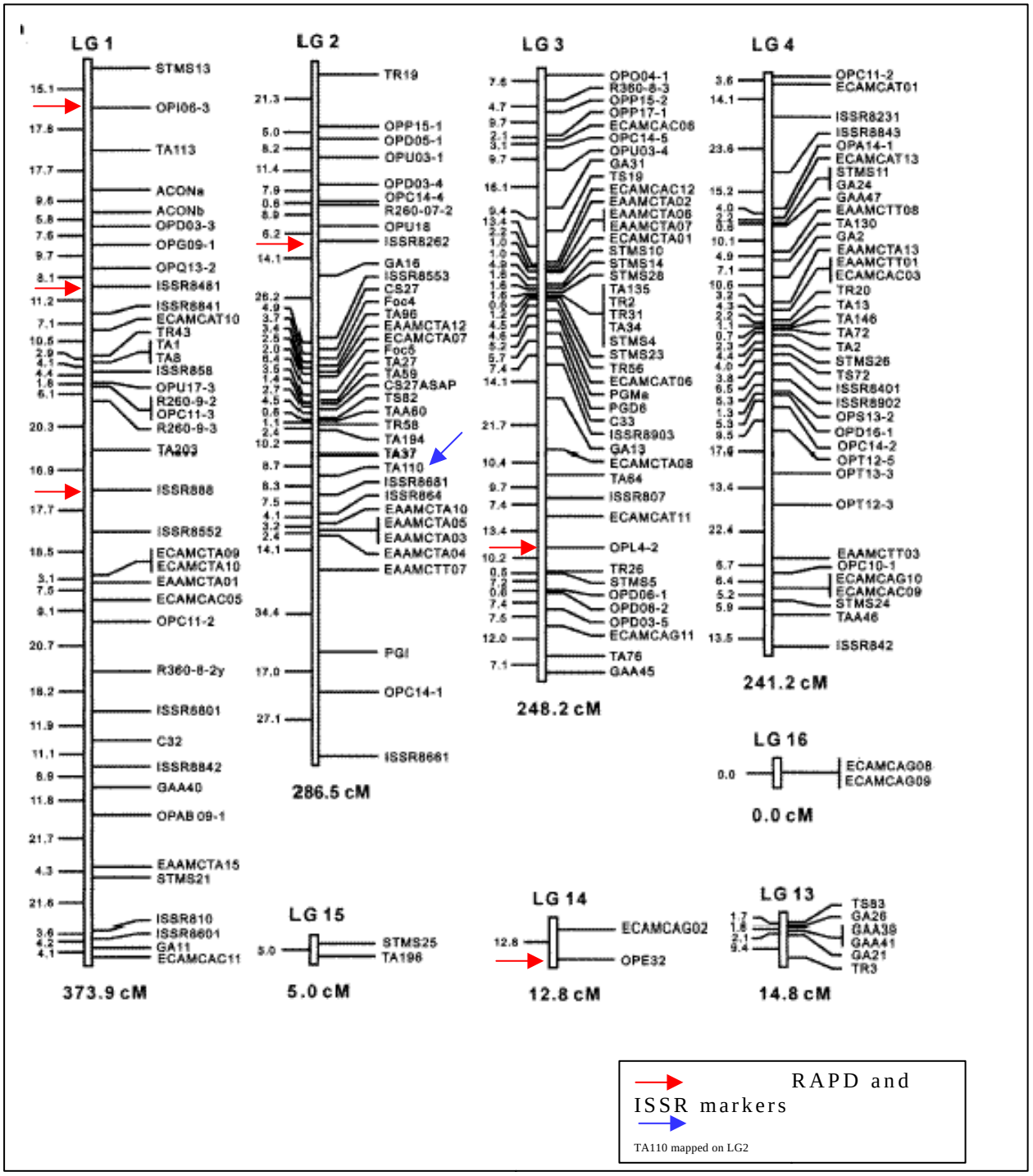


Figure 4.2 ***C. arietinum integrated genome map represented as 16 linkage groups consisting of 8 major and 8 minor linkage groups***

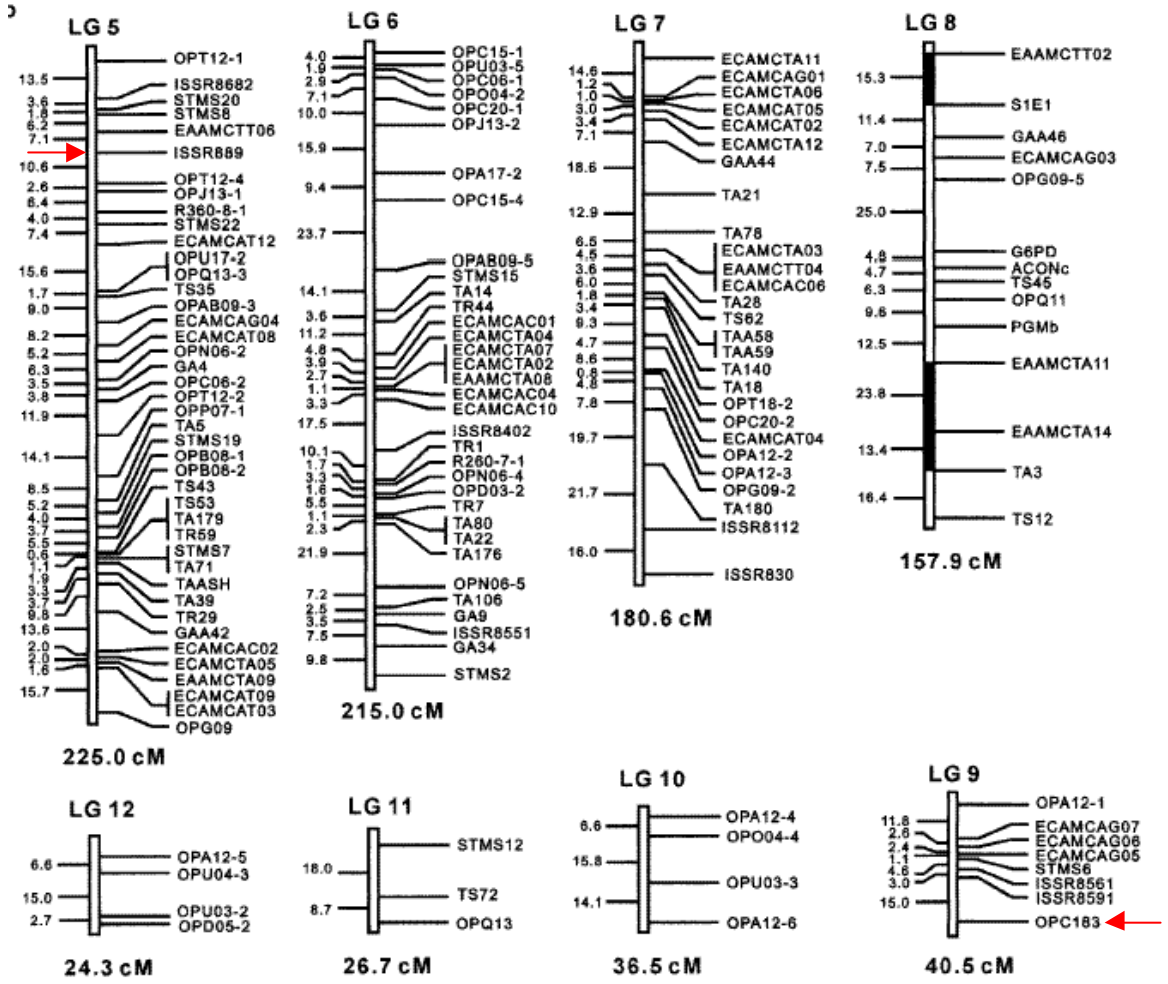


Figure 4.2. *C. arietinum* integrated genome map represented as 16 linkage groups consisting of 8 major and 8 minor linkage groups

4.3.2 Identification of a putative marker linked to Fusarium wilt resistance in chickpea

4.3.2.1 Primer selection for detecting parental polymorphism

Polymorphism between JG62 and Vijay was detected using RAPD, ISSR and RGA markers along with STMS and ASAP markers mapped on the *C. arietinum* genome from the studies in collaboration with a group at University of Frankfurt headed by Prof. Günter Kahl and at Washington State University headed by Dr. Fred Muehlbauer. Likewise, both the parents were screened with 740 random primers and 100 ISSR primers to reveal 34 random primers and 3 ISSR primers to be polymorphic. Two polymorphic STMSs were identified after attempting 9 STMS primers, while the ASAP primers detected no polymorphism. The RGA primers amplified number of bands but failed to reveal polymorphic band among the two parents. The identified polymorphic primers were later used for selective genotyping. Such low level of polymorphism has also been observed by Menendez *et al.* (1997) during cowpea map construction using an intraspecific cross where only 16 RAPD markers (9%) segregated out of 182 polymorphic markers amplified using 268 primers.

4.3.2.2 Analysis of phenotypic data

Field trials for resistance to Fusarium wilt race 1 were conducted on 166 F_4 and F_5 generation RILs along with the control plants which included ICC 4958 and the parents Vijay and JG62. ICC4958 was sown to observe late wilting reaction. The control plants showed wilting in the following duration from days after sowing (DAS): JG62 (17 to 20DAS), ICC4958 (71 to 90DAS) and Vijay was completely resistant. The RILs were scored simultaneously and were classified into nine groups as detailed in 4.2.4. The resistance and susceptibility reactions of the RILs during field trials at MPKV, Rahuri are as shown in figure 4.3. The figure shows resistant and susceptible RILs which are seen as green standing and yellow wilted lines, respectively on the 40th DAS. The reaction of the control plants is seen in figure 4.4a and b. Figure 4.4a shows the completely wilted JG62 on 20 DAS with Vijay and ICC4958 showing no wilting while figure 4.4b shows the resistant parent Vijay totally unwilted and the completely wilted ICC4958 on 90 DAS.

If two recessive genes control a qualitative trait, the expected segregation ratio is 1:3 (1= resistant; 3= susceptible) for two discrete classes of the trait in the recombinant inbred line population. However, the pattern did not follow such

segregation and therefore, lines with the score 1 and 2 were arbitrarily grouped into resistant category and the remaining as susceptible. Accordingly, 19 lines were resistant and 147 lines were susceptible. If the expected segregation ratio for wilt resistance is 1:3, the chi-square value obtained would be 16.264 which is significant at $P<0.05$. However, if 3 recessive genes are considered to be controlling the qualitative trait, the expected segregation ratio would be 1:7 (1=resistant and 7=susceptible). Considering the expected ratio of 1:7, the chi-

square value obtained would be 0.168, which is not significant at P<0.05. This suggested involvement of three recessive genes conferring resistance to Fusarium wilt in the population obtained from JG62 X Vijay. Figure 4.5 shows

the frequency distribution of the disease scores of the RILs thereby indicating three genes conferring resistance to Fusarium wilt with a possibility of some modifiers.

4.3.2.3 Selective genotype analysis

Analysis of the RILs developed from the cross of JG62 and Vijay was performed using the identified polymorphic primers to amplify marker bands by selective genotyping as described in 4.2.4. The markers co-segregating in either resistant or susceptible phenotype with 60% or greater number of RILs were chosen for further studies. The previously identified polymorphic ISSR and random primers generated 5 ISSR markers and 25 RAPD markers which failed to segregate in selective RILs as mentioned above. The STMS primer TA110 amplified fragments of 0.234 kb and 0.220 kb in JG62 and Vijay, respectively. The allele from the resistant parent Vijay (0.220 kb) showed its presence in 86% of the resistant RILs while absence in 87.5% of the susceptible RILs in selective genotyping. This primer was, therefore, further used for studying the segregation in all 166 RILs.

4.3.2.4 Identification of a putative STMS marker linked to Fusarium wilt resistance race 1

Segregation of the alleles of STMS marker amplified by TA110 primer namely $TA110_{0.234}$ and $TA110_{0.220}$ was analyzed in the 166 RILs. The marker segregated in an expected 1:1 ratio (X^2 =0.433, P<0.05). Figure 4.6 is a

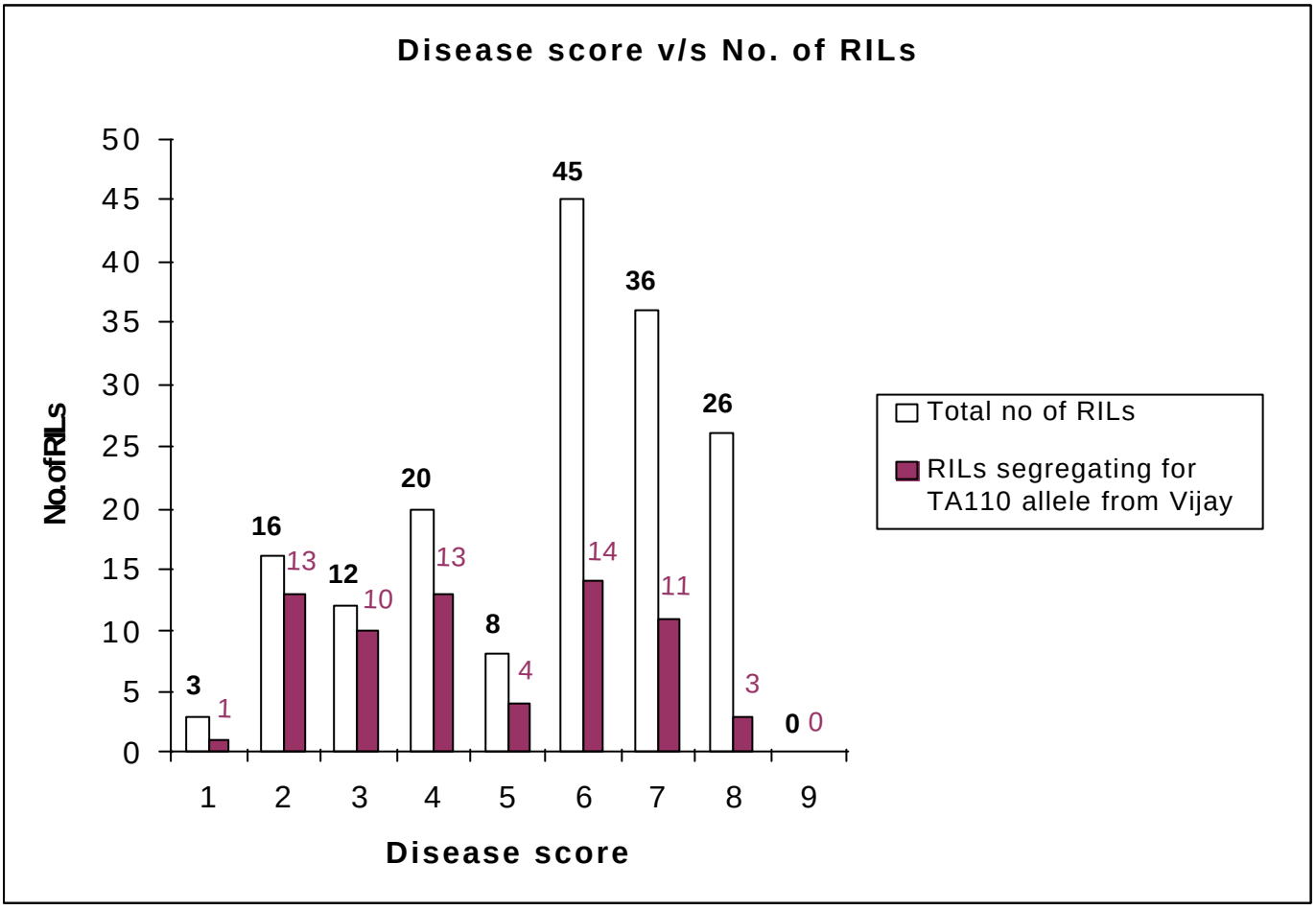


Figure 4.5 Frequency distribution of disease score for the RILs along with the segregation of the TA110 allele amplified from Vijay (resistant). Disease was scored on a 1-9 scale where 1 indicated resistance and 9 indicated complete wilting (in 0-10 DAS). RILs with disease score 1-2 were considered resistant and 2-9 as susceptible.

representative picture of the segregation of the marker TA110 in 45 F_5 generation RILs along with the resistant and susceptible parents. STMS markers being co-dominant in nature, heterozygosity for the locus can be detected. The expected residual heterozygosity in $F_{4:5}$ generation RILs is 6.2% $(1/2)^4$. In my study out of 166 RILs, 8 were heterozygous for TA110, which amounted to 4.8% heterozygosity. The segregating marker and the phenotype data scored were analyzed by simple linear regression using the software Q-gene. The R^2 value indicated an estimated 20.62% contribution to the phenotypic variation for resistance to Fusarium wilt at P<0.01. This indicated the presence of a putative

quantitative locus contributing for resistance to race 1 of *F. oxysporum* f. sp. *ciceri*. The distribution of the STMS marker TA110 in the RILs with differential disease score is seen in figure 4.5. A few recombinants at locus TA110 with respect to their genotype and phenotype were observed in RILs. Therefore, the marker was placed at a distance of 23 cM from the locus conferring resistance to Fusarium wilt at LOD 6.0.

Thus TA110 marker was identified in this study first by selective genotype analysis and was later confirmed in the complete set of RILs on the basis of the Fusarium wilt resistance reaction obtained in the field trials in two consecutive years 1998-99 and 1999-2000. However, it is essential to have multiple replications of wilt resistance testing in the field and also in the glass house by single spore inoculation of the pathogen. This is essential to confirm the linkage of TA110 and to further identify markers more tightly linked to other major and minor genes conferring complete resistance to wilt in chickpea.

The STMS marker TA110 used in this study was mapped on linkage group 2 of the integrated linkage map of *C. arietinum* genome developed from an interspecific cross between *C. arietinum* (ICC4958) and *C. reticulatum* (PI489777) during collaborative efforts among the three laboratories from India, USA and Germany as detailed in 4.3.2.1. As seen in figure 4.3, TA110 is mapped at a distance of 35.1 cM from the *Foc* 4 and *Foc* 5 resistance genes on linkage group 2. In my study, this marker showed linkage to a major locus conferring resistance to *F. oxysporum* f. sp. *ciceri* race 1 (*Foc*1). This indicates that the locus conferring resistance to race 1 may also be located on the same linkage group but at a different position than *Foc* 4 and *Foc* 5.

4.4 Discussion

4.4.1 Segregation distortion in RILs of the interspecific cross

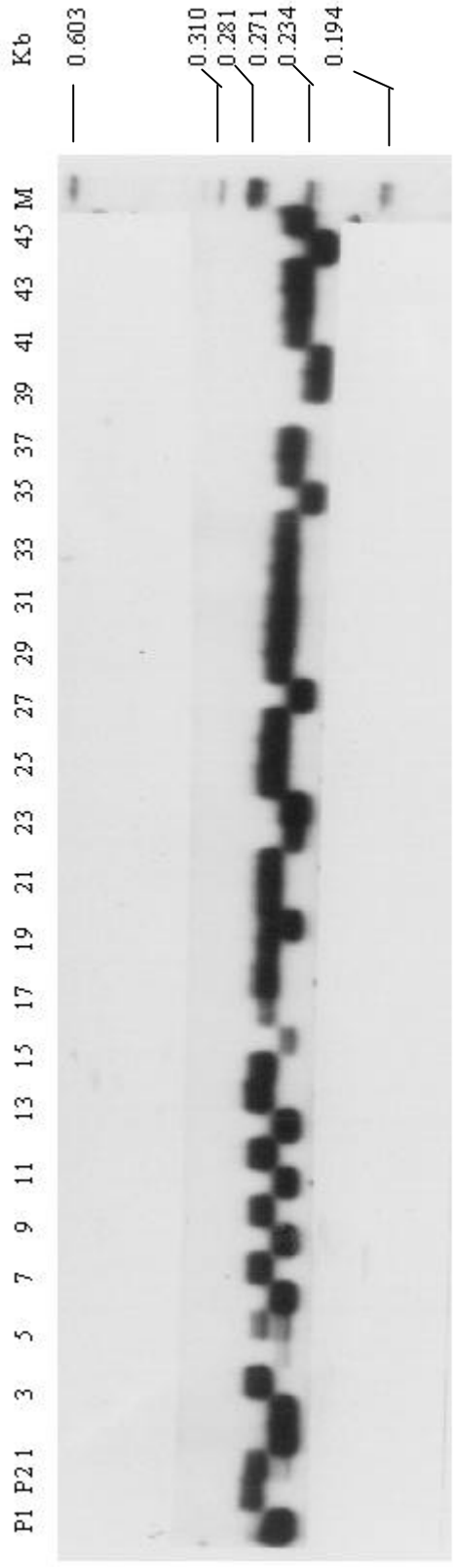


Figure 4.6 Segregation of the alleles $TA110_{0.220}$ and $TA110_{0.234}$ amplified in Vijay (P1) and JG62 (P2) respectively, in the RILs developed from the cross of JG62 X Vijay. M= ϕX 174 *Hae* III digest.

The *C. arietinum* genome map, generated using the interspecific cross, included different types of markers with each marker type showing different amounts of segregation distortion. Here, RAPD and ISSR markers showed 47%

and 37.5% distortion, respectively as against the AFLP and DAF markers which showed 28.5% and 42.1% distortion, respectively. The average segregation distortion was 38% considering all the different types of markers mapped on the *C. arietinum* genome. This value is very close to that reported for RILs (39.4±2.5%) by Xu *et al.* (1997), who surveyed 53 different populations with a known number of distortedly segregating markers. The extent of segregation distortion was not dependant on the marker type but probably on the segregation distortion of the region where they resided. Similar observation has been reported by Korzun *et al.* (1998) where 4 chromosomes carrying most of the RFLP loci showed distorted segregation during linkage map construction using F_2 population in rye. Clustering of markers with distorted segregation was also previously reported for several crops including common bean (Nodari *et al.* 1993), barley (Heun *et al.* 1991) and potato (Bonierbale *et al.* 1988). Zamir and Tadmor (1986) indicated that distorted segregation might be the result of linkages between markers and the genes operating in the pre and post zygotic phases of reproduction. The other possible explanations included preferrential chromosome elimination, preferrential fertilization and the selective elimination of particular zygotes (Tanksley 1984, Koenig and Gepts 1989)

4.4.2 Status of RAPD markers in linkage mapping

In figure 4.2, all the types of markers, other than RAPDs, showed a tendency for clustering. The linkage groups revealed clustering of different types of markers at specific regions and further subclustering of the markers of the same type. These linkage groups included specific core regions where different marker types, mostly STMS and AFLPs, and a few of isozyme and ISSR markers were closely linked. These clusters were mostly located in central regions of linkage groups, whereas marker density in distal regions was low. Tanksley *et al.* (1992) identified high marker density regions on the linkage groups in advanced linkage maps of tomato and potato and suggested that these regions corresponded to centromeric areas and in some instances, telomeric regions. A similar clustering of markers was also observed in soybean, where in the first study, random distribution of SSRs was observed (Akkaya *et al.* 1995) but later clustering became visible when more markers were added. This clustering was due to the suppression of recombination in the heterochromatic region around the centromeres (Cregan *et al.* 1999).

In my study, higher percentage of RAPD markers remained unlinked as compared to the other types of markers studied. Similar case was observed for genetic linkage mapping in peach where only 7 out of 12 RAPD markers were mapped, amounting to 41.6% unlinked markers. RFLP and morphological markers were also scored in the same population, wherein 9 RFLP (19.6%) markers remained unlinked out of 46 RFLP loci scored (Rajapakshe *et al.* 1995). Later one more map of peach was developed using a different population characteristic for four more agronomic characters where 92 additional RAPD markers were scored, out of which only 10 (10.8%) were unlinked (Dirlewanger *et al.* 1998). This indicates that addition of more number of RAPD markers in chickpea may assign the unlinked markers to different linkage groups further suggesting that large regions of the chickpea genome are yet to be covered by markers.

4.4.3 Differential wilting in chickpea genotypes

Genetics to Fusarium wilt was studied by Kumar and Haware (1982) in chickpea suggesting that resistance might be conferred due to the existence of major genes or polygenic complexes acting as modifiers for wilt resistance. In another study, the difference in the number of days to wilt by JG62 (early wilter) and C104 (late wilter) was shown to be controlled by a single gene with early wilting partially dominant to late wilting (Upadhyaya *et al.* 1983a). Smithson *et al.* (1983) also indicated the presence of three recessive genes for resistance to race 1 and identified genes for very late wilting, late wilting and early wilting in different genotypes. In subsequent studies crosses among late wilters, and between early wilters and completely resistant lines indicated that resistance to race 1 was controlled by at least three independent loci designated as H_1, H_2 and H_3 (Upadhyaya *et al.* 1983b, Singh *et al.* 1987a). These studies reported the phenomenon of early wilting to be caused by two dominant genes at the first two loci and recessive gene at the third locus (Upadhyaya *et al.* 1983b). These studies suggested that partially recessive alleles in homozygous form at either the H_1 or H_2 locus or the dominant allele at the third locus (H_3) separately delay wilting but any two of these loci together confer complete resistance to race 1 of *Fusarium oxysporum*. Girase *et al.* (1999) reported the resistance in Vijay to be conferred by 4 recessive genes. The susceptible parent JG62 and resistant parent Vijay differed at the first, second and fourth loci, the third loci being

common in both the parents. Thus the genotype for Vijay was reported as $h_1h_1h_2h_2h_3h_3h_4h_4$ and for JG62 as $H_1H_1H_2H_2h_3h_3H_4H_4$

A few chickpea genotypes were demonstrated to differ in times of wilting (Haware and Nene 1980) which might indicate different degrees of resistance and might reflect in differential behavior in crosses (Upadhyaya *et al.* 1983a). In order to dissect the differential time of wilting of the RILs in this work, I treated the data in quantitative manner. However, it is necessary to study the differential wilting in different genotypes under identical conditions and to standardize a scale for disease scoring.

4.4.4 Genes for resistance to Fusarium wilt are clustered in chickpea

Several workers have identified molecular markers linked to genes conferring resistance to Fusarium wilt caused by race 1 and 4. Muehlbauer *et al.* (1994) identified linkage of $UBC170_{550}$ in coupling and $Cs27_{700}$ in repulsion to race 4 of *F. oxysporum*. Later Ratnaparkhe *et al.* (1998) reported an ISSR marker $UBC855_{500}$ linked in repulsion to race 4 resistance. Mayer *et al.* (1997) studied the segregation of markers $UBC170_{550}$ and $Cs27_{700}$ for resistance to race 1 and found both the markers to be linked to resistance in coupling and repulsion, respectively. These markers were converted into allele specific associated primer (ASAP) pairs and both the ASAP generated markers were located on the same side of the resistance gene with 6% recombination between the two markers and 7% recombination between these loci and the locus that controlled resistance. Tullu *et al.* (1999) reported the linkage of $Cs27_{700}$ to one of the recessive genes out of the two conferring resistance to race 4. Later $Cs27_{700}$ was mapped on the integrated chickpea map on linkage group 2 by the concerted efforts of India, USA and Germany. Genes conferring resistance to *Foc* 4 and *Foc* 5 were also mapped on linkage group 2 where $Cs27_{700}$ was linked to *Foc* 4 and *Foc* 5 at a distance of 3.7cM and 18 cM, respectively. In my study, the co-dominant STMS marker TA110 is linked in coupling to the locus conferring resistance to race 1 and which is also located on the linkage group 2. On the basis of the above efforts, it can be concluded that the genes for resistance to *Fusarium oxysporum* f.sp.*ciceri* races 1, 4 and 5 (*Foc*1, *Foc*4, and *Foc* 5) form a cluster in chickpea.

Clustering of resistance genes for different races of a pathogen has been demonstrated in several plants including legumes (Kanazin *et al.* 1996, Yu *et al.* 1996). Close linkage of microsatellite markers with resistance genes has also been documented for rice bacterial leaf blight (Blair and McCouch 1997). Microsatellites occur in high density in the chickpea genome (Hüttel *et al.* 1999) which might result in close linkage of these markers with the wilt resistance genes.

In conclusion, chickpea genome map needs to be constructed using an intraspecific cross to represent and map the *Cicer arietinum* genome completely. Furthermore, the differential wilting phenomenon observed in my studies as well as in other studies in chickpea needs to be dissected in more details to confirm that the Fusarium wilt resistance is governed by quantitative resistance loci (QRL).

CHAPTER 5

General Discussion

5.1 Introduction

Most successful plant breeding programs consist of several components like: 1) identifying clear and reasonable goals 2) surveying and selecting germplasm (with specific DNA sequences, genes, pure lines, clones, accessions, cultivars and populations); 3) identifying and creating genetic variation; 4) defining the target environment for production; 5) carefully selecting and creating test environments, representative of the target environment; 6) assembling genes into genotypes; and 7) matching genotypes with environment to optimize production (Lee 1995). Preliminary investigations in a crop system are required to be made in each of these components. In my thesis work, I have attempted to throw light on some of these aspects such as genetic diversity and identification of important fungal disease resistant gene locus in chickpea. In this chapter, I discuss briefly each of these aspects in comparison with some of the well studied crops to highlight the future research prospectives in chickpea.

5.2 Prevention of genetic erosion is important to conserve genetic diversity

Modern agricultural practices, which generally emphasize maximum productivity with acceptable quality and uniformity, have resulted into a reduction in the genetic diversity of the primary gene pool under cultivation with similar fates for the secondary and tertiary gene pools of most of the major crops. The consequence of narrow genetic base of major crops is often accompanied with agro-economic losses which is a serious problem considering their increasing demand due to increasing population. The range of genetic variation is much greater in exotic germplasm than among cultivated types, as cultivated types are usually derived from a small number of ancestors, and have been selected intensely over many centuries (Hawkes 1977, Vaughan 1989). Consequently, the current breeding pool for many crops, especially self-pollinated species (such as wheat, rice, soybean, cotton and tomato) contains only a small fraction of the existing genetic variation. Soybean is an excellent example in this regard where a limited number of accessions were used to derive the cultivars commonly grown in the US today. An analysis of the pedigrees of cultivars in the Northern US germplasm

collection indicated that 88% of their collective genome was derived from just 10 accessions, while in the Southern germplasm collection 70% of the genome of the commonly grown Southern cultivars was shown to be contributed by as few as seven accessions (Delannay et al 1983, Specht and Williams 1984). This narrow genetic base created during breeding and cultivar development in soybean limited the amount of genetic diversity found among elite lines.

In chickpea wide genetic diversity is expected as the genetic resources include the primitive land races or cultivars, genetic and mutant stocks of the cultivated species and wild species of *Cicer* (Malhotra *et al* 1987). The Indian subcontinent, has area under chickpea cultivation involving land races, thus conserving the wide genetic diversity. However, there are some factors causing serious genetic erosion over a period of time. For example, during 1968/69 and again in 1978/79 in the Punjab province of India and adjacent Pakistan, Ascochyta blight developed in an epiphytotic form, killing almost all the chickpea crop. Similarly, loss caused by drought in Ethiopia during the mid-1970s reduced chickpea cultivation by half. Also, in Spain, low net returns from chickpea reduced production drastically. These calamities force the replacement of landraces with improved cultivars, thereby resulting in the continuous erosion of the genetic diversity. Besides, the wild species of *Cicer* which occur in small isolated areas, are also reduced in number and distribution due to overgrazing or changes in land use (Malhotra et al 1987).

It is important to identify the existing genetic variation in chickpea using various tools and use it in breeding programs. In my work on genetic diversity in the elite chickpea germplasm microsatellite repeats could detect large amounts of variation which could be exploited for making further progress in the improvement of this crop.

5.3 Plant genomic research: Bridging the basic and applied science

Genomic research is a branch of science that seeks to understand the structure, function and evolution of genes and genomes. Plant genomics is oriented towards the use of genomic research to enhance the productivity, quality and sustainability of our food production systems. The use of agricultural species with their rich history of genetic improvement, their

extensive germplasm collections, and their intimate associations with human history and culture throughout the world bring in new challenges and possibilities. (McCouch 1998).

In the past decade, complete genome sequencing has been undertaken in model plants like rice and *Arabidopsis*. Information obtained is based on sophisticated computational analysis of large data sets generated by high throughput sequencing and chip reading robots. The first step towards plant genomic studies is understanding the genome based on the organization of repetitive DNA present in the heterochromatic region or interspersed with genes. Only recently, it has been shown that most of these interspersed repetitive DNA sequences are mobile DNAs, mainly LTR-retrotransposons (Bennetzen 2000). Studies of plants with large genomes have shown the presence of mostly interspersed repetitive DNAs (Flavell et al. 1974), where some of these repetitive DNAs are interspersed with genes, but much of it is intermixed primarily with other repeats in centromere-associated regions (Schmidt et al. 1995). Most of this interspersed repetitive DNA (at least in monocots) is composed of LTR-retrotransposons (Bennetzen, 2000). Some of the intact and/or fragmented LTR retrotransposons are associated with genes but most lie in intergenic blocks that are methylated and presumably heterochromatic. In chickpea, both the LTR retrotransposon Ty1-*copia* and Ty3-*gypsy* are present in the pericentric heterochromatin block on all chromosomes with extensions into the euchromatin regions (Staginnus *et al*. 1999). Small plant genomes, like *Arabidopsis*, have fairly few repetitive DNAs and most of these are found in large blocks like satellites (Pelissier et al. 1995), centromeres (Jiang et al. 1996), telomeres (Richards and Ausubel 1988) and centromere associated regions. Although the same types of elements may be found in different plant species, the specific families will largely be distinctive to a genus and some types may predominate in some species, while others may be more abundant in other species (e.g LTR-retrotransposons in yeast versus LINEs in man) (Bennetzen 2000). The rules of arrangement of any plant genome (except the plant genomes which have been sequenced completely, e.g *Arabidopsis*) are not yet known and we have little idea about the nature of most of such exceptional regions.

Besides studies using the markers like microsatellites and ISSRs in chickpea, the detailed characterization of the non-LTR retrotransposons like LINEs as well as of LTR retrotransposons can help in investigating the diversity generated due to these elements in the genome. My studies describe the potential of these elements and the variation caused by them in chickpea genome. Other types of retrotransposable elements are predicted to be present in this crop and as yet these elements and their role remain uninvestigated.

5.4 Genetic mapping: The link between genomics and genetics

The genetic map has maintained its central importance as a basic tool that links information in the nucleotide sequence to phenotypic traits. It makes available genomic data interpretable in ways that are useful to agricultural scientist and geneticists. In the United States, one of the major national research initiatives in agriculture has been the Plant Genome Project which has supported the development of molecular genetic maps for more than 50 crop species. These maps have been used to localize genes and quantitative trait loci (QTL) and to select agriculturally important genes for positional gene cloning. Over the last decade, extensive species-specific maps have been developed for many members of the family Gramineae. Comparative mapping among these genera by using a common set of cDNA clones suggests that the same genes are arranged in identical order along large tracts of chromosome (Chen *et al* 1998). This puts forth the concept that all these crops contain essentially the same set of ancestral genes duplicated extensively in some genomes (especially in maize and wheat) and rearranged through translocation and other mutational events over time. Van Deynze *et al* (1998) have developed a set of "anchor" cDNA probes which can align the species specific maps and have distributed these probes to more than 50 research groups for mapping experiments during the last two years. Linkage among anchored loci has allowed the identification of homeologous regions of distantly related genomes and has provided instantaneous co-ordination of mapping results among independent investigators. It is now attempted to localize these grass anchor probes on maps of more distantly related plant species like dicots and gymnosperms to determine how extensive the

homeologous relationships are across larger evolutionary distances. Such evaluation of macrosynteny, based on the chromosomal segments, can be complemented with evaluation of microsynteny based on sequencing of large tracts of genomic DNA. Comparative sequencing of multiple genomic regions anchored by homologous cDNAs, both within the grass family and between evolutionarily more divergent groups of plants will provide critical data for determining the most efficient way to align sequence information across diverse plant species (McCouch, 1998).

Construction of chickpea genome map is still at a preliminary stage and I have given information about my inputs in this area which involves placing come markers on the map. It is evident that more regions from the cultivated chickpea need to be represented. Comparison of chickpea genome with evolutionarily related, well studied legumes will help to identify syntenic regions. In my work I have identified a putative DNA marker linked to agronomically important trait Fusarium wilt resistance, which is also placed on the chickpea map.

In summary, it is necessary to bridge the sequence information generated and integrate it with the classical and molecular data of the genome in evolutionarily diverse organisms. One of the greatest challenges posed by the Plant Genome Initiative is to make instructive use of (i) the rich genetic reservoirs of crop plant germplasm that has accumulated over many decades and (ii) exploit and link together the knowledge of the agricultural community gained over time and the performance characteristics, crossing histories and environmental adaptations of the crop species (McCouch 1998).