# GENETIC DIVERSITY STUDY AND MOLECULAR MARKER DEVELOPMENT IN *CYAMOPSIS TETRAGONOLOBA*

**A THESIS**

***Submitted in partial fulfilment of the requirements for the award of the degree of***

DOCTOR OF PHILOSOPHY

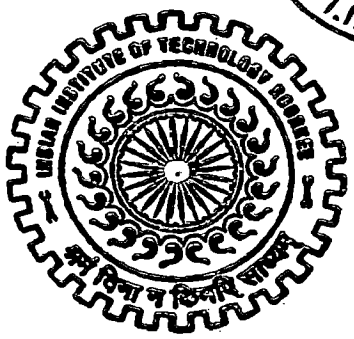
***in***

BIOTECHNOLOGY

***by***

**NAGESH K. A.**






DEPARTMENT OF BIOTECHNOLOGY
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247 667 (INDIA)

DECEMBER, 2011

# INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
# ROORKEE

## CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **GENETIC DIVERSITY STUDY AND MOLECULAR MARKER DEVELOPMENT IN *CYAMOPSIS TETRAGONOLOBA*** in partial fulfilment of the requirements for the award of the Degree of Doctor of Philosophy and submitted in the Department of Biotechnology of the Indian Institute of Technology Roorkee, Roorkee is an authentic record of my own work carried out during a period from January 2009 to December 2011 under the supervision of Prof. G. S. Randhawa, Department of Biotechnology and, Prof. S. K. Tripathi, Department of Water Resources Development and Management , Indian Institute of Technology Roorkee, Roorkee, India.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

K.A. Nagesh

**(NAGESH K. A.)**

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

(S. K. Tripathi)
Supervisor

(G.S. Randhawa)
Supervisor

Date: 23/12/2011

The Ph. D. Viva-Voce Examination of **Mr. Nagesh K. A.** Research Scholar, has been held on 25.06.2012

25.6.2012
Signature of Supervisor

25.6.2012
Signature of Chairman SRC

Sunita Singh Dhawa 25.06.2012
Signature of External Examiner

28.6.12
Head of the Department/Chairman ODC

# ABSTRACT

Cluster bean (*Cyamopsis tetragonoloba* [L.] Taub.), commonly known as guar, is used as fodder, vegetable and green manure. In recent times it has become a major industrial crop due to the need of the guar galactomannan/gum present in the endosperm of its seeds. Guar gum is useful in various industries like paper, textile, petroleum, drilling, pharmaceutics, food, cosmaceutics, explosives, etc. Guar is a drought resistant, hardy, deep rooted annual legume. In India the crop is mainly grown in the dry habitats of Rajasthan, Haryana, Gujarat, Punjab and to a limited extent in Uttar Pradesh and Madhya Pradesh. Outside India, guar is grown in Pakistan, South Africa, Brazil, Australia and Oklohoma planes of North Texas in USA. Guar is a cultivated crop not found in wild conditions and hence its available landraces are the main source of genetic variability.

Guar is strictly a self-pollinated diploid legume with chromosome number (2n) equal to 14 and genome size approximately 2.45 Giga Bases/C. Cross pollination is prevented due to the cleistogamous nature of flowers. Thus, the heterosis available is reduced, which makes commercial hybrid seed production difficult and non-economical. This limiting factor of yield gap can be overcome by production of improved varieties of guar through molecular marker based selection and breeding programs. The DNA based molecular markers reveal natural variation at the DNA sequence level; these markers are used in plant genotyping, diversity studies, genetic linkage studies, quantitative trait mapping and marker-assisted selection during plant breeding. Hence, an overview of the genetic diversity and the development of molecular markers are very important for breeding and crop improvement in guar.

In the present study genetic diversity in 19 commercial varieties and 29 landraces of cluster bean belonging to Gujarat, Rajasthan, Haryana and Delhi regions of India were analyzed using 13 RAPD (Randomly Amplified Polymorphic DNA) and 7 ISSR (Inter Simple Sequence Repeat) markers.

The amplification using RAPD primers produced a total of 118 bands, out of which 103 were polymorphic and 15 monomorphic. Out of the 13 primers used OPQ-09 produced the highest number of bands (12); the average percentage polymorphism for RAPD markers was 87.63. UPGMA tree was constructed using Jaccard's similarity. The accessions of cluster bean distinguished into two major clusters at 75% similarity and a third cluster at lower similarity. The observed number of alleles, effective number of alleles, Nei's genetic diversity, Shannon's

information index for landraces and commercial varieties using 13 RAPD markers were found to be 1.872±0.335, 1.589±0.351, 0.333±0.170, and 0.490±0.230, respectively. The value of total genotypic diversity among population (Ht) was 0.333±0.029, whereas diversity within population (Hs) was found to be 0.283±0.026. Mean coefficient of gene differentiation (Gst) value was 0.148 which indicated that 86.2% of genetic diversity was present within the population. AMOVA was used to analyze molecular variance among and within the population. Percentage of molecular variance was found to be 27% among populations and 73% variance was attributed to variance within the population.

Seven ISSR markers used in the study produced 64 bands out of which 50 were polymorphic. Among the ISSR primers used UBC-868 produced highest number of bands (13); the average percentage polymorphism for ISSR markers was 77.82. The dendrogram from ISSR data showed one major cluster at 75% similarity and five minor clusters at lower level of similarity. The major cluster possesses six sub-clusters. The dendrogram did not differentiate between landraces and commercial varieties. The observed number of alleles, effective number of alleles, Nei's genetic diversity, Shannon's information index for landraces and commercial varieties using 7 ISSR markers were found to be 1.7812±0.4167, 1.4627±0.3844, 0.267±0.1939, and 0.3988±0.2681, respectively. The value of total genotypic diversity among population (Ht) was 0.2639±0.0378 whereas diversity within population (Hs) was found to be 0.253±0.035. Mean coefficient of gene differentiation (Gst) value was 0.041 and the estimated gene flow in the population was 11.549. AMOVA was used to analyze variation among and within the populations. Molecular variances were 8% and 92% among and within the population, respectively.

Pooled RAPD+ISSR data of cluster bean distinguished into two major clusters at 75% similarity and 3 minor clusters at lower similarity in the dendrogram constructed. The observed number of alleles, effective number of alleles, Nei's genetic diversity, Shannon's information index for landraces and commercial varieties using 13 RAPD and 7 ISSR markers were found to be 1.8407±0.367, 1.5446±0.3671, 0.3130±0.1817, and 0.4584±0.2478, respectively. The value of total genotypic diversity among population (Ht) was 0.3089±0.0333 whereas diversity within population (Hs) was found to be 0.272±0.029. Mean coefficient of gene differentiation (Gst) value was 0.116 and the estimated gene flow in the population was found to be 3.787. AMOVA was used to analyze variation among and within the populations. Molecular variance among populations was found to be 21% and that within the population was 79% indicating higher variation within the population.

The Mantle test revealed a significant correlation between the molecular data and the geographic data. The correlation value (R) for RAPD, ISSR, RAPD+ISSR data with geographic data were 0.5252, 0.3144, 0.5303, respectively. This indicates that molecular variation corresponds to differences in geographic distribution of landraces.

A recombinant inbred population is essential for various purposes in plant genetics. Hence a cross was made between guar cultivars M-83 X RGC-1066. The parental lines used were unbranched type with 0-3 branches. M-83 is a vegetable variety with low gum content, glabrous leaves and white flowe. While, RGC-1066 is a gum producing variety with high gum content, hairy leaves and purple flowers. The leaf pubescence is associated with uneasiness in handling the plants, yet no information on the morphology of trichomes in guar is available. Scanning Electron Microscopy (SEM) was done to observe the fine structure of leaf pubescence. The SEM pictures showed long, slender, porous hair like structures in the hairy plant. The F1 hybrid plants showed hairy leaf, purple flowers and 0-3 branches. The F2 population segregated in Mendelian ratio showing ~3:1 ratio of hairy and glabrous plants. No observable difference in the density of the pubescence was observed in hairy plants. Similar ratio was observed for purple and white flowered phenotypes, showing single locus control for these two phenotypes. But, the branching behavior showed transgressive segregation leading to very high branching in some F2 plants. In the population majority of the plants showed less than 10 branches per plant.

Expressed Sequence Tags (EST) are considered as a quick and inexpensive source for obtaining Simple Sequence Repeat (SSR) markers. Available *C. tetragonoloba* EST (16,476) were downloaded from dbEST of NCBI. The EST sequences were trimmed using EST trimmer. The sequences so obtained (16,108) were assembled into contigs using CAP3. The candidate SSR containing sequences in the EST sequences (16,108), assembled contigs (1755) and singlet sequences (4320) were mined using the PERL script MISA. The program showed a total of 1568 microsatellite repeats in the complete EST sequences; among them 91 were in compound formation and remaining were perfect microsatellites. The contigs showed 327 microsatellite repeats from 276 sequences of which 28 were in compound formation. Singlets had 580 microsatellite repeats in 506 sequences of which 41 were in compound formation. Mononucleotide repeats (435) were the most abundant among the SSR types in *C. tetragonoloba* ESTs, followed by di-nucleotide repeats (189). Only one hexa-nucleotide repeat and seven penta-nucleotide repeats were found. A/T repeats were the most abundant form of nucleotide repeats. The sum of the microsatellite repeats from the contigs and singlets (907)

were used for designing primers. Primers flanking the SSR regions were designed using Primer3.

To test the EST based SSRs, 226 primers were synthesized and tested on 3 accessions of *C. tetragonoloba,* viz., M-83, RGC-1002, RGC-1066 and one accession each of *C. serrata* and *C. senegalensis.* The amplification, transferability and polymorphism of these markers were analyzed. Out of the 226 primers used 190 amplified to produce the product in the expected range. The polymorphism between the *C. tetragonoloba* accessions was very low; only 5 of the markers were polymorphic, showing very high homogeneity in the cultivated genotypes. Euclidean similarity coefficient was used to generate a dendrogram using the SSR data. The dendrogram showed that *C. senegalensis* was very diverse from the other accessions. The accession of *C. serrata* was very close to the cultivated guar cultivars.

DNA from the F2 population was extracted and used to study the inheritance of the markers developed. Bulk segregant analysis was also done to check the linkage of leaf pubescence and flower color traits with the polymorphic markers derived. The amplification in bulks showed both the marker alleles of parents. This shows that the traits are independent of the markers tested.

This research work showed significant variation between the landraces and commercially grown cultivars. In the landraces the diversity correlates with the geographic distribution. The work also showed very high variation in the genomes of the wild *Cyamopsis* species in comparison to guar. In this work a total of 362 primers flanking the SSR regions in guar were designed; this would be a useful resource for genetic studies and breeding in guar.

# ACKNOWLEDGEMENTS

It is the contribution of one or more persons towards the task, which makes it successful. It gives me immense pleasure in acknowledging all the help that I have received during the period of research.

First and foremost I express my deep sense of gratitude to my supervisor **Prof. G. S. Randhawa**, Department of Biotechnology, Indian Institute of Technology, Roorkee, who is not only a teacher and guide to me, but my mentor and well wisher too. His persistent encouragement, perpetual motivation, everlasting patience, constructive criticism and valuable technical inputs in research have benefitted me to an extent which is beyond expression. He has not only trained me in science but in all aspects of life. I would like to take this opportunity to also thank **Dr. Mrs. Surinder Randhawa** who has been supportive in all the efforts during my research.

I would also like to thank **Prof. S. K. Tripathi,** Department of Water Resource Development and Management, Indian Institute of Technology, Roorkee, for his constant support and guidance throughout the period of research.

I wish to acknowledge my deep sense of gratitude for **Prof. Ritu Barthwal,** HOD, Department of Biotechnology, Indian Institute of Technology, Roorkee, for providing all the necessary resources and lab facility. I also express my gratitude to **all the faculty members** of Department of Biotechnology and Institute Instrumentation Facility, IIT Roorkee for their support, guidance and timely help for my research.

I would like to give warm compression of thanks to **Dr. Kanwarpal S. Dhugga,** and **Dr. Rajeev Gupta,** Scientists, Crop Genetics and Research and Development, Pioneer Hi-Bred International, Johnston, USA for their valuable help and guidance. Without their contribution the completion of this thesis would have been impossible. I express my deep sense of gratitude to them for sparing

time from their busy schedule. I would also like to thank **Pioneer Hi-Bred International Company** for the materials provided for my research work.

It gives me immense pleasure to acknowledge my heartiest thanks to **Dr. Kulvinder S. Gill** , Professor & Vogel Endowed Chair in Wheat Breeding and Genetics, Washington State University, Pullman for his valuable guidance and constant support. I would like to thank him for sparing time from his busy schedule.

I would also like to thank **Dr. Anupam Dixit** for providing his lab facilities and guidance for the completion of my work. I gratefully acknowledge his help and support.

I would like to thank **Dr. D. Kumar**, CAZRI, Jodhpur, **Dr. Sukhpal Singh Choudary**, Jaipur, Head, Forage section, CCS HAU, Hisar and NBPGR, New Delhi for providing guar accessions. I also express my gratitude for the support by **Dr. Rakesh pathak**, CAZRI, Jodhpur, **Dr. Surender Pahuja** and **Dr. Anshul Bajaj**, CCS HAU, Hissar.

I would like to take this opportunity to extend my gratitude to my seniors **Dr. Durga Prasad Panigrahi, Dr. Vijay Tiwari, Dr. Megha Agrawal, Dr. Nidhi Rawat, Dr. Kumari Neelam, Ms. Naincy Girdharwal and Ms. Pranita Bhatele Tiwari** for their timely advices, encouragement, guidance and goodwill which enabled me to start and successfully complete my work.

I express my thanks to my friends **Shailender Kumar Verma, Satish Khatkar, Anjali Chaturvedi, Priyanka Paul, Rajbala Yadav, Satya Tapas, Pradeep Kumar T. P., Prashant A.S., Govardhan and Keshv** for supporting me at every stage of my work.

I would like to express profound thanks and make a special mention of all my lab mates **Umesh Kumar Tanwar, Manisha Choudhary, Swati Verma, Shilpi Kumari, Shalini Pareek, Pallavi Gahlot, Navneet Kaur Sekhon** and **Deepa Dewan** for assisting me at every stage of work and maintaining a friendly environment in the lab.

I express my thanks to **Rakesh Kumar Gowswami, Yashodhara, Manjari, Vibhor Goyal, Gurdeep, Reema, Anita, Aakriti Tayal** who were the project students in lab for the scientific discussions and friendly interaction during their project period.

I express my gratitude to **Sanjay, Praveen, Pradeep and Mohsin** who have assisted in travel, material collection and growing of plants at IIT Roorkee.

I also express my gratitude to **Shashi Prabha, Saini, Jain** and **Sandeep**, the non teaching staff at the department, for helping me in processing all the documents and also for their supportive nature.

I express my gratitude to my previous teachers **Prof. N. Satyamurthy, Dr. Geetha Samak, Prof. M. R. N. Murthy, Ms. Rashimi, Ms. Pallavi, Prof. K. R. Sridhar, Prof. Krishnamurthy, Mr. Asif, Ms. Rochell, Dr. Anu Appaya, Mrs. Vanajakshi** who inspired me to be in the field of science and research.

I express my gratitude to **Department of Biotechnology** (DBT), Government of India for the financial support in the form of JRF and SRF during the period of this research.

I express my unbound gratitude to **my parents Mr. K. Aswathnarayana and Mrs. K. Sumitra** for their blessings and dedicated efforts to educate me to this level. I also express my gratitude to all my family members for their constant support and encouragement.

Finally I thank the Almighty for leading me all the way towards successful completion of this work.

K. A. Nagesh

Date : 23 - 12 - 2011 **(Nagesh K. A.)**

# Table of Contents

# LIST OF ABBREVIATIONS USED

| | |
|---|---|
| °C | Degree centigrade |
| µl | Microlitre |
| µm | Micrometer |
| AFLP | Amplified fragment length polymorphism |
| $AgNO_3$ | Silver nitrate |
| am | Ante meridiem |
| AMOVA | Analysis of Molecular Variance |
| Ba | Barium |
| bp | Base pair |
| CAPS | Cleaved amplified polymorphic sequences |
| CAZRI | Central Arid Zone Research Institute |
| CCSHAU | Chaudhary Charan Singh Haryana Agriculture University |
| cDNA | Complementary DNA |
| cm | Centimeter |
| CMHPG | O-carboxymethyl-O-hydroxypropyl guar gum |
| CMHTPG | O-carboxymethyl-O-2-hydroxy-3-(trimethylammonio) propyl guar |
| Csl | Cellulose synthase like |
| Cu | Copper |
| DAF | Days after flowering |
| DaRT | Diversity array technology |
| dbEST | Database EST |
| DNA | Deoxyribose nucleic acid |
| dNTPs | Deoxy nucleotide tri phosphates |
| DP | Dominant parent |
| EC no. | European commission number |
| EDTA | Ethylenediaminetetracetic acid |
| eg. | For example |
| EMBL | European Molecular Biology Lab |
| ESNP | Electronic SNP |
| EST | Expressed sequence tags |
| et al. | et alia |
| F1 | Filial 1 |
| Fru | Fructose |
| g | Gram |
| G/Gal | Galactose |
| GB | Giga bases |
| GDP-MP | GDP-mannose pyrophosphorylase |
| GDP-MS | GDP mannan dependent-mannosyl transferase |
| GenALEx | Genetic AnaLysis in Excel |
| Glu | Glucose |
| GMGT | Galactomannan galactosyltransferase |
| GST | Mean coefficient of gene differentiation |
| GT | Glycosyl transferase |

| | |
|---|---|
| GT- UDP | UDP galactose dependent-galactosyl transferase |
| GUS | β-glucuronidase |
| HTPG | O-2-hydroxy-3-(trimethylammonio) propyl guar gum |
| HXK | Hexokinase |
| IARI | Indian Agricultural Research Institute |
| IITR | Indian Institute of Technology Roorkee |
| ISSR | Inter simple sequence repeats |
| KB | Kilo bases |
| kDa | Kilo Daltons |
| L. | Linnaeus |
| m | Meter |
| M/G ratio | Mannose/Galactose ratio |
| M/Man | Mannose |
| ManS | Mannan synthase |
| mg | Milligram |
| $MgCl_2$ | Magnesium chloride |
| min | Minute |
| MISA | MIcroSAtellite |
| ml | Millilitre |
| mm | Milli meter |
| mM | Millimolar |
| mRNA | messenger RNA |
| MS | Mannan synthase |
| NBPGR | National Bureau of Plant Genetic Resources |
| NCBI | National Center for Biotechnology Information |
| ng | Nanogram |
| nm | Nanometer |
| PAST | PAleontological STatistics |
| PCR | Polymerase chain reaction |
| PHGG | Partially hydrolysed guar gum |
| PIC | Polymorphic information content |
| pm | Post meridiem |
| pM | Picomolar |
| PMI | Phosphomanno isomerase |
| PMM | Phosphomanno mutase |
| POPGENE | POPulation GENEtic analysis |
| p-value | Probability value |
| R | puRine |
| R value | Coefficient of correlation |
| RAPD | Random amplified polymorphic DNA |
| RB | Recessive bulk |
| RNA | Ribose nucleic acid |
| RNaseA | RibonucleaseA |
| Rp | Resolving power |
| RP | Recessive parent |
| rpm | Revolutions per minute |
| Rs | Rupees |

| | |
|---|---|
| SCAR | Sequence characterized amplified regions |
| sec | Second |
| SNP | Single nucleotide polymorphism |
| SS | Sucrose synthase |
| SSCP | Single stranded conformational polymorphism |
| SSR | Simple sequence repeats |
| SSRIT | Simple Sequence Repeat Identification Tool |
| Taq | *Thermus aquaticus* |
| TBE | Tris borate EDTA |
| $TE/T_{10}E_1$ | Tris EDTA |
| TRA | Tandem repeats analyser |
| TROLL | Tandem Repeat Occurrence Locator |
| U | Unit |
| UDP-GE | UDP- galactose 4-epimerase |
| UPGMA | Unweighted Pair-Group Average |
| USA | United States of America |
| UV | Ultraviolet |
| V | Volt |
| w/v | Weight/volume |
| Y | pYrimidine |

## LIST OF TABLES

## LIST OF FIGURES

# *Chapter I*

# *Introduction*

## 1. Introduction

Cluster bean (*Cyamopsis tetragonoloba* [L.] Taub.), commonly known as guar, is being grown in India since ancient times. It has been useful as fodder, vegetable and green manure in the past; in the recent times guar has become a major industrial crop due to the need of the guar gum contained in the endosperm of its seeds [36]. Guar is a drought hardy, deep rooted, annual legume grown in Kharif season. The crop is mainly grown in the dry habitats of Rajasthan, Haryana, Gujarat, and Punjab and to a limited extent in Uttar Pradesh and Madhya Pradesh. In addition to India, guar is also grown in Pakistan and as a cash crop to a limited extent in other parts of the world like South Africa, Brazil, Australia and Oklahoma planes of North Texas in USA. Guar is generally 50-100 cm tall and bears 4-10 branches (branched types); however, unbranched type varieties having main stem only are also available [77]. The flowers are borne in axillary raceme on long pedicels. The flowers are small, typically papilionaceaous. The pods are oblong, 5-12 cm in length, normally contain 5-12 oval or cube shaped seeds. The crop matures in 90-150 days depending on the genotype. It prefers warm climates and can also be grown in subtropical areas during summer. Guar is known for its exceptionally high adaptation towards poor rainfall conditions, low inputs and less care, soil enrichment properties and multiple uses in cropping systems.

Guar seed consists of seed coat (14-17%), endosperm (35-42%) and germ (43-47%). The endosperm contains about 80% galactomannan, 12% water, 5% protein, 2% acid insoluble ash, 0.7% ash and 0.7% fat [107]. Galactomannan is a water soluble polysaccharide made up of linear chains of a β-1,4-mannan as backbone to which galactosyl residues are attached through an α-1,6 linkage [130]. Guar gum is used as a binder, disintegrant, suspending agent, thickening agent and stabilizing agent. It is soluble in cold water, hydrating quickly to produce viscous pseudo plastic solutions that generally have greater low-shear viscosity than other hydrocolloids [107]. These properties have made guar gum useful in various industries like paper, textile, petroleum, drilling, pharmaceutics, food, cosmaceutics and explosives [140].

Guar gum is used as an edible thickening agent and has cholesterol lowering effect. The cholesterol and glucose lowering effects are most often associated with gelling, mucilaginous, viscous and fibers properties of guar gum. The demand for guar galactomannan is growing rapidly because, in addition to its indispensable role in lowering serum cholesterol and glucose levels, it is also considered helpful in weight loss [20]. Dietary intake of high viscosity of guar gum decreases the protein efficacy [105]. Guar gum also decreases lipid utilization by interfering with digestion and absorption of nutrients when it is dissolved in water. This results

in a slower gastric emptying [156]. Partially hydrolyzed guar gum (PHGG) is a water-soluble, non-gelling fiber that has therapeutic benefits. It is used in the treatment of irritable bowel syndrome [47].

Guar gum is a very important foreign exchange earner for India. For instance, guar gum export reached 183.57 thousand tons providing foreign exchange worth Rs. 10384 million during the year 2004-05. In 2005-06, 189.11 thousand ton guar gum costing Rs. 11249 million was exported to western countries. Thus the demand for guar gum is rising in the global market and is becoming an opportunity for the farmers and the industrialists as well [76].

Guar gum is synthesized in the endosperm part of the developing seeds. In the recent years many efforts towards the understanding of molecular mechanism of galactomannan synthesis have shown that galactomannans are synthesized in Golgi lumen by the combined action of two enzymes: mannan synthase (ManS), which makes β-1,4-linked mannan backbone, and α-galactosyltransferase, which adds galactosyl residues to the mannan backbone [32]. Several genes from cellulose synthase like (Csl) gene family have been found to be involved in synthesis of galactomannan [130]. The stored galactomannans are utilized for plant development by galactomannan degrading enzymes such as α-galactosidase which hydrolyses the galactose side chain from mannose backbone and β-mannanase which cleaves the mannose polymer to oligo mannans and β-mannosidase which hydrolyses the oligomannans. The enzymes are synthesized during the seed development to nourish the developing seedling [95].

Apart from guar many other leguminous plants also produce galactomannans in their seeds. Naturally in legume seeds galactomannans perform energy storage, water-retaining and defense function [132]. The degree of galactose (G) addition to the mannan (M) backbone differ depending upon the source of the gum which is represented by M/G ratio; like M/G ratio for guar gum is 1.3-2 and for locust bean gum is 3.75-4. The change in M/G ratio changes the structure of the gum thereby changing the various properties and applications of the galactomannans [140]. Higher viscosity of the gum is essential for use in oil drilling and other industries. The locust bean gum is better in terms of its viscosity characteristics but it is difficult to obtain since carob tree flowers only after 10 years of vegetative growth. Moreover these trees are grown in Mediterranean climates. The cost of locust bean gums are high in comparison to guar gum because of the long maturation period of the plant, labour-intensive harvesting of fruits and competition from other cash crops [19]. Many other legumes have galactomannans in their seeds, but the extraction of gum from them is not economical because

of the low yield in comparison to guar and locust bean. To improve the characteristics of guar gum various chemical modifications are employed like O-carboxymethyl-O-hydroxypropyl guar gum (CMHPG), with anionic character, O-2-hydroxy-3-(trimethylammonio) propyl guar gum (HTPG), with cationic character, and O-carboxymethyl-O-2-hydroxy-3-(trimethylammonio) propyl guar (CMHTPG), with amphoteric character [160], polyacrylamide grafted guar gum [139] and sulfated guar gum [151]. These modifications require higher input and lead to higher cost of the modified guar gum. Even then modified guar gum has found various uses in capsule preparation for controlled release of drugs and oil drilling. Biotechnological approaches to produce guar gum with wide range of M/G ratio has been reported through application of α-galactosidase in the preprocessing steps to remove excess of galactose [19]. An effort on improving the M/G ratio of galactomannan through transformation and expression of α-galactosidase gene of senna in developing guar seed has shown slight increase in M/G ratio [64]. Yet there are no reports on any commercial operations based on these processes. Hence, improving the yield and M/G ratio of guar gum could be highly to the guar growers and the industries which use these galactomannans.

Even though guar gum synthesis is understood to a certain extent, guar is still considered a neglected crop due to the limitation of the genomic resources and lack of the understanding of its genomics. Cluster bean belongs to the genus *Cyamopsis* which has three sub species *C. tetragonoloba, C. serrata, C. senegalensis.* Among the species *C. tetragonoloba* being a cultivated crop, is not found in wild conditions [58], which makes the available landraces as the main source of genetic variability. Apart from variability, landraces are also known for adaptations to the selective environments in the region of cultivation [23]. Conventionally guar accessions have been characterized on the basis of their phenotypes [91]. In India, National Bureau of Plant Genetic Resources (NBPGR), New Delhi is working on the collection, maintenance and characterization of the guar landraces. A total of 4901 accessions are available with NBPGR which have been characterized for various phenotypic traits like pubescence, days to 50% flowering, days to 50% maturity, plant height, number of branches, number of clusters per plant, seed yield per plant, gum content, etc. [30]. But the limitation of this approach lies with the phenotype being influenced by the environmental factors during the developing stage of plant. This available genetic resource is yet to be utilized to transfer useful traits to elite cultivars of guar. Hence, there is a need to study these landraces for genetic diversity based on the geographic origin and phenotype. Only few studies towards studying the diversity of commercially grown varieties using RAPD have been carried out [100, 109]. An

overview of the genetic diversity is essential for any crop improvement programme. Hence the study of the genetic diversity available in the landraces of guar is very essential for selecting the source of genetic variation which could be applied for production of cultivars with various beneficial traits.

Guar is a self-pollinated diploid legume with chromosome number (2n) equal to14. The genome size of guar is estimated to be 2.45 Giga Bases/C [101]. Guar due to the cleistogamous nature is strictly self pollinated, but outcrossing, to the extent of 0.5% to 9% has been reported [48]. Thus, lesser heterosis is available. The optimal time for pollinating the buds is between 8:00 a.m-9:00 a.m. [26]. Generally the pods formed after hybridization are relatively smaller having 2-3 seeds. This makes commercial hybrid seed production non economical in guar. This limiting factor of yield gap can be overcome by production of improved varieties of guar through modern breeding programme; this requires the application of modern genomic approaches to increase the speed and efficiency.

The DNA-based molecular markers reveal natural variation at the DNA sequence level; these markers are used in plant genotyping, diversity studies, genetic linkage studies, quantitative trait mapping and marker-assisted selection during plant breeding [69]. DNA-based molecular markers offer numerous advantages over conventional phenotype based alternatives because they are stable and can be detected in all tissues regardless of growth, differentiation, development, or defense status of the plant. They are not confounded by the environment, pleiotropic and epistatic effects [1]. Molecular markers do not have any biological effect; they are identifiable DNA sequences, found at specific locations of the genome and are transmitted by the standard laws of inheritance from one generation to the next. They can be considered as constant landmarks in the genome. The existence of various molecular techniques and differences in their principles and methodologies require careful consideration in choosing one or more of such marker types. The diversity and implications of the markers have increased in the recent times with the discovery of various PCR based marker types like RAPD, ISSR, SSR, SCAR, SSCP, AFLP, SNP, DaRT, etc. These marker techniques differ from each other with respect to important features such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and cost. One of the widely used co-dominant marker for studying genetic linkage and marker-assisted selection is Simple Sequence Repeat (SSR) marker. SSRs have been developed and successfully applied for various applications in crop plants. This marker system has not been applied to guar. Very less attention has been paid to development of molecular markers in guar.

With the above information in view, the current research work on application of the molecular markers to understand the genetic diversity in guar landraces and utilization of the available DNA sequence resources to develop markers for application in guar was undertaken

The objectives of this work were as follows:-

1. to study the genetic variability in cluster bean accessions through application of RAPD and ISSR markers,
2. to screen publically available Expressed Sequence Tags of guar for Simple Sequence Repeat (SSR) regions and design primers flanking SSR regions,
3. to synthesize and test EST-SSR markers for amplification and polymorphism studies among guar accessions,
4. to cross guar genotypes for developing a population to study the inheritance of phenotypic traits and genetic linkage,
5. to conduct bulk segregant analysis for phenotypic traits and study the linkage of the markers with specific traits.

# Chapter II

# Review of Literature

## 2. Review of Literature

Guar, also known as cluster bean (*Cyamopsis tetragonoloba* (L.) Taub.), is being grown in India since ancient times. Earlier, it was used mainly as a forage, green manure and vegetable crop, but now it is an important industrial crop due to galactomannan gum contained in the endosperm of its seeds [36]. Cluster bean is mainly grown in semi-arid regions of Rajasthan, Haryana, Gujarat, Punjab, Madhya Pradesh and Uttar Pradesh in India. It is also cultivated to some extent in Pakistan and has been introduced as a cash crop in Oklahoma and Texas states of USA. It is also cultivated to a limited extent in Italy, Morocco, Spain, France, Greece, and Germany [109].

The name guar comes from the Sanskrit word 'gau ahaar' which means the fodder of cow. Guar is also known as gawaar in Hindi and Marathi, goruchikkudu kaya /orgokarakaya in Telugu, gorikayie/javalikayie in Kannada, and kotthavarai in Tamil [58]. Guar belongs to the family Fabaceae. Earlier *Cyamopsis tetragonoloba* was known as *Dolichos fabaeformis* or *C. psoralioides*. Guar is a self-pollinated diploid (2n = 14) plant. It is generally 50-100 cm tall and bears 4-10 branches (branched types); however, unbranched type varieties are also available. Guar plant may produce around 30-90 pods per plant. The leaves are alternate trifoliate, which are born on long petioles. The stem is slender and tall [77]. The flowers are purple, pink to white in colour, typically papilionaceous and form axillary raceme inflorescence. The calyx has five unequal linear teeth type sepals. The corolla has orbicular standard wing petals, which are oblong, while the keel petals are long and broad as the wings. All the ten stamens are fertile and the filaments form a tube while the anthers are apiculate. The pollen grains are circular and 40–43 μm in diameter. The stigma expands into a head-shape while the style is short and slender [91]. The pods are oblong, 5-12 cm in length and normally contain 5-12 oval or cube shaped seeds of variable shape [104] as shown in Fig 2.1.

Cluster bean prefers warm climates and can also be grown in semi-arid conditions in subtropical areas. It is known for its exceptionally high adaptation towards poor rains, low inputs and less care, soil enrichment properties and multiple uses in cropping systems. Guar has high transpiring ability leading to the cooling of canopy and increasing the soil moisture extraction from deep layers [77].

Green pods of guar are a very good source of vitamin A, calcium, iron, phosphorous and ascorbic acid. It is a rich source of protein (3.2%), fiber (3.2%) and carbohydrate (10.8%).

Guar pods are also used in traditional medicine for controlling constipation, anorexia, diabetes, arthritis, colic pain, hair fall and body pain [49].

## 2.1 History of guar

Cluster bean is considered to be originated by domestication of the African wild species, *C. senegalensis* which appear to be the ancestor of the *C. tetragonoloba*. The domestication process could have been taken place in the dry areas of the northwestern region of the Indo-Pakistan Subcontinent [58]. It was cultivated as a minor crop in India during ancient times as a vegetable and feed for cattle.

Guar was introduced in USA in 1903 for experimentation in the southwest region, where the climate is hot and have long growing seasons to suit its adaptation [59]. The objective was to use guar as a soil improving legume and forage for cattle. Before world war II the carob (locust bean) seed (*Ceratonia siliqua*) from Mediterranean was used to extract carob gum for extensive use in paper industry. During World War II the supply of imported carob seed from the Mediterranean region was cut off, as a result search for domestic source of galactomannan gum was initiated in USA by Institute of Paper Chemistry. This study revealed guar as a alternative source for galactomannan [5]. Further studies were done on milling the guar seeds for gum production and application in the manufacturing of paper revealed the beneficial effects of guar gum on paper processing. This information helped in the adoption of guar gum in the different manufacturing process.

## 2.2 Germplasm

Germplasm is a basic tool for crop any improvement programme. Natural variation represents a huge and largely untapped resource, which has been subjected to selection over millions of years of evolution, with both basic and practical value, as well as the potential to break yield barriers of agricultural plants [66, 143, 157] The variation available in the germplasm is utilized as a source of useful genes to improve the cultivars. As the importance of guar was realized in 1950's germplasm collection was initiated. The collection initially began from Maharashtra for vegetable varieties. Later the Pant Introduction Division of the Indian Agricultural Research Institute (IARI), New Delhi continued the collection and maintenance work. There are about 4,901 accessions in National Bureau of Plant Genetic Resources (NBPGR), New Delhi. These accessions have been catalogued based on the accession numbers

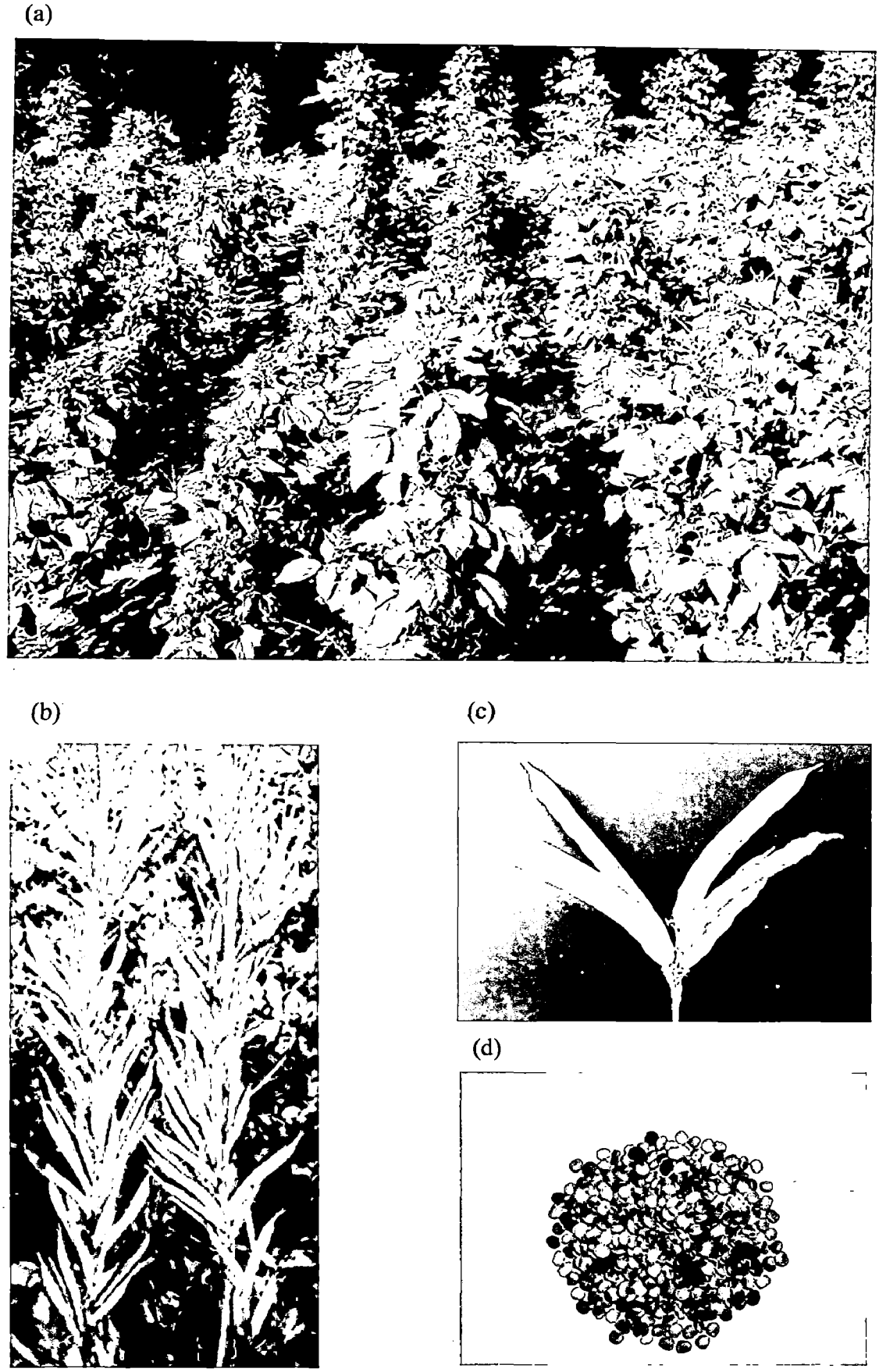



**Fig. 2.1. a) Guar plants grown in a field at IIT Roorkee, b) Green pods of guar c) Dried pods of guar, d) Seeds of guar**

and characterized for phenotypic traits like pubescence, days to 50% flowering, days to 50% maturity, plant height, branch number, total number of pods per plant, number of seeds per pod, seed yield per plant, seed color, gum content and disease resistance under field conditions [30]. These studies have identified the accessions that can be donors to traits like dwarfing, branching/unbranched, pod length, pods per plant, seed size, days to maturity, gum content and disease resistance. Classical approaches using the donor cultivars as sources have lead to the development of certain elite cultivars for cultivation in the previous decade. The development of elite cultivars and their widespread use in breeding programmes have reduced the utilization of available genetic resources. It has been estimated that for most crop species, less than 5% of the biodiversity known to exist has been utilized in agriculture, particularly in the case of self-pollinated crops [143]. Much of the diversity present in living systems is probably adaptive [66].

## 2.3 Breeding and Genetics

Crossing leads to the hybridization of DNA from two plants with different genetic makeups but belonging to the same genus. It is difficult with small flowers or less consistent pollen production. As in guar the flower is only 8 mm long and requires magnifying lens to emasculate the flowers [48]. Once the anthesis begins ten anthers can be seen encircling the stigma. The pollen is viable from two hours before to eleven hours after anthesis [141]. The flower morphology leads to self pollination and hence considered cleistogamous and only exhibit outcrossing to an extent of 9% [48]. Chaudhary et al. in 1974 developed a new crossing technique which had a success rate of 7%. The method was based on manual emasculation of anthers and pollination by bringing the mature anthers in contact of the stigma. The optimal time for pollinating the buds is between 8:00 a.m - 9:00 a.m. [26]. Generally the pods formed after hybridization are relatively smaller with 2-3 seeds in each of the pod. To ease the problem of manual breeding, male sterile lines of guar are necessary. There are some reports on nuclear sterility and partial male sterile system [142]. Recently other methods like caging have been applied in guar to speed up the process of hybrid production [48]. However, very little success has been achieved in improving efficiency in making crosses in guar.

Mutation breeding is considered as a useful tool to enrich the variation in crops where useful genetic variability is meager [6]. Many mutants with useful traits like early flowering, increased yields and gum content have been produced by this approach [137].

Through these procedures of hybridization plant breeders can manipulate the alleles to obtain recombination for various genetic traits. Hybridization allows conducting genetic studies on control of phenotypes and their inheritance pattern. The resulting population allows prediction of quantitative traits, epistasis and transgressive control of phenotypes.

The inheritance of five characteristics namely branching behavior, clustering pattern, growth habit, leaf size and hairiness have been studied in cluster bean. The results of the study showed that all the traits were controlled by single pair of genes except branching behavior which exhibited digenic inheritance. The study also showed that the alleles governing branching, discontinuous clustering, indeterminate growth habit, small leaf size and hairiness were dominant over the alleles controlling non-branching, continuous clustering, determinate growth habit, broad leaf and glabrous leaf and stem [25].

## 2.4 Galactomannans

Galactomannans are heterogeneous polysaccharides widely distributed in nature. These belong to the hemicellulosic cell wall component of the plant cell. These galactomannans assume the role of storage polysaccharides in seeds of some plants, which is analogous to that of starch in cereal grains [32]. Galactomannans also perform water-retaining and defense functions in plants [132]. Apart from plants galactomannans are also found in microbial sources such as yeast and fungi [35, 82]. Galactomannan is accumulated in the form of secondary wall thickenings in the endosperm of guar (*C. tetragonoloba*) and locust bean/carob (*Ceratonia siliqua*) which are used for commercial production of galactomannan gum. Other plants like tera (*Caesalpinia spinosa*) [148], fenugreek (*Trigonella foenum-graecum*) [17], mesquite (*Prosopis pallida*) [15, 16] and coconut palm (*Cocus nucifera*) also contain galactomannan in their seeds.

### 2.4.1 Structure

Galactomannans are natural nonionic high molecular weight (100–1000 kDa) polymers. Galactomannan is made of linear chains of a β-1,4-mannan as backbone to which galactosyl residues are attached through α-1,6 linkages [151]. The degree of galactose (G) substitution to the mannan (M) backbone differs depending upon the source of the gum which is represented by M/G ratio; like M/G ratio for guar gum is 1.3-2 and for locust bean gum is 3.75-4. The change in M/G ratio changes the structure of the gum thereby determining various properties

and applications of the galactomannans [140]. The Fig 2.2 shows the general structure of galactomannans.

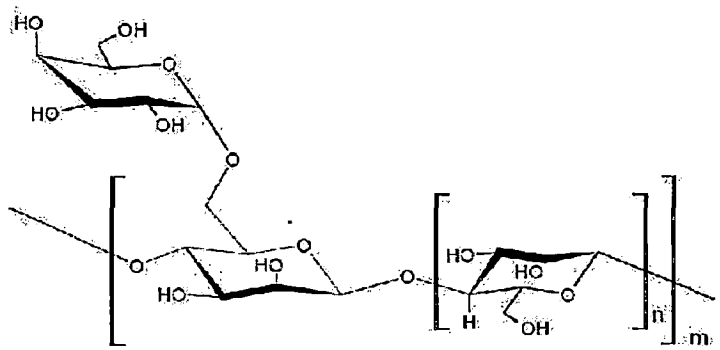


**Fig. 2.2. General structure of galactomannans [148], n indicates the number of unsubstituted mannose, m indicates the degree of polymerization**

Homo-polymer made of long chains of mannose is insoluble in cold water and the increase in galactosylation leads to an increase in solubility [18]. Locust bean gum with M/G ratio of 4 is considered superior in its viscosity characteristics in comparison to guar gum. Even though superior in quality it is difficult to produce [19]. Individual galactomannans differ from each other in distribution pattern of the galactose residues along the main chain. It is found that in guar galactomannan the distribution is blockwise whereas locust bean gum may contain random, blockwise, and ordered distributions [103]. Fig 2.3 shows different kinds of distribution of galactose residues in galactomannans. In a recent study the degree of substitution and the degree of polymerization of galactomannan was studied. Unmodified guar galactomannan was found to contain aggregates, but partial removal of galactose side units significantly decreased this tendency. The chain length seemed to dictate the solubility of the molecules together with the degree of substitution.

Regular

M – M – M – M – M – M – M – M – M – M – M – M – M – M – M – M

G G G G G G G G

Random

M – M – M – M – M – M – M – M – M – M – M – M – M – M – M – M

G G G G G

Blockwise

M – M – M – M – M – M – M – M – M – M – M – M – M – M – M – M

G G G G G G G G

**Fig. 2.3. Arrangement of the galactosyl side chains in galactomannan [148]**

The sample with a relatively low degree of polymerization and low degree of substitution formed large assemblies but stayed in solution while the solubility of the samples with longer chain length was reduced [103].

### 2.4.2 Galactomannan biosynthesis in plants

Galactomannans are synthesized in endosperm cells by the combined action of two enzymes, mannan synthase (ManS) which makes β-1, 4-linked mannan backbone, and α-galactosyltransferase, which adds galactosyl residues to the mannan backbone [32, 39, 116]. Mannan synthase, the enzyme that forms the β-1,4-linked backbone of galactomannan, assayed *in vitro* using membrane particles derived from developing endosperm, shows independent activity leading to mannose polymer [38]. It has a high affinity for the substrate, GDP-mannose, and requires divalent cations for activity [130]. But galactosyltransferase has been found to be dependent on mannan synthase for activity; it shows no activity in presence of the substrate UDP-galctose alone.

*In vitro* studies have shown that preformed long mannan chains cannot be galactosylated by the enzyme. However the combined action of both mannan synthase and galactosyl tranferase has produced polymers similar to galactomannan [38]. They have proposed a model for the interaction of GDP mannose mannosyltransferase and UDP-galactose galactosyltransferase in galactomannan biosynthesis [38]. The study on galactosyltransferase from fenugreek has shown that the enzyme acts on D-manno-oligosaccharides of length greater than or equal to 5 acts as acceptor. But the D-manno-oligosaccharides longer than 9 monomers require heating for dissolving [41].

The degree of galactosylation of natural galactomannans is also believed to be determined by two methods. One in which the control is at biosynthesis level i.e. in guar. Second kind of regulation involves the alteration of galactomannan formed by the action of α-galactosidase later in seed development [39] like that in locust bean gum. However both the mechanisms of mannose:galactose ratio control in plants are genetic [115]. The pathway for biosynthesis of guar galactomannans is shown in Fig 2.4.

The direct precursors for galactomannan biosynthesis, GDP-D-mannose and UDP-D-galactose, are formed by the actions of GDP mannose phosphorylase (EC 2.7.7.22) and UDP-galactose 4-epimerase (EC 5.1.3.2). *In vitro* experiments have shown that the relative concentrations of these precursors also effect the M:G ratio of the galactomannan [39].

The biodegradative hydrolysis of galactomannans during seed development requires the presence of three enzymes α-galactosidase, β-mannanase and β-mannosidase. α -galactosidase hydrolyses the galactose side chain from mannose backbone, β-mannanase cleaves the mannose polymer to oligo mannose and β-mannosidase hydrolyses the oligomannanans. These enzymes are synthesized during the seed development to nourish the developing seedling [95].

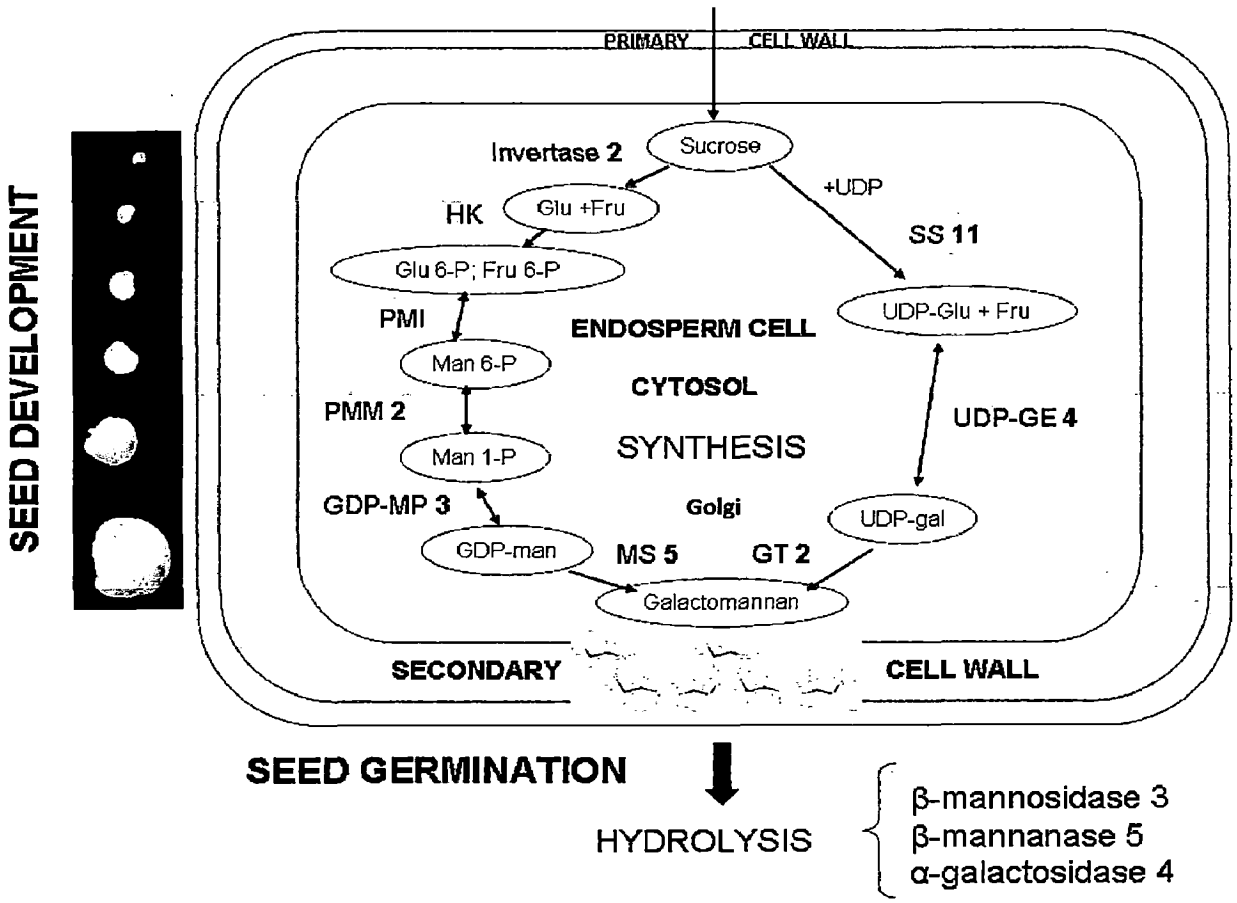



**Fig. 2.4. Schematic representation of galactomannan metabolism in guar seeds [95]**

Abbreviations: Glu, glucose; Fru, fructose; Man, mannose; Gal, galactose; HXK, hexokinase; PMI, phosphomanno-isomerase; PMM, phosphomanno-mutase; GDP-MP, GDP-mannose pyrophosphorylase; MS, GDPman- dependent mannosyl-transferase; GT, UDP-gal-dependent galactosyl transferase; SS, sucrose synthase; UDP-GE, UDPgalactose 4-epimerase. .

### 2.4.3 Genes in galactomannan biosynthesis

Identification of genes involved in the biosynthetic pathway of cell wall polysaccharides has been considered difficult till recent times because of various reasons. Plant cell wall synthesizing enzymes are integral membrane proteins with one or few transmembrane domains. They are quite challenging for biochemical studies as they tend to be labile, are

present in multimeric complexes and are encoded by large gene families whose members may have overlapping function [32].

UDP-galactose epimerase catalyzes the reversible conversion of UDP-D-glucose to UDP-D-galactose. This compound is a precursor for the biosynthesis of various cell wall polymers, including galactomannan[65]. Two cDNA clones encoding two different UDPG epimerases were isolated. Using functional complementation of a UDP-galactose epimerase deficient *E. coli* mutant by a cDNA expression library from immature guar seeds, galactose metabolising colonies with UDPG epimerase activities comparable to wild type level were obtained [65].

Success in identifying cDNA encoding mannan synthase (ManS) was achieved through the identification of a candidate gene via transcriptional profiling followed by functional expression in a heterologous system such as soybean somatic embryos. Further analysis showed the gene belongs to cellulose synthase like (*Csl*) gene family [32, 130]. The cellulose synthase-like (*Csl*) genes were first identified in the model plant Arabidopsis [119] and rice [56] which are grouped under the family based on certain sequence homology [29].

The enzyme responsible for the transfer of galactose (Gal) residues to a mannose (Man) on the mannan backbone is a member of the glycosyl transferase (GT) family of proteins called galactomannan galactosyltransferase (GMGT). The putative cDNA encoding galactosyl transferase from fenugreek seed was cloned and expressed in yeast *Pichia pastoris*. The expressed galactosyl transferase showed transfer of galactose to D-manno-oligosaccarides with chain length 5 or more [41]. The mRNA sequence for galactosyl transferase form guar is available in NCBI database with sequence id AJ938067.1 (gmgt1 gene).

The mannan synthase (ManS) and galactomannan galactosyl transferase (GMGT) enzymes are localized in the membrane of Golgi vesicles and are believed to work together very closely to determine the statistical distribution of galactosyl residues along the mannan chain [40].

### 2.4.4 Genes for biodegradative hydrolysis of galactomannans in guar

The gene encoding α-galactosidase, an enzyme that hydrolyses the galactose side chain from mannose backbone, was identified using oligo-nucleotide mixed probes based on the terminal amino acid sequence and the sequence of an internal peptide. The nucleotide sequence of the cDNA clone showed that the enzyme is synthesized in the form of a precursor

with a 47 amino acid $NH_2$ terminal extension. This pre-sequence mainly functions to target the protein outside the aleurone cells into the endosperm [99].

## 2.5 EST sequences in guar

A database of 16,476 guar seed ESTs was constructed from two cDNA libraries consisting of 8,163 and 8,313 ESTs sequences. A library constructed from seeds at an early developmental stage (15–25 days after flowering, DAF) formed the first set of sequences, and a library from seeds at 30–40 DAF formed the second set. Approximately 27% of the clones were not similar to known sequences, suggesting that these ESTs represent novel genes or may represent non-coding RNA. Further analyses have shown that the high input of energy into carbohydrate and storage protein synthesis in guar seeds was reflected by a high representation of genes annotated as involved in signal transduction, carbohydrate metabolism, chaperone and proteolytic processes, and translation and ribosome structure. Among the seed storage proteins, the most abundant contig represented a conglutin accounting for 3.7% of the total ESTs from both libraries [95].

## 2.6 Efforts for obtaining optimal M/G ratio in galactomannans

There is high importance, application and commercial benefit to obtaining galactomannans with good viscosity characteristics. This has lead to various biotechnological applications to improve M/G ratio of galactomannans both *in vivo* and *in vitro*. Initial studies were carried out with the use of α-galactosidase for enzymatic removal of side-chain galactose residues from guar gum to yield galactomannans similar in chemical composition and functional properties to locust bean gum. Galactomannans with varying galactose content were prepared by manipulating reaction time, temperature and enzyme/guar gum ratio. Enzymatically modified guar galactomannans with 22-24% galactose contents were found to reproduce the rheological and stabilization properties of locust bean gum [19].

Later the available information on genes involved in synthesis and biodegradation of galacotmannans lead to transformation studies. Fenugreek seed galactomannan is almost fully substituted by galactose, whereas galactomannan in tobacco seed (*Nicotiana tabaccum*) contains very low level of galactose substitution. Fenugreek galactoamannan galactosyl transferase (GMGT) was expressed under a strong constitutive promoter in tobacco to modify galactomannan. The results showed production of galactomannans with significantly increased level of galactose substitution in transformed tobacco similar to fenugreek galactomannan

[117]. This study also showed that the exogenously introduced fenugreek GMGT dominated over the endogenous tobacco GMGT and can operate mutually with the endogenous mannan synthase in tobacco [117]. In another approach, the GMGT gene in *Lotus japonicus* was down-regulated by sense and antisense expression of GMGT. The *Lotus japonicus* plant has native galactomannan with high galactose substitution (Man/Gal ratio of 1.2-1.3). The transformation of galactomannan galactosyltranferase constructs, resulted in the modification of its galactomannan to one with a lower galactose content at a Man/Gal ratio of 6 [40]. This study clearly showed the potential of modifying low value galactomannan to a more valuable galactomannan producing plants through genetic engineering methods [40]. An obvious target of such engineering would be guar which is already used for production of galactomannan. In addition guar is transformable and is an annual crop, to evaluate the potential role of α-galactosidase for the control of the final galactose content. An α-galactosidase gene expressed in immature senna seeds was cloned and transformed into guar using a wheat high-molecular-weight glutenin promoter by *Agrobacterium tumefaciens*-mediated gene transfer. About 30% of the guar transformants produced endosperm with galactomannans where the galactose content was significantly reduced [64].

Development of suitable endosperm- specific promoters for use in guar is desirable for metabolic engineering of the seed gum. A ~1.6 kb guar mannan synthase (MS) promoter region has been cloned and characterized by studying the quantitative expression of β-glucuronidase (GUS) directed by MS promoter. GUS expressed specifically in endosperm of transgenic alfalfa [94]. Thus, the guar MS promoter could prove useful for directing endosperm-specific expression of transgenes in legume species. This promoter can be used for driving high expression of foreign genes in transgenic guar endosperm as well as in other species [94].

## 2.7 Process of guar gum production

Guar gum is obtained by grinding the seed endosperms and is sold as whitish flour. Fig 2.5 shows the manufacturing process of guar gum from seeds. Depending on the degree of purification, various amounts of other seed tissues, such as residues of germ, seed coat and endosperm cell walls may be present in the above obtained guar gum.

After harvesting the pods are dried in sun and then threshed mechanically so that the seeds are freed from the pods. The endosperm contains most of the pure galactomannan, and so it must be separated from the seed coat [148]. On commercial scale, seeds are roasted in a

furnace to loosen the seed coat, after heat treatment the hull is easy to separate by attrition milling or various other types of impact milling [72]. The separated endosperm usually sticks with the hull and has to be resubjected to milling process. Then, the endosperm is powdered by means of pulverisers to desirable mesh size. The pulverization step is carried out in humid condition and special care is taken to avoid degradation of high molecular weight polymers. The gum so obtained is called crude gum or commercial grade gum [148].

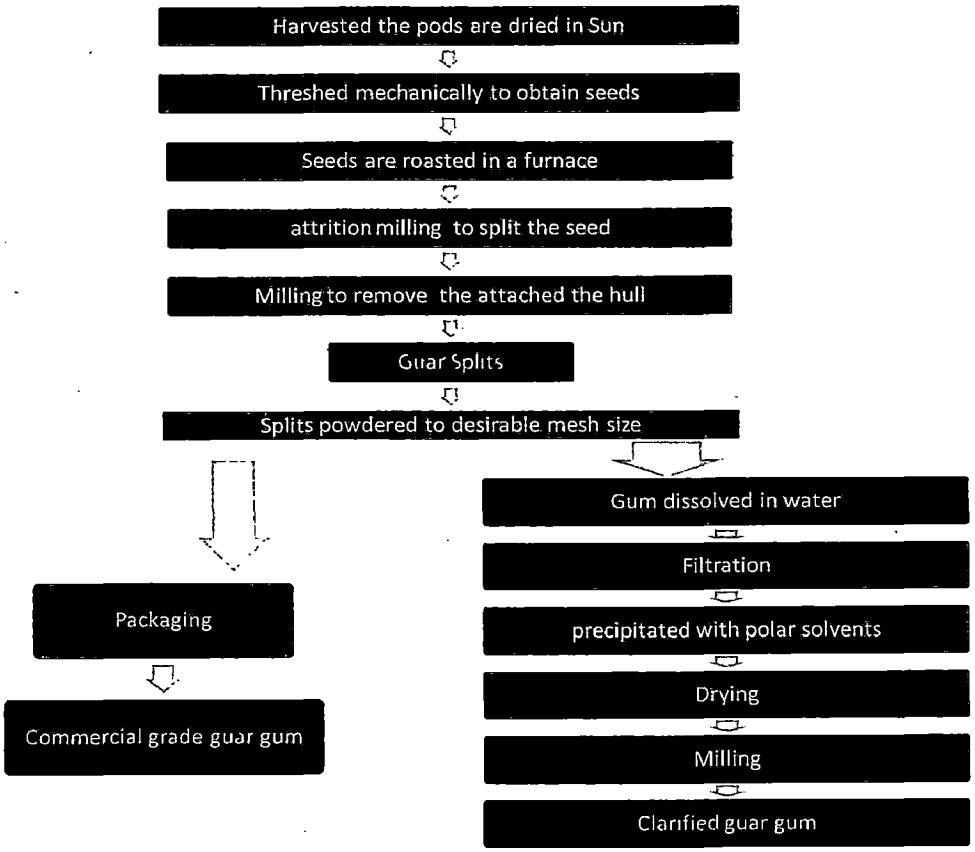



**Fig. 2.5. Flow chart showing the process of guar gum extraction from guar seeds [72]**

Crude gum obtained by milling process contains proteins, fibers, etc, as impurities. To determine physico-chemical properties the gum is dissolved in water and partially precipitated with polar solvents. Gum can also be purified by complexation with $Cu^{2+}$ and $Ba^{2+}$ salts or by dialysis for advanced studies [140].

## 2.8 Modifications

As galactomannans produce relatively high viscosity, they may not be suitable for certain applications; hence, depolymerized gum can be used to reduce the viscosity [156]. Some modifications make the gum less susceptible to microbial degradation [148] while certain

modifications increase the viscosity of the galactomannan [77]. Table 2.1 shows the applications of modified guar gum in various industries [77].

## 2.9 Applications of guar gum.

Guar gum has numerous applications. Fig 2.6 shows the applications of guar gum in different industries.

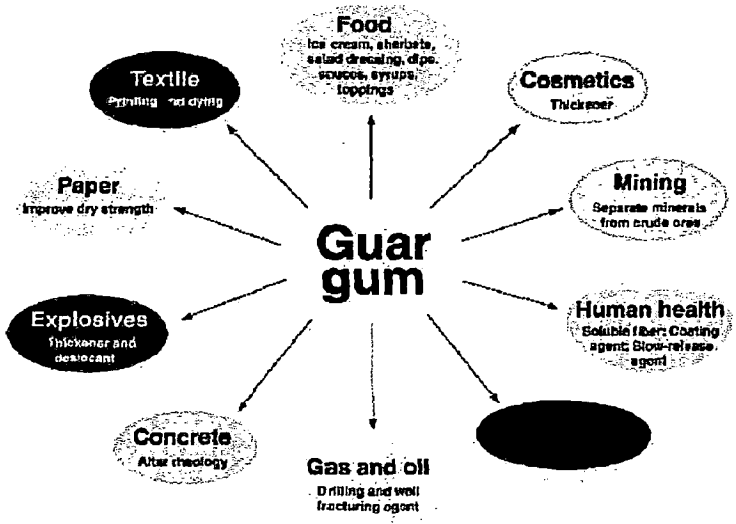



**Fig. 2.6. Applications of guar gum in different industries [31]**

**Table 2.1. Applications of guar gum and its modified forms in various industries [77]**

| Sr. No | Form of gum | Functions |
|---|---|---|
| 1 | **Explosives Industry** | |
| | Crosslinked guar gum | Water remover |
| | Guar gum | Plasticity improver |
| | Hydratable guar gum and self complexing guar gum | Gelling agent |
| | Nitrate ester of guar gum | Thickening agent |
| | Guar gum | Increases viscosity |
| | Cyanoethyl ether of guar gum | Thickening agent |
| | Guar gum with inorganic oxidizer salts | Thickening agent |
| | Guar gum with transition metal ions | Increases viscosity |
| 2 | **Petroleum Industry** | |
| | Sulfonated guar gum and $Me_3 N^+$ guar gum | Thickener |
| | Alkali refined guar gum crosslinked with borax | Gelling agent |
| | Hydroxy propyl guar gum | Thickening agent |
| | Guar gum and hydroxy ethyl cellulose | Gelling agent |
| | Cationic guar gum | Viscosity and thermally stable |
| | Guar gum with boric acid and magnesium-oxide | Provides viscosity stability |
| | Hydroxyl alkyl ether derivatives of guar gum | Reduces friction and increases permeability |
| | Guar gum with mineral salts and PEG | Improve water loss |
| | Hydroxy alkyl guar gum | Good fluid loss properties |

| | | |
|---|---|---|
| | Guar gum with methylene blue | Indicator of fluid life |
| | Guar gum with laminar silicate starch and swelling clay | For reduction of fluid loss |
| | Borate cross linked guar gum | Stable super-elastic liquid with lessened temperature sensitivity for control of lost circulation in oil field drilling operation |
| | Guar gum xanthomonas campestris mixture | For plugging leaks and stabilization of likages |
| 3 | **Textiles Industry** | |
| | Guar gum | Thickener |
| | Guar gum and xanthomonas gum | Printing paste thickener |
| | Guar gum derivatives (polyacrylic acid, polyacrylamide, carboxy methyl guar gum & guar gum mixture grafted with acrylic acid or acrylonitrite or acrylamide) | Thickener |
| | Guar gum PEG lauryl ether | Foam composition |
| | Sulfonated guar gum derivatives | Pigment retention aid |
| | Guar gum with acrylamide | Stabilizer |
| 4 | **Paper Industry** | |
| | Guar gum | Increase fines retention |
| | Quaternary ammonium guar gum | Imparts dry strength |
| | Derivatives of guar gum (mey proid, SFA, mey pro bend PA109, meyprofilm PA 116) | Enhance surface and sizing |
| | Cationic guar gum | Improves retention for filter barrier |
| | Guar gum formate | Retention & drainage providing agent |
| | Sodium salt carboxy methylated guar gum | Flocculent and sizing agent |
| 5 | **Ore refining/metal Industry** | Dry strength |
| | Amino ethyl gum | In settling fine particles colloidal |
| | Guar gum | Flocculent |
| | Guar gum with polyacrylamide | Less mechanical work |
| | Guar gum and its derivatives | Binder |
| 6 | **Coal mining Industry** | |
| | Esterified guar gum | Stabilizer |
| | Guar gum | Dispersant |
| | Guar gum with boric acid/borax | For shock impregnation of coal seams |
| 7 | **Tobacco Industry** | |
| | Guar gum | Adhesive |
| | Guar gum | Reduces irritation & strengthening agent |
| 8 | **Electricals and telephone** | |
| | Guar gum | Lubricant for installation of electric and telephone cable |
| | Guar gum with isobutylene and Carboxy methyl cellulose | Electrical insulator |
| 9 | **Fire fighting Industry** | |
| | Carboxy methylated guar gum | Provides storing stability |
| | Guar gum with glycerol and ethylene glycol | Dispersions |
| | Guar gum with decyl sulphate & ammonium phosphate solution | Provide viscosity and stability |
| 10 | **Building and construction Industry** | |
| | Guar gum | Gelling agent, foam stabilizer and thickening agent |
| | Guar gum | Water proofing |
| 11 | **Pollution Control/water purifier Industry** | |
| | Trimethyl ammonio, triethyl ammonio, diethyl amino and sulfo derivatives of guar gum | Adsorbent |

| | | |
|---|---|---|
| | Cationic guar gum derivatives | Flocculating and exchanging agent |
| 12 | **Analytical Industry** | |
| | Guar gum derivatives (Glycin hydroxamate in guaran, acetic acid hydroxamate in guaran, imino diacetic acid dihydroxamate in guaran) | Separation of metal ions |
| | Modified guar gum | Support for immobilization of ligands |
| | Guar gum | Purification of lectins |
| | Crosslinked guar gum | Chromatographic separation and selective resin for boron |
| 13 | **Photography** | |
| | Hydrolysed guar gum | Binder |
| 14 | **Food Industry** | |
| | Guar gum | Thickener, Binder, |
| | Guar gum | Stabilizer and gelling agent |
| | Guar gum with carragenan and | Modify visco-elastic behavior of wheat flour |
| | O-carboxy methyl cellulose | Stable thixotropic stabilizer, emulsifier system |
| | Guar gum with carboxymethyl cellulose and tamarind seed gum | Thickening and improver |
| | Glydyl trimethyl ammonium denatured guar gum | Improves sedimentation |
| | Guar gum with xanthan and carob gum | Thickening agent and stabilizer |
| | Guar gum | Freeze thaw and heat stabilizer |
| | Guar gum with whipable albumin and edible fat | To create pudding & cream dressing of cakes |
| 15 | **Dairy** | |
| | Guar gum | Thickener and calcium binder |
| 16 | **Pharmaceutical Industry** | |
| | Guar gum | Stabilizer, Suspending agent, Binder/disintegrant |
| | Sodium carboxymethyl guar | Binds medicine |
| | Guar gum | Drug targeting to colon |
| | Modified guar gum | Sustained release of diltiazem-hydrochloride |
| | Guar gum | Insulinogenic & blood glucose lowering agent |
| | Guar gum | Cholesterol lowering agent |
| | Guar gum | Decreases transit time of colon |
| | Partially hydrolysed guar gum | Blood glucose and blood lipid lowering agent |
| | Guar gum | Malnutrition treatment |
| | Guar gum | High resorptivity |
| | Guar gum | Synergistic activity with bismuth salt |
| 17 | **Agriculture** | |
| | Guar gum or guar gum + polyvinyl alcohol + borax | Increased water retention capacity of soil |
| | Guar gum derivatives | Improves water holding capacity |
| | Guar gum | Anticrusting agent and adhesive of azotobacter |
| | Guar gum polyacrylamide | Water retaining agent |
| | Guar gum | Decreasing cholesterol level, decreases nitrogen retention, fat absorption and metabolizable energy |
| | Guar gum with fatty acid derivatives and kaolin | Prevention of granules |
| 18 | **Cosmetics** | |
| | Guar gum with hydroxyethyl-cellulose and Zn pyrithione etc | Suspending agent |
| | Guar gum/urea an sulphite | Thickener |

| 19 | **Soap Industry** | |
|---|---|---|
| | Hydroxy-G-6 alkyl ether derivatives of guar gum | Thickener |
| | Guar gum | Removes soil but retains lipid |
| | Guar gum or hydroxy propyl guar gum | Thickener |

## 2.10 Market

Guar gum is a very important foreign exchange earner for India. For instance, guar gum export reached 183.57 thousand tons providing foreign exchange worth Rs. 10384 million during the year 2004-05. In 2005-06, 189.11thousand tons guar gum costing Rs. 11249 million was exported to western countries. Thus the demand for guar gum is rising in the global market and is becoming an opportunity for the farmers and the industrialists as well [76].

## 2.11 Major areas of interest in guar

Guar is a versatile crop used in various ways and is essential for various industries. However, the crop is facing number of challenges affecting its productivity and market potential. There is a need to overcome various technical bottlenecks to solve these problems.

### 2.11.1 Productivity of guar

The productivity of guar is fluctuating over the years and also among different growing regions. In India Rajasthan has the highest area under guar cultivation yet the productivity is lowest. Thus sustained productivity has to be achieved through the production of improved cultivars and the use of better agronomic practices [77].

### 2.11.2 Susceptibility to diseases

The crop suffers from various diseases like Alternaria leaf spot, bacterial blight, Anthracnose, root rot and wilts. There is an urgent need to identify resistance genes for these diseases and transfer them to commercially grown cultivars [77].

### 2.11.3 Adaptable varieties for all season and regions

With constant increasing demand for guar gum, the cultivation area under the crop has to be increased. This is possible by producing cultivars that are suited for off-season cultivation and growth in irrigated regions where it is not commonly cultivated [77].

The major challenges faced by the crop can be overcome by application of molecular approaches to understand the genetics and by producing better cultivars through molecular breeding.

## 2.12 Molecular Markers

The concept of genetic markers is not a new one. Since long, plant scientists have employed a diverse array of molecules in an effort to resolve relationships among plant species [28] and study genetic inheritance. Gregor Mendel used phenotype-based genetic markers to study the inheritance of traits in the nineteenth century [1]. Phenotype based genetic markers for *Drosophila* led to the establishment of the theory of genetic linkage. In 1960s and 1970s the use of secondary chemical compounds, such as flavonoids and terpenoids, played a major role in efforts to resolve plant relationships at many taxonomic levels [28, 138]. Isozymes were also used as markers in various studies. But, the limitations of phenotypic and biochemical markers led to the development of more general and useful direct DNA based markers that became known as molecular markers.

The DNA based molecular markers reveal natural variation at the DNA sequence level. These markers are used in plant genotyping, diversity studies, genetic linkage studies, quantitative trait mapping and marker-assisted selection during plant breeding [69]. DNA-based molecular markers offer numerous advantages over conventional phenotype based alternatives because they are stable and can be detected in all tissues regardless of growth, differentiation, development, or defense status of the plant. They are not confounded by the environment, pleiotropic and epistatic effects [1]. Molecular markers do not have any biological effect; they are identifiable DNA sequences found at specific locations of the genome, and transmitted by the standard laws of inheritance from one generation to the next. They can be considered as constant landmarks in the genome. Since the markers and the genes they mark are close together on the same chromosome, they tend to stay together in each generation of plants produced. As scientists learn where markers occur on a chromosome, and how close they are to specific genes, they can create a genetic linkage map [134]. The first theory about the construction of genetic linkage map in man using Restriction Fragment Length Polymorphism (RFLP) was given in 1980 [12]. Today the existence of various molecular techniques and differences in their principles and methodologies require careful consideration in choosing one or more of such marker types.

Molecular markers are considered ideal if, they have the following characteristics [1]:

1. Polymorphic and evenly distributed throughout the genome
2. Provide adequate resolution of genetic differences
3. Generate multiple, independent and reliable markers
4. Simple, quick and inexpensive
5. Need small amounts of tissue and DNA samples
6. Have linkage to distinct phenotypes
7. Require no prior information about the genome of an organism

Unfortunately none of the molecular marker techniques is ideal for every situation. Techniques differ from one another with respect to important features such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and cost. Depending on the need, modifications in the molecular markers have been made, leading to their advanced versions. Fig 2.7 shows the development of various molecular markers in the previous two decades [1].

### 2.12.1 Restriction fragment length polymorphism (RFLP)

In RFLP, DNA polymorphism is detected by hybridizing a chemically labeled DNA probe to a Southern blot of DNA digested by restriction endonucleases, resulting in differential DNA fragment profile [1]. The RFLP markers are highly polymorphic, co-dominantly inherited and highly reproducible. RFLPs are present throughout the genome, heritable and locus specific. The method can also be used to simultaneously screen numerous samples. The DNA blots can be analyzed repeatedly by stripping and re-probing (usually eight to ten times) with different RFLP probes [1]. RFLP was used widely in early 1980's for a wide range of plant species [106]. This technique is not very widely used now because it is time consuming, involves expensive and radioactive/toxic reagents [111] and requires large quantity of high quality genomic DNA. The requirement of prior sequence information for probe generation increases the complexity of the methodology. These limitations led to the conceptualization of a new set of less technically complex techniques known as PCR-based techniques [1].

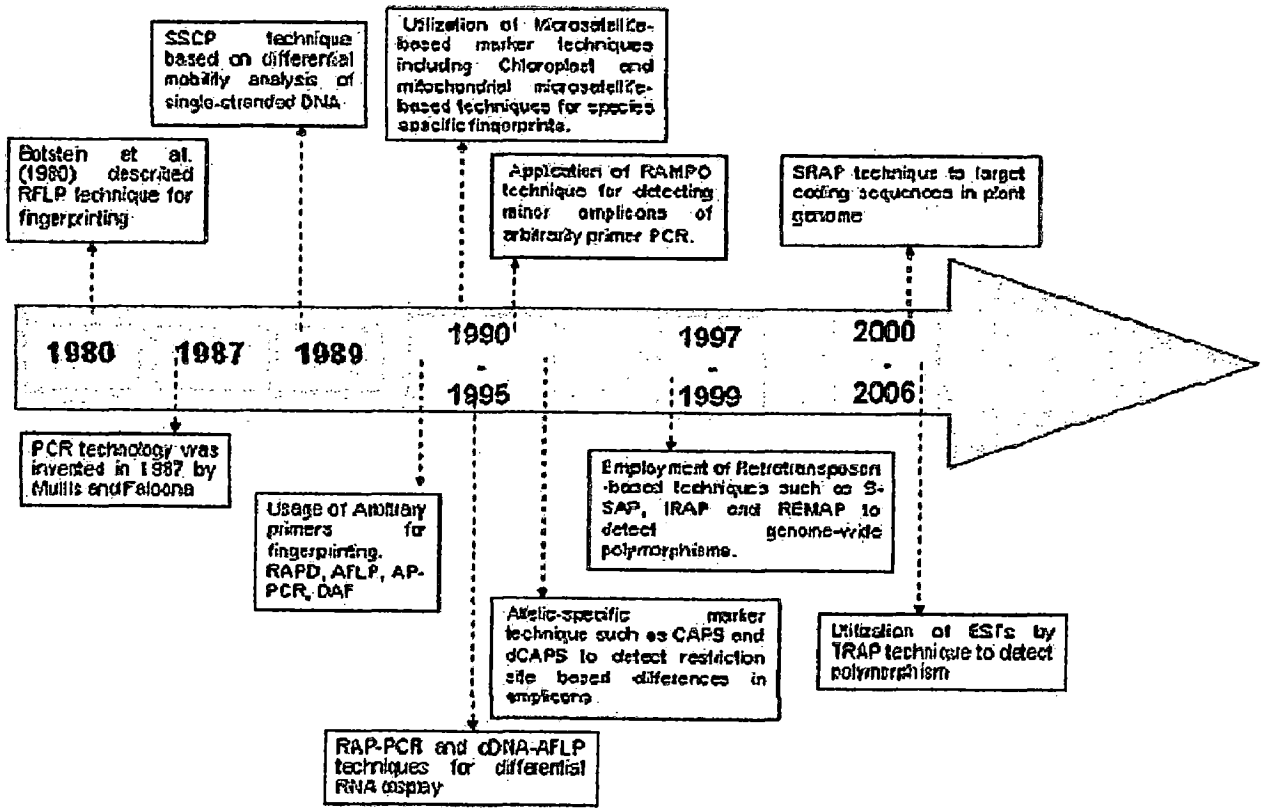



**Fig. 2.7. Schematic diagram describing the development and utilization of molecular marker techniques over last two decades [1]**

The invention of polymerase chain reaction (PCR) technology [92, 126, 127] has lead to the development of various marker systems for the study of genetic diversity. Following are the various PCR-based marker techniques mainly used in plants.

### 2.12.2 Amplified fragment length polymorphism (AFLP)

AFLPs are generated by complete restriction endonuclease digestion of total genomic DNA, followed by selective PCR amplification and electrophoresis of a subset of the fragments. This results in a unique, reproducible fingerprint/profile for each individual. The fingerprint allows an assessment of genome-wide variation. These anonymous markers consist largely of non-coding DNA [89].

AFLP is useful in a wide range of applications including linkage mapping [128, 147, 155], measuring genetic diversity [2, 44], identifying hybrids [136] and cultivars [45], population genetics [83], population assignment and developing single-locus sequence-characterized amplified region (SCAR) markers [70, 96].

### 2.12.3 Random amplified polymorphic DNA (RAPD) markers

Randomly amplified polymorphic DNA markers technique was introduced in the year 1990 [154]. It is an effective and powerful technique for determining genetic variation and has the capacity to generate markers that span the genome without prior knowledge of the sequence [9, 61].

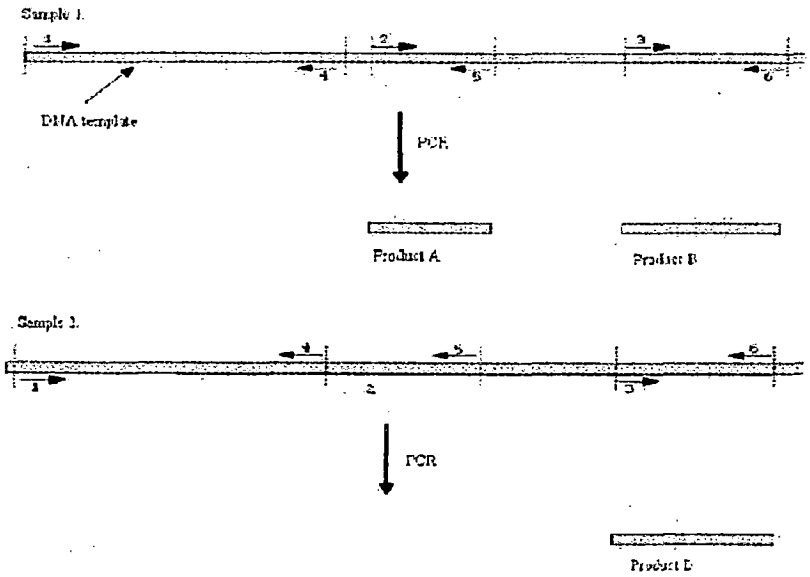



**Fig. 2.8. Schematic diagram showing RAPD technique**

The RAPD protocol usually uses a single 10 bp arbitrary primer at a constant low annealing temperature. Although the sequences of RAPD primers are arbitrarily chosen, two basic criteria must be met: a minimum of 40% GC content (50 - 80% GC content is generally used) and the absence of palindromic sequence (a base sequence that reads exactly the same from right to left as from left to right) [154]. The basis of RAPD technique is differential PCR amplification of genomic DNA. It deduces DNA polymorphisms produced by rearrangements or deletions at or between oligonucleotide primer binding sites in the genome [1]. Fig 2.8 shows schematic diagram of RAPD reaction with polymorphism.

RAPD has been applied for various purposes in many plant species. It has been used to construct genetic maps [74, 75, 97]. It has also been used for genome fingerprinting and characterization [114, 131, 153]. Markers linked to important plant genes like *Pseudomonas* resistance gene in tomato [84], *Uromyces appendiculatus* (Pers.) resistance in common bean [52], etc. have been developed using random primers. Sequence characterized amplified

regions (SCARs) markers can also be developed from RAPD [33, 87]. There are few reports of RAPD analysis in commercial varieties of cluster bean [81, 100, 109].

### 2.12.4 DNA amplification fingerprinting

A modification of the RAPD technique called DNA amplification fingerprinting (DAF) was introduced in 1993 where one or more very short (less than or equal to 5 nt) arbitrary oligonucleotides are used to direct the enzymatic amplification of DNA. Polyacrylamide gel electrophoresis and silver staining is used to resolve the complex patterns. It is suitable for DNA fingerprinting [21]. DAF is used for studying relationships among closely related species and cultivars in plants [24] and bacteria [7].

### 2.12.5 Sequence characterized amplified regions (SCAR)

The arbitrary marker techniques are sensitive to changes in the reaction conditions and are not very easy to use for various studies. In SCAR, pairs of 20-25 bp oligonucleotide primers specific to the sequence of polymorphic bands are used to amplify the characterized regions from genomic DNA under stringent conditions, which makes these markers more specific and dependable as compared to RAPD markers [33].

In order to bridge the gap between the ability to obtain linked markers to a gene of interest in a short time and the use of these markers for map-based cloning approaches or for routine screening procedures, SCAR marker technique was developed and applied [1, 3]. The difficulty of developing SCARs varies according to the technique used to produce the original profile. Conversion to SCAR is considered easy in comparison to AFLP and ISSR [3]. The derived SCAR markers can be applied for marker assisted selection [63], map-based cloning, linkage mapping [155], etc.

### 2.12.6 Cleaved amplified polymorphic sequences (CAPS)

CAPS are also known as PCR-RFLP markers. The CAPS deciphers the restriction fragment length polymorphisms caused by single base changes like SNPs, insertions/deletions, which modify restriction endonuclease recognition sites in PCR amplicons [1]. The results are highly reproducible with different DNA extraction methods. The stability and simplicity of CAPS analysis makes it a good tool for the identification of cultivars [78].

CAPS markers can be developed by comparing the sequence difference between two known regions and designing a combination of primers and restriction enzymes for assay [78]. These markers can also be generated form arbitrary marker techniques [150].

### 2.12.7 Single strand conformation polymorphism (SSCP)

Single strand conformation polymorphism is the mobility shift analysis of single-stranded DNA sequences on neutral polyacrylamide gel electrophoresis. These are used to detect polymorphisms produced by differential folding of single-stranded DNA due to subtle differences in sequence (often a single base pair) [1, 98].

### 2.12.8 Single nucleotide polymorphism (SNP)

Single nucleotide polymorphisms and insertions/deletions, which are the basis of most differences between alleles, have been simplified by many developments in sequencing technology. SNP discovery in many crop species, such as corn and soybean, is relatively straight forward because of the presence of high level of intraspecific nucleotide diversity, and the availability of many gene and expressed sequence tag (EST) sequences [110]. Several different strategies for the discovery of SNPs may be used. These include the re-sequencing of PCR amplicons with or without pre-screening, electronic SNP (eSNP) discovery in shotgun genomic libraries, and eSNP discovery in expressed sequence tag (EST) libraries [110].

### 2.12.9 Microsatellite or simple sequence repeats (SSR) markers

Microsatellite or short tandem repeats or simple sequences repeats are repetitions of very short (1-5) nucleotide motifs, which occur as interspersed repetitive elements in all eukaryotic genomes [144]. The variation in the number of repeated units is mainly due to strand slippage during DNA replication where the repeats allow matching via excision or addition of repeats [133]. As slippage in replication is more than a point mutation, microsatellite loci tend to be hypervariable. Microsatellite shows extensive length polymorphisms among individuals during PCR analysis of unique loci using specific primer sets. SSRs are actually considered the most efficient markers but their use is still limited because of the time consuming and laborious steps to develop them [113].

Microsatellites are popular genetic markers because of their co-dominant inheritance, high abundance, enormous extent of allelic diversity, and the ease of assessing SSR size variation by PCR with pairs of flanking primers [1]. Their potential for automation is an

additional advantage when compared with other types of molecular markers. SSRs are highly polymorphic, genome specific, abundant and co-dominant, and have recently become important genetic markers [60]. These are used in studying genetic diversity [4, 90], parentage analysis of clones [73], linkage mapping [50, 67], identifying hybrids [22] and marker assisted selection [118].

### 2.12.10 Inter simple sequence repeat (ISSR) markers

Microsatellites are usually more or less proportionally dispersed in the genome. However, regions with a greater abundance of these sequences have been found and are named as "SSR hot spots" [10, 11, 161]. Such regions can serve as a source of ISSR markers.

The ISSR technology is based on the amplification of regions (100-3000 bp) between inversely oriented closely spaced microsatellites. Single primer (16-18 bp) consisting of several simple sequence repeats is used for amplification of these regions. Primers can be based on any SRR motif along with 5' or 3' anchored bases usually 2-4 bases which are arbitrary selective nucleotides. However, nonanchored primers have also been used [11].

**Table 2.2. Characteristics of the molecular markers used in the study [14, 27, 113, 134]**

| Sl. No | Feature | RAPD | ISSR | SSR |
|---|---|---|---|---|
| 1 | Abundance | High | High | High |
| 2 | Locus specificity | No | No | Yes |
| 3 | Nature of polymorphism | Base changes (insertions, deletions) | Base changes (insertions, deletions) variation in SSR repeat length/number of motifs | Variation in repeat length/ number of motifs |
| 4 | Level of Polymorphism | High | High/medium | High/very high |
| 5 | Inheritance mode | Dominant | Dominant /codominant | Codominant |
| 6 | Reproducibility | Low/Medium | High/Medium | High |
| 7 | Sequence information | Not required | Not required | Required |
| 8 | Technical demands | Low/Medium | Low/Medium | Medium/Low(For use) High(For development) |
| 9 | Cost | Low | Low | Medium |
| 10 | Labor | Low | Low | Medium (For use) High (For development) |
| 11 | Time | Low | Low | High |
| 12 | Principle | Uses single random 10-mer primers | Uses single primer with sequence repeat motifs with anchors | Uses primers designed from sequences flanking SSR motif. |
| 13 | Applications | DNA fingerprinting of cultivars<br>Genetic correspondence of plant material from nursery<br>Detection of intra-cultivar variability<br>Construction of linkage map<br>Cultivar traceability<br>Phylogenetic studies | Phylogenetic studies<br>Detection of intra-cultivar variability<br>Cultivar traceability | DNA fingerprinting of cultivars<br>Construction of linkage map<br>Paternity analysis<br>Cultivar traceability in Phylogenetic studies<br>Gene tagging, |
| 14 | Advantages | Simple, No prior sequence information needed, Multiple loci analysis from single primer, Requires very less DNA, Less expensive | Simple, Reproducible, Multiple loci analysis from single primer, Requires very less DNA, Less expensive | Simple, Reliable, Highly reproducible, Transferable between population , Co-dominant, Highly polymorphic. |
| 15 | Disadvantages | Dominant mode of transfer, Reproducibility of banding pattern | Dominance of alleles, May not be homologous in case of small fragments. | relatively time consuming and labor-intensive, high complexity of amplification profiles may occur |

## 2.13 Development of SSR markers

SSRs are considered as the most efficient markers because of their co-dominant inheritance, high abundance, enormous extent of allelic diversity, and the ease of assessing SSR size variation by PCR with pairs of flanking primers [113]. However, high development cost is the major impediment for the routine application of SSRs for genetic studies [55, 123]. The major drawback of microsatellites is that they need to be isolated *de novo* from species that are examined for the first time [158]. Various approaches have been devised to ease the detection and development of SSRs. The two approaches that are commonly used for developing SSR markers are

1. Through library construction
2. SSR mining from sequence databases

### 2.13.1 Development of SSR through library construction

Construction of genomic libraries followed by screening for SSR containing sequences is the method of choice for new species of interest. The method varies depending upon the type of library construction. Depending on the fragmentation method, DNA fragments are ligated into a common plasmid vector either directly or after ligation to specific adaptors[158].

#### 2.13.1.1 Non-enriched libraries

The following steps are commonly followed for generating SSR markers from a library. The DNA is isolated and digested with appropriate restriction enzymes. Fragments between 300 and 1000 bp are selected by electrophoresis and ligated to a vector.

Transformation of bacterial cells with ligation vector generally results in thousands of recombinant clones which can be subsequently screened for the presence of microsatellite sequences. Screening for positive clones is generally carried out by Southern hybridization using repeat-containing probes, after blotting bacterial colonies onto nylon membranes. Colony transfer can be carried out either by classical replica plating or by picking single colonies and ordering them in new arrayed plates [158]. The positive clones are sequenced and primers complementary to both flanking regions of SSR are designed [124].

Although this approach has been applied in many cases, a number of disadvantages exist. It is difficult to construct large libraries for species with large genomes. Other problems

such as low effectiveness and specificity of hybridisation as well as the presence of one-side flanks in sequenced fragments exist. To improve this method pre-screening for insert length, repeat position and orientation by anchor PCR technique was described [112].

**2.13.1.2 Enriched libraries**

Several approaches have been developed to increase the SSR containing sequences in the libraries. Recently several kits for enriched protocols have been developed [158]. The easiest method of enriched library construction is selective hybridization of DNA fragments using streptavidin-coated magnetic beads or nylon membranes [53, 71]. The procedure of the construction of enriched libraries using streptavidin-coated magnetic beads or nylon membranes comprises the following steps [113]:

- DNA digestion followed by ligation of the resulting fragments to double-stranded adaptors
- DNA hybridization to biotinylated microsatellite probes, followed by binding to streptavidin-coated magnetic beads
- Elution of the DNA fragments from the beads, and PCR amplification with primers complementary to the adaptor sequence
- Cloning of the amplified products into the vector
- Transformation of *E.coli*
- Sequencing of the positive clones

The method used for generating enriched libraries containing SSRs is described in Fig 2.9 This method is widely applied in plants [68, 88, 121, 159]. The efficiencies of obtaining SSR containing sequences from enriched libraries were higher in all cases than in the traditional method. It ranged from 55% [25] up to 100% of the clones containing microsatellites suitable for primer designation [53, 113]. In spite of the sufficient progress in the efficiency of positive clone isolation, the procedure employing magnetic beads allows enrichment in a single or, in few SSR motifs only [113]. This problem can be solved by using nylon membranes with many bound microsatellite oligonucleotides [37].

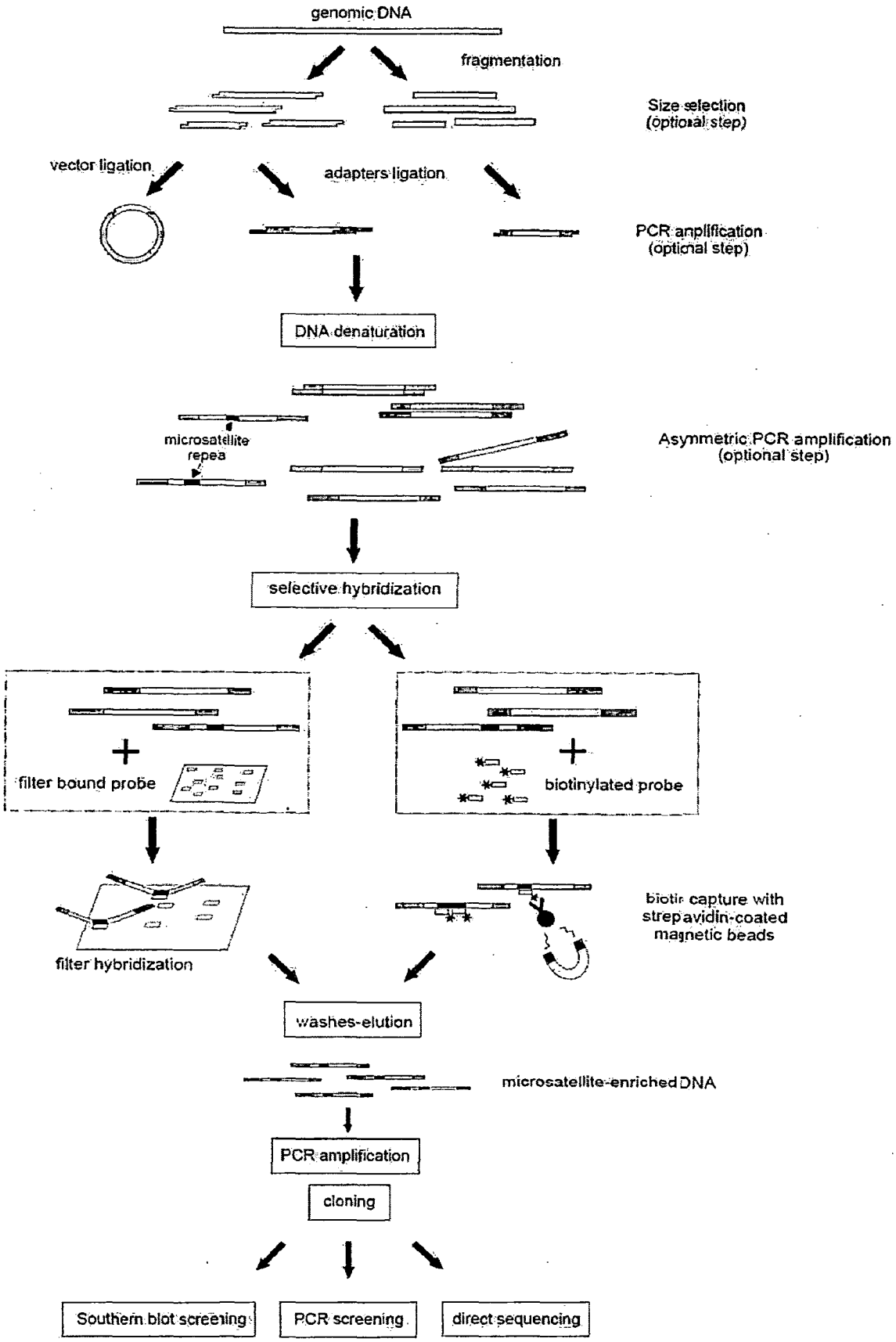



**Fig. 2.9. Schematic diagram representing selective hybridization protocols [158]**

### 2.13.2 SSR mining from sequence databases

Recently, expressed sequences tags (ESTs) have been shown to be good sources of SSR markers [42]. ESTs are parts of expressed genes, and the EST-derived SSRs can be considered as type I markers and used to map genes of known functions [152]. This strategy of developing SSR markers is based on bioinformatic screening for sequences containing microsatellites deposited in the data bases (EMBL, GenBank). This method is cost-effective, simple and relatively quick. However, it has some limitations. When exploring data from expressed sequences, a considerable amount of potential polymorphism can be lost as microsatellites are broadly present in the non-coding regions of genomes. Additionally, this strategy is limited to plants with high economical or scientific interest which are well represented in the databases [113].

Different tools available on public domain can be used for this purpose. The commonly used tools are MIcroSAtellite identification tool (MISA) [146], Tandem Repeats Analyzer (TRA)[8], TROLL [85], SSRIT [145] and SSR Primer [122]. Among these 5 tools MISA program has given maximum coverage of SSRs in both oil palm ESTs and Contigs [120]. However it has the advantage of detecting the mono to decamer repeats and also compound repeats but has the disadvantage of inability to detect above decanucleotide repeats.

The above literature review shows that guar is an important industrial corp. Guar gum is essential for various industries like paper, textile, petroleum, oil drilling, pharmaceutics, food, cosmaceutics, explosives, etc. The enzymes and the pathway of the galactomannan biosynthesis have been studied and the genes responsible for galactomannan production have been cloned and characterized. However, still the crop is facing various challenges like disease and low productivity. The genetic diversity in the landraces of the crop is not well understood. There are no reports of development of any co-dominant markers like SSR which are very useful in molecular breeding and crop improvement programmes. There is a need to develop molecular markers and understand the natural variation available in the landraces for quick and easy execution of crop improvement programs.

# *Chapter III*

*Studying genetic variability in cluster bean accessions through application of RAPD and ISSR markers*

# 3. Studying genetic variability in cluster bean accessions through application of RAPD and ISSR markers

## 3.1 Materials and methods

### 3.1.1 Plant material

Accessions of guar landraces were kindly provided by National Bureau of Plant Genetic Resources (NBPGR), New Delhi. Among 89 accessions provided 29 accessions were selected based on their place of origin according to Dwivedi et al. (1995)[36]. The regions of origin of selected accessions are diagrammatically represented in Fig 3.1. Seeds of 19 elite varieties of *C. tetragonoloba,* which are commonly grown in Gujarat, Rajasthan and Haryana or being tested for release were obtained from Central Arid Zone Research Institute (CAZRI), Jodhpur, Rajasthan. The details of the accessions are given in Table 3.1. Plants were grown in field conditions in Indian Institute of Technology Roorkee, India. Leaves were collected from 3 weeks old plants for DNA extraction.

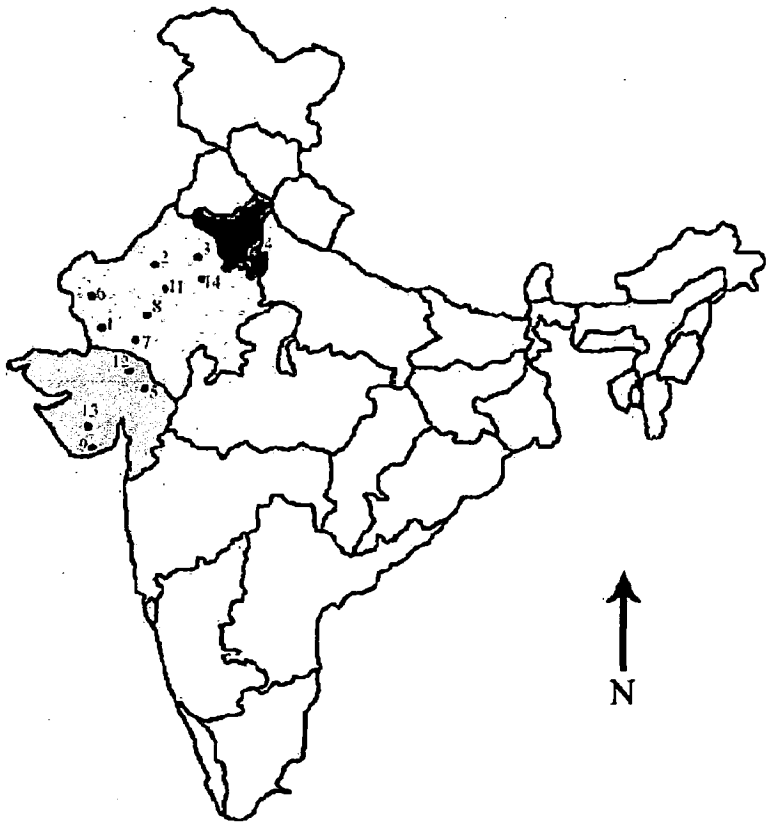



**Fig. 3.1. Diagrammatic representation of locations of collection of landraces as stated by Dwivedi et al. (1995) [36]: 1.Barmer, 2.Bikaner, 3.Churu, 4.Delhi, 5.Himmatnagar, 6.Jaisalmer, 7.Jalore, 8.Jodhpur, 9.Kachchh, 10.Mahendragarh, 11.Nagaur, 12.Palanpur, 13.Rajkot, 14.Sikar**

**Table 3.1. Details of cluster bean genotypes used for studying genetic diversity [36]**

| Sl.no | Accession Number | Location of collection |
|---|---|---|
| 1 | IC 8783 | Delhi |
| 2 | IC 11388 | Jalore |
| 3 | IC 103304 | Kachchh |
| 4 | IC 103323 | Himmat nagar |
| 5 | IC 116522 | Jodhpur |
| 6 | IC 116527 | Jodhpur |
| 7 | IC 116546 | Jodhpur |
| 8 | IC 116547 | Jodhpur |
| 9 | IC 116565 | Barmer |
| 10 | IC 116568 | Barmer |
| 11 | IC 116569 | Barmer |
| 12 | IC 116570 | Jaisalmer |
| 13 | IC 116577 | Jaisalmer |
| 14 | IC 116583 | Churu |
| 15 | IC 116595 | Jodhpur |
| 16 | IC 116601 | Bikaner |
| 17 | IC 116609 | Churu |
| 18 | IC 116659 | Churu |
| 19 | IC 116682 | Sikar |
| 20 | IC 116751 | Nagaur |
| 21 | IC116752 | Nagaur |
| 22 | IC 116767 | Nagaur |
| 23 | IC 116835 | Mahendra garh |
| 24 | IC 116836 | Mahendra garh |
| 25 | IC 116890 | Rajkot |
| 26 | IC 116926 | Palanpur |
| 27 | IC 116953 | - |
| 28 | IC 116958 | - |
| 29 | IC 116960 | - |
| 30 | M 83 | Commercially grown genotypes of cluster bean |
| 31 | RGC 197 | |
| 32 | RGC 936 | |
| 33 | RGC 986 | |
| 34 | RGC 1002 | |
| 35 | RGC 1003 | |
| 36 | RGC 1017 | |
| 37 | RGC 1031 | |
| 38 | RGC 1038 | |
| 39 | RGC 1055 | |
| 40 | RGC 1066 | |
| 41 | RGC 1076 | |
| 42 | RGC 1088 | |
| 43 | RGR 9 | |
| 44 | HG 563 | |
| 45 | HG 04-876 | |
| 46 | SRG 1058 | |
| 47 | GAUG 501 | |
| 48 | NSG369 | |

### 3.1.2 DNA extraction buffer

The composition of DNA extraction buffer is given in Table 3.2.

**Table 3.2. Composition of the DNA extraction buffer**

| S.No. | Name of the reagent | Concentration |
|---|---|---|
| 1 | Tris.HCl (pH 8.0) | 100mM |
| 2 | EDTA (pH 8.0) | 20mM |
| 3 | NaCl | 1.4M |
| 4 | CTAB | 5% |
| 5 | β-mercaptoethanol | 0.01% |

All the chemicals used were purchased from HiMedia (Molecular biology grade).

### 3.1.3 TE/$T_{10}E_1$ buffer

TE buffer was prepared by mixing 1ml of 1M Tris.HCl (pH 8) and 200µL of 0.5M EDTA. The volume was made up to 100ml with milliQ water [129].

### 3.1.4 RNase solution

RNase solution was prepared by dissolving 10mg of RNase (Genei, Bangalore) in 1ml of autoclaved milliQ water. The solution was incubated in water bath at 70°C for 20 min [129].

### 3.1.5 TBE buffer

The composition of TBE buffer is given in Table 3.3.

**Table 3.3. Composition of 1X TBE buffer [129]**

| S.No. | Name of the reagent | Composition |
|---|---|---|
| 1 | Tris | 10.9g |
| 2 | Boric acid | 5.56g |
| 3 | EDTA | 0.98g |
| 4 | Distilled water | 1L |

### 3.1.6 DNA gel loading buffer

The loading dye was prepared with 0.25% bromophenol blue, 0.25% xylene cyanol dissolved in 50% glycerol made in TE buffer.

### 3.1.7 DNA extraction and purification

DNA was extracted from leaves collected from field grown plants using CTAB method as described by Doyle and Doyle (1990) with slight modifications [34]. Approximately 0.5g leaves were ground to fine powder in liquid nitrogen using sterile, pre-chilled pestle and mortar. The pulverized leaf powder was transferred to a 2ml micro-centrifuge tube containing 1ml of pre-warmed DNA extraction buffer. The contents were mixed well followed by incubation at 65°C for 1 hour. The tubes were then kept at room temperature for 10 min. Equal volume of chloroform:isoamyl alcohol (24:1) mixture was added and mixed gently. The tubes were centrifuged at 8000rpm for 10 min at room temperature for phase separation. The upper aqueous layer was pipetted out carefully into a fresh tube. The DNA was precipitated by adding equal volume of ice cold iso-propanol and kept at 4°C for 2 hours. The precipitated DNA was pelletted by centrifugation at 8000rpm for 10 min. The pellet was washed with 70% ethanol, air dried and dissolved in TE buffer. RNA was removed by treatment with 2µl of RNase (10mg/ml) and incubation at 37°C in a water bath for 1 hour followed by treatment of chloroform:isoamyl alcohol (24:1). The tubes were centrifuged and supernatant was transferred to another tube. DNA was precipitated using 100% ethanol. The tubes were centrifuged at 8000 rpm for 10 min to pellet DNA. The supernatant was discarded and the pellet was air dried. Finally DNA was dissolved in TE buffer.

### 3.1.8 Quantification and dilution of DNA

The quality of the extracted DNA was checked by gel electrophoresis on 0.8% agarose [129]. The quantity of DNA was estimated by measuring absorbance at 260nm in a spectrophotometer (Varian) and diluted to approximately 100ng/µl. The diluted DNA samples were stored at -20°C until use.

### 3.1.9 DNA amplification conditions for RAPD and ISSR markers

All accessions, along with a negative control without DNA, were used for RAPD and ISSR analysis. Out of the 30 RAPD primers initially screened 13 decamer primers (Table 3.4) were selected on the basis of polymorphism and reproducibility. Polymerase chain reaction

(PCR) was performed in 20µl reaction volume containing 1U Taq DNA polymerase (Biotools), 2.5mM $MgCl_2$ (Biotools), 5mM dNTPs, 1µL primer and 50ng of template DNA. Amplification was carried out in a Mastercycler gradient programmable thermal cycler (Eppendorf). The annealing temperature for each primer was determined by gradient PCR. The PCR was programmed with initial denaturation step at 94°C for 4 min, followed by 35 cycles of 1 min at 94°C, 1 min at annealing temperature (Table 3.4) and 1 min at 72°C. A final extension was carried out at 72°C for 10 min and a hold temperature of 4°C at the end [109, 129].

**Table 3.4. Details of the RAPD and ISSR primers used in the present study**

| Marker | Sequence (5'-3') | Annealing temperature Tm(°C) |
|---|---|---|
| **RAPD** | | |
| OPA 1 | CAGGCCCTTC | 40 |
| OPD 12 | CACCGTATCC | 40 |
| OPM 2 | ACAACGCCTC | 38.4 |
| OPM 12 | GGGACGTTGG | 40 |
| OPM 15 | GACCTACCAC | 38.4 |
| OPN 1 | CTCACGTTGG | 37.9 |
| OPN 2 | ACCAGGGGCA | 38.4 |
| OPN 3 | GGTACTCCCC | 38.4 |
| OPN 4 | GACCGACCCA | 39 |
| OPN 5 | ACTGAACGCC | 38.4 |
| OPQ-09 | GGCTAACCGA | 40 |
| OPU 15 | ACGGGCCAGT | 39 |
| OPX 12 | TCGCCAGCCA | 39 |
| **ISSR** | | |
| UBC808 | AGAGAGAGAGAGAGAGC | 55 |
| UBC818 | CACACACACACACACAG | 56 |
| UBC820 | GTGTGTGTGTGTGTGTC | 57 |
| UBC854 | TCTCTCTCTCTCTCTCRG | 49 |
| UBC856 | ACACACACACACACACYA | 57 |
| UBC868 | GAAGAAGAAGAAGAAGAA | 44.5 |
| UBC879 | CTTCACTTCACTTCA | 47 |

Y = pYrimidine (C, T)
R = puRine (A, G)

ISSR primers were custom synthesized by Ocimum Biosciences, Hyderabad according to University of British Columbia catalogue and used for screening polymorphism. Out of 16 primers initially screened, 7 primers which produced clear and unambiguous bands (Table 3.4)

were used for diversity study. PCR was performed in 20µl reaction volume containing 1U Taq DNA polymerase (Biotools, Spain), 2.5mM $MgCl_2$, 5mM dNTPs, 1µL of 20pM primer and 50ng of template DNA. Amplification was performed in Mastercycler gradient programmable thermal cycler (Eppendorf). The annealing temperature for each primer was determined by gradient PCR. The PCR was programmed with initial denaturation step at 94°C for 4 min, followed by 30 cycles of 1 min at 94°C, 1 min at annealing temperature (Table 3.4) and 1 min at 72°C. A final extension was carried out at 72°C for 10 min and hold temperature of 4°C at the end.

### 3.1.10 Gel electrophoresis and recording

PCR amplified products were electrophoresed on 1.5% (w/v) agarose gels containing 0.02µg/ml ethidium bromide in 1X TBE buffer. The gel was run at 5V/cm for 2.5 to 3 hours. A 100 bp DNA ladder (Genei, Bangalore) was used as a molecular marker to determine the approximate size of the fragments. The gel was visualized under UV light and documented in gel documentation system (Bio-Rad).

### 3.1.11 Analysis of RAPD and ISSR markers data

The RAPD and ISSR bands were scored according to their positions. The binary data were recorded using '1' for presence of band and '0' for absence of band at a particular position. A similarity matrix was computed for RAPD, ISSR and RAPD+ISSR combined data using Jaccard's coefficient by the software package PAST [54]. A dendrogram was obtained by Unweighted Pair-Group Average (UPGMA) method using the similarity matrix by Jaccard's similarity coefficient [62] to determine the relatedness of 48 genotypes under study. Bootstrapping was done with 1000 repetitions to evaluate the clusters formed. Principle component analysis (PCA) of the landrace data was done using PAST to reveal the genetic relations among the landraces according to their geographic origin. Population genetic analysis (POPGENE Ver 1.32) software was used to determine the values of Nei's genetic diversity (Nei, 1978), Shannon's information index, total genetic diversity among populations (Ht), genetic diversity within population (Hs) and mean coefficient of gene differentiation (Gst) for two populations, i.e. landraces and commercial varieties. RAPD and ISSR data were also subjected to Analysis of Molecular Variance (AMOVA) at individual and population level, using GenALEx software [102]. Correlation was determined between molecular, geographical and morphological data [36] by performing Mantel tests between their Euclidian distance matrices using the software package PAST.

### 3.1.12 Polymorphic information content (PIC)

Polymorphic Information Content (PIC) for dominant marker system was calculated according to Roldan-Ruiz et al. (2000)[125] as

$$PIC_i = 2f_i (1-f_i)$$

Where, $PIC_i$ is polymorphic information content of marker 'i', $f_i$ is frequency of the amplified allele (band present), and $(1-f_i)$ is frequency of the null allele (band absent). PCI was averaged for all the fragments for each primer.

### 3.1.13 Resolving power (Rp)

Resolving power (Rp) for each primer was calculated according to Prevost and Wilkinson (1999) [108] as

$$Rp = \sum I_b$$

Where, $I_b$ is the informativeness of the bands scored which can be calculated as

$$I_b = 1-(2*|0.5-p|)$$

Where 'p' is the proportion of 48 genotypes containing bands.

## 3.2 Results

### 3.2.1 Genetic diversity study using RAPD and ISSR markers

A set of 13 RAPD markers (Table 3.5) was used for analysis of genetic diversity among 29 landraces and 19 commercial varieties of cluster bean. The total number of bands, number of polymorphic bands, percentage of polymorphic bands, polymorphic information content (PIC) and resolving power (Rp) obtained for each primer are shown in Table 3.5. Among the RAPD primers used OPN-1 showed highest resolving power whereas OPX-12 showed lowest.

### 3.2.2 RAPD analysis

RAPD PCR amplification produced a total of 118 bands, out of which 103 were polymorphic and 15 were monomorphic. Out of the 13 primers used OPQ-09 produced highest number of bands (12) with 100% polymorphism, whereas the primer OPX-12 produced only 5 bands. Fig 3.2 and Fig 3.3 show the RAPD fingerprint patterns produced with OPM-02 and OPQ-09 primers, respectively; the sequence of accessions in sample lanes of the photographs are according to the serial number in Table 3.1.

**Table 3.5. Properties of 13 RAPD primers used and their percentage polymorphism, polymorphic information content and resolving power**

| Marker | Total number of bands | Polymorphic bands | Monomorphic bands | Percentage polymorphism | Polymorphic Information Content(PIC) | Resolving power |
|---|---|---|---|---|---|---|
| OPA-1 | 7 | 5 | 2 | 71.42 | 0.143 | 1.208 |
| OPD-12 | 6 | 6 | 0 | 100.00 | 0.419 | 3.625 |
| OPM-2 | 9 | 5 | 4 | 55.55 | 0.141 | 2.000 |
| OPM-12 | 8 | 8 | 0 | 100.00 | 0.199 | 2.000 |
| OPM-15 | 11 | 7 | 4 | 63.64 | 0.189 | 3.083 |
| OPN-1 | 11 | 10 | 1 | 90.90 | 0.363 | 6.333 |
| OPN-2 | 10 | 10 | 0 | 100.00 | 0.252 | 3.500 |
| OPN-3 | 10 | 10 | 0 | 100.00 | 0.371 | 5.417 |
| OPN-4 | 10 | 9 | 1 | 90.00 | 0.199 | 2.542 |
| OPN-5 | 10 | 9 | 1 | 90.00 | 0.186 | 2.500 |
| OPQ-09 | 12 | 12 | 0 | 100.00 | 0.296 | 4.875 |
| OPU-15 | 9 | 7 | 2 | 77.78 | 0.227 | 3.208 |
| OPX-12 | 5 | 5 | 0 | 100.00 | 0.151 | 0.917 |

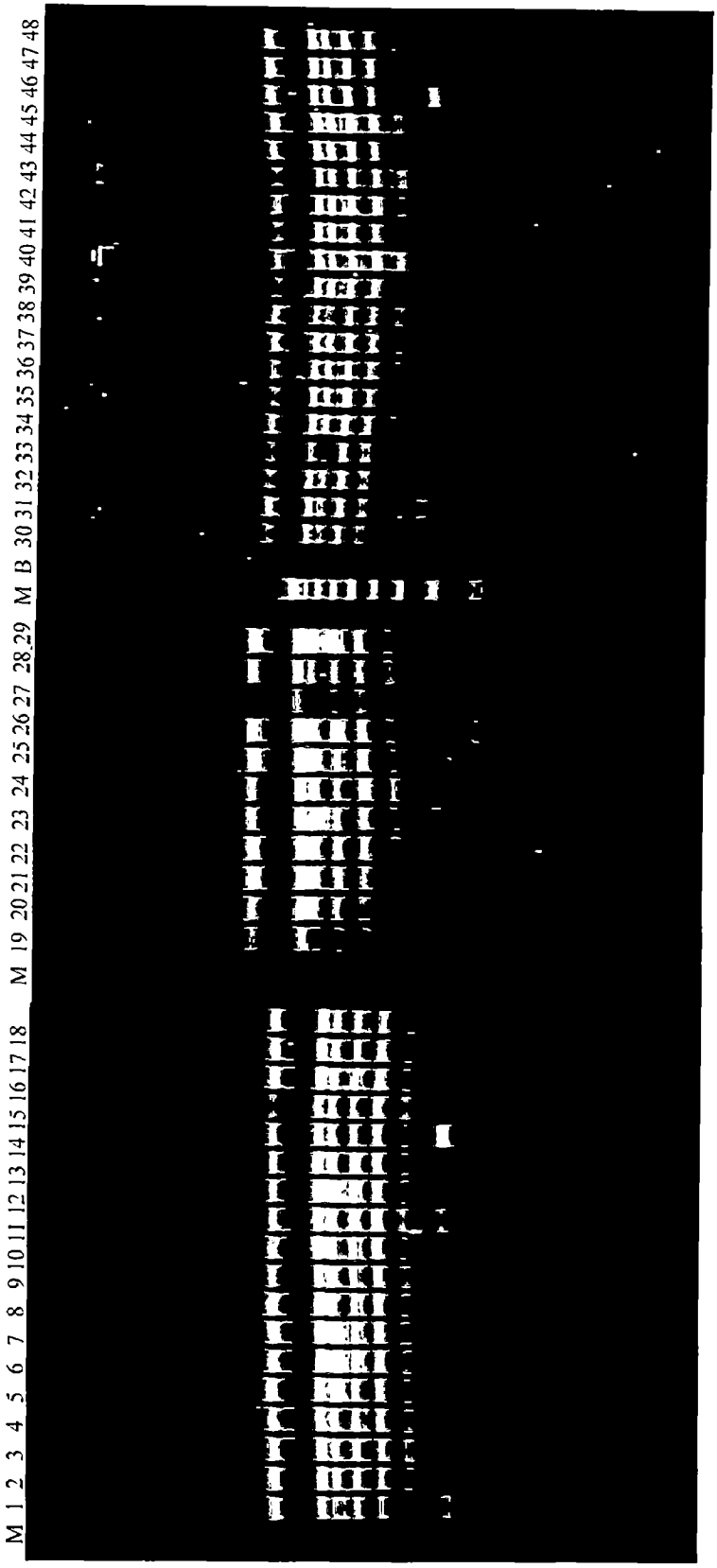



**Fig. 3.2. Ethidium bromide stained agarose gel showing RAPD profile from OPM-02. The sample numbers are as represented in Table 3.1. M represents 100 bp marker and B represents negative control**

KDCL 00297362 23
03.JPG N N N-2 AS+02SA100

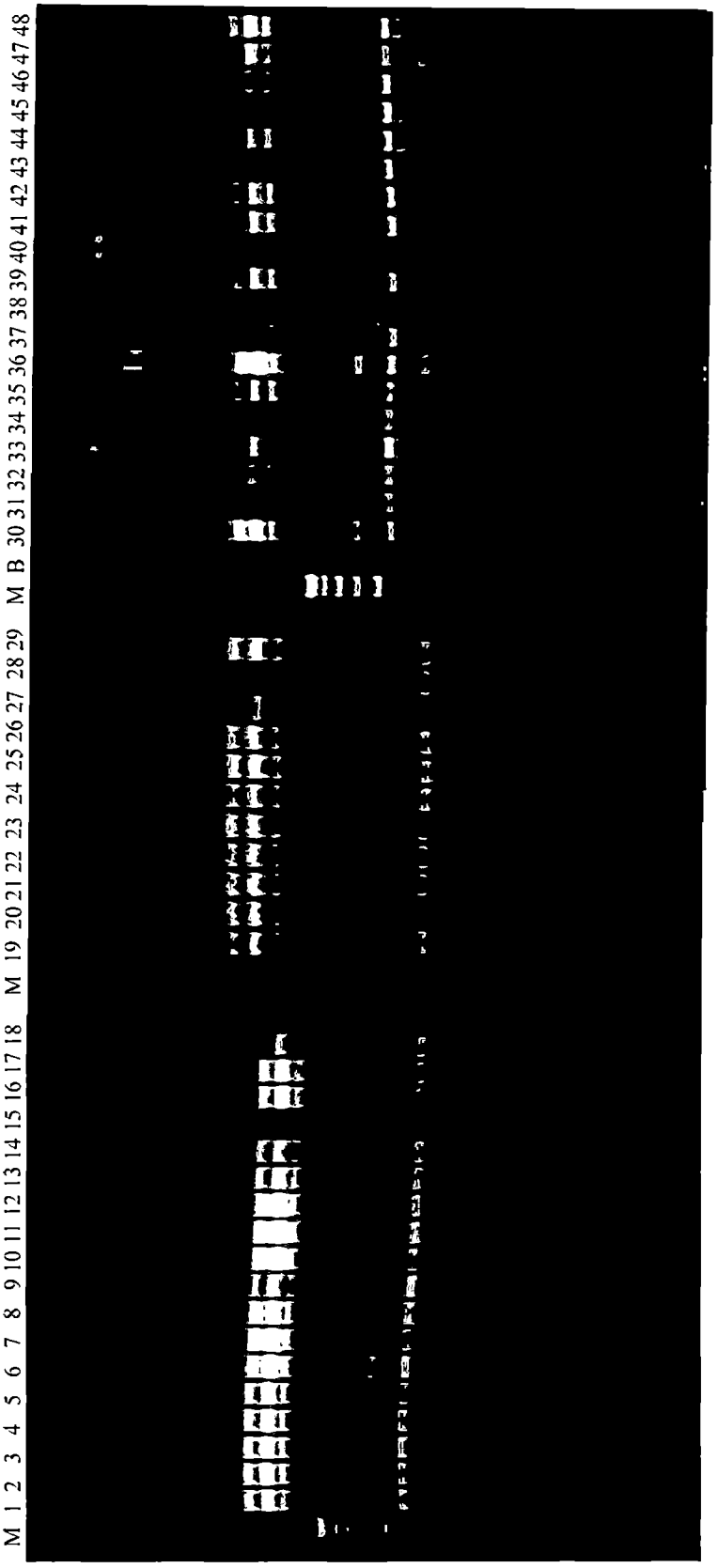



**Fig. 3.3. Ethidium bromide stained agarose gel showing RAPD profile from OPQ-09. The sample numbers are as represented in Table 3.1. M represents 100 bp marker and B represents negative control**

The accessions of cluster bean distinguished into two major clusters at 75% similarity and a third cluster at lower similarity in the dendrogram constructed using RAPD band data (Fig 3.4), the first major cluster consisted only landraces and had four sub clusters. The second and third major clusters had only commercial genotypes; the second cluster could be clearly differentiated into two sub clusters. Three landraces namely IC-116953, IC-116958 and IC-116595 did not fall into any cluster. PCA analysis (Fig 3.5) of landraces showed clustering of landraces into a single major group containing the landraces from central Rajasthan like Jodhpur, Barmer, Bhilwara and two loosely bound groups.

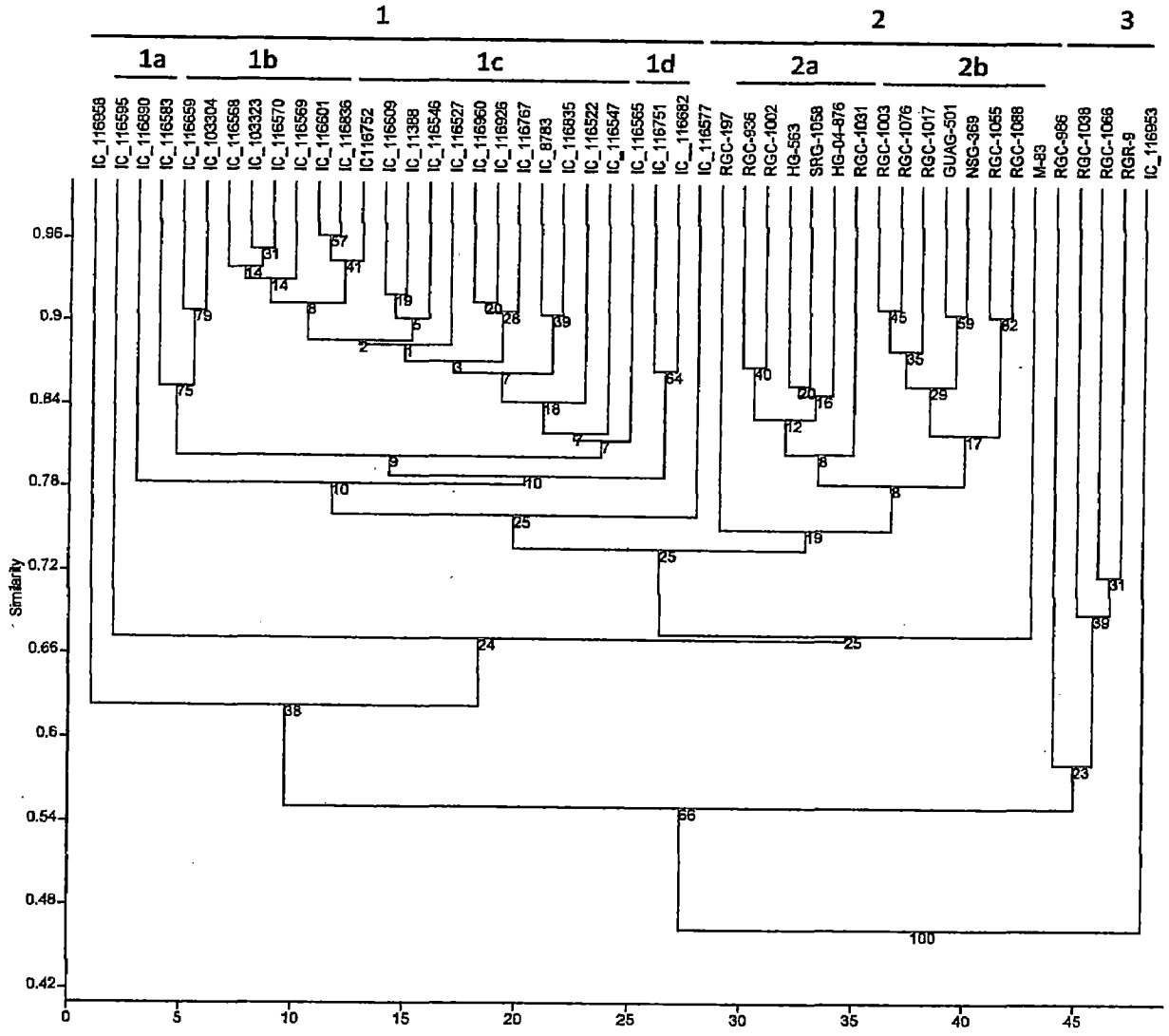



**Fig. 3.4. Dendrogram generated by UPGMA method, showing relationship between 48 accessions of cluster bean based on genetic profile from RAPD data. The numbers at the forks shows the confidence limits for the grouping of those accessions in the branch, based on 1,000 cycles of bootstrap analysis**

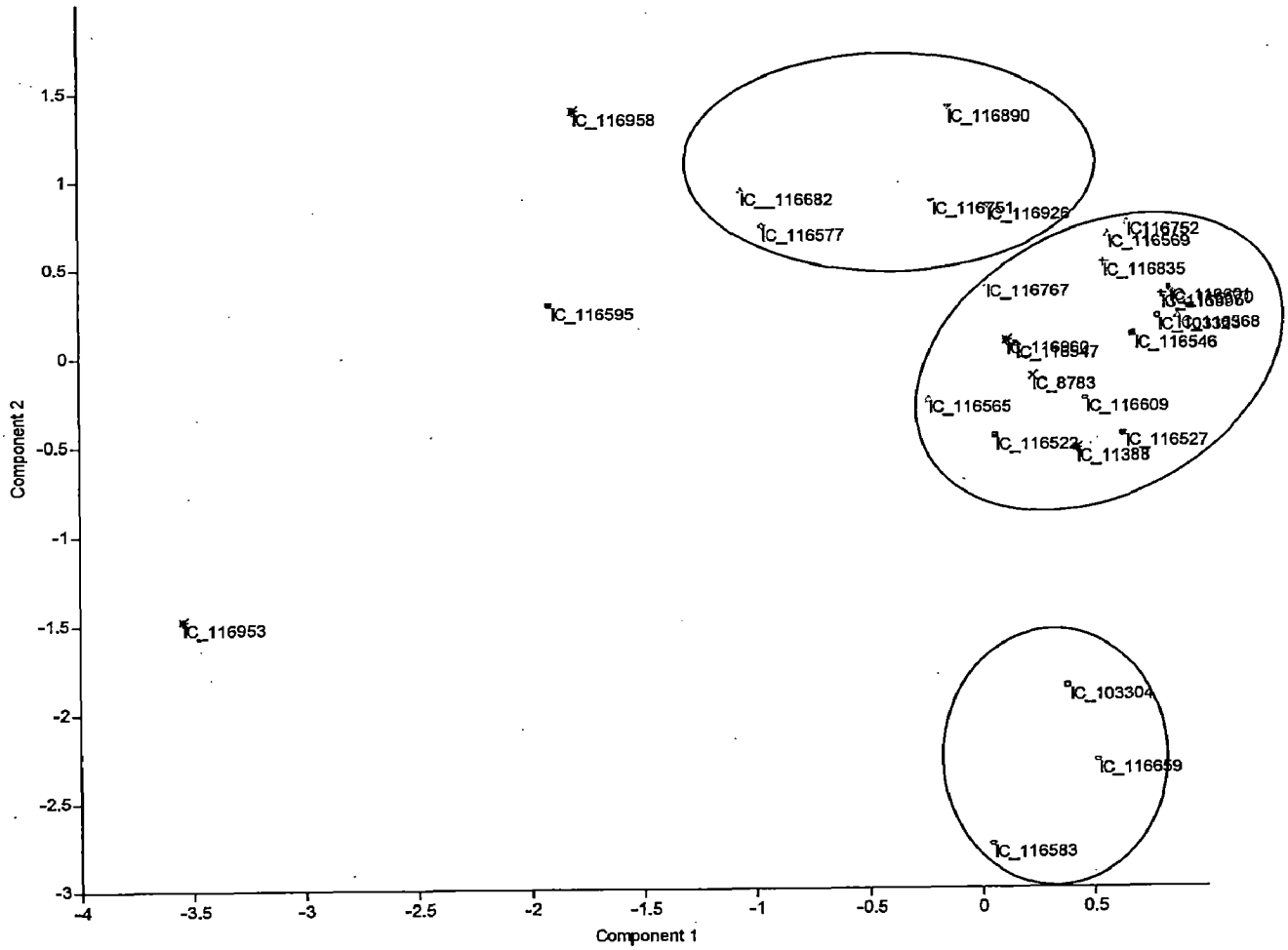



**Fig. 3.5. Two-dimensional plot of principle component analysis (PCA) of cluster bean landraces using RAPD analysis. The circles indicate the group of accessions which are similar to each other in the PCA analysis**

The observed number of alleles, effective number of alleles, Nei's genetic diversity, Shannon's information index for landraces and commercial varieties using 13 RAPD markers were found to be 1.872±0.335, 1.589±0.351, 0.333±0.170, and 0.490±0.230, respectively. The value of total genotypic diversity among population (Ht) was 0.333±0.029, whereas diversity within population (Hs) was found to be 0.283±0.026. Mean coefficient of gene differentiation (Gst) value 0.148 indicated 86.2% of genetic diversity present within the population. The estimated gene flow in the population was 2.857 (Table 3.6). AMOVA helps in understanding RAPD variation among and within the populations. Percentage of molecular variance was found to be 27% among populations and 73% variance was attributed to variance within the

population. Table 3.7 shows the $\varphi_{pt}$ value, estimated variance, percentage variance among and within the populations.

**Table 3.6. A comparative list of genetic variability factors across the accessions using RAPD primers**

| | | Observed number of alleles | Effective number of alleles | Nei's gene diversity | Shannon's Information index | Ht | Hs | Gst | estimate of gene flow |
|---|---|---|---|---|---|---|---|---|---|
| RAPD | Mean | 1.8729 | 1.589 | 0.3339 | 0.4908 | 0.3333 | 0.283 | 0.148 | 2.857 |
| | St. Dev | 0.3345 | 0.351 | 0.1709 | 0.2309 | 0.0295 | 0.026 | | |

**Table 3.7. Details of Analysis of Molecular Variance (AMOVA) based on RAPD marker data within and among landraces and commercial varieties (level of significance based on 999 iteration steps)**

| Primers | Source of variance | Estimated variance | Percentage (%) | $\varphi_{pt}$ | P(rand >= data) |
|---|---|---|---|---|---|
| RAPD | Among Pops | 4.594 | 27% | 0.271 | 0.001 |
| | Within Pops | 12.355 | 73% | | |

### 3.2.3 ISSR Analysis

The primers $(AG)_8$, $(CA)_8$ and $(AC)_8YA$ produced more bands as compared to $(GT)_8C$ and $(TC)_8RG$ primers (Table 3.8). The primers $(GAA)_6$ and $(CTTCA)_3$ produced 13 and 11 bands, respectively. The primer $(GA)_8YC$ produced unclear bands; the primer $(AT)_8C$ did not produce any band. The 7 primers used produced a total of 64 bands out of which 50 were polymorphic and 14 were monomorphic. Three unique bands were observed in accession number IC116752 with 2 bands in primer UBC-856 of about 320bp and 350bp and a single band with primer UBC-868 of about 275bp. Fig 3.6 and Fig 3.7 show the ISSR fingerprint patterns produced with UBC-808 and UBC-879 primers; the sequence of accessions in the samples in lanes is according to Table 3.1.

The dendrogram (Fig 3.8) from ISSR data showed one major cluster at 75% similarity and five minor clusters at lower level of similarity. The major cluster possessed six sub-clusters. The dendrogram did not differentiate between landraces and commercial varieties. The PCA analysis (Fig 3.9) of the landraces formed a loose group. The accessions IC116953, IC116595 and IC116682 were distant from the group.

**Table 3.8. Properties of 7 ISSR primers used and their percentage polymorphism, polymorphic information content and resolving power**

| Marker | Total number of bands | Polymorphic bands | Monomorphic bands | Percentage polymorphism | Polymorphic Information Content(PIC) | Resolving power |
|---|---|---|---|---|---|---|
| UBC808 | 10 | 7 | 3 | 70.00 | 0.179 | 2.208 |
| UBC818 | 12 | 9 | 3 | 75.00 | 0.258 | 4.500 |
| UBC820 | 6 | 6 | 0 | 100.00 | 0.266 | 2.250 |
| UBC854 | 3 | 2 | 1 | 66.67 | 0.174 | 0.625 |
| UBC856 | 9 | 6 | 3 | 66.67 | 0.081 | 1.083 |
| UBC868 | 13 | 11 | 2 | 84.62 | 0.258 | 4.708 |
| UBC879 | 11 | 9 | 2 | 81.82 | 0.273 | 4.208 |

**Table 3.9. A comparative list of genetic variability factors across the accessions using ISSR primers**

| | | Observed number of alleles | Effective number of alleles | Nei's gene diversity | Shannon's Information index | Ht | Hs | Gst | Estimate of gene flow |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |
| ISSR | Mean | 1.7812 | 1.4627 | 0.267 | 0.3988 | 0.2639 | 0.253 | 0.041 | 11.549 |
| | St. Dev | 0.4167 | 0.3844 | 0.1939 | 0.2681 | 0.0378 | 0.035 | | |

**Table 3.10. Analysis of Molecular Variance (AMOVA) based on ISSR marker data within and among landraces and commercial varieties (level of significance based on 999 iteration steps)**

| Primers | Source of variance | Estimated variance | Percentage (%) | $\varphi_{pt}$ | P(rand >= data) |
|---|---|---|---|---|---|
| ISSR | Among Pops | 0.608 | 8% | 0.082 | 0.001 |
| | Within Pops | 6.783 | 92% | | |

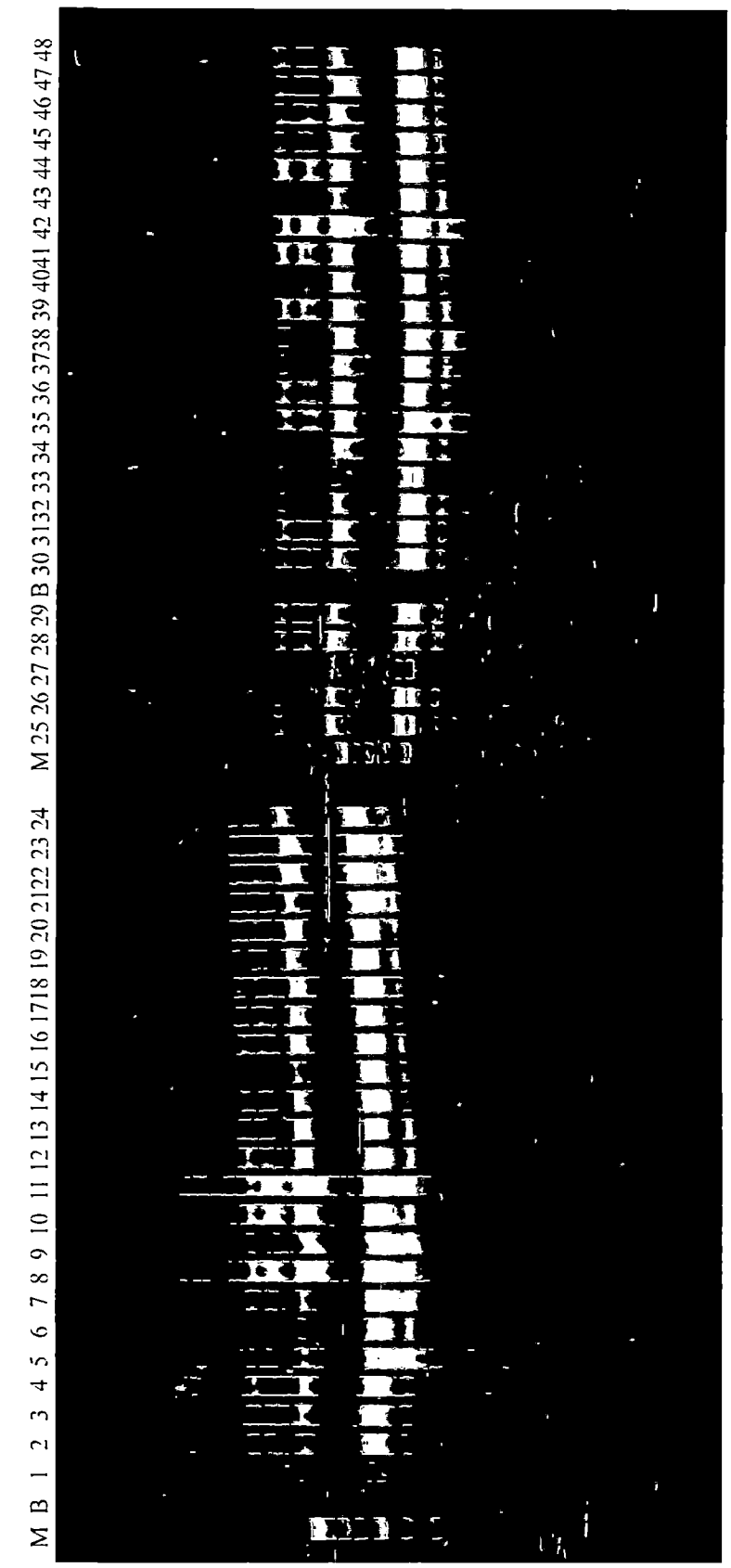



**Fig. 3.6. Ethidium bromide stained agarose gel showing ISSR profile from UBC-808. The sample numbers are as represented in Table 3.1. *M represents 100 bp marker and B represents negative control***

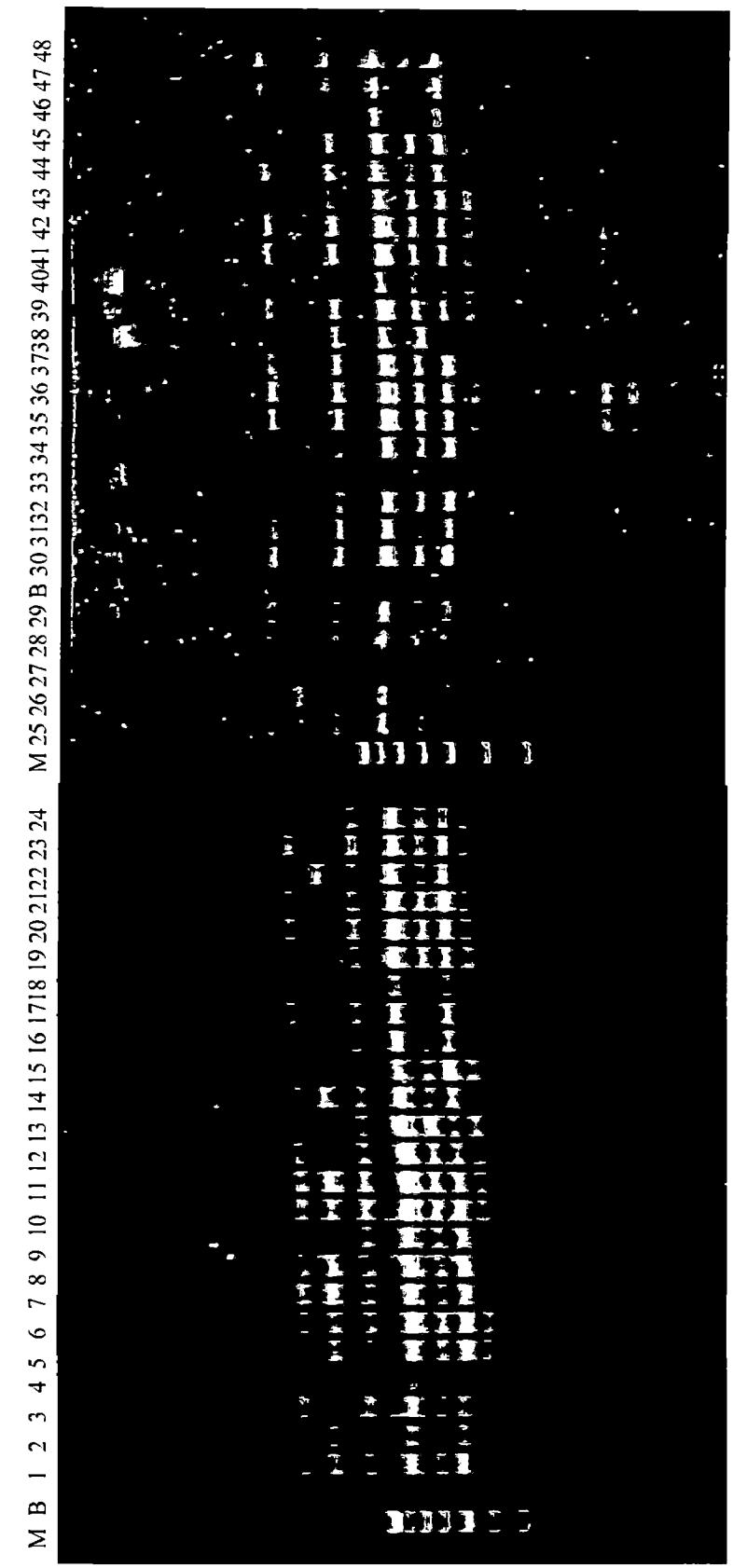



**Fig. 3.7. Ethidium bromide stained agarose gel showing ISSR profile from UBC-879. The sample numbers are as represented in Table 3.1. *M represents 100 bp marker and B represents negative control***

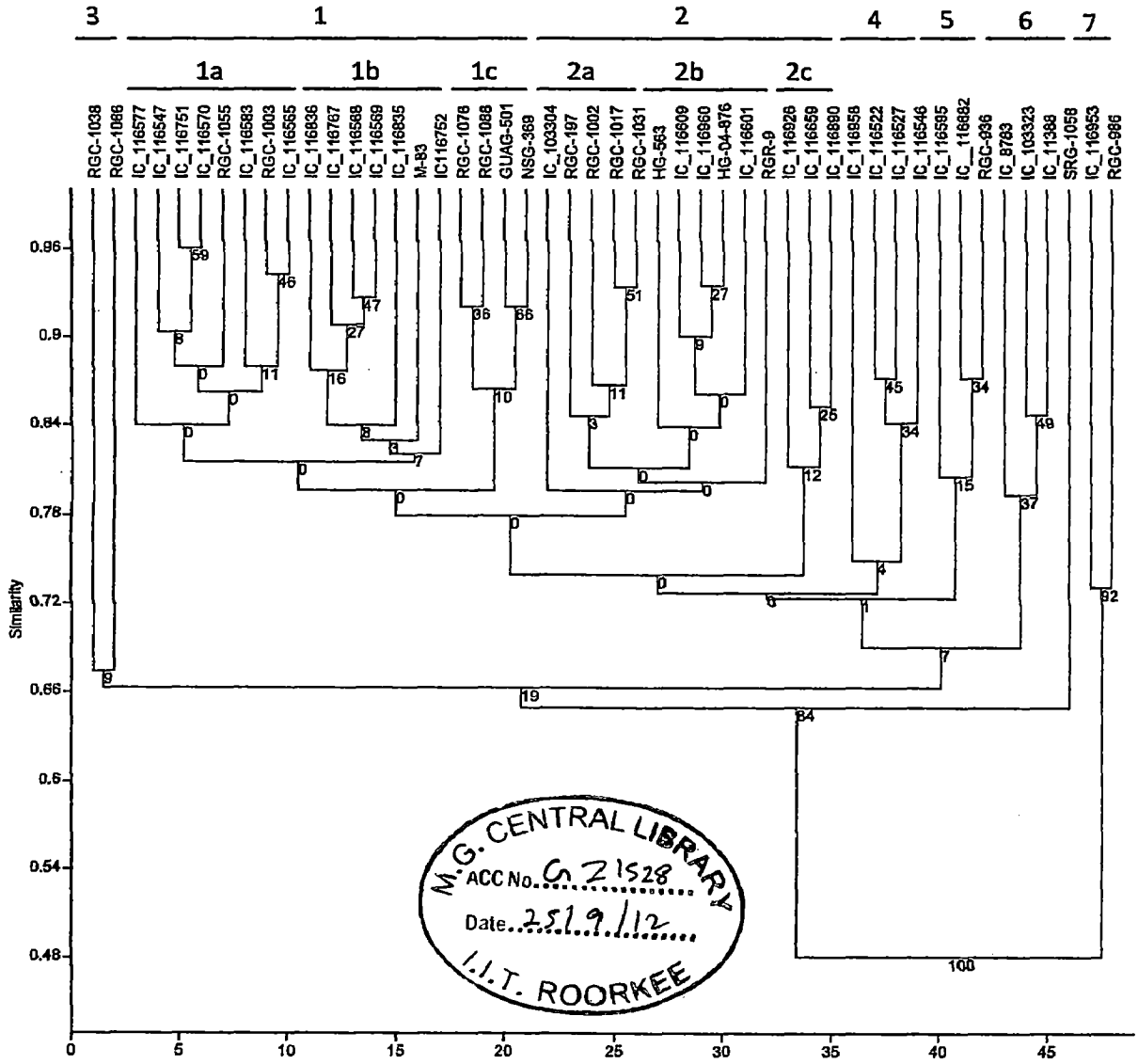



**Fig. 3.8. Dendrogram generated by UPGMA method, showing relationship among 48 accessions of cluster bean based on genetic profile from ISSR data. The numbers at the forks show the confidence limits for the grouping of those accessions in the branch, based on 1,000 cycles of bootstrap analysis**

The observed number of alleles, effective number of alleles, Nei's genetic diversity, Shannon's information index for landraces and commercial varieties using 7 ISSR markers were found to be 1.7812±0.4167, 1.4627±0.3844, 0.267±0.1939, and 0.3988±0.2681, respectively. The value of total genotype diversity among population (Ht) was 0.2639±0.0378 whereas diversity within population (Hs) was found to be 0.253±0.035. Mean coefficient of gene differentiation (Gst) value was 0.041 and the estimated gene flow in the population was

11.549 (Table 3.9). AMOVA was used to analyze variation among and within the populations. Molecular variances were 8% and 92% among and within the population, respectively (Table 3.10).

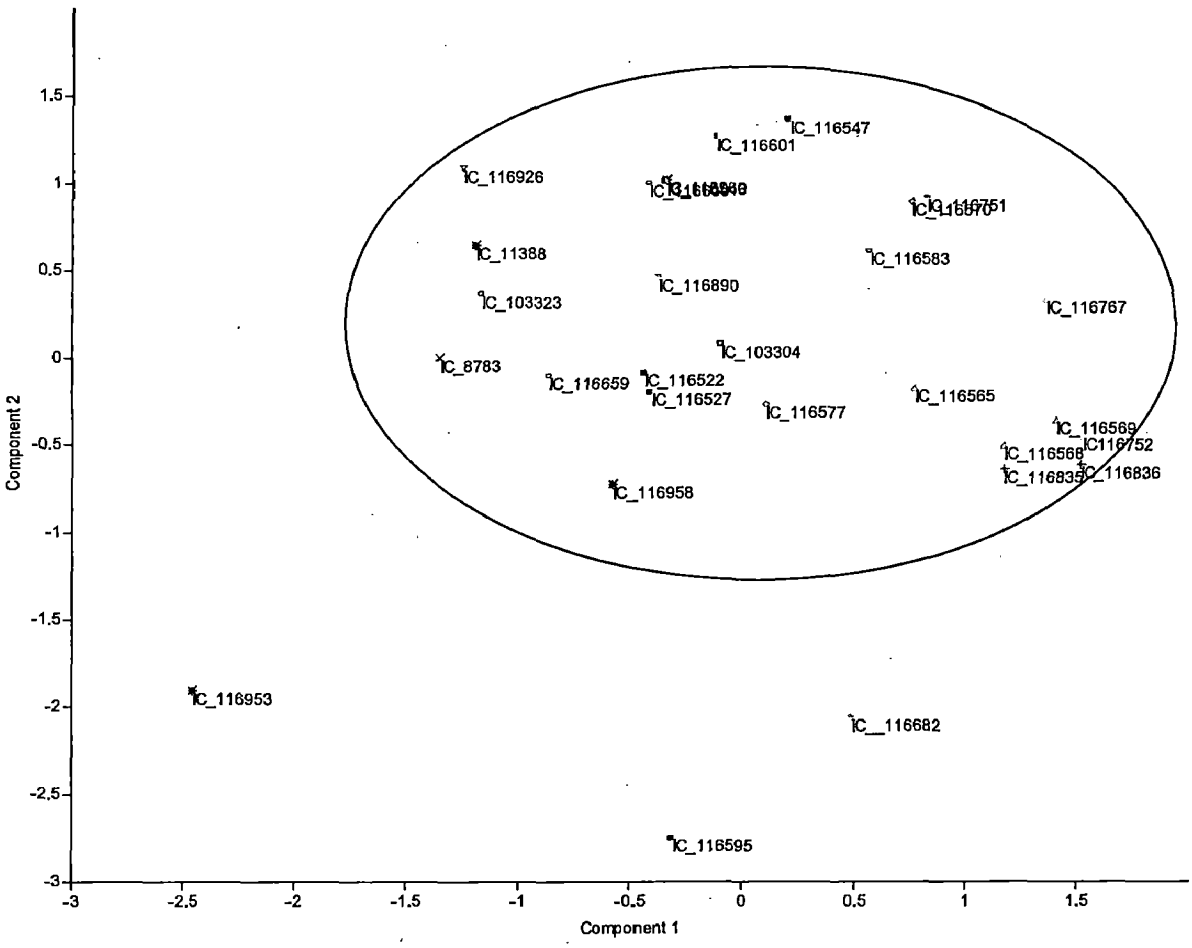



**Fig. 3.9. Two-dimensional plot of principal component analysis (PCA) of cluster bean landraces using ISSR and analysis. The circles indicate the group of accessions which are similar to each other in the PCA analysis**

### 3.2.4 RAPD and ISSR combined analysis

The accessions of cluster bean distinguished into two major clusters at 75% similarity and 3 minor clusters with lower similarity in the dendrogram constructed using the combined data of RAPD and ISSR (Fig 3.10); the 1st major cluster had only landraces and consisted of 5 sub-clusters. The second major cluster had only commercial genotypes and it consisted of 3 sub-clusters. PCA analysis (Fig 3.11) of landraces showed clustering of landraces into a two major

groups. A tightly linked group contained the landraces from Jodhpur, Jalore, Churu, Palanpur and one loosely bound group with accessions from Barmer, Mahendragarh, Nagaur, Bikaner and Rajkot.

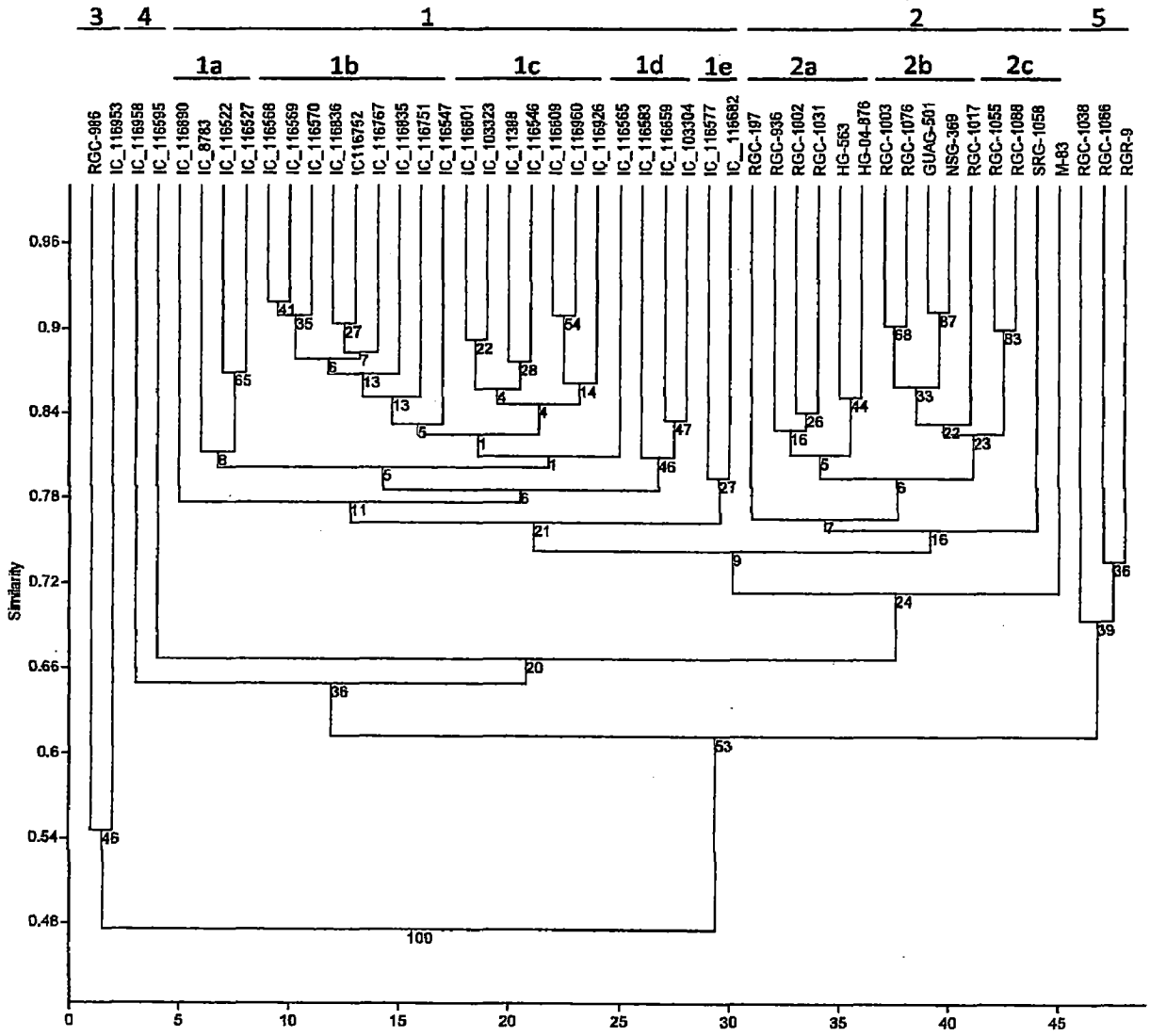



**Fig. 3.10. Dendrogram generated by UPGMA method, showing relationship among 48 accessions of cluster bean based on genetic profile from RAPD+ISSR data. The numbers at the forks show the confidence limits for the grouping of those accessions in the branch, based on 1,000 cycles of bootstrap analysis**

The observed number of alleles, effective number of alleles, Nei's genetic diversity, Shannon's information index for landraces and commercial varieties using 13 RAPD and 7 ISSR markers were found to be 1.8407±0.367, 1.5446±0.3671, 0.3130±0.1817, and 0.4584±0.2478, respectively. The value of total genotypic diversity among population (Ht) was 0.3089±0.0333 whereas diversity within population (Hs) was found to be 0.272±0.029. Mean

coefficient of gene differentiation (Gst) value was 0.116 and the estimated gene flow in the population was found to be 3.787 (Table 3.11) . AMOVA was used to analyze variation among and within the populations. Molecular variance among populations was found to be 21% and that within the population was 79% indicating higher variation within the population (Table 3.12).

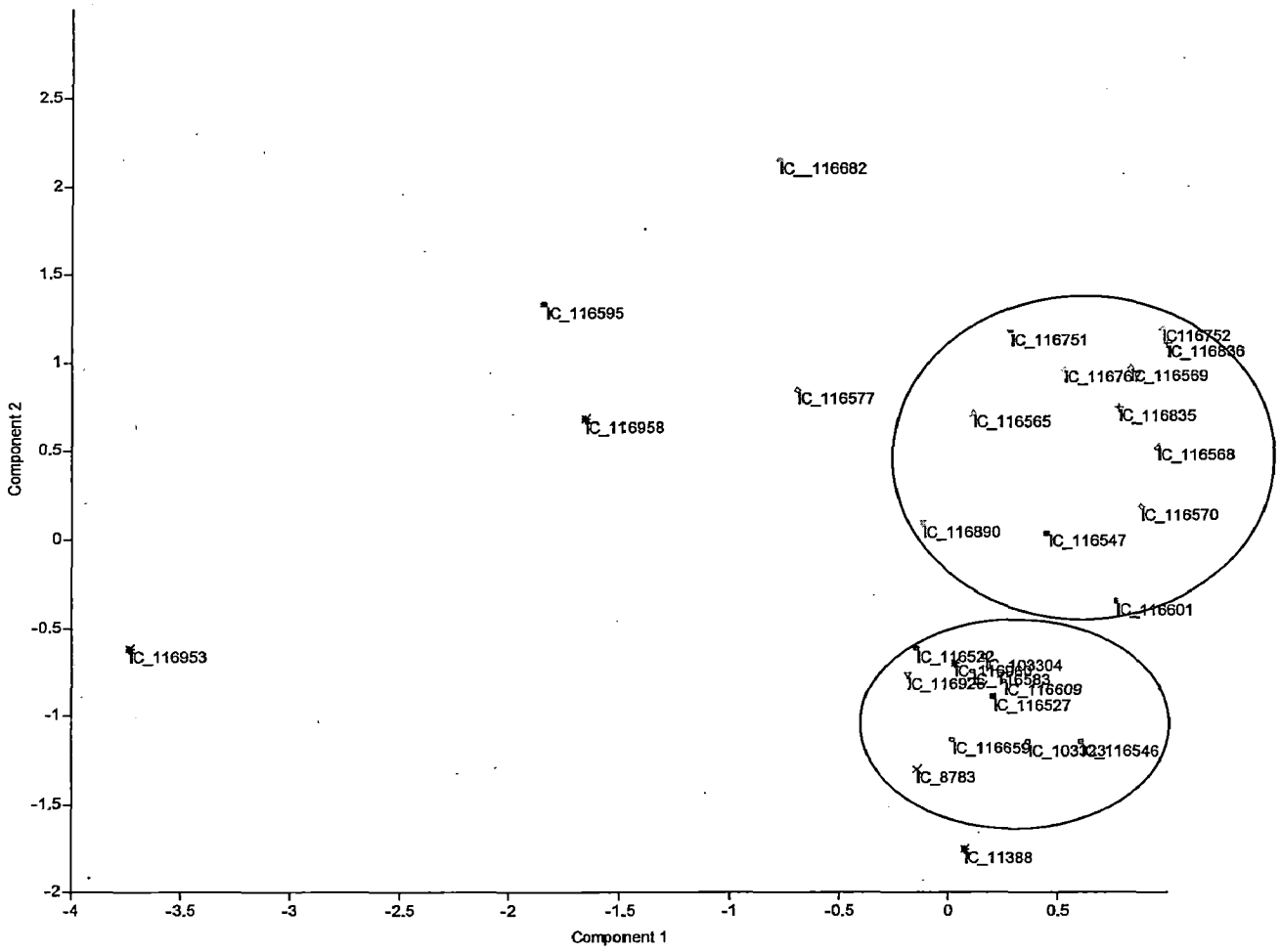



**Fig. 3.11. Two-dimensional plot of principal component analysis (PCA) of cluster bean landraces using RAPD+ISSR analysis. The circles indicate the group of accessions which are similar to each other in the PCA analysis**

**Table 3.11. A comparative list of genetic variability factors across the accessions using RAPD+ISSR primers**

| | | Observed number of alleles | Effective number of alleles | Nei's gene diversity | Shannon's Information index | Ht | Hs | Gst | Estimate of gene flow |
|---|---|---|---|---|---|---|---|---|---|
| RAPD + ISSR | Mean | 1.8407 | 1.5446 | 0.3103 | 0.4584 | 0.3089 | 0.272 | 0.116 | 3.787 |
| | St. Dev | 0.367 | 0.3671 | 0.1817 | 0.2478 | 0.0333 | 0.029 | | |

**Table 3.12. Analysis of Molecular Variance (AMOVA) based on RAPD+ISSR marker data within and among landraces and commercial varities (level of significance based on 999 iteration steps)**

| Primers | Source of variance | Estimated variance | Percentage (%) | $\varphi_{pt}$ | P(rand >= data) |
|---|---|---|---|---|---|
| RAPD + ISSR | Among Pops | 5.202 | 21% | 0.214 | 0.001 |
| | Within Pops | 19.138 | 79% | | |

### 3.2.5 Correlation between molecular, geographical and morphological data in landraces

The Mantle test revealed a significant correlation between the molecular data and the geographic data. For RAPD data the correlation values were R=0.5252, P=0.0066, for ISSR and the geographic data the values were R=0.3144, P= 0.0266 while the RAPD+ISSR data showed R=0.5303, P= 0.0076. This indicated that molecular variation corresponds to differences in geographic distribution of landraces.

There was no significant correlation between the molecular data and the morphological data. For RAPD data the values were R=0.1021, P= 0.188, ISSR data gave the correlation R=0.04437, P=0.2934. The RAPD+ISSR data with morphological data gave R=0.08897, P=0.1598.

## 3.3 Discussion

PCR using 13 RAPD primers produced a total of 118 bands whereas 7 ISSR primers produced 64 bands. Both the markers revealed genetic variability among the accessions studied. The average level of polymorphism revealed by RAPD was 87.63 which is higher than 77.82 for ISSR. Similarly, higher polymorphism for RAPD primers as compared to ISSR primers was reported by Muthusamy et al. (2008) while assessing genetic variation in rice bean [93]. PIC and Rp values of markers revealed the discriminating power of markers. PIC values ranged between 0.141 (OPM-02) to 0.419 (OPD-12) for RAPD and 0.081 (UBC-856) to 0.273(UBC-879) for ISSR markers. Rp of a marker gives a moderate idea about the number of genotypes that could be resolved by that marker [108]. The values of resolving powers for

RAPD markers were between 0.917 (OPX-12) and 6.333 (OPN-1) while that of ISSR markers ranged between 0.625 (UBC-854) and 4.708 (UBC-868); indicating better resolving capability of RAPD over ISSR. However, in cowpea (*Vigna unguiculata* (L.) Walp.) landraces more polymorphic loci were detected with ISSR than with RAPD fingerprinting [46].

Shannon's Information Index is a commonly used index to characterize gene diversity in a population. The values of the Shannon's information index were similar for RAPD and RAPD+ISSR data but were low for ISSR data. The values of total genetic diversity among population (Ht) also showed similar pattern as that of Shannon's information index. Nei's measure of the average gene diversity per locus within population (Hs) was similar for RAPD (0.283±0.026), ISSR (0.253±0.035) and RAPD+ISSR (0.275±0.029) data showing that both the markers revealed similar level of genetic difference within the genotypes.

AMOVA analysis showed higher polymorphism within the population (73% for RAPD, 92% for ISSR and 79% for RAPD+ISSR combined data). AMOVA results also showed that RAPD was more efficient in detecting the variation between populations.

Cluster analysis was done on RAPD, ISSR and RAPD+ISSR data. The results based on RAPD and RAPD+ISSR data grouped landraces and commercial varieties separately showing the presence of distinguishable genetic difference existing between them. The accession IC116958 formed an out-group singlet with RAPD data. Similar results with singlet out-group in dendrogram were reported by Punia et al. (2009) [109]. ISSR data produced a cluster with commercial varieties along with the landraces. The difference in the clustering pattern can be partially attributed to the number of amplified loci analyzed (118 for RAPD and 64 for ISSR). Similar difference in clustering pattern between RAPD and ISSR markers was observed in genotypes of *Jatropha curcas* because RAPD and ISSR detected polymorphic DNA by amplifying different regions of genome [51].

Principle component analysis (PCA) of the landraces showed that the accessions belonging to central Rajasthan grouped together indicating genetic similarity. The Mantle test between the molecular data and the geographic data of the landraces showed significant correlation. This indicated that the molecular variation corresponded to geographic distribution of landraces.

# *Chapter IV*

## *Mining of EST-SSR markers for amplification and polymorphism among guar accessions*

# 4. Mining of EST-SSR markers for amplification and polymorphism among guar accessions

## 4.1 Materials and methods

### 4.1.1 Preparation of EST sequences for analysis

The EST accession numbers EC616385.1 to EG991296.1 were downloaded from dbEST (http: / /www.ncbi.nlm.nih.gov/dbEST/) of NCBI. All the EST's were from developing guar embryos. Available *C. tetragonoloba* EST (16,476) were trimmed using EST trimmer. The filtered sequences (16108) were assembled into contigs using CAP3 [57] to remove redundancy in the sequences.

### 4.1.2 Plant material and DNA extraction

The accessions of *C. serrata* and *C. senegalensis* were kindly provided by Chaudhary Charan Singh Haryana Agriculture University (CCSHAU), Hisar. The cultivars of *C. tetragonoloba* were obtained from Central Arid Zone Research Institute (CAZRI), Jodhpur, Rajasthan. DNA was extracted from the field grown plants as described in section 3.1.7.

### 4.1.3 Mining simple sequence repeats and primer designing

The sequences so obtained (16,108) were assembled into contigs using CAP3. The candidate SSR containing sequences in the ESTs, assembled contigs and singlet sequences were mined using the PERL script MIcroSAtellite analyzer (MISA). The following default definements were set for microsatellites (unit size / minimum number of repeats) : (1/10) (2/6) (3/5) (4/5) (5/5) (6/5) in MISA. Maximal number of bases interrupting 2 SSRs in a compound microsatellite were set to 100. Fig 4.1 shows the work flow followed for screening microsatellites. The sum of the microsatellite repeats from the contigs and singlets were used for designing primers. Primers flanking the SSR regions were designed using Primer3.

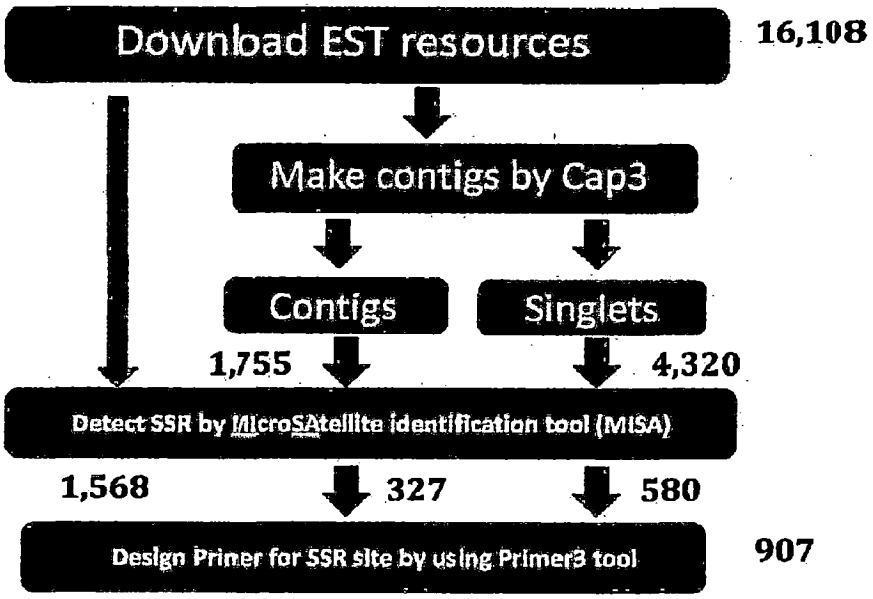



**Fig. 4.1. Schematic work flow followed for screening microsatellites**

### 4.1.4 Polyacrylamide gel electrophoresis of DNA

#### 4.1.4.1 Preparation of 40% Acrylamide-bis acrylamide solution

Acrylamide (38g) and bis acrylamide (2g) were weighed and dissolved in 50ml of double distilled water. The volume was made up to 100 ml by double distilled water.

#### 4.1.4.2 Preparation of 10X TBE

The composition of TBE buffer is given in Table 4.1.

**Table 4.1. Composition of 10X TBE buffer [129]**

| S.No. | Name of the reagent | Composition |
|---|---|---|
| 1 | Tris | 10.9g |
| 2 | Boric acid | 5.56g |
| 3 | EDTA | 0.98g |
| 4 | Distilled water | 100ml |

#### 4.1.4.3 Procedure

The composition of PAGE gel is given in Table 4.2.

**Table 4.2. Composition of 8% PAGE gel**

| S.No. | Name of the reagent | Composition for 150 ml |
|---|---|---|
| 1 | 40% acrylamide-bis acrylamide solution | 30ml |
| 2 | 10X TBE | 7.5ml |
| 3 | Ammonium persulphate | 0.105g |
| 4 | TEMED | 125µL |
| 5 | Double distilled water | 92.38ml |

All the constituents as mentioned in Tabel 4.2 were taken. Ammonium persulphate and TEMED were added just before pouring. After pouring the comb was fixed in gel and allowed to solidify for about 1 hour [129].

### 4.1.5 Silver staining

#### 4.1.5.1 Preparation of Solution 1 (Fixative solution)

The composition of fixative solution is given in Table 4.3.

**Table 4.3. Composition of fixative solution**

| S.No. | Name of the reagent | Composition |
|---|---|---|
| 1 | Methanol | 20ml |
| 2 | Glacial acetic acid | 1ml |
| 3 | Distilled water | 179ml |

#### 4.1.5.2 Preparation of Solution 2 (Staining solution)

The composition of staining solution is given in Table 4.4.

**Table 4.4. Composition of staining solution**

| S.No. | Name of the reagent | Composition |
|---|---|---|
| 1 | Methanol | 20ml |
| 2 | Glacial acetic acid | 1ml |
| 3 | $AgNO_3$ | 0.2g |
| 4 | Distilled water | 179ml |

#### 4.1.5.3 Preparation of Solution 3 (Developing solution)

The composition of developing solution is given in Table 4.5.

**Table 4.5. Composition of developing solution**

| S.No. | Name of the reagent | Composition |
|---|---|---|
| 1 | NaOH | 5.1g |
| 2 | Formaldehyde | 600µL |
| 3 | Distilled water | 199.4ml |

#### 4.1.5.4 Procedure for silver staining

After gel electrophoresis, PAGE plates were disassembled. The gel was carefully placed in the staining tray. The gel was treated with fixative solution for 5 min with gentle rocking. After fixing the DNA the fixing solution was decanted. Similarly the gel was incubated in staining solution for 5 min. The staining solution was decanted and the gel was washed gently with distilled water to remove excess silver nitrate on the gel and tray. The gel was then treated with developing solution for visualizing the bands. After the visualization of bands the developing solution was replaced by fixative solution for increasing the depth and sharpness of bands.

### 4.1.6 DNA amplification conditions for SSR markers

Five accessions viz., M-83, RGC-1066, RGC-1002, C. *serrata* and C. *senegalensis* were used for SSR analysis. Polymerase chain reaction (PCR) was performed in 5µl reaction volume containing 0.5U Taq DNA polymerase (Biotools), 2.5mM $MgCl_2$ (Biotools), 5mM dNTPs, 0.5µL of 20pM primer each and 50ng of template DNA. Amplification was carried out in a Mastercycler gradient programmable thermal cycler (Eppendorf). The PCR was programmed with initial denaturation step at 94°C for 4 min, followed by 30 cycles of 30 sec at 94°C, 30 sec at 55°C and 30 sec at 72°C. A final extension was carried out at 72°C for 5 min and a hold temperature of 4°C at the end [129].

### 4.1.7 Gel electrophoresis and recording

PCR amplified products were electrophoresed on 8% PAGE gels and silver stained. The gel was run at 5V/cm for 5 to 6 hours. A 100 bp DNA ladder was used as a molecular marker to

determine the approximate size of the fragments. The gel was visualized under white light and documented in gel documentation unit (Bio-Rad).

### 4.1.8 Data analysis of SSR markers

The SSR bands were scored according to their allele size. The binary data were recorded using '0' for absence of band at expected product size and '1' , '2' or '3' for different alleles. A similarity matrix was computed for the data using Euclidean coefficient by the software package PAST [54]. A dendrogram was obtained by Unweighted Pair-Group Average (UPGMA) method using the similarity matrix by Euclidean similarity coefficient [62] to determine the relatedness of 5 accessions under study. Bootstrapping was done with 1000 repetitions to evaluate the clusters formed.

## 4.2 Results

### 4.2.1 Data mining for simple sequence repeats

A total of 16,108 sequences were analyzed for SSRs. The contig making program CAP3 produced 1,755 contigs and 4,320 singlets. The contigs showed 327 microsatellite repeats from 276 sequences of which 28 of them were in compound formation. Singlets had 580 microsatellite repeats in 506 sequences of which 41 were in compound formation. The sum of the microsatellite repeats from the contigs and singlets (907) from 3.7 million base pairs. The results suggest that there is about one SSR in every 4.1 kb of expressed sequences in cluster bean. The 782 sequences that contained the SSRs were used for designing primers.

The abundance of different repeat types were also analyzed. Mononucleotide repeats (435) were the most abundant among the SSR types in *C. tetragonoloba* ESTs, followed by di-nucleotide repeats (189). Fig 4.2 shows the abundance of different kind of repeat types in cluster bean EST's, contigs and singlests. Table 4.6 shows the frequency of different kinds of repeats in cluster bean ESTs.

**Table 4.6. Frequency of different kinds of simple sequence repeats in cluster bean ESTs**

| | Total EST | Contigs | Singlets |
|---|---|---|---|
| **A/T** | 821 | 167 | 262 |
| **C/G** | 7 | 2 | 4 |
| **AC/GT** | 17 | 3 | 13 |
| **AG/CT** | 157 | 39 | 83 |
| **AT/AT** | 100 | 17 | 34 |
| **AAC/GTT** | 23 | 9 | 5 |
| **AAG/CTT** | 108 | 29 | 60 |
| **AAT/ATT** | 172 | 24 | 39 |
| **ACC/GGT** | 24 | 6 | 12 |
| **ACG/CGT** | 1 | | 1 |
| **ACT/AGT** | 10 | 1 | 8 |
| **AGC/CTG** | 23 | 6 | 12 |
| **AGG/CCT** | 12 | 3 | 8 |
| **ATC/ATG** | 44 | 12 | 14 |
| **CCG/CGG** | 4 | 2 | |
| **AAAC/GTTT** | 2 | 1 | |
| **AAAG/CTTT** | 9 | 1 | 4 |
| **AAAT/ATTT** | 8 | 1 | 6 |
| **AAGC/CTTG** | 1 | | 1 |
| **AAGG/CCTT** | 1 | | 1 |
| **AATC/ATTG** | 1 | | 1 |
| **AATT/AATT** | 6 | 1 | 1 |
| **ACAT/ATGT** | 3 | | 3 |
| **AGAT/ATCT** | 1 | | 1 |
| **AGCC/CTGG** | 2 | 1 | |
| **AGCT/AGCT** | 1 | | 1 |
| **AAAAT/ATTTT** | 2 | 1 | |
| **AAAGG/CCTTT** | 1 | | 1 |
| **AAGAG/CTCTT** | 2 | | 2 |
| **AATCG/ATTCG** | 1 | | 1 |
| **AATTC/AATTG** | 2 | 1 | |
| **ACAGC/CTGTG** | 1 | | 1 |
| **AGCCTG/AGGCTC** | 1 | | 1 |

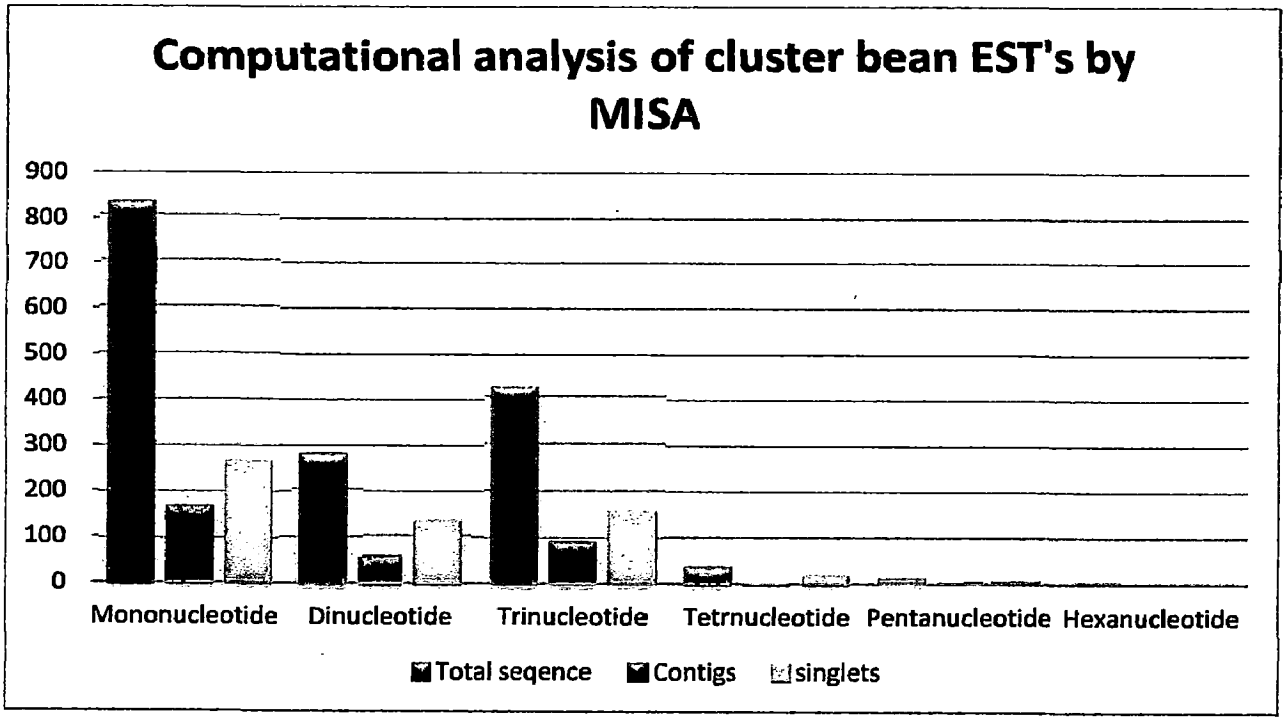



**Fig. 4.2. Graph showing the abundance of different kinds of repeat types in cluster bean EST's, contigs and singlets**

### 4.2.2 Primer designing

Flanking primers were designed for SSR containing sequences (except mononucleotide repeats) using the online tool Primer3. The details of the sequence, motif type, SSR length and primer sequences are given in Appendix I. A total of 362 primers were designed out of which 226 were synthesized. Table 4.7 shows the names and sequences of the primers synthesized.

**Table 4.7 Name and sequence of primers synthesized**

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| IITR N1F | CGAGGCAACAATATCATCTTCA | IITR N1R | TTTGGGATCGATCAAAGTGA |
| IITR N2F | CTTTCCCTTCCTTGCTTCCT | IITR N2R | TGAACCCAAAACACCAAACC |
| IITR N3F | CCACCACTTAGGTCGTGCTT | IITR N3R | CAAAAGCAAGGGAAAGGTCA |
| IITR N4F | CTCCCCTTCCACTTTCCTTC | IITR N4R | AAGAACAAGATGTTGGGCATTT |
| IITR N5F | TGCAAGTGACAACAGATTTGC | IITR N5R | ACGGCATTATCCGCTACAAA |
| IITR N6F | AAGGAAATTGCTCAGGCTCA | IITR N6R | CCCAAAGGCCAACATAATTG |
| IITR N7F | AGGAGAAGGCTTGGGATGAC | IITR N7R | TTCCACCATGTGAAAAAGCA |
| IITR N8F | CGGAAATCAACATCAACGAA | IITR N8R | TGTGACTTGGAAGCCCTTTT |
| IITR N9F | TCAGCCACAATGACACCAGT | IITR N9R | CTTCGTCGTCCGACACTACA |
| IITR N10F | ACTATGAATGCCGTGGAAGC | IITR N10R | AAAACAAACGCACCACACAA |
| IITR N11F | AGGAGAAGATGCTCCCAACA | IITR N11R | TGAGCATCCATGAAAGCAAG |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| IITR N12F | TACCGTGTTGGTGACCTTGA | IITR N12R | ATGTCCAAACCCATCTCCAA |
| IITR N13F | CACGAGGATTTCACTCACACA | IITR N13R | AGGAGGGAGTTCAGGTTGGT |
| IITR N14F | AAAGAAGAAAGATCCTTGTTTGGA | IITR N14R | AAACAAAAATAACCGCTTGATGA |
| IITR N15F | CTTTGCTTCAAAGGGTTTCG | IITR N15R | GCTCTGTCATTGTGGGAAAAA |
| IITR N16F | TCGAATTGAAGGCAAAGAGC | IITR N16R | GTAGCTGAGTGCCCTGGAAG |
| IITR N17F | CAACCAGAAACATGAAGCACA | IITR N17R | CCTGTTATGCCTGCTGAGTG |
| IITR N18F | CGAGGGAAAGAGTTAAGAGAAAA | IITR N18R | CCATAGCCAAAAACCCAGAA |
| IITR N19F | CTGCCTCAACAACCTCAACA | IITR N19R | CGAGCAAACCAAGGAAGAAG |
| IITR N20F | CCAAAGATCAACAACAACCAGA | IITR N20R | CCAACAGTGGCTTGCTTGTA |
| IITR N21F | TCACCGAGTTATGTGGGTTTC | IITR N21R | CCCACAACCAAATACCCAAC |
| IITR N22F | GCCGGAGTATCCATCAAAAC | IITR N22R | AATCTTGTCAACCCCCTGTG |
| IITR N23F | CATGATTGCGTCAAAACTCG | IITR N23R | CCCAGGGGAGGGTTACTAGA |
| IITR N24F | TACCACCCTAGGCCCTAACC | IITR N24R | GGCCCCATTTTTATTTTGTTT |
| IITR N25F | TAAGTGCAATTGTGCGAAGC | IITR N25R | AAGCACTTAGAGTGTGCTTTGTTG |
| IITR N26F | TAATCCATGTGCCAACAAGC | IITR N26R | TCTCCTTGCTCACCCAAAAT |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| CT001F | CTGACCAAAGCCTTTTCCTG | CT001R | TGATGACGATGGAGATGGAA |
| CT002F | ACTGAAATGGTGGCAGAAGG | CT002R | TGTGAAAGCGAAGAAAGGTG |
| CT003F | TCCCTGGCTCTTCAATCACT | CT003R | TTTGGGGACGACCTATTTGA |
| CT004F | CACGAGGTTCACACTTTCCTT | CT004R | TGATCCGAGAAATCCTCCAC |
| CT005F | CCAGAGGAGTGAATTAAGAAACAAA | CT005R | GGGTTCATTCCAGATTGCAT |
| CT006F | GGTTCCCTCTGTTCAAGCAA | CT006R | TGGTCACAGAGCTTGTCAGG |
| CT007F | CACGAGGCAGAGACTTCACTC | CT007R | GTCAAGGGGCATTTGTCAGT |
| CT008F | GAGGCATAAAGTCCCTTCCA | CT008R | CTGCCCTTCACACTTTCGAT |
| CT009F | TGAGAATTCAAAGGAGCAAGTG | CT009R | TGCTGCTTCTTGTTCTGCTC |
| CT010F | AATGCAGAAGCAACGGAAAT | CT010R | CCTCTCGCCTTCTGTTCTGT |
| CT011F | CCCCACACTCACTCACACAC | CT011R | GTTCGGTCAAACCTTCATGG |
| CT012F | ACAAGGAGCCACTCGTTCAG | CT012R | CACGCACGCATGATACATAG |
| CT013F | CACGAGGCAGGAACAGAACT | CT013R | TCTCGCTTGGACATTCTTCA |
| CT014F | TTCCGATCCTGTTTGTTTCC | CT014R | CGTGTCTGTTACCGGATGTG |
| CT015F | TGGCAGAAATGGAAAATTGA | CT015R | TGAGGTTCATTCTCCACCAA |
| CT016F | CAAAACAAAACAAAACAAGAAACAA | CT016R | TACCGCTTGAAGACCAAACC |
| CT017F | CACCACCACAAGCACAGATT | CT017R | TTCCAGTCCGAAAATTCCAC |
| CT018F | TTGCTGTTGCTGTTGAGACC | CT018R | AGCAAAGCCAGCATCAGC |
| CT019F | TGATGGAGGCACCTTCAGAT | CT019R | CCAAGCCGAGTCTTTCTTGA |
| CT020F | CCACCTTCCATGAACCTTGT | CT020R | CACCAACAATGATGCTTCCA |
| CT021F | GAGAAGACGATCCGTTACGC | CT021R | CACATGCATGATTCAGACCA |
| CT022F | CTCAGCCTCAGCCACTCACT | CT022R | TCAAGGTCACCAACACGGTA |
| CT023F | GCCATCAAGGACATCTCCAG | CT023R | GATCACACACACAAATAACTAAGATTT |
| CT024F | TTTCCGAGATGGTGAGAAGG | CT024R | CCTCCCGTTTGTTTTCTTCA |
| CT025F | GAAGAGTGCCCTTGCTATGC | CT025R | CACCACAAAATGTGGAGATACA |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| CT026F | GATTGGGACCCACAATTCAC | CT026R | CATGAACAAGGCTCGTGAGA |
| CT027F | TCGATCATCCAAGAGGGACT | CT027R | TCCAATACCACCAACCCTTG |
| CT028F | ATAAGGGGTCCCCATCATCA | CT028R | GCCAAACAAGAAGTGACAAACA |
| CT029F | AGCAACTGCTGATGATGGTG | CT029R | AGGCAGATGAACTTGCCATC |
| CT030F | TTCCAGCTCTTGTTGTCCTCT | CT030R | TGGCTGGATGGATTGTCTTA |
| CT031F | GCCCTGCACTTTTGTGTTTT | CT031R | ATTTATGCTCTTCCCGCAAA |
| CT032F | TCGCGACATGAACTTGTTTC | CT032R | TCGCCTTGAGACAACCTACC |
| CT033F | AGATCATGGCAAGGCTTTTG | CT033R | CCGCTATAATGGGCATCTGT |
| CT034F | CACCAACACCAAGACTGCTT | CT034R | CGCGCAGTTTAATTTCATCA |
| CT035F | GCACCATTTGGGCTTAATTATC | CT035R | CCTGATTATTTCCTCACACCCTA |
| CT036F | CCCATGTGCCTTTGTTTTGT | CT036R | CAGGGCTCCCTCTTTTCTCT |
| CT037F | CACTGCCAGAATGGAATGAA | CT037R | CCATTCATTTTCGAATTCACC |
| CT038F | CATCATGTTTGAGCCACGAC | CT038R | GCAATGGACCCTATCCTCCT |
| CT039F | AGTTGCATTGCACAGGTTTG | CT039R | TGGCCAAATTACAAGTAGAACAA |
| CT040F | TTTTATCCGCATCCCAACAT | CT040R | GAGAGGATGGAAGGTTGAAGC |
| CT041F | TTCCTTTGGAATCCACATCC | CT041R | ACCCAAATCACATCCACACA |
| CT042F | GGCTCTGATGCTTTCCAACT | CT042R | AGGAAATGGATGACGACGAC |
| CT043F | CTTTTACCCATTCGGCCATT | CT043R | GTCATTGGGTCCATGGAAGT |
| CT044F | TCCCACCCTCATCTTCTTTC | CT044R | TCACTTCCTCGTCGGTTCTC |
| CT045F | CACGAGGCTTCTTCTTCATCA | CT045R | GCTTTGGTGAGTGGTTGGTT |
| CT046F | CCCTGGTGCTAGAAGAGTGG | CT046R | TAGGTAGGGGGCTGGACTTT |
| CT047F | CCGTGGCCAATAGAAAGAGA | CT047R | GCTGTTCCATTCACCTCCTC |
| CT048F | GCCGAGAGCAGAGAAGAAGA | CT048R | TCCCTGGAAGTGATGGTAGG |
| CT049F | CTGGTGACAGTCGTTGCATT | CT049R | AGGCAACTTAACCCATGTCG |
| CT050F | AGGGGTCAAGGGTGAAGACT | CT050R | GTTCCCATCCAATCCAACAC |
| CT051F | TTTTCTACACAAATGAAGATATTAGCA | CT051R | GCTCCTGCTAGTCTGTCAATCA |
| CT052F | GCTCTTCCCTCCCTCAATTC | CT052R | GGGTGACAGCGAAGAGTAGG |
| CT053F | ATGGAATTCACTTGGCTTGC | CT053R | GCCAAAGTTGCGCAATAAAG |
| CT054F | GTCATGGCTTGTTGGGATCT | CT054R | GTGGGAACAAGAGGACCAAA |
| CT055F | TTCCCTCAGGCTGACAAAGT | CT055R | CCCCATGTCCTCATCAGACT |
| CT056F | GATGCCATGAGTGGGAAAGT | CT056R | CACGAGGCTCTGTTTCTGTG |
| CT057F | TCGCTCTACTGGCAGACCTT | CT057R | TAGAGGCAGTCAAGCCGAAT |
| CT058F | TCCCGCTAATTTAACCTCCA | CT058R | CGCGATCAGAATAACAAGCA |
| CT059F | AAAGCAATTTGCCCTTGCTA | CT059R | AGGAGCCCATAAACACAAGG |
| CT060F | GCACTTCGAGGGGATTCA | CT060R | GGCTTATGGCTGTTGTGGTT |
| CT061F | CGAGGCTTCTAAGTTCATGG | CT061R | AAAGAGCCATACCCACATGC |
| CT062F | ACCGTCTTGAGCCAACTGTC | CT062R | CAAAGGATCCAATGCCAAGT |
| CT063F | CCCTCTCTTTCGAGGGTTTT | CT063R | AACCTCTGTGGCTGCATCTT |
| CT064F | ATGCCTAATGGAGGGTCCTT | CT064R | TTTACCCTCCTCCTCCCCTA |
| CT065F | ACCCTACTGCTTCCCCATTT | CT065R | CCCTTATGGCTCTTGCTACG |
| CT066F | CCCCATTTAAGATATGAAAGAACTC | CT066R | TCAACATGGAGCCAGCAGTA |
| CT067F | CCTTTGTGGAGGCAGAACAT | CT067R | GAAGGGTATGGTGTCCCAGA |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| CT068F | CTGGCTCCATGTTGATGATG | CT068R | CAGCCTTGAGGCACTTCTCT |
| CT069F | TGGCTGAAAAGCTGGTTCTT | CT069R | CAGGATGCTTTTACATTGTCCA |
| CT070F | TGTTTGGGCAATAACAAGGA | CT070R | TGGGAAATATCCGAAGCAAG |
| CT071F | TTCTGATTTTCTTTTGGCTCAAG | CT071R | TGCTGTGAACTTGGGTGAAG |
| CT072F | TTTGACATTGGACTCGCTCA | CT072R | TGAAAAACAGGGTGGCATCT |
| CT073F | GCTGCTGCAACCAACAAGTA | CT073R | CAGCAACGTTTTCACTTCCA |
| CT074F | TCCAAACATTGGCAAAACAA | CT074R | TCAAGCATGAAGCTCTTGGA |
| CT075F | GGAAAGGCAGCACCATTG | CT075R | GAGAGTGGACTTCCCGTGAT |
| CT076F | GCGAGCGATCTCACTCTTCT | CT076R | AGTTCGAAGCTCCGATAGCA |
| CT077F | ATGTTGCAGTTTCCCCACTC | CT077R | CAGCCAATGGAGTGAGATGTT |
| CT078F | TTCACAATTCACCTCCGTTCT | CT078R | TCATTCAAACCAGCTGTGGA |
| CT079F | CGCGTGGAAACAAAACTGAT | CT079R | AGAAGCTTCGTCAGCTCTGC |
| CT080F | TTCACTCAGATCCACCACCA | CT080R | TGGATCAGGGACCAGAGAAG |
| CT081F | CCGCCGGTAGTACCATCAC | CT081R | CAGCTTCAACGTGTCAATCG |
| CT082F | TCAAGACCACACTGCACTCC | CT082R | TGCAAATGCTGGTTCTTTTC |
| CT083F | TTGCTTCAATGCTTTCCTGA | CT083R | TGCCCAAAAGTCATTCACAG |
| CT084F | AGCATTGCATAGCAGCCTCT | CT084R | TTTGGGAATTGGTTTGGAAG |
| CT085F | CTTCCAAACCAATTCCCAAA | CT085R | CTTGAACTCAACGTGCCTGA |
| CT086F | CACGAGGATTCTCTCTCTCTTTC | CT086R | GTATGTGTGGTGGTGGTGGA |
| CT087F | TGTTGTTGGTGGGAAATTGA | CT087R | TAATGAGCAGGCTTCCCTTC |
| CT088F | AGAAAACGTCCATGGCTGAC | CT088R | TCCCCCAACATAAGGAATGA |
| CT089F | GATGGCTCTTCCCTCTCTGA | CT089R | GGCTTCTCTGTTGGCTTCAC |
| CT090F | CACAAGGGAAAGGCAATGTT | CT090R | TCTGAACCTCCAATCACAGG |
| CT091F | TCAGCAGCTCTTACCAAGCA | CT091R | TTTCGCTTCTTCCTGCACTT |
| CT092F | CCAGGTTCTGTGGATTGGAT | CT092R | TCTCGAAGCAGTTGTAAATAGAAAA |
| CT093F | TCTGGAGTTGCAAGGTGTTG | CT093R | GGGATCCAGAGAGAATGCAG |
| CT094F | TCTTTCATGGTGGTTTTGATTG | CT094R | ACCTTGTGGATGGGTCAGAA |
| CT095F | ACGAGGTTGAAGCCTCTGAA | CT095R | ACATCTGAGGGGACAACAGG |
| CT096F | CAATTTGCTTTGTGCCCTTT | CT096R | GGTCACAGAAGACGCTACCC |
| CT097F | CCCTAGCCTTCTCGTTCCTT | CT097R | CTTCATTAGCGCCCTTTTTG |
| CT098F | CAGACGCTAGCAAGGAAACC | CT098R | TAATTGGATCGATCGGAAGC |
| CT099F | CACGAGGAAGCCATGTTAAT | CT099R | TGCAACCCAAATTTGTGAAA |
| CT100F | CCCGTGTGTGTGAGAGAAAG | CT100R | TCGACAATGCTGAGCAAAAC |
| CT101F | TTCATTAGCAACACCCACGA | CT101R | CATCATCAACCATCCAACCA |
| CT102F | CATCTCCATAGCCACGGTCT | CT102R | ACTGTTAACGAGGCGCAGTC |
| CT103F | CAGGGTAGGGACCAGTGAAG | CT103R | CAGGGTAGGGACCAGTGAAG |
| CT104F | TGGTTCACCAACTCCAATCA | CT104R | AAGCCAGAGCCATTTCTTCA |
| CT105F | CCTGCAGATAAGGTGCATGA | CT105R | GCGTGCAATTGGAAATTCTT |
| CT106F | CACGAGGAATTGGTTACATTCTT | CT106R | TGGCTGCATGCATCATAAAC |
| CT107F | ACTTTTTGTCCCGCAGCTAA | CT107R | GGTGAAGAATATGGCGGAAA |
| CT108F | TGGAAACCATGAGAACAGGA | CT108R | TCATTTCCGGATTTTGTTTTG |
| CT109F | TGGTGGTAACAGCAACAGGA | CT109R | GCTTTCATCACCCAAGATGG |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| CT110F | GCACTGATCCATACCCACAA | CT110R | CATCTCCTGTGTCAGCCTCA |
| CT111F | CGGGAATTTTCCATCATCAC | CT111R | GCCATTTGGAATTCAAGCTC |
| CT112F | GCAGCTGAAGATGTTGGACA | CT112R | CGAGGGTTGACTAGCTGACG |
| CT113F | GGTGAAGGTGATTGCTTGGT | CT113R | CGAGGGTTGCATTTCTCAAT |
| CT114F | TCGGGCTAGCTAAGTTCCTG | CT114R | TCCTGATGTGTTCCCCTCTC |
| CT115F | ATCAATTTGGTGGGTTCCAA | CT115R | TTGGGCAATACAGTCAGCAA |
| CT116F | ACTCAAGAAGCGGTGCTGAT | CT116R | TTCACTGAGAGGCTATACATAAATGA |
| CT117F | TTGGGGAAAACTGATGAACC | CT117R | TGAGCAGGGCTATATATATGTGTGA |
| CT118F | AAGCGAAGCTCAAACACATT | CT118R | TGGCAAACTACAACTGTGTGG |
| CT119F | ATCCAAACATGAGGCCAGAA | CT119R | TGTGCCCATCAATCAACTTC |
| CT120F | TTGCAATTAGGATCATTTTGTGA | CT120R | TGGTGCCTTTTCTGATCTCC |
| CT121F | AACATAGCAAAGGAGGTTGACA | CT121R | TTGTTGAATTATCATTTGGGTTTG |
| CT122F | TTCAACTCTCAGAAGGAACCTG | CT122R | TTGTTGTTACGATGATGAAGCTA |
| CT123F | CCGTTGTGCAAATTGATGAG | CT123R | TCAAGTGCATTTTGCTAGGC |
| CT124F | CGAGGCTGAGGAGAGAATGT | CT124R | ATTGGGCAGCGCTTATACTG |
| CT125F | CGCTTCTTCACTGCCTCTTC | CT125R | AAAAGCTTCAGTTGCGGAAA |
| CT126F | TTGATGTGTTTGCACTTGAGAA | CT126R | TGATGATAGAAGTCTACAAATTAGGG |
| CT127F | CTTTTCTCCCCATCCCTCTC | CT127R | CCGGCAACCTCCACTAATAA |
| CT128F | TGGAAGCTCTCAGCTTTGGT | CT128R | TTGTTCCAGCAGACCATTGA |
| CT129F | ACGAGGGTCATTGTTCCCTA | CT129R | GCCACCCTCCTCTCTGTTCT |
| CT130F | TCTTCTCGTAAACCACACATACAAA | CT130R | AAGTTCTTCCTCCGGCATCT |
| CT131F | GCCCTGCACTTTTGTGTTTT | CT131R | ATTTATGCTCTTCCCGCAAA |
| CT132F | TGTGGGTTTGGAACAGAACA | CT132R | ATCTGCTGAAGCGTCGTTTT |
| CT133F | AATTTGGACTTGGTGCAAGG | CT133R | GCAAAGGCAAGGATTGTAGG |
| CT134F | TGGATTGCTCCTTCTCATCC | CT134R | CATCCAAAGCCCTAGCTATTTTT |
| CT135F | TTTGTGTCATGGGAGACAGC | CT135R | CTCTTGCAACCAGCATCAAA |
| CT136F | TGAATGGGAGGCTGTAGAGG | CT136R | AAATCTCAAATTTGCAACCTATTT |
| CT137F | TGCAGAAGAGTAGCCATCCAT | CT137R | GCCCCATCTTAATTGGTTAGC |
| CT138F | TCATAACAAATAACCATTGTTTGAGAA | CT138R | ATGCAGCACGTTTCTGATTG |
| CT139F | ATCAGAAACGTGCTGCATGA | CT139R | TTCAGCCTAACCAAGGTACGA |
| CT140F | GTAGCCTCGAATTGGACGAG | CT140R | TGCATTAGATCTGGAATTGAAGC |
| CT141F | TTCACAGGGCACCATACAAA | CT141R | TTACCATGCAAGGTGACAGC |
| CT142F | CGTCGGATTATCCAAGGTGA | CT142R | TTTCCGAAAAAGAAAATTGC |
| CT143F | GAGGGGGAGCTTCTGTTTCT | CT143R | GGCTTCTTTTGGTTGGTGAA |
| CT144F | GTAGCCGGAGAAATGGGTCT | CT144R | AAGGGTATATTATCCAACCAACCA |
| CT145F | AGGTCACTGCCAGAATGGAA | CT145R | GAATGAAAAAGAAAAAGGTATTCATGT |
| CT146F | CCACTGCCATTGAAGTTCCT | CT146R | AAAATGAAACAGGTCAAATCATACA |
| CT147F | CGTCCATAAATCAAGTCTTAGCA | CT147R | ATGGAAGCTGCAAGATCGTC |
| CT148F | GCAGAACAGAGAGTGCGAAA | CT148R | CCTCTAAGGCAGCATTCACC |
| CT149F | CCCCTTTTGACATACCATCC | CT149R | CCCCTTTTGACATACCATCC |
| CT150F | TCTAGATCATTCATAACAAGTTGCTG | CT150R | ATTGGATCCCGAGTCAAACA |
| CT151F | GCTGCCTTGTCTACTTTTCCA | CT151R | CCAACAAAACTTGGGTGCAT |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| CT152F | AAAATTTCAGAGAGAGAAACACTCA | CT152R | GTCGATTTCCCATCAACACC |
| CT153F | GCACACTCACAAATCCTCCA | CT153R | TTTGCAGAAAGCAAACTAAACTACA |
| CT154F | CACGAGGGGAAGAGTTATCAA | CT154R | TGTACATGGCATTGGTAGATTTG |
| CT155F | AAGGCAAGAGTCAGCCTTGA | CT155R | AATGAAAGATGATTATGACTAAAGCA |
| CT156F | TTGATTTTAATGGAAGGGATGG | CT156R | AACCGGATTTTGTAGGAAGTGA |
| CT157F | CAACCTTCCTCTCCCGAACT | CT157R | CAACCTTCCTCTCCCGAACT |
| CT158F | CCTGCAGCAATATGTGCCTA | CT158R | AGAGGGAGTGCCTTGGAATG |
| CT159F | AGGGTAAAAGGCTTGGAGGA | CT159R | TCGGTTCATTTTCCAATAGGTC |
| CT160F | AGATGGATTCGGTTGGAATG | CT160R | GTCCTCCCTTCCATGTTCCT |
| CT161F | TCTGAACCTCCAATCACAGG | CT161R | CACAAGGGAAAGGCAATGTT |
| CT162F | CACGAGGGCCTAATAAACATACA | CT162R | TCACCTTCATTCAGCCATCA |
| CT163F | CCCGTCTCAATGAAGCATTT | CT163R | TCAACGCCCACACACATAAT |
| CT164F | GGGTTTCGCAGATAGCTTCA | CT164R | GGATTTTGCAGAAAGCAAGC |
| CT165F | TGAAGGGTTAGCTGCTGTCA | CT165R | GCTCGGCAAAATTCTTTTCA |
| CT166F | TGGATCGATCTTCCTGGTTC | CT166R | AGCTCCACAGTTGATTGATTGA |
| CT167F | TCAATGCAAAGATCGGTCAA | CT167R | ACCCCAACGGAGCTAATCTT |
| CT168F | TTTCTCTGGGATCCATCCTG | CT168R | CCAAGCTAACCCTCCAACAA |
| CT169F | ATTCCCAAATGGCAAGACAC | CT169R | ACCACATGTGATGGTGCAAA |
| CT170F | CAAAAACGACGACAACATCG | CT170R | GCCTTACACGCAGCTCTTCT |
| CT171F | CACGAGGTCTCTTTCTACTCTGC | CT171R | GTGGTGACCGAGAACAAGGT |
| CT172F | CAGTCACGGTCATCTTCAGG | CT172R | AAACAAGCATCGGAAACAGC |
| CT173F | CGGTGAGGGAAGATGTTAGG | CT173R | CCTATTTTGCAAGACCCATAAA |
| CT174F | GAGGGTTTCGTTGTCCAGAA | CT174R | GCCAGGAAAAGCACGAGATA |
| CT175F | GTGGATTTGGAGCAAGGAAA | CT175R | CCCTCCCAGAAACAGAATCA |
| CT176F | CTTTCCTCCTACCATCAACCA | CT176R | GGAACATTCTGAAGATAATCGAG |
| CT177F | GAGGAACCATTCCACGTCAC | CT177R | TCATGGCGGTGGTTACACTA |
| CT178F | TTTTCTTCACCCGTTTCAGC | CT178R | TGATCGATGGTGGGGATACT |
| CT179F | AACCACCGTCTCGTTCTCTG | CT179R | TCCTTAAACGATCCCGACAC |
| CT180F | TTTCTCCCTCCAACTCGAAA | CT180R | TGAAGCACCCCAACTAGGAC |
| CT181F | ATGGGTGATGGATTTGTGCT | CT181R | GCAGCAGATGAAGGAGGAAC |
| CT182F | GCCACTGCGGAAAATATCTC | CT182R | CCGTACGATTGGAGAAACGA |
| CT183F | TCTGTAAAAATTCAGACTTTGACTCAC | CT183R | CATGTCAACGCAGGAAGAAA |
| CT184F | ATGGCTTCCATGATGAGACC | CT184R | TCCCATACTGCAGCTCCTCT |
| CT185F | TGGATCATGAAATAATGGATGG | CT185R | TGATGATGGGTGTCAAATCG |
| CT186F | CGAGGCATCTGTTTCTCTCA | CT186R | ATGGTTCTGCTGGAACAGGT |
| CT187F | AATGGGTAAGGCCTCAAGGT | CT187R | AGCGTTAGAAAGGGGATGGT |
| CT188F | CCACCACCACATTCCTCTTT | CT188R | TACTGGGAAGACCCATTTGC |
| CT189F | ACTGACGCGAAAAGCAGATT | CT189R | ACAACCAATCTTGCACCTGA |
| CT190F | GATTCTGAGCGAAGCCAAAC | CT190R | TGGAGGCGAAAGAGAGAAGA |
| CT191F | CAACTCATCATCCTTGAATGAAA | CT191R | GGTCACTGCCAGAATGGAAT |
| CT192F | CGAGGTTCTTCTTCAATTCCA | CT192R | GTTCCGATCCCGTAAACCTT |
| CT193F | ACCACCAACCAGATCCAAAA | CT193R | CGTAAGGGAAGGCGACATTA |

| Name | Primer sequence 5'-3' | Name | Primer sequence 5'-3' |
|---|---|---|---|
| CT194F | CCACCCCTTGATCCCTTTAT | CT194R | TCCGATGTGAATCTATTGGTCT |
| CT195F | ACTCCTCCGTCTGGAATCCT | CT195R | GGTTCTTGCTCTTCGTCGTC |
| CT196F | ACTTTGGCCCCTAAGGTGAT | CT196R | TCCTGGAGCAGTTAAATCAGG |
| CT197F | TGTGTAATGCTGGGGTCAGA | CT197R | CTCATCCTCATCATCTTCAATTT |
| CT198F | TGCAAAGCGTTGACAGTTTC | CT198R | GGCAGCCTTTCCAAACTTAT |
| CT199F | GAGGGGTCCCTCTCTCAGTC | CT199R | ACCCTTTGGGATTTGGATTC |
| CT200F | CACGAGGCAAAGCTGAAGA | CT200R | GCGAAATTCTTCGCAATCTC |

### 4.2.3 Analysis of the synthesized primers

The primers were tested for amplification, transferability and polymorphism among five accessions of *Cyamopsis*. Out of the 226 primers tested 190 primers showed amplification. Forty four markers were monomorphic and 67 of the markers were polymorphic only in *C. senegalensis*; 63 were non transferable in *C. senegalensis.* Only 5 were polymorphic in commercial cultivars tested. Fig 4.3 shows the amplification, transferability, monomorphic and polymorphic markers among five accessions of *Cyamopsis*. The binary data from the amplification of SSRs were used for generating a dendrogram based on Euclidean similarity coefficient for SSRs. Fig 4.4 shows the dendrogram based on SSR data.

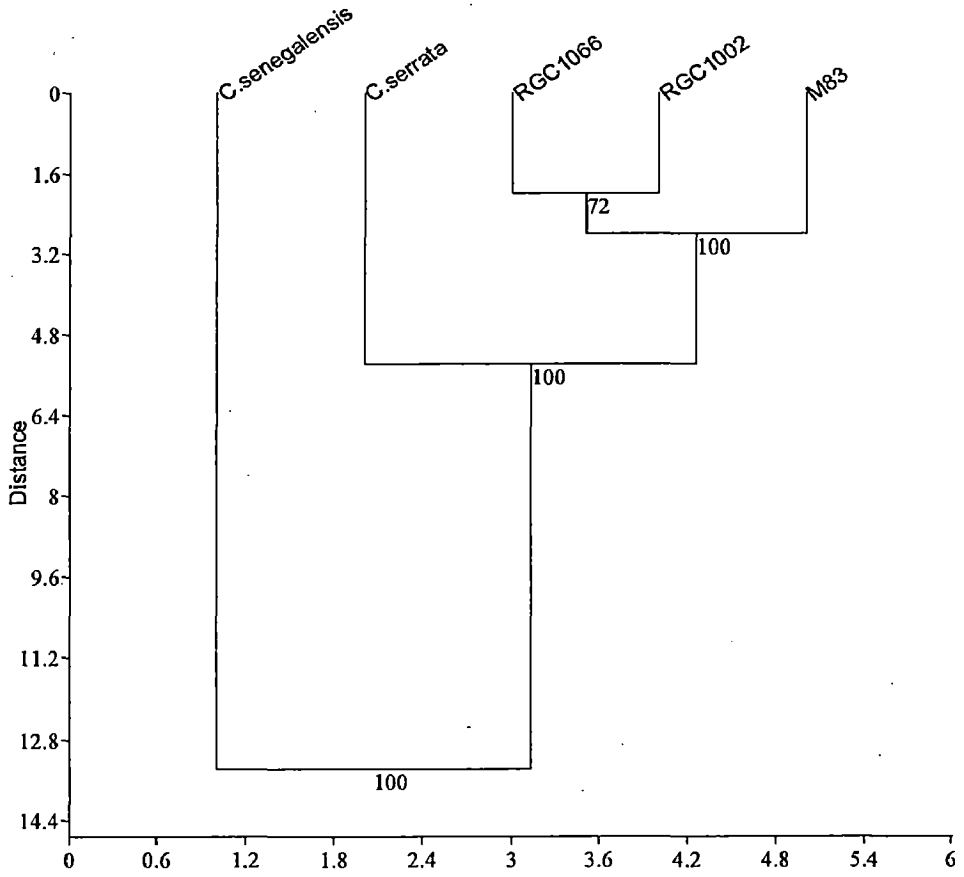



**Fig. 4.4. Dendrogram generated by UPGMA method showing relationship between 5 accessions of *Cyamopsis* based on genetic profile from SSR data. The numbers at the forks show the confidence limits for the grouping of those accessions in the branch, based on 1,000 cycles of bootstrap analysis**

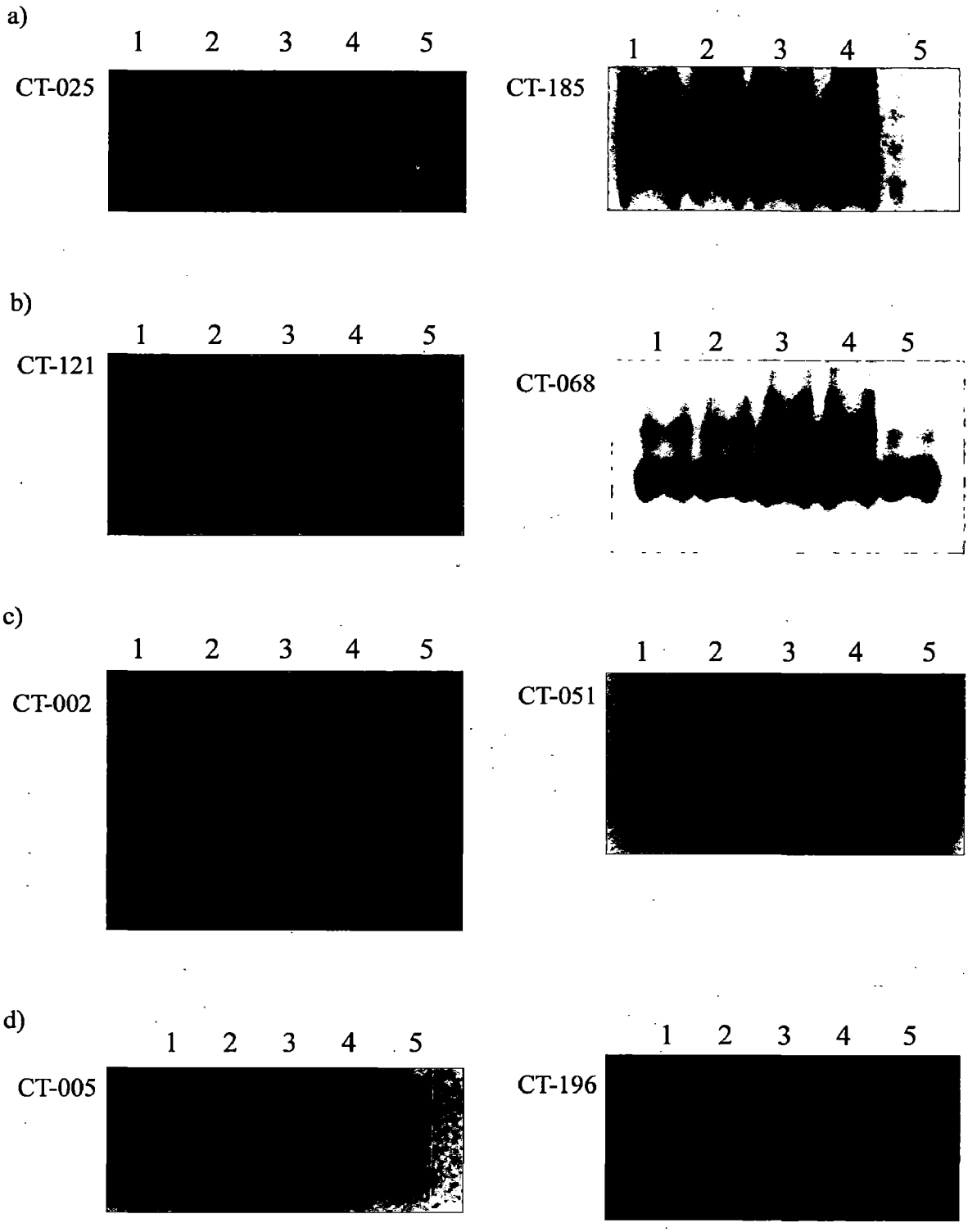



**Fig. 4.3. Banding patterns from amplification of SSR markers tested**
**1. M83, 2. RGC 1066, 3. RGC 1002, 4. *C. serrata* and 5. *C. senegalensis***
**a) Non transferable to *C. senegalensis*, b) Monomorphic,**
**c) Polymorphic in *C. senegalensis*, d) Polymorphic in accessions of *C. tetragonoloba***

KDCL 00297362 55
07-JPG N N N N AS+02SA108

The dendrogram shows that RGC-1066 and RGC-1002 are in close relation; these cultivars are used for gum production. While the accession M-83 is a vegetable variety which shows very high similarity to other accessions of *C. tetragonoloba.* The wild plant *C. senegalensis* seems to be very distant from *C. tetragonoloba* and *C. serrata* and is in the middle showing relatedness to both *C. tetragonoloba* and *C. senegalensis.*

### 4.3 Discussion

The present study was designed to develop SSR markers from the available ESTs of cluster bean in databases. All the EST sequences available in NCBI were downloaded and assembled using Cap3 to produce 1,755 contigs and 4,320 singlets. The formation of contigs reduces the redundancy in the sequences. MISA was used for mining the SSR containing sequences. MISA is considered as a good program as it detects compound SSRs and have shown maximum coverage [120]. In some studies lowering the threshold has resulted in additional SSR containing sequences than with default assignments [152]. However, in the present study the default definements for microsatellites (unit size / minimum number of repeats) : (1/10) (2/6) (3/5) (4/5) (5/5) (6/5) were used in MISA for obtaining optimal results.

The results suggest that there is approximately one SSR in every 4.1 kb of expressed sequences in cluster bean. This is comparable to that of barley which has one SSR in every 6.3 kb [146] and very high when compared to rice which has one SSR in every 157 kb [86].

Out of 362 markers designed 226 SSR markers were synthesized and tested. These markers can be a good resource for a variety of purposes including pedigree analysis, study of marker-trait associations and comparative mapping. To the best of our knowledge, this is the first set of SSR markers developed for cluster bean. In this study, 84% of the primer pairs were successfully amplified which is very high in comparison to genomic SSRs [80, 121] and EST-SSRs in other species [41]. As the EST-SSR markers are derived from expressed genes these markers could be useful for studying functional diversity in cluster bean. A study in six diverse barley genotypes has shown that the gene-derived markers are a good resource for sampling the functional diversity in natural germplasm collection or breeding material [149].

The wild genotypes showed more polymorphism than cultivated or elite genotypes. This may be due to the presence of unique alleles in wild genotypes, which have been lost during the adoption process in case of cultivated genotypes. These results are similar to earlier studies in other plants [149]. The loss of alleles may also be due to a smaller genome size of elite genotypes in comparison to the wild [101]. These results demonstrate the diversity of wild

germplasm with respect to cultivated genotypes. However, the utility of wild germplasm for crop improvement programmes is not possible because of incompatibility of *C. tetragonoloba* with the wild species. The incompatibility is mainly due to the failure of pollen germination on the foreign stigma [77].

# *Chapter V*

*Crossing of guar genotypes for developing population to study the inheritance of phenotypic traits and their linkage with markers*

# 5. Crossing of guar genotypes for developing population to study the inheritance of phenotypic traits and their linkage with markers

## 5.1 Material and methods

### 5.1.1 Plant material

Plants of the parents i.e., RGC 1066 and M 83 were grown in field condition in the campus of Indian Institute of Technology Roorkee during the summer of 2010. The plants were grown in rows; the space between two rows was 0.5Meter. The plant to plant distance was maintained at 0.2Meter within a row.

### 5.1.2 Crossing of guar plants

The crosses were made during summer of year 2010, according to the method described by Chaudhary et al. (1974) [26]. The cultivar M83 was used as female. The buds which were likely to open in one or two days were selected in the afternoon between 4:00 p.m to 6:00 p.m. The flowers below the selected bud were removed; thus the lowest bud was automatically the one used for emasculation. However, the buds above the selected buds were left intact to avoid any damage to the inflorescence. The front sepals were removed through the help of forceps by gently pulling. Petals were also removed by pulling in forward direction by forceps and necessary support to the bud was provided by fingers in such a way that only petals along with almost all the stamens were removed together. Care was taken to avoid any damage to stigma or style. The whole inflorescence was bagged using light butter paper bag to avoid any chance of cross pollination.

The buds were pollinated in the next morning between 8:00 a.m. to 9:00 a.m. The mature anthers (from RGC 1066) were directly brought into contact with the stigma. Care was taken to avoid any damage to stigma. After pollination the bud was labeled properly and bagged again.

The pollinated flowers were checked after 2-5 days for pod formation. The pods formed following pollination were properly tagged and recorded. Fig. 5.1 shows the method followed for developing the inbred population.

### 5.1.3 Population development

The seeds formed after cross pollination were grown to produce $F_1$ plants in rainy season of year 2010. The phenotype was observed to confirm the cross. The seeds of $F_1$ plants were harvested. The seeds from a single $F_1$ plant were grown in a separate plot to produce $F_2$ population in the summer of year 2011. The phenotypes of the F2 plants were observed and a leaf was taken from each plant for DNA extraction. DNA was extracted by using the method described in 3.1.7. The quality and quantity of DNA was measured according to the method described in 3.1.8

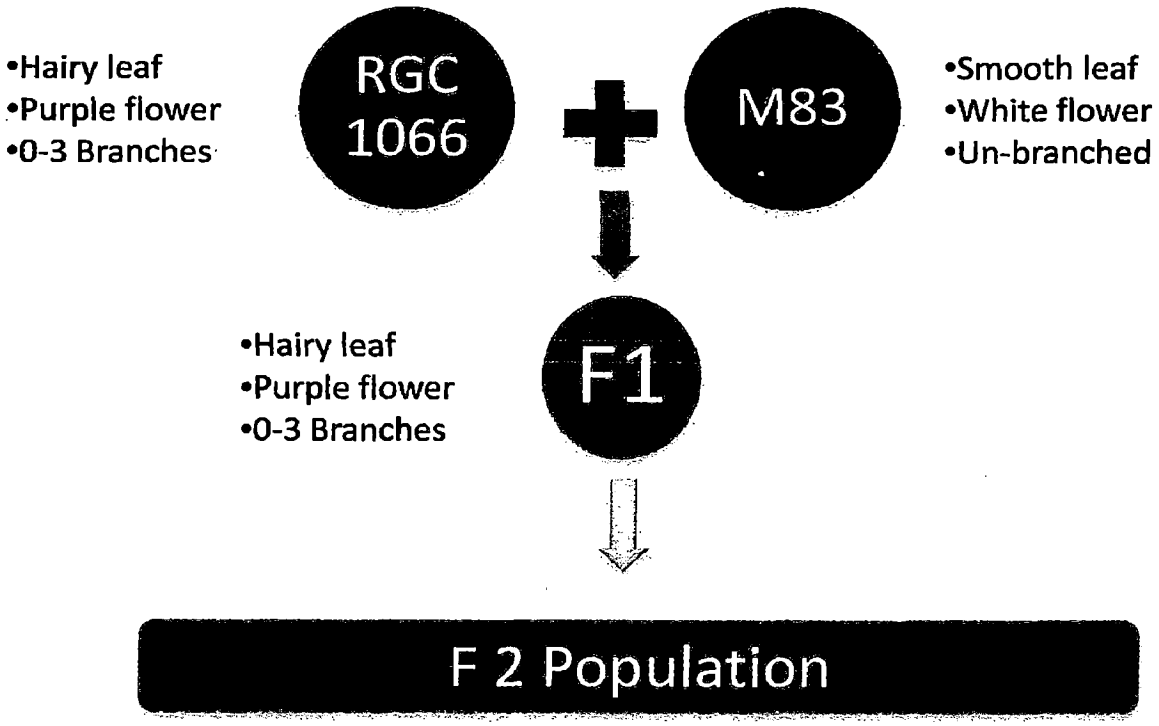



**Fig 5.1 Schematic diagram showing the method followed for developing population.**

## 5.2 Scanning electron microscopy to observe the morphology of trichome and its variation in parents of population

### 5.2.1 Preparation of fixative solution

The composition of fixative solution is given in Table 5.1.

**Table 5.1. Composition of fixative solution**

| S.No. | Name of the reagent | Composition |
|---|---|---|
| 1 | 95% Ethanol | 60ml |
| 2 | Glacial acetic acid | 5ml |
| 3 | 20% Formaldehyde | 20ml |
| 4 | Distilled water | 15ml |

### 5.2.2 Sample preparation scanning electron microscopy

Leaf samples were prepared by following the protocol described by Feeney et al. (2003), with slight modification [43]. The leaf samples from field grown plants were collected in fixative solution and incubated for 24 hours at room temperature. The samples were transferred to 2.5% glutaraldehyde for 2 hours and incubated for 5 minutes in 50% glutaraldehyde. These were then treated with 70% glutaraldehyde for 30 minutes followed by incubation in 90% glutaraldehyde twice. Finally the samples were kept in absolute ethanol for 30 minutes.

The prepared leaf sample was cut to a suitable size using a sharp blade and fixed on metal disc for microscopy. The samples were bombarded with gold metal [135] in sputter coater (BAL-TEC SCD005) for 60 sec at 30mA current for depositing 5-7nm thick gold coating and observed in scanning electron microscope (FEI Quanta 200F).

## 5.3 Bulk segregant analysis

### 5.3.1 Preparation of bulk

The quantity of DNA from all the $F_2$ plants were estimated and diluted to 100ng/μL concentration as described in 3.1.8. Two bulks were prepared for each trait. For leaf pubescence 10 plants each containing hairy leaf and smooth leaf were selected. Equal amount of DNA from each plant was taken to prepare the bulk DNA. Similarly for flower color, 10

plants, each with purple and white colored flower, were selected and the bulk DNA was prepared.

### 5.3.2 Analysis

DNA from the parents and the bulks were amplified according the method described in 4.1.6. The products were separated by polyacraylamide gel electrophoresis and silver stained as described in 4.1.7.

## 5.4 Results

### 5.4.1 Phenotypes of the plants in $F_2$ generation

The phenotypes of the plants in $F_2$ were observed and recorded. The details of the phenotypic data collected are given in Table 5.2.

**Table 5.2. Phenotypic characteristics of $F_2$ plants**

| Plant no | Leaf Pubescence | Flower Color | No. of Branches | Plant height |
|---|---|---|---|---|
| P1 | Hairy | Purple | 13 | 180 |
| P2 | Smooth | White | 0 | 75 |
| P3 | Hairy | Purple | 8 | 220 |
| P4 | Smooth | Purple | 0 | 160 |
| P5 | Hairy | - | 2 | 58 |
| P6 | Smooth | Purple | 3 | 210 |
| P7 | Hairy | - | 2 | 250 |
| P8 | Hairy | Purple | - | - |
| P9 | Smooth | White | 9 | 155 |
| P10 | Hairy | Purple | 2 | 290 |
| P11 | Hairy | Purple | 4 | 144 |
| P12 | Smooth | White | 2 | 220 |
| P13 | Smooth | Purple | 0 | 190 |
| P14 | Hairy | White | 3 | 95 |
| P15 | Hairy | Purple | 14 | 195 |
| P16 | Hairy | Purple | 1 | 200 |
| P17 | Hairy | - | 2 | 170 |
| P18 | Hairy | - | 3 | 210 |
| P19 | Smooth | Purple | 14 | 200 |
| P20 | Hairy | Purple | 0 | 290 |
| P21 | Hairy | Purple | - | - |
| P22 | Smooth | Purple | 2 | 170 |
| P23 | Smooth | Purple | 3 | 165 |
| P24 | Hairy | Purple | 5 | 170 |

| Plant no | Leaf Pubescence | Flower Color | No. of Branches | Plant height |
|---|---|---|---|---|
| P25 | Hairy | - | 1 | 230 |
| P26 | Smooth | Purple | 4 | 190 |
| P27 | Smooth | Purple | 4 | 170 |
| P28 | Hairy | - | - | - |
| P29 | Hairy | White | 0 | 150 |
| P30 | Hairy | White | 1 | 270 |
| P31 | Smooth | Purple | 2 | 60 |
| P32 | Hairy | Purple | 2 | 210 |
| P33 | Smooth | White | 2 | 70 |
| P34 | Hairy | Purple | 3 | 235 |
| P35 | Hairy | Purple | 0 | 100 |
| P36 | Hairy | White | 3 | 200 |
| P37 | Smooth | White | 4 | 220 |
| P38 | Hairy | Purple | 14 | 220 |
| P39 | Hairy | Purple | 0 | 170 |
| P40 | Hairy | White | 8 | 180 |
| P41 | Hairy | Purple | 10 | 175 |
| P42 | Hairy | - | 0 | 130 |
| P43 | Hairy | Purple | 3 | 210 |
| P44 | Hairy | Purple | 1 | 260 |
| P45 | Hairy | Purple | 1 | 270 |
| P46 | Hairy | Purple | 1 | 60 |
| P47 | Smooth | White | 3 | 171 |
| P48 | Hairy | Purple | 6 | 155 |
| P49 | Hairy | Purple | 0 | 200 |
| P50 | Hairy | White | 0 | 120 |
| P51 | Hairy | White | 4 | 250 |
| P52 | Hairy | - | 18 | 170 |
| P53 | Hairy | White | 0 | 210 |
| P54 | Smooth | Purple | 0 | 100 |
| P55 | Hairy | Purple | - | - |
| P56 | Hairy | Purple | 0 | 178 |
| P57 | Hairy | Purple | 4 | 150 |
| P58 | Hairy | - | 5 | 220 |
| P59 | Smooth | Purple | 0 | 80 |
| P60 | Smooth | Purple | 4 | 185 |
| P61 | Hairy | Purple | 0 | 125 |
| P62 | Hairy | - | 0 | 160 |
| P63 | Hairy | - | 0 | 160 |
| P64 | Hairy | - | 5 | 230 |
| P65 | Hairy | Purple | 9 | 90 |
| P66 | Smooth | Purple | 4 | 150 |
| P67 | Hairy | Purple | 3 | 80 |

| Plant no | Leaf Pubescence | Flower Color | No. of Branches | Plant height |
|---|---|---|---|---|
| P68 | Smooth | Purple | 12 | 220 |
| P69 | Hairy | White | 12 | 130 |
| P70 | Hairy | White | 5 | 110 |
| P71 | Hairy | Purple | 6 | 210 |
| P72 | Smooth | Purple | 0 | 160 |
| P73 | Smooth | Purple | 7 | 215 |
| P74 | Hairy | Purple | 5 | 240 |
| P75 | Hairy | White | 2 | 220 |
| P76 | Hairy | White | 5 | 110 |
| P77 | Smooth | White | 0 | 240 |
| P78 | Hairy | Purple | 0 | 120 |
| P79 | Hairy | - | 14 | 210 |
| P80 | Hairy | - | - | - |
| P81 | Hairy | Purple | 0 | 110 |
| P82 | Smooth | Purple | 0 | 150 |
| P83 | Hairy | - | 12 | 200 |
| P84 | Smooth | White | 2 | 160 |
| P85 | Hairy | Purple | 3 | 140 |
| P86 | Hairy | - | 3 | 130 |
| P87 | Smooth | - | - | - |
| P88 | Smooth | White | 3 | 250 |
| P89 | Smooth | - | 0 | 140 |
| P90 | Hairy | White | 2 | 220 |
| P91 | Hairy | - | 2 | 215 |
| P92 | Hairy | - | 3 | 40 |
| P93 | Hairy | Purple | 3 | 110 |
| P94 | Smooth | White | 0 | 170 |
| P95 | Hairy | Purple | 10 | 170 |
| P96 | Hairy | White | 7 | 170 |
| P97 | Hairy | Purple | 0 | 147 |
| P98 | Hairy | Purple | 0 | 80 |
| P99 | Hairy | Purple | 17 | 210 |
| P100 | Hairy | Purple | 0 | 70 |
| P101 | Smooth | - | 2 | 140 |
| P102 | Smooth | White | 0 | 75 |

### 5.4.2 Segregation pattern of phenotypes

The segregation of leaf pubescence, flower color and branching trait was calculated. The dominance pattern and genetic ratios of leaf pubescence and flower color characteristics for cross in cluster bean is given in Table 5.3.

**Table 5.3. The dominance pattern and genetic ratios of characteristics for cross in cluster bean**

| Character | F1 Phenotype | F2 Segregation | | Total | Ratio |
|---|---|---|---|---|---|
| **Leaf pubiscence** | | Hairy | Smooth | | |
| | Hairy | 71 | 31 | 102 | 2.29:1 |
| **Flower color** | | Purple | White | | |
| | Purple | 56 | 25 | 81 | 2.24:1 |

The branching trait was not present in both the parents but in $F_2$ many plants showed very high branching. The frequency distribution of various levels of branching is shown in Fig 5.2.

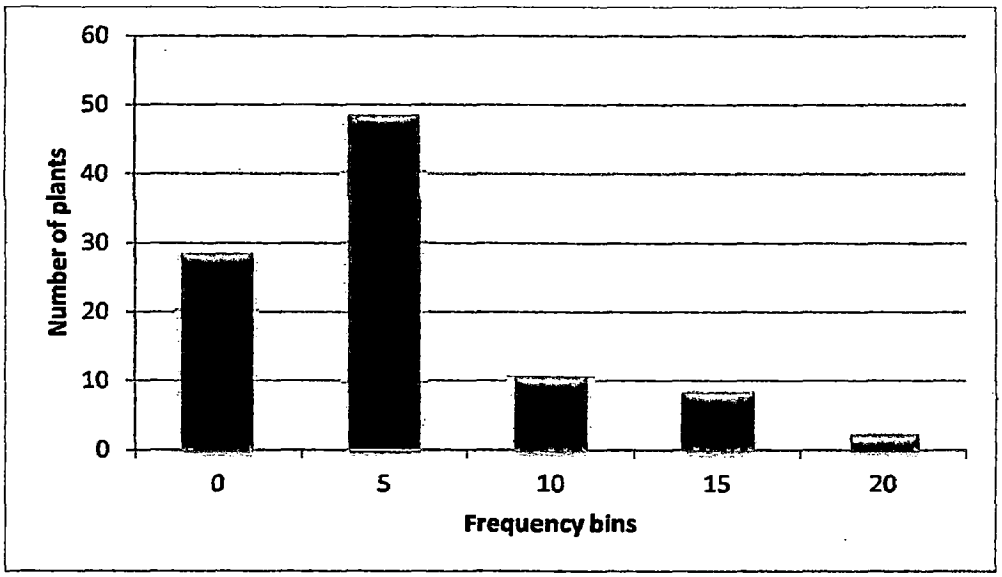



**Fig. 5.2. Graph showing the number of plants with various levels of branching**

### 5.4.3 Variation of leaf pubescence in parental plants

The variation in the pubescence in leaf and the ultrastructure of the trichome in guar were studied by scanning electron microscopy. The SEM pictures showed long, slender, porous hair like structures in the hairy plant. While there was complete absence of such structures on the leaves of glabrous plants. Fig 5.3 and Fig 5.4 show the difference in leaf hairiness of parent plants and the ultrastructure of the trichomes. The Figure 5.3 also shows higher hairiness in lower leaf surface of RGC 1066.

### 5.4.4 Bulk segregant analysis

The DNA from parents and the bulks were amplified to study the pattern of inheritance of the SSR markers and to find out if any linkage to any of the traits exists. Fig 5.5 shows the banding pattern of the PCR products from parents and the bulks for leaf pubescence and flower color.

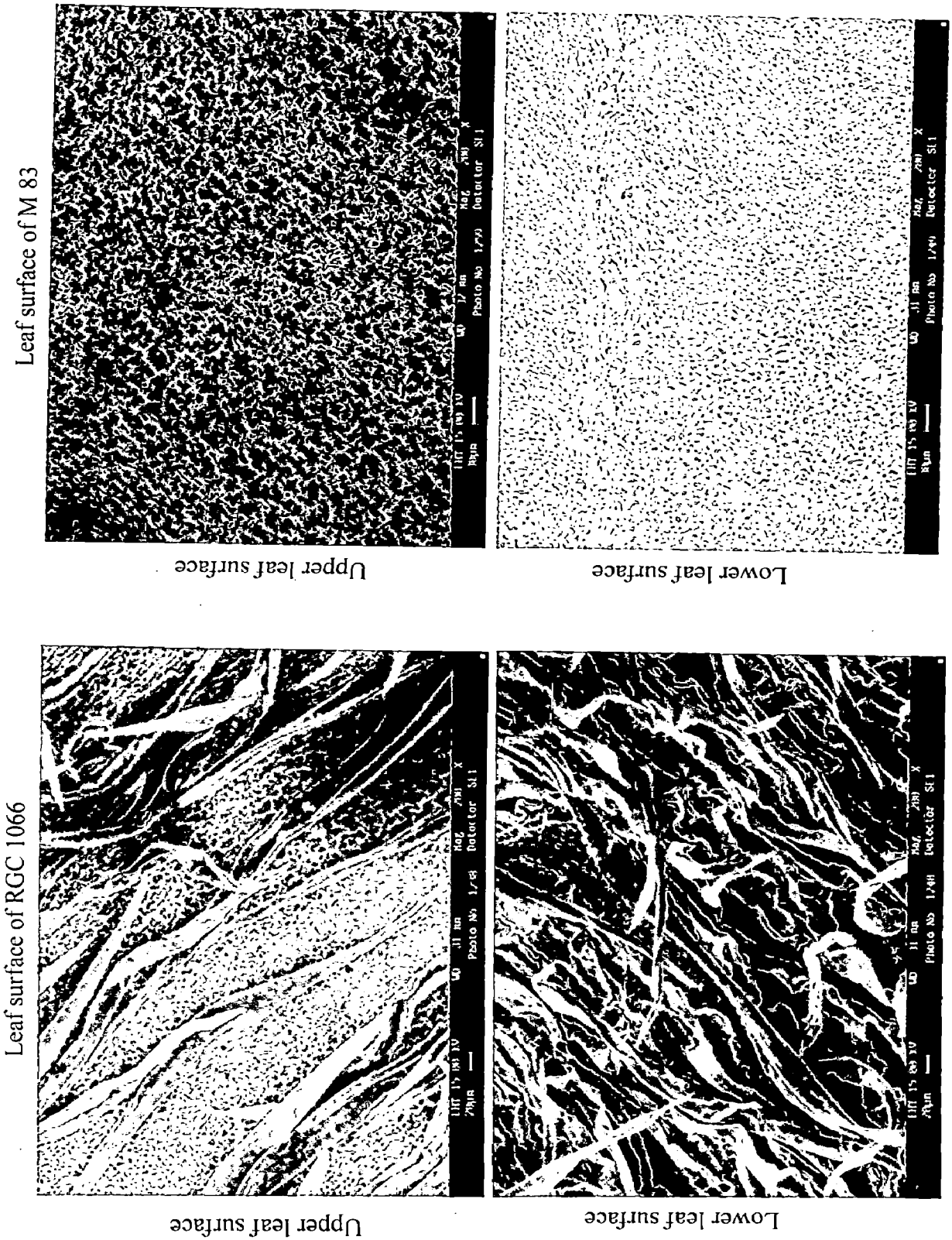



**Fig. 5.3. Scanning electron microscopic images of upper and lower leaf surface of RGC 1066 and M 83 showing contrast in pubescence**

<DCL 00297362 63
38.JPG N N N N AS+02SA108

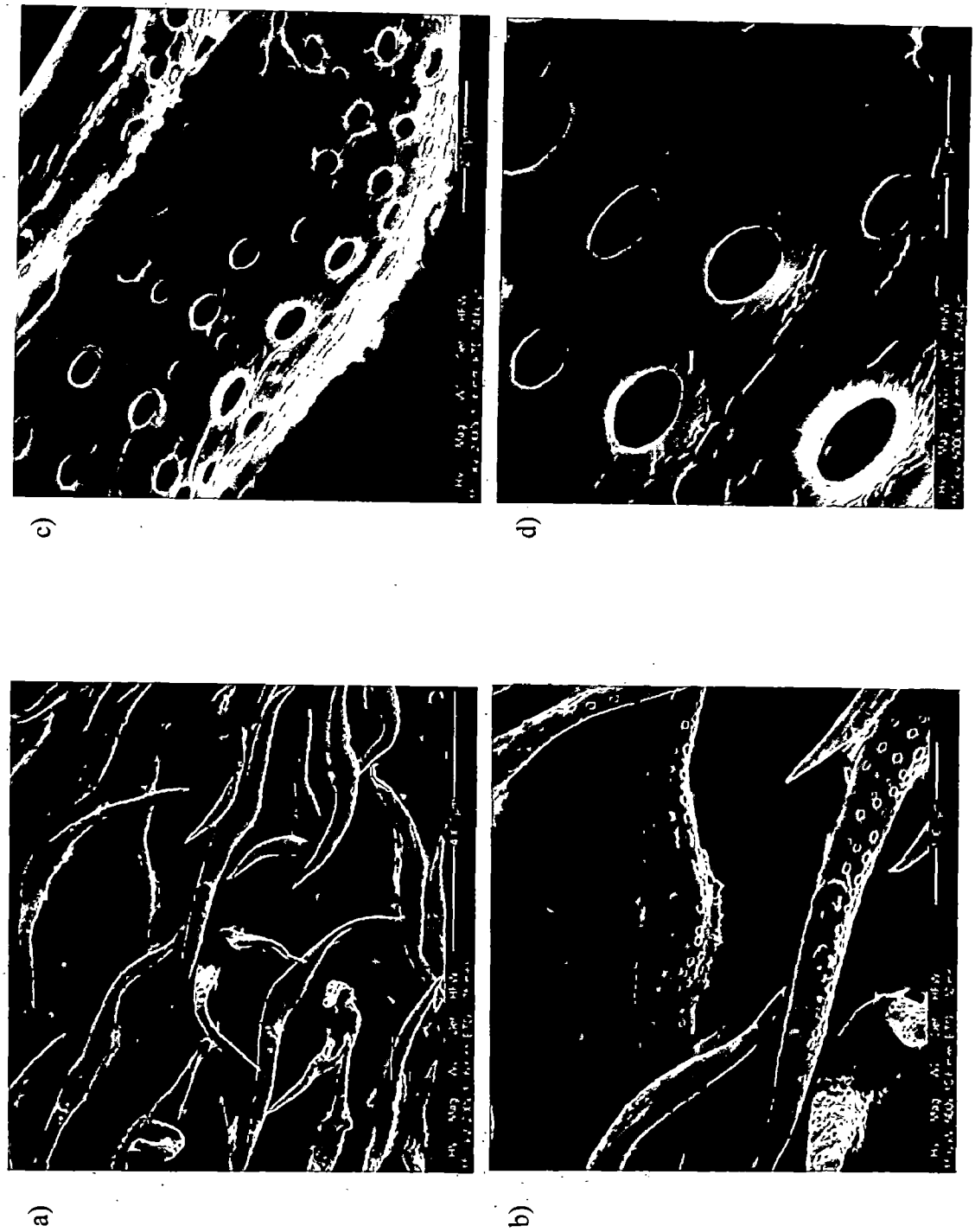



**Fig. 5.4. Scanning electron microscopic images of guar trichome a) at 200X, b) at 500X, c) at 2000X, d) at 5000X**

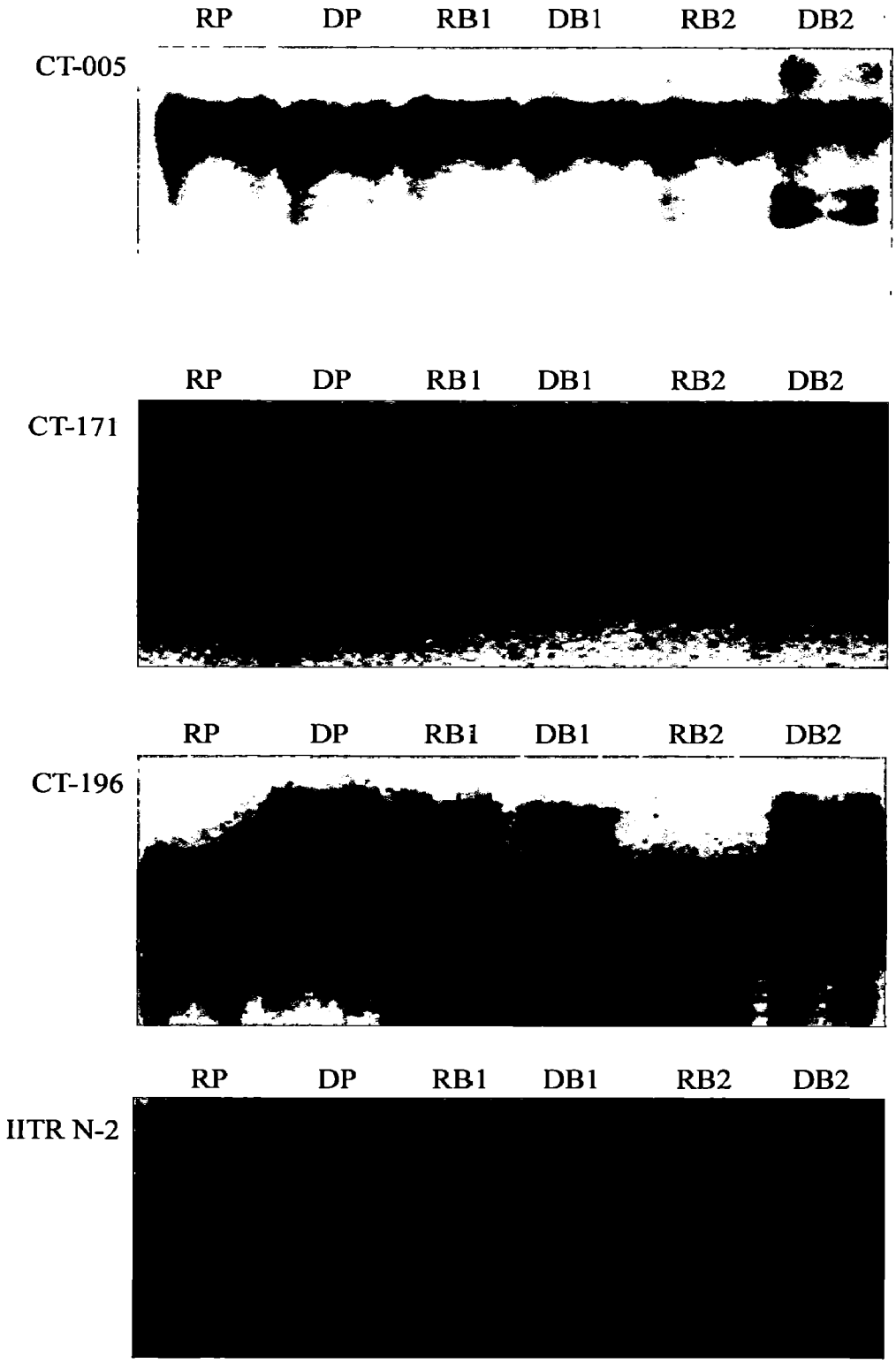



**Fig. 5.5. Banding pattern from ampification of polymorphic markers. RP-Recessive parent, DP-Dominant parent, RB1-Recessive bulk for leaf pubescence, DB1-Dominant bulk for leaf pubescencce, RB2-Recessive bulk for flower color, DB2- Dominant bulk for flower color**

<DCL B0297362 79
[0.JPG N N N N AS+02SA100

## 5.5 Discussion

Leaf pubescence is a quantitative trait in some species and has a potential for insect pest management [79] but its absence in vegetable varieties of guar is an advantage. The population which is segregating for pubescence trait can be used to find markers linked to leaf pubescence. Leaf pubescence inherited in Mendelian ratio in the $F_2$ population of guar. This was similar to the previous report in guar [25] and shows that the trait is qualitative and controlled by single locus. Even though guar is known to have pubescence, its structure and function is not well understood. It is previously known that trichomes occur on the surfaces of many plants and can make a contribution to plant resistance against herbivores. Glandular trichomes have heads containing various sticky and/or toxic exudates that may be secreted onto the plant surface or may rupture on contact with herbivores, causing irritation [13]. Non-glandular trichomes do not have heads and affect herbivores by mechanically obstructing their movement across the plant surface [13]. The observation of guar trichomes by scanning electron microscopy showed that the guar trichomes are non-glandular in nature. Trichomes were long, slender, porous hair like structures. The structure of trichome in guar suggests that it might be helpful for the plant to maintain a cool area surrounding the leaf. This shows the adaptation of the plant towards dry environments where it is cultivated commonly.

Inheritance of flower color also showed Mendelian ratio. The purple color of flower was found to be dominant over the white colored flower. Flower color is not an economically important trait. However it can be used as a phenotypic marker for testing the crosses made for genetic studies.

Both parents of the population were nonbranching. However, few of the $F_2$ plants were branched type. This may be due to the control of the trait by multiple loci as reported previously [25].

As expected, all polymorphic SSR markers tested here segregated and also showed recombinant alleles. The SSR markers that were polymorphic between the parental accessions of the population developed were applied on the bulks. Two bulks that differ in leaf pubescence and flower color were used in the study. The amplification in bulks showed both the marker alleles of parents. This shows that the traits are independent of the markers tested. Further studies with large number of markers are required to find the markers linked to these phenotypes.

# *Chapter VI*

# *Conclusions*

## 6. Conclusions

The study of genetic diversity in guar through RAPD and ISSR markers showed that the overall genetic variability in the germplasm collection is limited. However landraces show significant variation and is different from commercially grown cultivars. Correlation between the molecular data from marker genotyping and the geographic coordinates for the region of landrace collection showed that the molecular data was consistent with the geographic distribution pattern of the landraces. The study also showed that RAPD is a better technique to study genetic variation in guar as the resolving powers of RAPD markers were higher the ISSR markers.

Guar EST sequences available in NCBI were downloaded and mined for SSRs. Analysis of the SSR containing sequences showed very high number of mononucleotide repeats while, negligible number of penta and hexanucleotide repeats were present. A total of 362 primer pairs were designed and 226 primer pairs were tested. Eighty four percent of the tested EST-SSR markers showed amplification in guar. The study showed high polymorphism between *C. tetragonoloba and C. senegalensis*. This shows *C. senegalensis* is a more distant relative to *C. tetragonoloba* and *C. serrata.*

Cross between M83 X RGC 1066 was made successfully to produce an F2 population. The observation showed that leaf pubescence and flower color are single locus controlled phenotypes and segregate with Mendelian ratio in the population. The phenotype of branching showed transgressive behavior in the population. It shows that, the trait is controlled by more than one loci. The SSR markers that were polymorphic between the parental accessions of the population developed were applied on the bulk DNA to study the linkage of markers to leaf pubescence and flower color. The amplification in bulks showed both the marker alleles of parents. This shows that the traits are independent of the markers tested.

# *Chapter VII*

# *Future scope*

## 7. Future Scope

- SCAR markers can be developed using the polymorphic data from RAPD and ISSR
- Linkage analysis of the polymorphic markers could be studied in other populations for commercially important traits
- Genetic linkage map can be developed in the population by generation of large number of sequences from NGS technology for SNP mining
- Metabolomic studies to understand various differences between parents of the population can be carried out

## 8. References


1. Agarwal, M., Shrivastava, N. and Padh, H. Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports* 27(4):617-631 (2008).

2. Ajmone Marsan, P., Castiglioni, P., Fusari, F., Kuiper, M. and Motto, M. Genetic diversity and its relationship to hybrid performance in maize as revealed by RFLP and AFLP markers. *TAG Theoretical and Applied Genetics* 96(2):219-227 (1998).

3. Albani, M. C., Battey, N. H. and Wilkinson, M. J. The development of ISSR-derived SCAR markers around the Seasonal Flowering Locus (SFL) in *Fragaria vesca*. *TAG Theoretical and Applied Genetics* 109(3):571-579 (2004).

4. Ali, M. L., Rajewski, J. F., Baenziger, P. S., Gill, K. S., Eskridge, K. M. and Dweikat, I. Assessment of genetic diversity and relationship among a collection of US sweet sorghum germplasm by SSR markers. *Molecular Breeding* 21(4):497-509 (2008).

5. Anderson, E. Endosperm mucilages of legumes. *Industrial & Engineering Chemistry* 41(12):2887-2890 (1949).

6. Arora, R. N. and Pahuja, S. K. Mutagenesis in Guar [*Cyamopsis tetragonoloba* (L.) Taub.]. *Plant Mutation Reports* 2(1):7-9 (2008).

7. Bassam, B. J., Caetano-Anolles, G. and Gresshoff, P. M. DNA amplification fingerprinting of bacteria. *Applied Microbiology and Biotechnology* 38(1):70-76 (1992).

8. Bilgen, M., Karaca, M., Onus, A. N. and Ince, A. G. A software program combining sequence motif searches with keywords for finding repeats containing DNA sequences. *Bioinformatics* 20(18):3379-3386 (2004).

9. Bocianowski, J. and Seidler-Lozykowska, K. The relationship between RAPD markers and quantitative traits of caraway ( *Carum carvi* L.). *Industrial Crops and Products* 36(1):135-139 (2011).

10. Bornet, B., Goraguer, F., Joly, G. and Branchard, M. Genetic diversity in European and Argentinian cultivated potatoes (*Solanum tuberosum* subsp. tuberosum) detected by inter-simple sequence repeats (ISSRs). *Genome* 45(3):481-484 (2002).

11. Bornet, B., Muller, C., Paulus, F. and Branchard, M. Highly informative nature of inter simple sequence repeat (ISSR) sequences amplified using tri-and tetra-nucleotide primers from DNA of cauliflower (*Brassica oleracea* var. botrytis L.). *Genome* 45(5):890-896 (2002).

12. Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32(3):314 (1980).

13. Boughton, A. J., Hoover, K. and Felton, G. W. Methyl jasmonate application induces increased densities of glandular trichomes on tomato, *Lycopersicon esculentum. Journal of Chemical Ecology* 31(9):2211-2216 (2005).

14. Bracci, T., Busconi, M., Fogher, C. and Sebastiani, L. Molecular studies in olive (*Olea europaea* L.): overview on DNA markers applications and recent advances in genome analysis. *Plant Cell Reports* 30(4):449-426 (2011).

15. Bravo, L., Grados, N. and Saura-Calixto, F. Characterization of syrups and dietary fiber obtained from mesquite pods (*Prosopis pallida* L). *Journal of Agricultural and Food Chemistry* 46(5):1727-1733 (1998).

16. Bravo, L. and Saura-Calixto, F. Composition and potential uses of mesquite pods (*Prosopis pallida* L): comparison with carob pods (*Ceratonia siliqua* L). *Journal of the Science of Food and Agriculture* 65(3):303-306 (1994).

17. Brummer, Y., Cui, W. and Wang, Q. Extraction, purification and physicochemical characterization of fenugreek gum. *Food Hydrocolloids* 17(3):229-236 (2003).

18. Buckeridge, M. S., Pessoa dos Santos, H. and Tiné, M. A. S. Mobilisation of storage cell wall polysaccharides in seeds. *Plant Physiology and Biochemistry* 38(1-2):141-156 (2000).

19. Bulpin, P. V., Gidley, M. J., Jeffcoat, R. and Underwood, D. R. Development of a biotechnological process for the modification of galactomannan polymers with plant α-galactosidase. *Carbohydrate Polymers* 12:155-168 (1990).

20. Butt, M. S., Shahzadi, N., Sharif, M. K. and Nasir, M. Guar gum: a miracle therapy for hypercholesterolemia, hyperglycemia and obesity. *Critical Reviews in Food Science and Nutrition* 47(4):389-396 (2007).

21. Caetano-Anolles, G. and Bassam, B. J. DNA amplification fingerprinting using arbitrary oligonucleotide primers. *Applied Biochemistry and Biotechnology* 42(2):189-200 (1993).

22. Carlos de Oliveira, A., Novac Garcia, A., Cristofani, M. and Machado, M. A. Identification of citrus hybrids through the combination of leaf apex morphology and SSR markers. *Euphytica* 128(3):397-403 (2002).

23. Ceccarelli, S. Specific adaptation and breeding for marginal conditions. *Euphytica* 77(3):205-219 (1994).

24. Cerny, T. A., Caetano-Anolles, G., Trigiano, R. N. and Starman, T. W. Molecular phylogeny and DNA amplification fingerprinting of *Petunia* taxa. *TAG Theoretical and Applied Genetics* 92(8):1009-1016 (1996).

25. Chaudhary, B. S. and Lodhi, G. P. Studies on the inheritance of five qualitative characteristics in clusterbean (*Cyamopsis tetragonoloba* (L.) Taub). *Euphytica* 30(1):161-165 (1981).

26. Chaudhary, B. S., Paroda, R. S. and Solanki, K. R. A new crossing technique in cluster bean (*Cyamopsis tetragonoloba* [L.] Taub.). *Current Science* 43 (1974).

27. Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B. and Pang, E. C. K. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* 142(1):169-196 (2005).

28. Crawford, D. J. Plant molecular systematics: macromolecular approaches. New York: John Wiley & Sons (1990).

29. Cutler, S. Cellulose synthesis: cloning in silico. *Current Biology* 7(2):R108-R111 (1997).

30. Dabas, B. S., Mandal, S., Phogat, B. S., Bisht, I. S. and Agrawal, R. C. Guar (*Cyamopsis tetragonoloba*)- A resume of research at NBPGR. New Delhi: National Bureau of Plant Genetic Resources (2001).

31. Dhugga, K. S. Food additives. Gums the word. *Asia Pacific Food Industry*:57-58 (2005).

32. Dhugga, K. S., Barreiro, R., Whitten, B., Stecca, K., Hazebroek, J., Randhawa, G. S., Dolan, M., Kinney, A. J., Tomes, D., Nichols, S. and Anderson, P. Guar seed beta-mannan synthase is a member of the cellulose synthase super gene family. *Science* 303(5656):363-366 (2004).

33. Dnyaneshwar, W., Preeti, C., Kalpana, J. and Bhushan, P. Development and application of RAPD-SCAR marker for identification of *Phyllanthus emblica* LINN. *Biological & Pharmaceutical Bulletin* 29(11):2313-2316 (2006).

34. Doyle, J. J. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15 (1990).

35. Duncan, C. J. G., Pugh, N., Pasco, D. S. and Ross, S. A. Isolation of a galactomannan that enhances macrophage activation from the edible fungus *Morchella esculenta*. *Journal of Agricultural and Food Chemistry* 50(20):5683-5685 (2002).

36. Dwivedi, N. K., Bhandari, D. C., Dubas, B. S., Agrawal, R. C., Mandal, S. and Rana, R. S. Catalogue on cluster bean (*Cyamopsis tetragonoloba* (L.) Taub ) germplasm part III. New Delhi: NBPGR (1995).

37. Edwards, K. J., Barker, J. H., Daly, A., Jones, C. and Karp, A. Microsatellite libraries enriched for several microsatellite sequences in plants. *BioTechniques* 20(5):758 (1996).

38. Edwards, M., Bulpin, P. V., Dea, I. C. M. and Reid, J. S. G. Biosynthesis of legume-seed galactomannans *in vitro*. *Planta* 178(1):41-51 (1989).

39. Edwards, M., Scott, C., Gidley, M. J. and Reid, J. S. G. Control of mannose/galactose ratio during galactomannan formation in developing legume seeds. *Planta* 187(1):67-74 (1992).

40. Edwards, M. E., Choo, T. S., Dickson, C. A., Scott, C., Gidley, M. J. and Reid, J. S. G. The seeds of *Lotus japonicus* lines transformed with sense, antisense, and sense/antisense galactomannan galactosyltransferase constructs have structurally altered galactomannans in their endosperm cell walls. *Plant Physiology* 134(3):1153-1162 (2004).

41. Edwards, M. E., Dickson, C. A., Chengappa, S., Sidebottom, C., Gidley, M. J. and Reid, J. S. Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. *The Plant Journal* 19(6):691-697 (1999).

42. Eujayl, I., Sorrells, M., Baum, M., Wolters, P. and Powell, W. Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. *Euphytica* 119(1):39-43 (2001).

43. Feeney, K. A., Heard, P. J., Zhao, F. J. and Shewry, P. R. Determination of the distribution of sulphur in wheat starchy endosperm cells using secondary ion mass spectroscopy (SIMS) combined with isotope enhancement. *Journal of Cereal Science* 37(3):311-318 (2003).

44. Ferriol, M., Pico, B. and Nuez, F. Genetic diversity of a germplasm collection of *Cucurbita pepo* using SRAP and AFLP markers. *TAG Theoretical and Applied Genetics* 107(2):271-282 (2003).

45. Geuna, F., Toschi, M. and Bassi, D. The use of AFLP markers for cultivar identification in apricot. *Plant Breeding* 122(6):526-531 (2003).

46. Ghalmi, N., Malice, M., Jacquemin, J. M., Ounane, S. M., Mekliche, L. and Baudoin, J. P. Morphological and molecular diversity within Algerian cowpea (*Vigna unguiculata* (L.) Walp.) landraces. *Genetic Resources and Crop Evolution* 57(3):371-386 (2010).

47. Giannini, E. G., Mansi, C., Dulbecco, P. and Savarino, V. Role of partially hydrolyzed guar gum in the treatment of irritable bowel syndrome. *Nutrition* 22(3):334-342 (2006).

48. Gill, S. L. Evaluation of reciprocal hybrid crosses in guar. Texas, USA: Texas Tech University; (2009).

49. Goyal, M. and Sharma, S. K. Traditional wisdom and value addition prospects of arid foods of desert region of North West India. *Indian Journal of Traditonal Knowledge* 8:581-585 (2009).

50. Gupta, P., Balyan, H., Edwards, K., Isaac, P., Korzun, V., Roder, M., Gautier, M. F., Joudrier, P., Schlatter, A. and Dubcovsky, J. Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. *TAG Theoretical and Applied Genetics* 105(2):413-422 (2002).

51. Gupta, S., Srivastava, M., Mishra, G. P., Naik, P. K., Chauhan, R. S., Tiwari, S. K., Kumar, M. and Singh, R. Analogy of ISSR and RAPD markers for comparative analysis of genetic diversity among different *Jatropha curcas* genotypes. *African Journal of Biotechnology* 7(23):4230-4243 (2010).

52. Haley, S. D., Miklas, P. N., Stavely, J. R., Byrum, J. and Kelly, J. D. Identification of RAPD markers linked to a major rust resistance gene block in common bean. *TAG Theoretical and Applied Genetics* 86(4):505-512 (1993).

53. Hamilton, M. B., Pincus, E. L., Di Fiore, A. and Fleischer, R. C. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques* 27(3):500-507 (1999).

54. Hammer, O., Harper, D. A. T. and Ryan, P. D. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4(1):1-9 (2001).

55. Hayden, M. J. and Sharp, P. J. Targeted development of informative microsatellite (SSR) markers. *Nucleic Acids Research* 29(8):e44 (2001).

56. Hazen, S. P., Scott-Craig, J. S. and Walton, J. D. Cellulose synthase-like genes of rice. *Plant Physiology* 128(2):336-340 (2002).

57. Huang, X. and Madan, A. CAP3: A DNA sequence assembly program. *Genome Research* 9(9):868-877 (1999).

58. Hymowitz, T. The trans-domestication concept as applied to guar. *Economic Botany* 26(1):49-60 (1972).

59. Hymowitz, T. and Matlock, R. S. Guar in the United States. *Oklahoma Agricultural Experiment Station Technical Bulletin* 611:1-34 (1963).

60. Ijaz, S. Microsatellite markers: An important fingerprinting tool for characterization of crop plants. *African Journal of Biotechnology* 10(40):7723-7726 (2011).

61. Irzykowska, L. and Bocianowski, J. Genetic variation, pathogenicity and mycelial growth rate differentiation between *Gaeumannomyces graminis* var. tritici isolates derived from winter and spring wheat. *Annals of Applied Biology* 152(3):369-375 (2008).

62. Jaccard, P. Nouvelles recherché sur la distribution florale. 44 (1908).

63. Jiang, C. and Sink, K. C. RAPD and SCAR markers linked to the sex expression locus M in asparagus. *Euphytica* 94(3):329-333 (1997).

64. Joersbo, M., Marcussen, J. and Brunstedt, J. *In vivo* modification of the cell wall polysaccharide galactomannan of guar transformed with α -galactosidase gene cloned from senna. *Molecular Breeding* 7(3):211-219 (2001).

65. Joersbo, M., Pedersen, S. G., Nielsen, J. E., Marcussen, J. and Brunstedt, J. Isolation and expression of two cDNA clones encoding UDP-galactose epimerase expressed in developing seeds of the endospermous legume guar. *Plant Science* 149:147–154 (1999).

66. Johal, G. S., Balint-Kurti, P. and Weil, C. F. Mining and Harnessing Natural Variation: A Little MAGIC. *Crop Science* 49:2066-2073 (2008).

67. Jones, E., Dupal, M., Dumsday, J., Hughes, L. and Forster, J. An SSR-based genetic linkage map for perennial ryegrass (*Lolium perenne* L.). *TAG Theoretical and Applied Genetics* 105(4):577-584 (2002).

68. Jones, E. S., Dupal, M. P., Kolliker, R., Drayton, M. C. and Forster, J. W. Development and characterisation of simple sequence repeat (SSR) markers for perennial ryegrass (*Lolium perenne* L.). *TAG Theoretical and Applied Genetics* 102(2):405-415 (2001).

69. Jones, N., Ougham, H., Thomas, H. and Pasakinskiene, I. Markers and mapping revisited: finding your gene. *New Phytologist* 183(4):935-966 (2009).

70. Julio, E., Verrier, J. L. and Dorlhac de Borne, F. Development of SCAR markers linked to three disease resistances based on AFLP within *Nicotiana tabacum* L. *TAG Theoretical and Applied Genetics* 112(2):335-346 (2006).

71. Karagyozov, L., Kalcheva, I. D. and Chapman, V. M. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. *Nucleic Acids Research* 21(16):3911-3912 (1993).

72. Kawamura, Y. GUAR GUM Chemical and Technical Assessment. www.fao.org/ag/agn/agns/jecfa/cta/69/Guar_gum_CTA_69.pdf

73. Khasa, D. P., Nadeem, S., Thomas, B., Robertson, A. and Bousquet, J. Application of SSR markers for parentage analysis of Populus clones. *Forest Genetics* 10(4):273-282 (2003).

74. Kiss, G. B., Csanadi, G., Kalman, K., Kaló, P. and Ökrész, L. Construction of a basic genetic map for alfalfa using RFLP, RAPD, isozyme and morphological markers. *Molecular and General Genetics MGG* 238(1):129-137 (1993).

75. Kubisiak, T. L., Nelson, C. D., Nance, W. L. and Stine, M. RAPD linkage mapping in a longleaf pine x slash pine F 1 family. *TAG Theoretical and Applied Genetics* 90(7):1119-1127 (1995).

76. Kumar, D. Silver jubilee workshop annual progress report 2008-09. Jodhpur, Rajasthan: Central Arid Zone Research Institute (2009).

77. Kumar, D. and Singh, N. B. Guar in India. Jodhpur: Scientific Publishers (2002).

78. Kunihisa, M., Fukino, N. and Matsumoto, S. Development of cleavage amplified polymorphic sequence (CAPS) markers for identification of strawberry cultivars. *Euphytica* 134(2):209-215 (2003).

79. Lacape, J. M. and Nguyen, T. B. Mapping quantitative trait loci associated with leaf and stem pubescence in cotton. *Journal of Heredity* 96(4):441-444 (2005).

80. Li, G., Hubert, S., Bucklin, K., Ribes, V. and Hedgecock, D. Characterization of 79 microsatellite DNA markers in the Pacific oyster *Crassostrea gigas. Molecular Ecology Notes* 3(2):228-232 (2003).

81. Liu, W. X. and Hou, A. F. Genetic relationship of guar commercial cultivars. *Chinese Agricultural Science Bulletin* 25(2):133-138 (2009).

82. Lloyd, K. O. Isolation, characterization, and partial structure of peptido galactomannans from the yeast form of *Cladosporium werneckii. Biochemistry* 9(17):3446-3453 (1970).

83. Maheswaran, M., Subudhi, P. K., Nandi, S., Xu, J. C., Parco, A., Yang, D. C. and Huang, N. Polymorphism, distribution, and segregation of AFLP markers in a doubled haploid rice population. *TAG Theoretical and Applied Genetics* 94(1):39-45 (1997).

84. Martin, G. B., Williams, J. G. and Tanksley, S. D. Rapid identification of markers linked to a Pseudomonas resistance gene in tomato by using random primers and near-isogenic lines. *Proceedings of the National Academy of Sciences* 88(6):2336 (1991).

85. Martins, W., De Sousa, D., Proite, K., Guimaraes, P., Moretzsohn, M. and Bertioli, D. New softwares for automated microsatellite marker development. *Nucleic Acids Research* 34(4):e31 (2006).

86. McCouch, S. R., Teytelman, L., Xu, Y., Lobos, K. B., Clare, K., Walton, M., Fu, B., Maghirang, R., Li, Z. and Xing, Y. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research* 9(6):199-207 (2002).

87. Melotto, M., Afanador, L. and Kelly, J. D. Development of a SCAR marker linked to the I gene in common bean. *Genome* 39(6):1216-1219 (1996).

88. Metais, I., Hamon, B., Jalouzot, R. and Peltier, D. Structure and level of genetic diversity in various bean types evidenced with microsatellite markers isolated from a genomic enriched library. *TAG Theoretical and Applied Genetics* 104(8):1346-1352 (2002).

89. Meudt, H. M. and Clarke, A. C. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science* 12(3):106-117 (2007).

90. Mir, R. R., Rustgi, S., Sharma, S., Singh, R., Goyal, A., Kumar, J., Gaur, A., Tyagi, A. K., Khan, H. and Sinha, M. K. A preliminary genetic analysis of fibre traits and the use of new genomic SSRs for genetic diversity in jute. *Euphytica* 161(3):413-427 (2008).

91. Morris, J. B. Morphological and reproductive characterization of guar (*Cyamopsis tetragonoloba*) genetic resources regenerated in Georgia, USA. *Genetic Resources and Crop Evolution* 57(7):985-993 (2010).

92. Mullis, K. B. and Faloona, F. A. [21] Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 155:335-350 (1987).

93. Muthusamy, S., Kanagarajan, S. and Ponnusamy, S. Efficiency of RAPD and ISSR markers system in accessing genetic variation of rice bean (*Vigna umbellata*) landraces. *Electronic Journal of Biotechnology* 11(3):32-41 (2008).

94. Naoumkina, M. and Dixon, R. A. Characterization of the mannan synthase promoter from guar (*Cyamopsis tetragonoloba*). *Plant Cell Reports* 30(6):997-1006 (2011).

95. Naoumkina, M., Torres-Jerez, I., Allen, S., He, J., Zhao, P. X., Dixon, R. A. and May, G. D. Analysis of cDNA libraries from developing seeds of guar (*Cyamopsis tetragonoloba* (L.) Taub). *BMC Plant Biology* 7:62 (2007).

96. Negi, M. S., Devic, M., Delseny, M. and Lakshmikumaran, M. Identification of AFLP fragments linked to seed coat colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. *TAG Theoretical and Applied Genetics* 101(1):146-152 (2000).

97. Oh, T. J., Gorman, M. and Cullis, C. A. RFLP and RAPD mapping in flax (*Linum usitatissimum*). *TAG Theoretical and Applied Genetics* 101(4):590-593 (2000).

98. Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proceedings of the National Academy of Sciences* 86(8):2766 (1989).

99. Overbeeke, N., Fellinger, A. J., Toonen, M. Y., Wassenaar, D. and Verrips, C. T. Cloning and nucleotide sequence of the α-galactosidase cDNA from *Cyamopsis tetragonoloba* (guar). *Plant Molecular Biology* 13:541-550 (1989).

100. Pathak, R., Singh, S. K., Singh, M. and Henry, A. Molecular assessment of genetic diversity in cluster bean (*Cyamopsis tetragonoloba*) genotypes. *Journal of Genetics* 89(2):243-246 (2010).

101. Patil, C. G. Nuclear DNA amount variation in *Cyamopsis* DC (Fabaceae). *Cytologia* 69(1):59-62 (2004).

102. Peakall, R. and Smouse, P. E. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6(1):288-295 (2006).

103. Pitkänen, L., Tuomainen, P., Mikkonen, K. S. and Tenkanen, M. The effect of galactose side units and mannan chain length on the macromolecular characteristics of galactomannans. *Carbohydrate Polymers* 86:1230– 1235 (2011).

104. Poats, F. J. Guar, a summer row crop for the Southwest. *Economic Botany* 14(3):241-246 (1958).

105. Poksay, K. S. and Schneeman, B. O. Pancreatic and intestinal response to dietary guar gum in rats. *The Journal of Nutrition* 113(8):1544-1549 (1983).

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

107. Prabaharan, M. Prospective of guar gum and its derivatives as controlled drug delivery systems. *International Journal of Biological Macromolecules* 49(2):117-124 (2011).

108. Prevost, A. and Wilkinson, M. J. A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *TAG Theoretical and Applied Genetics* 98(1):107-112 (1999).

109. Punia, A., Arora, P., Yadav, R. and Chaudhury, A. Optimization and inference of PCR conditions for genetic variability studies of commercially important cluster bean varieties by RAPD analysis. *Asia Pacific Journal of Molecular Biology and Biotechnology* 17:33-38 (2009).

110. Rafalski, A. Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology* 5(2):94-100 (2002).

111. Rafalski, J. A. and Tingey, S. V. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends in Genetics* 9(8):275-280 (1993).

112. Rafalski, J. A., Vogel, J. M., Morgante, M., Powell, W., Andre, C. and Tingey, S. V. Generating and using DNA markers in plants. San Diego, USA: Academic Press Inc. (1996).

113. Rakoczy-Trojanowska, M. and Bolibok, H. Characteristics and a comparison of three classes of microsatellite-based markers and their application in plants. *Cellular and Molecular Biology Letters* 9(2):221-238 (2004).

114. Ranade, S. A., Verma, A., Gupta, M. and Kumar, N. RAPD profile analysis of betel vine cultivars. *Biologia Plantarum* 45(4):523-527 (2002).

115. Reid, J. S., Edwards, M. and Dea, I. Biosynthesis of galactomannan in the endosperms of developing fenugreek (*Trigonella foenum-graecum* L.) and guar (*Cyamopsis tetragonoloba*[L.] Taub.) seeds. *Food Hydrocolloids* 1(5):381-385 (1987).

116. Reid, J. S. G., Edwards, M., Gidley, M. J. and Clark, A. H. Enzyme specificity in galactomannan biosynthesis. *Planta* 195(4):489-495 (1995).

117. Reid, J. S. G., Edwards, M. E., Dickson, C. A., Scott, C. and Gidley, M. J. Tobacco transgenic lines that express fenugreek galactomannan galactosyltransferase constitutively have structurally altered galactomannans in their seed endosperm cell walls. *Plant Physiology* 131(3):1487-1495 (2003).

118. Ribaut, J. M., Hu, X., Hoisington, D. and González-de-León, D. Use of STSs and SSRs as rapid and reliable preselection tools in a marker-assisted selection-backcross scheme. *Plant Molecular Biology Reporter* 15(2):154-162 (1997).

119. Richmond, T. A. and Somerville, C. R. The cellulose synthase superfamily. *Plant Physiology* 124(2):495-498 (2000).

120. Riju, A. and Arunachalam, V. Data mining for simple sequence repeats in oil palm expressed sequence tags. http://hdl.handle.net/10101/npre.2009.3581.1 (2009).

121. Ritschel, P., Lins, T., Tristan, R., Buso, G., Buso, J. and Ferreira, M. Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). *BMC Plant Biology* 4(1):9 (2004).

122. Robinson, A. J., Love, C. G., Batley, J., Barker, G. and Edwards, D. Simple sequence repeat marker loci discovery using SSR primer. *Bioinformatics* 20(9):1475-1476 (2004).

123. Roder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M. H., Leroy, P. and Ganal, M. W. A microsatellite map of wheat. *Genetics* 149(4):2007-2023 (1998).

124. Roder, M. S., Plaschke, J., Konig, S. U., Borner, A., Sorrells, M. E., Tanksley, S. D. and Ganal, M. W. Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular and General Genetics MGG* 246(3):327-333 (1995).

125. Roldan-Ruiz, I., Calsyn, E., Gilliland, T. J., Coll, R., Van Eijk, M. J. T. and De Loose, M. Estimating genetic conformity between related ryegrass (Lolium) varieties. 2. AFLP characterization. *Molecular Breeding* 6(6):593-602 (2000).

126. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. and Erlich, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239(4839):487-491 (1988).

127. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and Arnheim, N. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230(4732):1350-1354 (1985).

128. Saliba-Colombani, V., Causse, M., Gervais, L. and Philouze, J. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. *Genome* 43(1):29-40 (2000).

129. Sambrook, J. and Russell, D. W. Molecular cloning: a laboratory manual. Cold Spring Hrbor, New York: CSHL press (2001).

130. Sandhu, A. P. S., Randhawa, G. S. and Dhugga, K. S. Plant cell wall matrix polysaccharide biosynthesis. *Molecular Plant* 2(5):840-850 (2009).

131. Sangwan, R. S., Sangwan, N. S., Jain, D. C., Kumar, S. and Ranade, S. A. RAPD profile based genetic characterization of chemotypic variants of *Artemisia annua* L. *Biochmistry and Molecular Biology International* 47(6):935-944 (1999).

132. Scherbukhin, V. D. and Anulov, O. V. Legume seed galactomannans. *Applied Biochemistry and Microbiology* 35(3):229-244 (1999).

133. Schlötterer, C. and Tautz, D. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* 20(2):211-215 (1992).

134. Semagn, K., Bjornstad, Ã. and Ndjiondjop, M. N. An overview of molecular marker methods for plants. *African Journal of Biotechnology* 5(25):2540-2568 (2006).

135. Seopsi, L., Larsson, L. I., Bastholm, L. and Nielsen, M. H. Silver-enhanced colloidal gold probes as markers for scanning electron microscopy. *Histochemistry and Cell Biology* 86(1):35-41 (1986).

136. Shasany, A. K., Darokar, M. P., Dhawan, S., Gupta, A. K., Gupta, S., Shukla, A. K., Patra, N. K. and Khanuja, S. P. S. Use of RAPD and AFLP markers to identify inter- and intraspecific hybrids of Mentha. *Journal of Heredity* 96(5):542-549 (2005).

137. Singh, V. P. Induced high yielding mutants in clusterbean. *Indian Journal of Agricultural Sciences* 56:695-700 (1986).

138. Soltis, D. E., Moore, M. J., Burleigh, G. and Soltis, P. S. Molecular markers and concepts of plant evolutionary relationships: progress, promise, and future prospects. *Critical Reviews in Plant Sciences* 28(1-2):1-15 (2009).

139. Soppimath, K. S., Kulkarni, A. R. and Aminabhavi, T. M. Chemically modified polyacrylamide-g-guar gum-based crosslinked anionic microgels as pH-sensitive drug delivery systems: preparation and characterization. *Journal of Controlled Release* 75(3):331-345 (2001).

140. Srivastava, M. and Kapoor, V. P. Seed galactomannans: An overview. *Chemistry & Biodiversity* 2(3):295-317 (2005).

141. Stafford, R. E. Yield stability of guar breeding lines and cultivars. *Crop Science* 2(5):1009-1011 (1982).

142. Stafford, R. E. Inheritance of partial male-sterility in guar. *Plant Breeding* 103(1):43-46 (1989).

143. Tanksley, S. D. and McCouch, S. R. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277(5329):1063-1066 (1997).

144. Tautz, D. and Renz, M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* 12(10):4127-4138 (1984).

145. Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S. and McCouch, S. Computational and experimental analysis of microsatellites in rice

(*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. *Genome Research* 11(8):1441-1452 (2001).

146. Thiel, T., Michalek, W., Varshney, R. and Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *TAG Theoretical and Applied Genetics* 106(3):411-422 (2003).

147. van der Voort, J. R., Van Zandvoort, P., Van Eck, H. J., Folkertsma, R. T., Hutten, R. C. B., Draaistra, J., Gommers, F. J., Jacobsen, E., Helder, J. and Bakker, J. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. *Molecular and General Genetics MGG* 255(4):438-447 (1997).

148. Vandamme, E. J., De Baets, S. and Steinbuchel, A. Biopolymers, Vol. 6: polysaccharides II, polysaccharides from Eukaryotes: Wiley-VCH Weinheim, Germany (2002).

149. Varshney, R. K., Chabane, K., Hendre, P. S., Aggarwal, R. K. and Graner, A. Comparative assessment of EST-SSR, EST-SNP and AFLP markers for evaluation of genetic diversity and conservation of genetic resources using wild, cultivated and elite barleys. *Plant Science* 173(6):638-649 (2007).

150. Wang, C. and Roberts, P. A. Development of AFLP and derived CAPS markers for root-knot nematode resistance in cotton. *Euphytica* 152(2):185-196 (2006).

151. Wang, X., Wang, J., Zhang, J., Zhao, B., Yao, J. and Wang, Y. Structure-antioxidant relationships of sulfated galactomannan from guar gum. *International Journal of Biological Macromolecules* 46(1):59-66 (2009).

152. Wang, Y., Shi, Y. and Guo, X. Identification and characterization of 66 EST-SSR markers in the eastern oyster *Crassostrea virginica* (Gmelin). *Journal of Shellfish Research* 28(2):227-234 (2009).

153. Welsh, J., Chada, K., Dalal, S. S., Cheng, R., Relph, D. and McClelland, M. Arbitrarily primed PCR fingerprinting of RNA. *Nucleic Acids Research* 20(19):4965-4970 (1992).

154. Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18(22):6531-6535 (1990).

155. Yan, Z., Denneboom, C., Hattendorf, A., Dolstra, O., Debener, T., Stam, P. and Visser, P. B. Construction of an integrated map of rose with AFLP, SSR, PK, RGA, RFLP, SCAR and morphological markers. *TAG Theoretical and Applied Genetics* 110(4):766-777 (2005).

156. Yoon, S. J., Chu, D. C. and Juneja, L. R. Chemical and physical properties, safety and application of partially hydrolized guar gum as dietary fiber. *Journal of Clinical Biochemistry and Nutrition* 42(1):1-7 (2008).

157. Zamir, D. Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics* 2(12):983-989 (2001).

158. Zane, L., Bargelloni, L. and Patarnello, T. Strategies for microsatellite isolation: a review. *Molecular Ecology* 11(1):1-16 (2002).

159. Zavodna, M., Arens, P., Van Dijk, P. J. and Vosman, B. Development and characterization of microsatellite markers for two dioecious *Ficus* species. *Molecular Ecology Notes* 5(2):355-357 (2005).

160. Zhang, L. M., Zhou, J. F. and Hui, P. S. A comparative study on viscosity behavior of water-soluble chemically modified guar gum derivatives with different functional lateral groups. *Journal of the Science of Food and Agriculture* 85(15):2638-2644 (2005).

161. Zietkiewicz, E., Rafalski, A. and Labuda, D. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20(2):176-183 (1994).

## Appendix II : Characteristics of the SSR motifs and the sequences of the primers designed

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Contig1001 | 1 | p2 | (AC)6 | 12 | 167 | 178 | CT001F | CTGACCAAAGCCTTTTCCTG | CT001R | TGATGACGATGGAGATGGAA | 238 |
| Contig1004 | 2 | p3 | (ATA)5 | 15 | 266 | 280 | CT002F | ACTGAAATGGTGGCAGAAGG | CT002R | TGTGAAAGCGAAGAAAGGTG | 232 |
| Contig1006 | 1 | p2 | (TC)6 | 12 | 351 | 362 | CT003F | TCCCTGGCTCTTCAATCACT | CT003R | TTTGGGGACGACCTATTTGA | 250 |
| Contig1012 | 1 | p3 | (CTT)5 | 15 | 46 | 60 | CT004F | CACGAGGTTCACACTTTCCTT | CT004R | TGATCCGAGAAATCCTCCAC | 180 |
| Contig1019 | 1 | c | (TAA)5tggcatctgg(ATC)7 | 46 | 106 | 151 | CT005F | CCAGAGGAGTGAATTAAGAAACAAA | CT005R | GGGTTCATTCCAGATTGCAT | 195 |
| Contig1029 | 1 | p3 | (CTT)6 | 18 | 55 | 72 | CT006F | GGTTCCCTCTGTTCAAGCAA | CT006R | TGGTCACAGAGCTTGTCAGG | 155 |
| Contig1068 | 1 | p3 | (TTC)5 | 15 | 132 | 146 | CT007F | CACGAGGCAGAGACTTCACTC | CT007R | GTCAAGGGGCATTTGTCAGT | 239 |
| Contig1074 | 1 | p3 | (ATT)5 | 15 | 151 | 165 | CT008F | GAGGCATAAAGTCCCTTCCA | CT008R | CTGCCCTTCACACTTTCGAT | 248 |
| Contig1112 | 1 | p3 | (AAT)6 | 18 | 67 | 84 | CT009F | TGAGAATTCAAAGGAGCAAGTG | CT009R | TGCTGCTTCTTGTTCTGCTC | 184 |
| Contig1125 | 1 | p3 | (GCA)5 | 15 | 785 | 799 | CT010F | AATGCAGAAGCAACGGAAAT | CT010R | CCTCTCGCCTTCTGTTCTGT | 215 |
| Contig114 | 1 | p3 | (ATA)6 | 18 | 267 | 284 | CT011F | CCCCACACTCACTCACACAC | CT011R | GTTCGGTCAAACCTTCATGG | 229 |
| Contig1167 | 1 | p2 | (AT)8 | 16 | 365 | 380 | CT012F | ACAAGGAGCCACTCGTTCAG | CT012R | CACGCACGCATGATACATAG | 232 |
| Contig1182 | 1 | p3 | (CAA)5 | 15 | 104 | 118 | CT013F | CACGAGGCAGGAACAGAACT | CT013R | TCTCGCTTGGACATTCTTCA | 200 |
| Contig1190 | 1 | c | (T)15caagtgtatttttgttta(ATT)7 | 54 | 253 | 306 | CT014F | TTCCGATCCTGTTTGTTTCC | CT014R | CGTGTCTGTTACCGGATGTG | 225 |
| Contig1195 | 1 | p3 | (ATG)5 | 15 | 145 | 159 | CT015F | TGGCAGAAATGGAAAATTGA | CT015R | TGAGGTTCATTCTCCACCAA | 169 |
| Contig1197 | 1 | p2 | (TA)7 | 14 | 118 | 131 | CT016F | CAAAACAAAACAAAACAAGAAACAA | CT016R | TACCGCTTGAAGACCAAACC | 159 |
| Contig120 | 1 | p3 | (CAA)5 | 15 | 108 | 122 | CT017F | CACCACCACAAGCACAGATT | CT017R | TTCCAGTCCGAAAATTCCAC | 220 |
| Contig1242 | 1 | p3 | (GCT)6 | 18 | 370 | 387 | CT018F | TTGCTGTTGCTGTTGAGACC | CT018R | AGCAAAGCCAGCATCAGC | 218 |
| Contig1248 | 1 | c | (ATA)5ttttctatttacaactgcttcgagacttaaatataatgttatatataataataaaaagttcacacacaactttact(A)25 | 116 | 594 | 709 | CT019F | TGATGGAGGCACCTTCAGAT | CT019R | CCAAGCCGAGTCTTTCTTGA | 232 |
| Contig1260 | 2 | p3 | (TTG)6 | 18 | 979 | 996 | CT020F | CCACCTTCCATGAACCTTGT | CT020R | CACCAACAATGATGCTTCCA | 153 |
| Contig1286 | 1 | p2 | (TA)6 | 12 | 400 | 411 | CT021F | GAGAAGACGATCCGTTACGC | CT021R | CACATGCATGATTCAGACCA | 241 |
| Contig129 | 1 | p2 | (CT)11 | 22 | 52 | 73 | CT022F | CTCAGCCTCAGCCACTCACT | CT022R | TCAAGGTCACCAACACGGTA | 160 |
| Contig131 | 1 | p2 | (TA)6 | 12 | 1168 | 1179 | CT023F | GCCATCAAGGACATCTCCAG | CT023R | GATCACACACACAAATAACTAAGATTT | 250 |
| Contig1320 | 1 | p3 | (GAA)5 | 15 | 441 | 455 | CT024F | TTTCCGAGATGGTGAGAAGG | CT024R | CCTCCCGTTTGTTTTCTTCA | 233 |
| Contig146 | 1 | p3 | (TAA)6 | 18 | 408 | 425 | CT025F | GAAGAGTGCCCTTGCTATGC | CT025R | CACCACAAAATGTGGAGATACA | 182 |
| Contig1495 | 1 | c | (A)14cacacgaaaaatacaagaaatag(A)12 | 49 | 105 | 153 | CT026F | GATTGGGACCCACAATTCAC | CT026R | CATGAACAAGGCTCGTGAGA | 193 |
| Contig1524 | 1 | p2 | (TA)6 | 12 | 812 | 823 | CT027F | TCGATCATCCAAGAGGGACT | CT027R | TCCAATACCACCAACCCTTG | 210 |
| Contig1524 | 2 | p3 | (AAG)6 | 18 | 925 | 942 | CT028F | ATAAGGGGTCCCCATCATCA | CT028R | GCCAAACAAGAAGTGACAAACA | 152 |
| Contig1537 | 1 | p3 | (TTC)5 | 15 | 308 | 322 | CT029F | AGCAACTGCTGATGATGGTG | CT029R | AGGCAGATGAACTTGCCATC | 222 |
| Contig1597 | 1 | p3 | (ATC)6 | 18 | 238 | 255 | CT030F | TTCCAGCTCTTGTTGTCCTCT | CT030R | TGGCTGGATGGATTGTCTTA | 250 |
| Contig1602 | 1 | p3 | (TAA)6 | 18 | 520 | 537 | CT031F | GCCCTGCACTTTTGTGTTTT | CT031R | ATTTATGCTCTTCCCGCAAA | 175 |
| Contig1610 | 1 | p3 | (GAT)5 | 15 | 214 | 228 | CT032F | TCGCGACATGAACTTGTTTC | CT032R | TCGCCTTGAGACAACCTACC | 249 |
| Contig1632 | 1 | c | (T)13attttgttattttaccgatgtgtattgtgtaatataaaacttattctgtatc(A)13 | 78 | 522 | 599 | CT033F | AGATCATGGCAAGGCTTTTG | CT033R | CCGCTATAATGGGCATCTGT | 230 |
| Contig1636 | 1 | c | (ATC)5ataagtattcattagta(ACT)5 | 47 | 735 | 781 | CT034F | CACCAACACCAAGACTGCTT | CT034R | CGCGCAGTTTAATTTCATCA | 162 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Contig1639 | 1 | c | (A)51nant(A)14 | 69 | 698 | 766 | CT035F | GCACCATTTGGGCTTAATTATC | CT035R | CCTGATTATTTCCTCACACCCTA | 356 |
| Contig164 | 1 | p2 | (AG)11 | 22 | 170 | 191 | CT036F | CCCATGTGCCTTTGTTTTGT | CT036R | CAGGGCTCCCTCTTTTCTCT | 209 |
| Contig1665 | 1 | p2 | (AT)6 | 12 | 451 | 462 | CT037F | CACTGCCAGAATGGAATGAA | CT037R | CCATTCATTTTCGAATTCACC | 208 |
| Contig1687 | 1 | p3 | (CAC)6 | 18 | 593 | 610 | CT038F | CATCATGTTTGAGCCACGAC | CT038R | GCAATGGACCCTATCCTCCT | 179 |
| Contig1731 | 2 | p2 | (TA)6 | 12 | 1067 | 1078 | CT039F | AGTTGCATTGCACAGGTTTG | CT039R | TGGCCAAATTACAAGTAGAACAA | 190 |
| Contig1736 | 1 | p3 | (CAT)7 | 21 | 363 | 383 | CT040F | TTTTATCCGCATCCCAACAT | CT040R | GAGAGGATGGAAGGTTGAAGC | 227 |
| Contig1747 | 1 | p3 | (ATA)5 | 15 | 513 | 527 | CT041F | TTCCTTTGGAATCCACATCC | CT041R | ACCCAAATCACATCCACACA | 150 |
| Contig187 | 1 | p3 | (ATC)6 | 18 | 264 | 281 | CT042F | GGCTCTGATGCTTTCCAACT | CT042R | AGGAAATGGATGACGACGAC | 150 |
| Contig195 | 1 | p2 | (CT)6 | 12 | 92 | 103 | CT043F | CTTTTACCCATTCGGCCATT | CT043R | GTCATTGGGTCCATGGAAGT | 227 |
| Contig222 | 1 | p3 | (TCT)5 | 15 | 31 | 45 | CT044F | TCCCACCCTCATCTTCTTTC | CT044R | TCACTTCCTCGTCGGTTCTC | 231 |
| Contig253 | 1 | p3 | (TCT)9 | 27 | 28 | 54 | CT045F | CACGAGGCTTCTTCTTCATCA | CT045R | GCTTTGGTGAGTGGTTGGTT | 163 |
| Contig255 | 1 | p2 | (AT)13 | 26 | 351 | 376 | CT046F | CCCTGGTGCTAGAAGAGTGG | CT046R | TAGGTAGGGGGCTGGACTTT | 160 |
| Contig262 | 1 | p3 | (ATT)5 | 15 | 1279 | 1293 | CT047F | CCGTGGCCAATAGAAAGAGA | CT047R | GCTGTTCCATTCACCTCCTC | 233 |
| Contig269 | 1 | p3 | (AAG)5 | 15 | 679 | 693 | CT048F | GCCGAGAGCAGAGAAGAAGA | CT048R | TCCCTGGAAGTGATGGTAGG | 249 |
| Contig320 | 1 | p3 | (GAG)5 | 15 | 68 | 82 | CT049F | CTGGTGACAGTCGTTGCATT | CT049R | AGGCAACTTAACCCATGTCG | 214 |
| Contig324 | 1 | p2 | (AG)8 | 16 | 34 | 49 | CT050F | AGGGGTCAAGGGTGAAGACT | CT050R | GTTCCCATCCAATCCAACAC | 229 |
| Contig350 | 1 | c | (T)11acaaggaaattagcaatctatacatgaaaaattagcctttcaggaatgagtaaccaaaatgaagggaaaaaaaaccttattcctgca(T)10 | 108 | 273 | 380 | CT051F | TTTTCTACACAAATGAAGATATTAGCA | CT051R | GCTCCTGCTAGTCTGTCAATCA | 230 |
| Contig372 | 1 | p2 | (CT)10 | 20 | 97 | 116 | CT052F | GCTCTTCCCTCCCTCAATTC | CT052R | GGGTGACAGCGAAGAGTAGG | 171 |
| Contig375 | 1 | c | (T)10gggtttaggccttgcttttgttcccaac(A)11 | 49 | 449 | 497 | CT053F | ATGGAATTCACTTGGCTTGC | CT053R | GCCAAAGTTGCGCAATAAAG | 180 |
| Contig397 | 1 | p3 | (AGC)5 | 15 | 420 | 434 | CT054F | GTCATGGCTTGTTGGGATCT | CT054R | GTGGGAACAAGAGGACCAAA | 221 |
| Contig404 | 1 | p3 | (GCT)5 | 15 | 434 | 448 | CT055F | TTCCCTCAGGCTGACAAAGT | CT055R | CCCCATGTCCTCATCAGACT | 168 |
| Contig417 | 2 | p3 | (AAG)6 | 18 | 725 | 742 | CT056F | GATGCCATGAGTGGGAAAGT | CT056R | CACGAGGCTCTGTTTCTGTG | 216 |
| Contig419 | 1 | p2 | (CT)6 | 12 | 355 | 366 | CT057F | TCGCTCTACTGGCAGACCTT | CT057R | TAGAGGCAGTCAAGCCGAAT | 212 |
| Contig432 | 1 | p3 | (TCT)6 | 18 | 78 | 95 | CT058F | TCCCGCTAATTTAACCTCCA | CT058R | CGCGATCAGAATAACAAGCA | 206 |
| Contig487 | 1 | p3 | (TCA)5 | 15 | 430 | 444 | CT059F | AAAGCAATTTGCCCTTGCTA | CT059R | AGGAGCCCATAAACACAAGG | 242 |
| Contig489 | 1 | p2 | (CT)7 | 14 | 25 | 38 | CT060F | GCACTTCGAGGGGATTCA | CT060R | GGCTTATGGCTGTTGTGGTT | 222 |
| Contig507 | 1 | p3 | (CTT)5 | 15 | 56 | 70 | CT061F | CGAGGCTTCTAAGTTCATGG | CT061R | AAAGAGCCATACCCACATGC | 174 |
| Contig53 | 1 | p3 | (GTG)5 | 15 | 230 | 244 | CT062F | ACCGTCTTGAGCCAACTGTC | CT062R | CAAAGGATCCAATGCCAAGT | 215 |
| Contig552 | 1 | c | (T)10ccttttcg(T)10 | 28 | 80 | 107 | CT063F | CCCTCTCTTTCGAGGGTTTT | CT063R | AACCTCTGTGGCTGCATCTT | 188 |
| Contig559 | 1 | p2 | (TC)6 | 12 | 48 | 59 | CT064F | ATGCCTAATGGAGGGTCCTT | CT064R | TTTACCCTCCTCCTCCCCTA | 219 |
| Contig587 | 1 | p3 | (AGA)5 | 15 | 915 | 929 | CT065F | ACCCTACTGCTTCCCCATTT | CT065R | CCCTTATGGCTCTTGCTACG | 172 |
| Contig602 | 1 | p3 | (ATT)6 | 18 | 75 | 92 | CT066F | CCCCATTTAAGATATGAAAGAACTC | CT066R | TCAACATGGAGCCAGCAGTA | 204 |
| Contig602 | 3 | p3 | (ATG)5 | 15 | 534 | 548 | CT067F | CCTTTGTGGAGGCAGAACAT | CT067R | GAAGGGTATGGTGTCCCAGA | 231 |
| Contig602 | 2 | p3 | (ATT)6 | 18 | 408 | 425 | CT068F | CTGGCTCCATGTTGATGATG | CT068R | CAGCCTTGAGGCACTTCTCT | 249 |
| Contig605 | 1 | p2 | (AG)7 | 14 | 507 | 520 | CT069F | TGGCTGAAAAGCTGGTTCTT | CT069R | CAGGATGCTTTTACATTGTCCA | 215 |
| Contig627 | 1 | p3 | (ATA)5 | 15 | 139 | 153 | CT070F | TGTTTGGGCAATAACAAGGA | CT070R | TGGGAAATATCCGAAGCAAG | 233 |
| Contig642 | 1 | p3 | (GAT)5 | 15 | 70 | 84 | CT071F | TTCTGATTTTCTTTTGGCTCAAG | CT071R | TGCTGTGAACTTGGGTGAAG | 242 |
| Contig659 | 1 | p2 | (TC)6 | 12 | 389 | 400 | CT072F | TTTGACATTGGACTCGCTCA | CT072R | TGAAAAACAGGGTGGCATCT | 150 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Contig676 | 1 | p2 | (AG)9 | 18 | 124 | 141 | CT073F | GCTGCTGCAACCAACAAGTA | CT073R | CAGCAACGTTTTCACTTCCA | 169 |
| Contig68 | 1 | p3 | (GAA)7 | 21 | 264 | 284 | CT074F | TCCAAACATTGGCAAAACAA | CT074R | TCAAGCATGAAGCTCTTGGA | 202 |
| Contig702 | 1 | p2 | (CT)6 | 12 | 35 | 46 | CT075F | GGAAAGGCAGCACCATTG | CT075R | GAGAGTGGACTTCCCGTGAT | 215 |
| Contig705 | 1 | p2 | (GA)6 | 12 | 125 | 136 | CT076F | GCGAGCGATCTCACTCTTCT | CT076R | AGTTCGAAGCTCCGATAGCA | 202 |
| Contig717 | 1 | p3 | (ATA)7 | 21 | 206 | 226 | CT077F | ATGTTGCAGTTTCCCCACTC | CT077R | CAGCCAATGGAGTGAGATGTT | 187 |
| Contig743 | 1 | p2 | (AG)7 | 14 | 127 | 140 | CT078F | TTCACAATTCACCTCCGTTCT | CT078R | TCATTCAAACCAGCTGTGGA | 178 |
| Contig750 | 1 | p2 | (AT)7 | 14 | 62 | 75 | CT079F | CGCGTGGAAACAAAACTGAT | CT079R | AGAAGCTTCGTCAGCTCTGC | 155 |
| Contig756 | 1 | c | (T)11c(T)12 | 24 | 442 | 465 | CT080F | TTCACTCAGATCCACCACCA | CT080R | TGGATCAGGGACCAGAGAAG | 242 |
| Contig761 | 1 | p3 | (CCG)5 | 15 | 573 | 587 | CT081F | CCGCCGGTAGTACCATCAC | CT081R | CAGCTTCAACGTGTCAATCG | 159 |
| Contig794 | 1 | p2 | (TA)7 | 14 | 875 | 888 | CT082F | TCAAGACCACACTGCACTCC | CT082R | TGCAAATGCTGGTTCTTTTC | 226 |
| Contig806 | 1 | p3 | (TGG)5 | 15 | 223 | 237 | CT083F | TTGCTTCAATGCTTTCCTGA | CT083R | TGCCCAAAAGTCATTCACAG | 250 |
| Contig810 | 1 | p2 | (AT)8 | 16 | 141 | 156 | CT084F | AGCATTGCATAGCAGCCTCT | CT084R | TTTGGGAATTGGTTTGGAAG | 157 |
| Contig810 | 2 | p3 | (TTG)5 | 15 | 284 | 298 | CT085F | CTTCCAAACCAATTCCCAAA | CT085R | CTTGAACTCAACGTGCCTGA | 186 |
| Contig814 | 1 | c | (CT)14cctctgtctctgtgcacttctgtcacccaccgctctttaatttctttctccatgaaaccgttgaacgtttttcaactctcactgccttccagt(TCA)5 | 136 | 27 | 162 | CT086F | CACGAGGATTCTCTCTCTCTTTC | CT086R | GTATGTGTGGTGGTGGTGGA | 201 |
| Contig823 | 1 | p3 | (GAA)5 | 15 | 245 | 259 | CT087F | TGTTGTTGGTGGGAAATTGA | CT087R | TAATGAGCAGGCTTCCCTTC | 225 |
| Contig85 | 1 | p3 | (ATG)5 | 15 | 429 | 443 | CT088F | AGAAAACGTCCATGGCTGAC | CT088R | TCCCCCAACATAAGGAATGA | 245 |
| Contig884 | 1 | p3 | (CCT)5 | 15 | 190 | 204 | CT089F | GATGGCTCTTCCCTCTCTGA | CT089R | GGCTTCTCTGTTGGCTTCAC | 231 |
| Contig886 | 1 | p3 | (TGG)5 | 15 | 480 | 494 | CT090F | CACAAGGGAAAGGCAATGTT | CT090R | TCTGAACCTCCAATCACAGG | 209 |
| Contig889 | 1 | p2 | (CT)6 | 12 | 544 | 555 | CT091F | TCAGCAGCTCTTACCAAGCA | CT091R | TTTCGCTTCTTCCTGCACTT | 169 |
| Contig90 | 1 | p3 | (ATA)5 | 15 | 776 | 790 | CT092F | CCAGGTTCTGTGGATTGGAT | CT092R | TCTCGAAGCAGTTGTAAATAGAAAA | 200 |
| Contig904 | 1 | p2 | (TG)7 | 14 | 255 | 268 | CT093F | TCTGGAGTTGCAAGGTGTTG | CT093R | GGGATCCAGAGAGAATGCAG | 225 |
| Contig905 | 1 | p3 | (AAT)5 | 15 | 110 | 124 | CT094F | TCTTTCATGGTGGTTTTGATTG | CT094R | ACCTTGTGGATGGGTCAGAA | 217 |
| Contig909 | 1 | p2 | (TC)7 | 14 | 38 | 51 | CT095F | ACGAGGTTGAAGCCTCTGAA | CT095R | ACATCTGAGGGGACAACAGG | 197 |
| Contig912 | 1 | p3 | (CTT)5 | 15 | 54 | 68 | CT096F | CAATTTGCTTTGTGCCCTTT | CT096R | GGTCACAGAAGACGCTACCC | 198 |
| Contig913 | 1 | p3 | (CAA)5 | 15 | 26 | 40 | CT097F | CCCTAGCCTTCTCGTTCCTT | CT097R | CTTCATTAGCGCCCTTTTTG | 158 |
| Contig917 | 1 | p2 | (AG)6 | 12 | 122 | 133 | CT098F | CAGACGCTAGCAAGGAAACC | CT098R | TAATTGGATCGATCGGAAGC | 184 |
| Contig924 | 1 | p2 | (CT)8 | 16 | 60 | 75 | CT099F | CACGAGGAAGCCATGTTAAT | CT099R | TGCAACCCAAATTTGTGAAA | 186 |
| Contig95 | 1 | p3 | (GAA)7 | 21 | 60 | 80 | CT100F | CCCGTGTGTGTGAGAGAAAG | CT100R | TCGACAATGCTGAGCAAAAC | 155 |
| Contig950 | 1 | c | (A)11gaaaaggaaaaagcag(A)10tatgaaaattgtgtattattgaattttcttttcttcttctgtgtatgtttttatcatta(TCT)5 | 110 | 394 | 503 | CT101F | TTCATTAGCAACACCCACGA | CT101R | CATCATCAACCATCCAACCA | 237 |
| Contig952 | 1 | p3 | (CCG)5 | 15 | 400 | 414 | CT102F | CATCTCCATAGCCACGGTCT | CT102R | ACTGTTAACGAGGCGCAGTC | 232 |
| Contig964 | 1 | p2 | (TC)6 | 12 | 79 | 90 | CT103F | CAGGGTAGGGACCAGTGAAG | CT103R | CAGGGTAGGGACCAGTGAAG | 154 |
| Contig965 | 1 | p2 | (AG)6 | 12 | 570 | 581 | CT104F | TGGTTCACCAACTCCAATCA | CT104R | AAGCCAGAGCCATTTCTTCA | 243 |
| Contig973 | 1 | p3 | (TGC)5 | 15 | 337 | 351 | CT105F | CCTGCAGATAAGGTGCATGA | CT105R | GCGTGCAATTGGAAATTCTT | 219 |
| EG974822.1 | 1 | p2 | (AC)6 | 12 | 87 | 98 | CT106F | CACGAGGAATTGGTTACATTCTT | CT106R | TGGCTGCATGCATCATAAAC | 212 |
| EG974867.1 | 1 | p3 | (TTC)5 | 15 | 192 | 206 | CT107F | ACTTTTTGTCCCGCAGCTAA | CT107R | GGTGAAGAATATGGCGGAAA | 219 |
| EG974899.1 | 1 | p3 | (AAT)5 | 15 | 199 | 213 | CT108F | TGGAAACCATGAGAACAGGA | CT108R | TCATTTCCGGATTTTGTTTTG | 224 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG975055.1 | 1 | c | (AGG)6agatcatcatcagctgaatgccatacaagcatcattagctg(CTT)5ctgcttcgcgtaatttcttca(ATC)5 | 110 | 378 | 487 | CT109F | TGGTGGTAACAGCAACAGGA | CT109R | GCTTTCATCACCCAAGATGG | 243 |
| EG975100.1 | 1 | p3 | (CAG)7 | 21 | 95 | 115 | CT110F | GCACTGATCCATACCCACAA | CT110R | CATCTCCTGTGTCAGCCTCA | 168 |
| EG975276.1 | 1 | p3 | (ATG)6 | 18 | 251 | 268 | CT111F | CGGGAATTTTCCATCATCAC | CT111R | GCCATTTGGAATTCAAGCTC | 174 |
| EG975421.1 | 1 | c | (AAT)19gataatgac(AAT)16 | 114 | 346 | 459 | CT112F | GCAGCTGAAGATGTTGGACA | CT112R | CGAGGGTTGACTAGCTGACG | 246 |
| EG975619.1 | 1 | p3 | (TAC)6 | 18 | 400 | 417 | CT113F | GGTGAAGGTGATTGCTTGGT | CT113R | CGAGGGTTGCATTTCTCAAT | 240 |
| EG975762.1 | 1 | p3 | (ATT)5 | 15 | 221 | 235 | CT114F | TCGGGCTAGCTAAGTTCCTG | CT114R | TCCTGATGTGTTCCCCTCTC | 183 |
| EG975806.1 | 1 | p2 | (AT)6 | 12 | 446 | 457 | CT115F | ATCAATTTGGTGGGTTCCAA | CT115R | TTGGGCAATACAGTCAGCAA | 237 |
| EG975855.1 | 1 | p2 | (TA)11 | 22 | 479 | 500 | CT116F | ACTCAAGAAGCGGTGCTGAT | CT116R | TTCACTGAGAGGCTATACATAAATGA | 168 |
| EG975858.1 | 1 | p2 | (AT)6 | 12 | 614 | 625 | CT117F | TTGGGGAAAACTGATGAACC | CT117R | TGAGCAGGGCTATATATATGTGTGA | 237 |
| EG976115.1 | 1 | p3 | (TTA)7 | 21 | 41 | 61 | CT118F | AAGCGAAGCTCAAACACATT | CT118R | TGGCAAACTACAACTGTGTGG | 195 |
| EG976297.1 | 1 | c | (TA)7(GA)6 | 26 | 253 | 278 | CT119F | ATCCAAACATGAGGCCAGAA | CT119R | TGTGCCCATCAATCAACTTC | 209 |
| EG976538.1 | 1 | p2 | (TA)6 | 12 | 278 | 289 | CT120F | TTGCAATTAGGATCATTTTGTGA | CT120R | TGGTGCCTTTTCTGATCTCC | 234 |
| EG976648.1 | 1 | p3 | (AAT)5 | 15 | 61 | 75 | CT121F | AACATAGCAAAGGAGGTTGACA | CT121R | TTGTTGAATTATCATTTGGGTTTG | 223 |
| EG976743.1 | 1 | p4 | (AATC)5 | 20 | 58 | 77 | CT122F | TTCAACTCTCAGAAGGAACCTG | CT122R | TTGTTGTTACGATGATGAAGCTA | 100 |
| EG976752.1 | 1 | p4 | (TTAT)5 | 20 | 198 | 217 | CT123F | CCGTTGTGCAAATTGATGAG | CT123R | TCAAGTGCATTTTGCTAGGC | 243 |
| EG976765.1 | 1 | p3 | (CCT)5 | 15 | 46 | 60 | CT124F | CGAGGCTGAGGAGAGAATGT | CT124R | ATTGGGCAGCGCTTATACTG | 222 |
| EG976795.1 | 1 | p3 | (TTC)7 | 21 | 40 | 60 | CT125F | CGCTTCTTCACTGCCTCTTC | CT125R | AAAAGCTTCAGTTGCGGAAA | 186 |
| EG976948.1 | 1 | p2 | (TA)6 | 12 | 634 | 645 | CT126F | TTGATGTGTTTGCACTTGAGAA | CT126R | TGATGATAGAAGTCTACAAATTAGGG | 233 |
| EG977357.1 | 1 | p3 | (TTC)5 | 15 | 122 | 136 | CT127F | CTTTTCTCCCCATCCCTCTC | CT127R | CCGGCAACCTCCACTAATAA | 206 |
| EG977463.1 | 1 | p2 | (TA)9 | 18 | 320 | 337 | CT128F | TGGAAGCTCTCAGCTTTGGT | CT128R | TTGTTCCAGCAGACCATTGA | 248 |
| EG978172.1 | 1 | p3 | (CAT)6 | 18 | 37 | 54 | CT129F | ACGAGGGTCATTGTTCCCTA | CT129R | GCCACCCTCCTCTCTGTTCT | 224 |
| EG978199.1 | 1 | p2 | (TC)6 | 12 | 43 | 54 | CT130F | TCTTCTCGTAAACCACACATACAAA | CT130R | AAGTTCTTCCTCCGGCATCT | 155 |
| EG978224.1 | 1 | p3 | (TAA)6 | 18 | 376 | 393 | CT131F | GCCCTGCACTTTTGTGTTTT | CT131R | ATTTATGCTCTTCCCGCAAA | 175 |
| EG978249.1 | 1 | p3 | (GAA)7 | 21 | 405 | 425 | CT132F | TGTGGGTTTGGAACAGAACA | CT132R | ATCTGCTGAAGCGTCGTTTT | 227 |
| EG978492.1 | 1 | p3 | (AGC)5 | 15 | 212 | 226 | CT133F | AATTTGGACTTGGTGCAAGG | CT133R | GCAAAGGCAAGGATTGTAGG | 210 |
| EG978513.1 | 1 | p4 | (AATT)5 | 20 | 468 | 487 | CT134F | TGGATTGCTCCTTCTCATCC | CT134R | CATCCAAAGCCCTAGCTATTTTT | 181 |
| EG978866.1 | 1 | p3 | (ACT)6 | 18 | 328 | 345 | CT135F | TTTGTGTCATGGGAGACAGC | CT135R | CTCTTGCAACCAGCATCAAA | 225 |
| EG978879.1 | 1 | p2 | (TG)6 | 12 | 465 | 476 | CT136F | TGAATGGGAGGCTGTAGAGG | CT136R | AAATCTCAAATTTGCAACCTATTT | 221 |
| EG978908.1 | 1 | c | (AGC)5agaagttgcttgcattgtcctgtactccgaggaagagcgttgatccctcgtgccgaattcggcacgaggg(T)18 | 103 | 311 | 413 | CT137F | TGCAGAAGAGTAGCCATCCAT | CT137R | GCCCCATCTTAATTGGTTAGC | 235 |
| EG979029.1 | 1 | p2 | (AT)14 | 28 | 63 | 90 | CT138F | TCATAACAAATAACCATTGTTTGAGAA | CT138R | ATGCAGCACGTTTCTGATTG | 210 |
| EG979029.1 | 2 | p4 | (ATTT)5 | 20 | 280 | 299 | CT139F | ATCAGAAACGTGCTGCATGA | CT139R | TTCAGCCTAACCAAGGTACGA | 239 |
| EG979058.1 | 1 | p3 | (TTC)5 | 15 | 387 | 401 | CT140F | GTAGCCTCGAATTGGACGAG | CT140R | TGCATTAGATCTGGAATTGAAGC | 181 |
| EG979231.1 | 1 | p2 | (AT)13 | 26 | 57 | 82 | CT141F | TTCACAGGGCACCATACAAA | CT141R | TTACCATGCAAGGTGACAGC | 174 |
| EG979338.1 | 1 | p3 | (ATT)9 | 27 | 314 | 340 | CT142F | CGTCGGATTATCCAAGGTGA | CT142R | TTTCCGAAAAAGAAAATTGC | 207 |
| EG979403.1 | 1 | c | (A)11t(A)12 | 24 | 61 | 84 | CT143F | GAGGGGGAGCTTCTGTTTCT | CT143R | GGCTTCTTTTGGTTGGTGAA | 188 |
| EG979559.1 | 1 | p3 | (TAT)5 | 15 | 326 | 340 | CT144F | GTAGCCGGAGAAATGGGTCT | CT144R | AAGGGTATATTATCCAACCAACCA | 208 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG980002.1 | 1 | p2 | (AT)6 | 12 | 587 | 598 | CT145F | AGGTCACTGCCAGAATGGAA | CT145R | GAATGAAAAAGAAAAAGGTATTCATGT | 150 |
| EG980080.1 | 1 | p3 | (ATT)6 | 18 | 261 | 278 | CT146F | CCACTGCCATTGAAGTTCCT | CT146R | AAAATGAAACAGGTCAAATCATACA | 237 |
| EG980192.1 | 1 | p3 | (ATT)5 | 15 | 76 | 90 | CT147F | CGTCCATAAATCAAGTCTTAGCA | CT147R | ATGGAAGCTGCAAGATCGTC | 205 |
| EG980378.1 | 1 | p2 | (AG)6 | 12 | 35 | 46 | CT148F | GCAGAACAGAGAGTGCGAAA | CT148R | CCTCTAAGGCAGCATTCACC | 221 |
| EG980524.1 | 1 | p3 | (ATA)7 | 21 | 32 | 52 | CT149F | CCCCTTTTGACATACCATCC | CT149R | CCCCTTTTGACATACCATCC | 245 |
| EG980584.1 | 1 | c | (T)10aatttttttcttgagaagcatgtgagaagagtttgtgcaactatcattgctttatggaatttatgtatagggttgtcaaactagg(A)23 | 120 | 328 | 447 | CT150F | TCTAGATCATTCATAACAAGTTGCTG | CT150R | ATTGGATCCCGAGTCAAACA | 248 |
| EG980589.1 | 1 | p3 | (GTT)5 | 15 | 337 | 351 | CT151F | GCTGCCTTGTCTACTTTTCCA | CT151R | CCAACAAAACTTGGGTGCAT | 184 |
| EG980783.1 | 1 | p2 | (GA)6 | 12 | 36 | 47 | CT152F | AAAATTTCAGAGAGAGAAACACTCA | CT152R | GTCGATTTCCCATCAACACC | 183 |
| EG980783.1 | 2 | p4 | (AATA)6 | 24 | 428 | 451 | CT153F | GCACACTCACAAATCCTCCA | CT153R | TTTGCAGAAAGCAAACTAAACTACA | 196 |
| EG981086.1 | 1 | p3 | (TAC)5 | 15 | 46 | 60 | CT154F | CACGAGGGGAAGAGTTATCAA | CT154R | TGTACATGGCATTGGTAGATTTG | 243 |
| EG981277.1 | 1 | c | (TA)8a(AT)11 | 39 | 213 | 251 | CT155F | AAGGCAAGAGTCAGCCTTGA | CT155R | AATGAAAGATGATTATGACTAAAGCA | 157 |
| EG981282.1 | 1 | p4 | (AAGC)5 | 20 | 82 | 101 | CT156F | TTGATTTTAATGGAAGGGATGG | CT156R | AACCGGATTTTGTAGGAAGTGA | 211 |
| EG981591.1 | 1 | p3 | (TTC)6 | 18 | 170 | 187 | CT157F | CAACCTTCCTCTCCCGAACT | CT157R | CAACCTTCCTCTCCCGAACT | 235 |
| EG981772.1 | 1 | p2 | (AT)7 | 14 | 559 | 572 | CT158F | CCTGCAGCAATATGTGCCTA | CT158R | AGAGGGAGTGCCTTGGAATG | 168 |
| EG981899.1 | 1 | c | (A)10gagagaatcatttattaatgtatactagtagagtttgttggatgtgtatttgttagggtttaa(T)10 | 85 | 68 | 152 | CT159F | AGGGTAAAAGGCTTGGAGGA | CT159R | TCGGTTCATTTTCCAATAGGTC | 226 |
| EG981971.1 | 1 | p3 | (TAA)5 | 15 | 347 | 361 | CT160F | AGATGGATTCGGTTGGAATG | CT160R | GTCCTCCCTTCCATGTTCCT | 215 |
| EG982001.1 | 1 | p3 | (CCA)5 | 15 | 174 | 188 | CT161F | TCTGAACCTCCAATCACAGG | CT161R | CACAAGGGAAAGGCAATGTT | 209 |
| EG982463.1 | 1 | p3 | (CAT)5 | 15 | 29 | 43 | CT162F | CACGAGGGCCTAATAAACATACA | CT162R | TCACCTTCATTCAGCCATCA | 237 |
| EG982475.1 | 1 | p3 | (TAT)6 | 18 | 194 | 211 | CT163F | CCCGTCTCAATGAAGCATTT | CT163R | TCAACGCCCACACACATAAT | 243 |
| EG982557.1 | 1 | c* | (T)17(TG)7* | 30 | 82 | 111 | CT164F | GGGTTTCGCAGATAGCTTCA | CT164R | GGATTTTGCAGAAAGCAAGC | 211 |
| EG982716.1 | 1 | p3 | (TCT)10 | 30 | 76 | 105 | CT165F | TGAAGGGTTAGCTGCTGTCA | CT165R | GCTCGGCAAAATTCTTTTCA | 194 |
| EG983111.1 | 1 | p3 | (TTC)5 | 15 | 287 | 301 | CT166F | TGGATCGATCTTCCTGGTTC | CT166R | AGCTCCACAGTTGATTGATTGA | 225 |
| EG983141.1 | 1 | p2 | (AG)6 | 12 | 162 | 173 | CT167F | TCAATGCAAAGATCGGTCAA | CT167R | ACCCCAACGGAGCTAATCTT | 174 |
| EG983145.1 | 1 | p3 | (GAT)5 | 15 | 249 | 263 | CT168F | TTTCTCTGGGATCCATCCTG | CT168R | CCAAGCTAACCCTCCAACAA | 176 |
| EG983168.1 | 1 | p3 | (TTA)5 | 15 | 399 | 413 | CT169F | ATTCCCAAATGGCAAGACAC | CT169R | ACCACATGTGATGGTGCAAA | 201 |
| EG983172.1 | 1 | p3 | (GCT)5 | 15 | 533 | 547 | CT170F | CAAAAACGACGACAACATCG | CT170R | GCCTTACACGCAGCTCTTCT | 171 |
| EG983196.1 | 1 | c | (TC)8tgcatcttc(CT)10 | 45 | 48 | 92 | CT171F | CACGAGGTCTCTTTCTACTCTGC | CT171R | GTGGTGACCGAGAACAAGGT | 162 |
| EG983258.1 | 1 | p2 | (CT)7 | 14 | 40 | 53 | CT172F | CAGTCACGGTCATCTTCAGG | CT172R | AAACAAGCATCGGAAACAGC | 211 |
| EG983263.1 | 1 | p3 | (TAT)5 | 15 | 406 | 420 | CT173F | CGGTGAGGGAAGATGTTAGG | CT173R | CCTATTTTGCAAGACCCATAAA | 230 |
| EG983329.1 | 1 | p2 | (CT)7 | 14 | 62 | 75 | CT174F | GAGGGTTTCGTTGTCCAGAA | CT174R | GCCAGGAAAAGCACGAGATA | 186 |
| EG983349.1 | 1 | p3 | (CAC)8 | 24 | 398 | 421 | CT175F | GTGGATTTGGAGCAAGGAAA | CT175R | CCCTCCCAGAAACAGAATCA | 211 |
| EG983420.1 | 2 | p2 | (AT)6 | 12 | 494 | 505 | CT176F | CTTTCCTCCTACCATCAACCA | CT176R | GGAACATTCTGAAGATAATCGAG | 250 |
| EG983420.1 | 1 | p3 | (AGA)6 | 18 | 119 | 136 | CT177F | GAGGAACCATTCCACGTCAC | CT177R | TCATGGCGGTGGTTACACTA | 199 |
| EG983480.1 | 1 | p3 | (AAT)8 | 24 | 105 | 128 | CT178F | TTTTCTTCACCCGTTTCAGC | CT178R | TGATCGATGGTGGGGATACT | 152 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Produc t size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG983648.1 | 2 | p2 | (TC)6 | 12 | 216 | 227 | CT179F | AACCACCGTCTCGTTCTCTG | CT179R | TCCTTAAACGATCCCGACAC | 213 |
| EG983656.1 | 1 | p2 | (CT)11 | 22 | 211 | 232 | CT180F | TTTCTCCCTCCAACTCGAAA | CT180R | TGAAGCACCCCAACTAGGAC | 183 |
| EG983676.1 | 1 | p3 | (AAG)5 | 15 | 338 | 352 | CT181F | ATGGGTGATGGATTTGTGCT | CT181R | GCAGCAGATGAAGGAGGAAC | 209 |
| EG983687.1 | 1 | p3 | (CTT)5 | 15 | 310 | 324 | CT182F | GCCACTGCGGAAAATATCTC | CT182R | CCGTACGATTGGAGAAACGA | 237 |
| EG983747.1 | 1 | p3 | (AAG)6 | 18 | 51 | 68 | CT183F | TCTGTAAAAATTCAGACTTTGA CTCAC | CT183R | CATGTCAACGCAGGAAGAAA | 207 |
| EG983774.1 | 1 | p3 | (TTC)7 | 21 | 126 | 146 | CT184F | ATGGCTTCCATGATGAGACC | CT184R | TCCCATACTGCAGCTCCTCT | 209 |
| EG983848.1 | 1 | p4 | (AAAG)5 | 20 | 485 | 504 | CT185F | TGGATCATGAAATAATGGATGG | CT185R | TGATGATGGGTGTCAAATCG | 212 |
| EG983878.1 | 1 | p4 | (TTTA)5 | 20 | 46 | 65 | CT186F | CGAGGCATCTGTTTCTCTCA | CT186R | ATGGTTCTGCTGGAACAGGT | 232 |
| EG983882.1 | 2 | p3 | (ACT)6 | 18 | 143 | 160 | CT187F | AATGGGTAAGGCCTCAAGGT | CT187R | AGCGTTAGAAAGGGGATGGT | 166 |
| EG983939.1 | 1 | p3 | (AAG)6 | 18 | 111 | 128 | CT188F | CCACCACCACATTCCTCTTT | CT188R | TACTGGGAAGACCCATTTGC | 222 |
| EG983941.1 | 1 | p2 | (AT)6 | 12 | 114 | 125 | CT189F | ACTGACGCGAAAAGCAGATT | CT189R | ACAACCAATCTTGCACCTGA | 205 |
| EG984108.1 | 1 | p2 | (CT)7 | 14 | 162 | 175 | CT190F | GATTCTGAGCGAAGCCAAAC | CT190R | TGGAGGCGAAAGAGAGAAGA | 191 |
| EG984110.1 | 1 | p2 | (TA)6 | 12 | 91 | 102 | CT191F | CAACTCATCATCCTTGAATGAA A | CT191R | GGTCACTGCCAGAATGGAAT | 164 |
| EG984246.1 | 1 | p2 | (AG)7 | 14 | 34 | 47 | CT192F | CGAGGTTCTTCTTCAATTCCA | CT192R | GTTCCGATCCCGTAAACCTT | 219 |
| EG984248.1 | 2 | p3 | (ACA)6 | 18 | 191 | 208 | CT193F | ACCACCAACCAGATCCAAAA | CT193R | CGTAAGGGAAGGCGACATTA | 247 |
| EG984320.1 | 1 | c | (T)10acttcactacttcacaaaagggc atgctggatcatgcatc(TA)6 | 62 | 461 | 522 | CT194F | CCACCCCTTGATCCCTTTAT | CT194R | TCCGATGTGAATCTATTGGTCT | 228 |
| EG984351.1 | 1 | p3 | (TGA)5 | 15 | 195 | 209 | CT195F | ACTCCTCCGTCTGGAATCCT | CT195R | GGTTCTTGCTCTTCGTCGTC | 223 |
| EG984412.1 | 1 | p4 | (TATT)6 | 24 | 278 | 301 | CT196F | ACTTTGGCCCCTAAGGTGAT | CT196R | TCCTGGAGCAGTTAAATCAGG | 185 |
| EG984698.1 | 2 | c | (A)13cacatacccttttgttggttagttc tgaaagtgaaagaaagaa(AG)8 | 72 | 582 | 653 | CT197F | TGTGTAATGCTGGGGTCAGA | CT197R | CTCATCCTCATCATCTTCAATTT | 234 |
| EG984713.1 | 1 | c | (T)10gtatttt(G)13 | 29 | 612 | 640 | CT198F | TGCAAAGCGTTGACAGTTTC | CT198R | GGCAGCCTTTCCAAACTTAT | 159 |
| EG984763.1 | 1 | p3 | (TTA)5 | 15 | 104 | 118 | CT199F | GAGGGGTCCCTCTCTCAGTC | CT199R | ACCCTTTGGGATTTGGATTC | 161 |
| EG984904.1 | 1 | p3 | (GCT)5 | 15 | 21 | 35 | CT200F | CACGAGGCAAAGCTGAAGA | CT200R | GCGAAATTCTTCGCAATCTC | 230 |
| EG984927.1 | 1 | p3 | (AAT)7 | 21 | 128 | 148 | CT201F | CGCTGTGTTCACAAGGAAAG | CT201R | TGGTGGTAAGGGGCAAAGTA | 232 |
| EG984984.1 | 1 | c | (T)11gttc(T)10 | 25 | 75 | 99 | CT202F | TTGCTTCCCTTTTCAAGCAT | CT202R | TCACACTGAAATGACGCTGA | 193 |
| EG984992.1 | 1 | p3 | (AGA)5 | 15 | 428 | 442 | CT203F | AGTTTCCAAGCTCGACCTCA | CT203R | CGTCCCCAAAGTGTAAATGG | 170 |
| EG985006.1 | 1 | p3 | (CTT)6 | 18 | 130 | 147 | CT204F | TCGCACTGCAAAATCAAAAA | CT204R | ATTGAAGAGAACGGCGAAGA | 229 |
| EG985039.1 | 1 | p3 | (TTC)5 | 15 | 521 | 535 | CT205F | CGCCCTTTGTGAAGAACATT | CT205R | CGCTCTGTAGGGGACAAGAC | 218 |
| EG985094.1 | 1 | p2 | (TC)7 | 14 | 21 | 34 | CT206F | TTCCACCCAATCGTAGACTC | CT206R | AAACCGACGAGGTAGCGTAA | 160 |
| EG985197.1 | 1 | p3 | (CTG)5 | 15 | 336 | 350 | CT207F | CCCCCGAAGGATTAACATCT | CT207R | CACACCCTTTTAGCTGCACA | 211 |
| EG985259.1 | 1 | p3 | (TGC)5 | 15 | 35 | 49 | CT208F | AGGGGTGAACCCTGAAGAAG | CT208R | TCTTGGTGTTGGTGGTGATG | 235 |
| EG985290.1 | 1 | p3 | (AGA)6 | 18 | 183 | 200 | CT209F | CAACAGCAACAACTGCGTTT | CT209R | AACGTTGGAACTTTGGTTCG | 250 |
| EG985336.1 | 1 | p3 | (CAC)5 | 15 | 458 | 472 | CT210F | CTGTTACCACCGCCATACCT | CT210R | TGCATGCGTGGAAACAATAG | 215 |
| EG985385.1 | 1 | p3 | (TCT)6 | 18 | 105 | 122 | CT211F | TGTGCATCTCAAGCCTCAAC | CT211R | AGAAGAACAAGGGGCTGTGA | 177 |
| EG985406.1 | 1 | p3 | (TTC)5 | 15 | 24 | 38 | CT212F | CGAGGCTCAACTCCTTCATC | CT212R | CCCTTTCTTCATTGCAAACG | 218 |
| EG985483.1 | 1 | p3 | (GTG)5 | 15 | 224 | 238 | CT213F | TCTTAAGTGGCCAGCAAGGT | CT213R | GGGAAGGGAGTATTTGCAGAC | 228 |
| EG985748.1 | 1 | p3 | (ATG)6 | 18 | 472 | 489 | CT214F | AGAAGAAGTGGGCGAGTCAA | CT214R | AGATTGCTTGCTCCTGGAAA | 244 |
| EG985769.1 | 1 | p3 | (CTT)9 | 27 | 72 | 98 | CT215F | TCCACGCACTCTCTACACAAA | CT215R | GAAAGAGCGAAGGCTTCTTG | 195 |
| EG985926.1 | 1 | p3 | (CTG)7 | 21 | 133 | 153 | CT216F | AGGCAGTGGAGACCAACAAC | CT216R | AAGGGGAAATAGGACGGAGA | 237 |
| EG985956.1 | 1 | p2 | (TC)6 | 12 | 38 | 49 | CT217F | CACGAGGACTCACTCCACAC | CT217R | CACCAACTTGAGACGCCATA | 185 |
| EG985970.1 | 1 | p3 | (CAT)5 | 15 | 106 | 120 | CT218F | AGGATTCAAGGGAGGGAAGA | CT218R | TTCGCTCCTGAACATTTGTG | 230 |
| EG986067.1 | 1 | p3 | (CTC)7 | 21 | 87 | 107 | CT219F | TAGGGTTCCCATCTGATCCA | CT219R | GCTTCAAGGACGCTGGACT | 198 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG986153.1 | 1 | p3 | (AGA)5 | 15 | 470 | 484 | CT220F | GAAGAATTCCTCGGTGACGA | CT220R | GTCGGAAAAGTCGCTGTCAT | 219 |
| EG986181.1 | 1 | p3 | (TTC)6 | 18 | 91 | 108 | CT221F | CTGCAGGCTGAAATTCACAA | CT221R | TTCGTAGTGGTTGCCATGAT | 250 |
| EG986243.1 | 1 | p3 | (TCT)8 | 24 | 106 | 129 | CT222F | CCTTTACTGCGGGAACAGAA | CT222R | GGATCGGATCGAATCAGAAA | 162 |
| EG986259.1 | 1 | p2 | (TC)6 | 12 | 33 | 44 | CT223F | CAAATTTTTGCAGAGGCAGA | CT223R | GGCTTTGATGAGACCAGCTC | 250 |
| EG986322.1 | 2 | p2 | (AG)7 | 14 | 599 | 612 | CT224F | CTGCACCATTGATTGCAAGA | CT224R | CGAATCGATCAGTCACAACC | 248 |
| EG986322.1 | 1 | p3 | (ACA)5 | 15 | 96 | 110 | CT225F | CTCCTGGCTTCTTCATCCAC | CT225R | CACGATTCCATTTCCAAACC | 186 |
| EG986332.1 | 1 | p3 | (TCA)5 | 15 | 401 | 415 | CT226F | AGCAGCTTGGGCACTTGTAT | CT226R | CCACCCAAAGAAGAGGATGA | 211 |
| EG986337.1 | 1 | p3 | (AAG)7 | 21 | 250 | 270 | CT227F | GACTTAAACGATGGGGTCCA | CT227R | CATCAATTAGCGGCTTCTCC | 215 |
| EG986339.1 | 1 | p4 | (CATA)5 | 20 | 58 | 77 | CT228F | CACGAGGGTTGTTCAATTTTC | CT228R | AAGAGCGCGATTTCACAGAG | 151 |
| EG986346.1 | 1 | p2 | (TC)6 | 12 | 34 | 45 | CT229F | AGGGTCCTTTCCAACACACA | CT229R | CACTGCCATTGAAGATCAGG | 162 |
| EG986480.1 | 1 | p2 | (TC)6 | 12 | 143 | 154 | CT230F | TCGTCCTTTCTTTCCACTGC | CT230R | GCTACAGCCATGCATCTTCA | 209 |
| EG986499.1 | 1 | p3 | (CCA)7 | 21 | 90 | 110 | CT231F | GTCCCAGAATGGTTCATTGG | CT231R | ATACTGCTCGAGGGAGCTGA | 199 |
| EG986512.1 | 1 | p3 | (TCA)6 | 18 | 433 | 450 | CT232F | TGGAACCTCGACAGACACTG | CT232R | AATTCTAGGGCCAGGGACAC | 211 |
| EG986539.1 | 1 | p3 | (TTA)6 | 18 | 565 | 582 | CT233F | GTGACAAGGCAAGGCAAATG | CT233R | TCCCTCCATCTGAAAGAGGTT | 241 |
| EG986580.1 | 1 | p3 | (ACC)5 | 15 | 471 | 485 | CT234F | ACGACGTTCCTGACCAAAAC | CT234R | CAAATCCTGGCTGCTAGCTC | 215 |
| EG986681.1 | 1 | p2 | (AG)7 | 14 | 331 | 344 | CT235F | TGTTGGGTTGATGGAGAACA | CT235R | TGACACGGACAACACCTCAT | 201 |
| EG986800.1 | 1 | p2 | (CT)7 | 14 | 63 | 76 | CT236F | CACGAGGCTGTCTCTCATCTT | CT236R | GGTCGACGGAGATCTGAAAA | 242 |
| EG986810.1 | 1 | p2 | (TC)7 | 14 | 53 | 66 | CT237F | AGCTGTTTTCTGGGTTCGAG | CT237R | GCTTTTGGTTGAAAGGGTGA | 170 |
| EG986849.1 | 1 | p3 | (CAT)5 | 15 | 320 | 334 | CT238F | TGGTGTGCATTAACGATGCT | CT238R | ATGCTGGAAATGACGAAGGT | 209 |
| EG986895.1 | 1 | p3 | (TGC)5 | 15 | 61 | 75 | CT239F | CACGAGGGAGCAGTGTCAT | CT239R | GAAAAATCAAGGCGTTCGAG | 240 |
| EG986974.1 | 1 | p3 | (AAG)6 | 18 | 40 | 57 | CT240F | CACCGAAAACACAGTCATCG | CT240R | ATAGGGACGAAGCAGAAGCA | 158 |
| EG987120.1 | 1 | p3 | (TTC)5 | 15 | 230 | 244 | CT241F | CGTTCCCCTTCTTCATTCCT | CT241R | GAGTACTTGCCGACGGTCAT | 231 |
| EG987140.1 | 1 | p3 | (AGA)5 | 15 | 194 | 208 | CT242F | CGAACACCCTCTTCCAAAAA | CT242R | AAACGTTGGCTTCGATCTTG | 216 |
| EG987212.1 | 1 | p3 | (TCT)5 | 15 | 150 | 164 | CT243F | GGCTGGAAACCCTAATTGGT | CT243R | GGCTCAAGGATTGATGAGGA | 152 |
| EG987212.1 | 2 | p3 | (AAT)6 | 18 | 507 | 524 | CT244F | GCAAGAAAGTGCAGGGAACT | CT244R | ATGATGATCTGGTGCCACTG | 211 |
| EG987222.1 | 1 | p2 | (AG)9 | 18 | 42 | 59 | CT245F | GCTGAAACTTGAAACTAGTGGTGA | CT245R | TTTCAGCGAACTCCCAGTCT | 172 |
| EG987327.1 | 1 | p2 | (CT)6 | 12 | 24 | 35 | CT246F | GAGGCTTCGAAACCAAATCA | CT246R | GGTCGAATCTGAGGGATTCA | 160 |
| EG987355.1 | 1 | p2 | (GA)6 | 12 | 44 | 55 | CT247F | CCTCGCTCTCTTTTCCACTC | CT247R | TGAGGTTCGGGAATCGATAG | 225 |
| EG987580.1 | 1 | p3 | (AAG)6 | 18 | 167 | 184 | CT248F | GGAGAAGAAGATCCCCAAGG | CT248R | GACGACTCCTGAGCGAGTTT | 198 |
| EG987600.1 | 1 | p3 | (TTC)5 | 15 | 34 | 48 | CT249F | CGTCCGTTTGTACTTGACGA | CT249R | TTCCCAAAAAGGGGGAATAC | 227 |
| EG987715.1 | 1 | p3 | (GAA)5 | 15 | 214 | 228 | CT250F | ACATGGCTGTCGAAAATTCC | CT250R | ACCCAATTTCACAGGACCAG | 230 |
| EG987728.1 | 2 | p2 | (TA)7 | 14 | 396 | 409 | CT251F | AACGCTTTGATGCCATGTTT | CT251R | AAAGTTATGACAATTTTGCACCA | 233 |
| EG987762.1 | 1 | p3 | (ATT)5 | 15 | 209 | 223 | CT252F | TTGCCAACATGATCTTCACC | CT252R | TCAGGCACTCTAACTCGAAGC | 206 |
| EG987778.1 | 1 | p3 | (ATT)5 | 15 | 430 | 444 | CT253F | TGGAATGAAAGTGGCAGTGA | CT253R | CAAACAAAATTGACACAAATGG | 224 |
| EG987794.1 | 1 | p3 | (ACG)5 | 15 | 96 | 110 | CT254F | GACTCGTCGTTTCGCTCTTC | CT254R | AACGAATTGATCGGGATCTG | 179 |
| EG987876.1 | 1 | p3 | (TAC)5 | 15 | 21 | 35 | CT255F | CACGAGGCCCTTCCTTAAT | CT255R | ACGAAGAAGTGCGGTTGAGT | 220 |
| EG987882.1 | 1 | p3 | (TTA)6 | 18 | 101 | 118 | CT256F | TGAGCAAGTGAGGTGGCTAA | CT256R | GCACAATCGGCCATTTTTAT | 241 |
| EG987893.1 | 1 | c | (TC)7aatttcatctcttc(CT)6 | 40 | 26 | 65 | CT257F | CACGAGGCACAGATTCTCTT | CT257R | CATCCTCCAGAGGGAAATCA | 221 |
| EG987921.1 | 1 | p2 | (TG)7 | 14 | 368 | 381 | CT258F | CACCCTCCATCCCTAGATGA | CT258R | GCACCACAATCTCTTAACATGAA | 231 |
| EG987980.1 | 1 | p2 | (TG)7 | 14 | 471 | 484 | CT259F | TTCTTTCTGTTGGCATGCAG | CT259R | TCACACCACGAGAAGGAAGA | 205 |
| EG987980.1 | 2 | p4 | (TATG)5 | 20 | 606 | 625 | CT260F | ATGGCATGTGTGTGTGTGTG | CT260R | AACAAGGAATCAGAAACTGCAA | 202 |
| EG988086.1 | 1 | p2 | (TC)6 | 12 | 39 | 50 | CT261F | GCTTGCAAAGGTGCTCTTTC | CT261R | GAATGAAATCCCGAACCAGA | 173 |
| EG988096.1 | 1 | p3 | (GAA)6 | 18 | 172 | 189 | CT262F | GCCACCATGTGTGTACAGGA | CT262R | TTCGCGAGGTTCTGTTTCTT | 227 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG988114.1 | 1 | p2 | (CT)11 | 22 | 35 | 56 | CT263F | CGAGGGTTAATAGTAACTCATTCTTC | CT263R | TCTGAGCAGAGACCAGAGCA | 195 |
| EG988232.1 | 1 | p3 | (TTG)5 | 15 | 343 | 357 | CT264F | TCTTAGCAGCAACCCCACTT | CT264R | AAAAGGTAGCAACCCCAACA | 219 |
| EG988288.1 | 2 | p3 | (CTC)6 | 18 | 297 | 314 | CT265F | AACTCCGCCTGTCCAAGTTA | CT265R | GGAATGAGGGTATTGGAGCA | 203 |
| EG988349.1 | 1 | p3 | (TAA)7 | 21 | 347 | 367 | CT266F | TGGTTCCAACCCAAAAAGAC | CT266R | TCCCTTTTCCTTCACACCAC | 227 |
| EG988416.1 | 1 | p3 | (TCT)5 | 15 | 56 | 70 | CT267F | CATCCAATTCGTGGACTCAA | CT267R | CCCACCTTCCATAATCCTCA | 232 |
| EG988418.1 | 1 | p3 | (AAG)5 | 15 | 71 | 85 | CT268F | GGCAAAATTAGCCACACCAG | CT268R | ATGCTCGAACCTCTGCAAGT | 192 |
| EG988422.1 | 1 | p2 | (TC)8 | 16 | 44 | 59 | CT269F | GGGTCGTTTTGATTCCTTCA | CT269R | TTTTCTTGTGCTCGCTACCC | 242 |
| EG988471.1 | 1 | p3 | (TTC)5 | 15 | 107 | 121 | CT270F | CGAGGCTCAGTTGGTACTCC | CT270R | GAATCAAACGCAGGAAGGTC | 214 |
| EG988480.1 | 1 | p2 | (AG)6 | 12 | 241 | 252 | CT271F | GGTTTCGTTTGCGTTCTGAG | CT271R | ATGCCCGAACACACATACAC | 235 |
| EG988495.1 | 1 | p4 | (CCTT)5 | 20 | 118 | 137 | CT272F | CCACCATGACCATTTGATGA | CT272R | TGAGCTGATTGGATGTCGAG | 238 |
| EG988552.1 | 1 | p2 | (GA)8 | 16 | 51 | 66 | CT273F | GAGGATCGAAGCGAAGGAAT | CT273R | GTTGGAGAGGCTGTCTTGGA | 202 |
| EG988682.1 | 1 | p4 | (TCTT)5 | 20 | 75 | 94 | CT274F | CATCTTTTACGAGGCCGAAC | CT274R | CCAGCTCGAGTGAAAGAACA | 248 |
| EG988701.1 | 1 | p2 | (AG)8 | 16 | 31 | 46 | CT275F | TGTGAACAAAGGAAGAGAAGAGAA | CT275R | CATAGGCTTCACCGTTGTCC | 227 |
| EG988745.1 | 1 | p2 | (TC)13 | 26 | 223 | 248 | CT276F | CCACCCTGTCAACAACACAC | CT276R | TGGTAAAGGTCGGTGGGATA | 233 |
| EG988750.1 | 1 | p3 | (ATA)7 | 21 | 416 | 436 | CT277F | CAGGGTGGTCATGATCTCAA | CT277R | TGGCTTAACATCTTGCATGG | 249 |
| EG988856.1 | 1 | p2 | (TC)9 | 18 | 77 | 94 | CT278F | CATCAAGCAAAGCATCCTCA | CT278R | CGACGAGAGGGAATTGAAGT | 167 |
| EG988902.1 | 1 | p3 | (AAT)6 | 18 | 65 | 82 | CT279F | GCCCTTTCCAATTTCAATTTC | CT279R | CCCTGAAACTCGAGGAACAA | 232 |
| EG988915.1 | 1 | p2 | (AT)8 | 16 | 519 | 534 | CT280F | CCAGTTGTTCAAGACGGTGA | CT280R | TGCTAAACATGAAAGGATCGAA | 152 |
| EG988922.1 | 1 | c | (CT)7tttataaccttcattcacctcttcctccctttcaaacac(CT)11 | 75 | 37 | 111 | CT281F | ACGAGGCTCCCTTTGACTTC | CT281R | GCGATTTGTTCATCCATCCT | 214 |
| EG989006.1 | 1 | p3 | (TCT)13 | 39 | 74 | 112 | CT282F | GGCTCTCCTTTAACCCACAA | CT282R | GCCAGAACCAGGGACTTGTA | 184 |
| EG989101.1 | 1 | p2 | (TC)6 | 12 | 515 | 526 | CT283F | TCCTTGATCACATTCGCTTCT | CT283R | CACAACGGAATACAGGATTCAA | 244 |
| EG989229.1 | 1 | p2 | (AC)8 | 16 | 32 | 47 | CT284F | CACGAGGCTTCATCTCTCTAGC | CT284R | AGATTGTGTCGCTTCCGTCT | 171 |
| EG989244.1 | 1 | p3 | (CAT)6 | 18 | 164 | 181 | CT285F | GCGTTGAATCCCAAACTGAT | CT285R | GTCCAATACCAAGCCCCTTT | 250 |
| EG989252.1 | 1 | p3 | (CTT)5 | 15 | 473 | 487 | CT286F | CTTGAGCAGAAGCAGCAATG | CT286R | TTGAAAGGAAAGGCGAAGGT | 206 |
| EG989357.1 | 1 | c | (AT)7aggtttactaagtttcaaacaacaaaggatcattgctgcctttactag(T)13 | 75 | 202 | 276 | CT287F | TTTGTGCAAGGTTTTATGGAAA | CT287R | CCCATTGGAGGACATGAAAT | 243 |
| EG989386.1 | 1 | p3 | (TCC)5 | 15 | 452 | 466 | CT288F | CGAATTCTCCACCAACACCT | CT288R | AAGCTGGTCTTGAAGCTGGA | 182 |
| EG989420.1 | 1 | p3 | (GCT)8 | 24 | 226 | 249 | CT289F | GTGGAAGCCAACCCATTTAG | CT289R | GTGCTGGCAGAGGAAAGAAC | 213 |
| EG989425.1 | 1 | p3 | (TGG)6 | 18 | 57 | 74 | CT290F | CACATCAACACCACTGAGCTG | CT290R | GGAGAGTGGCGATCGTAGAG | 249 |
| EG989599.1 | 1 | p2 | (GT)6 | 12 | 154 | 165 | CT291F | ACCACCATTTGGCAACTCTC | CT291R | TTCTTCGCCTATTCCGAGTG | 237 |
| EG989642.1 | 1 | p3 | (TTC)5 | 15 | 95 | 109 | CT292F | CTTCTTCCTCTCCTGCCTCA | CT292R | ATGCTCCAGCGTTATTGGTT | 238 |
| EG989692.1 | 1 | p2 | (TC)6 | 12 | 516 | 527 | CT293F | TTGGACCTCACAATGGAGAA | CT293R | CTGCACCTTTCAGGATGACA | 192 |
| EG989703.1 | 1 | p3 | (CCT)5 | 15 | 68 | 82 | CT294F | AGAACTGGTCACGCTCGTTT | CT294R | TCAGCGATTCCGTTTTTAGG | 173 |
| EG989710.1 | 1 | c | (TC)7tgccattgtgtagtaggagtaactctgtt(TC)6 | 55 | 30 | 84 | CT295F | CACGAGGCTTCACCTTCC | CT295R | CACGAGGCTTCACCTTCC | 182 |
| EG989764.1 | 1 | c | (TC)12tgaaaaggactgattctctctgctttcccctttcctcttttcttcttttatcccttttccccaacagatcctcctcctgctc(CTT)5ccctctcactctccctctcagcttcacagctatacacagtgagtgagatgggttttgtcctagaatggctcagcct(CAC)5 | 211 | 87 | 297 | CT296F | TTTCAATTGCAACGAAAGAAA | CT296R | ATGCCGAGAAAGTCATGGAG | 293 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG989798.1 | 1 | p2 | (TC)7 | 14 | 37 | 50 | CT297F | TTCTCTTCCTTCCTTCCTTCC | CT297R | CAAGCTTTTCATAATCAACCACA | 205 |
| EG989896.1 | 1 | p3 | (TTC)6 | 18 | 105 | 122 | CT298F | GAAAGAAGCAGCACCTCACC | CT298R | CTGCTATCGCGATGTGTACG | 208 |
| EG989897.1 | 1 | p3 | (CTA)5 | 15 | 173 | 187 | CT299F | CCCTCCATTCCTCCTTCTTC | CT299R | GTGAGCTTGAAAGCGGAGAG | 186 |
| EG989932.1 | 1 | p2 | (TG)6 | 12 | 578 | 589 | CT300F | CAAAGGCCCACATGGTTTTA | CT300R | CGGGAACTATGGCTCCAAAT | 163 |
| EG990011.1 | 1 | p3 | (AAT)6 | 18 | 553 | 570 | CT301F | ACGCAGAGAAGGGGAAAAGT | CT301R | TTCAACACAGGACTGGACCTT | 248 |
| EG990056.1 | 1 | p3 | (TGG)5 | 15 | 203 | 217 | CT302F | TTCAATGCTTTCCCGATTTC | CT302R | TCCTTGAGGTTGTGTCTCCA | 222 |
| EG990079.1 | 1 | c | (CAA)5ctgcagtagctgcatctgctattcctgctgccttcgccttttattcactttcc(TCT)5tcctcccaa(TCT)7 | 113 | 89 | 201 | CT303F | TGGCATTTCTAAGGTCACTCG | CT303R | TGAGGGTTTGGGAGATTGAG | 173 |
| EG990143.1 | 1 | p2 | (CT)6 | 12 | 23 | 34 | CT304F | GAGGCTTTCCCCTGTTTTG | CT304R | GGAGGTGGAAAAGTTGACGA | 152 |
| EG990175.1 | 1 | p2 | (AT)10 | 20 | 82 | 101 | CT305F | CTGCAACAATTGCGATGATTA | CT305R | TGAAGCCGGAGAAGAAGAAA | 157 |
| EG990202.1 | 3 | p3 | (ATT)5 | 15 | 383 | 397 | CT306F | CCACTGCTACTGTTCCCACA | CT306R | GGCAACATGATGATGAACCA | 154 |
| EG990281.1 | 1 | p3 | (ATG)5 | 15 | 324 | 338 | CT307F | GCACTAAGAAGTGGGTTTTTGG | CT307R | TCACTTTGTTCCCTCCCATT | 239 |
| EG990316.1 | 1 | c* | (CT)6(T)12* | 23 | 547 | 569 | CT308F | TCATCAAATGTAACCCCAACAA | CT308R | CTGATGTATAGGTGGAGTTGAGG | 246 |
| EG990344.1 | 1 | c | (A)14caaaggctaaaacaccaaacgaaggggtgaaatc(T)16 | 64 | 215 | 278 | CT309F | AAAAACATCCACACCCCAGA | CT309R | TCCAGTGGAGGATTTGAAGC | 236 |
| EG990472.1 | 1 | p3 | (TGG)5 | 15 | 46 | 60 | CT310F | TCTCAATGTTGTCCTCTGCCTA | CT310R | GAGCTGCAGTGTTTGTTCCA | 159 |
| EG990475.1 | 1 | p2 | (CT)15 | 30 | 233 | 262 | CT311F | ACCCGAAATAGGGACGAATC | CT311R | GGTGAACCTGAAACTGTTGTTG | 233 |
| EG990478.1 | 2 | p2 | (TC)10 | 20 | 173 | 192 | CT312F | AAAATGGGGTTCTGCTCCTT | CT312R | CGAAAATACGACGGGAAAAG | 159 |
| EG990478.1 | 1 | p3 | (TGC)5 | 15 | 28 | 42 | CT313F | CTTGCGAGCAAGGTCTTCTT | CT313R | CGAAAATACGACGGGAAAAG | 220 |
| EG990496.1 | 1 | c | (TC)6tgtgtctctgatctgagtctgacct(TTC)10 | 68 | 39 | 106 | CT314F | CACGAGGGTCTCTTCTGCTC | CT314R | GATGACGGTGAAGGAAAGGA | 216 |
| EG990548.1 | 1 | p2 | (TA)7 | 14 | 113 | 126 | CT315F | CCAGGCCATTTCTTCATCAG | CT315R | CATTCCCATTTGCAACCTCT | 155 |
| EG990549.1 | 1 | c | (T)13acctctctctttcttgatttctgtgt(TG)6cttcatcagattcacccccctgatgctgctgagtaccactgaagctgaggaattctttcttttttcttgttaacggaaaccaat(TTA)5 | 149 | 263 | 411 | CT316F | TTCATCATCTTAATTTCATTCCAA | CT316R | TCCAATGACACTGCAAGAGAA | 250 |
| EG990561.1 | 1 | p3 | (CTG)5 | 15 | 291 | 305 | CT317F | AAGCCATAGCTGCGTCTGAT | CT317R | AGTTCCGTGGTTGTTGGAAG | 218 |
| EG990565.1 | 1 | p2 | (CT)6 | 12 | 134 | 145 | CT318F | CGAGGGTCCTCTGTCTCTCA | CT318R | ATTTTCGGATCCACAAGCAG | 217 |
| EG990601.1 | 1 | p3 | (TTC)5 | 15 | 421 | 435 | CT319F | CCAATTGAAAAAGCACTTGGA | CT319R | CTTTTCTGCCCTTTGATGGA | 170 |
| EG990613.1 | 1 | p2 | (TC)6 | 12 | 32 | 43 | CT320F | CACCACCACCCTTCATCTTT | CT320R | GAAGACACCGAATCCATGCT | 201 |
| EG990636.1 | 1 | p3 | (CCT)7 | 21 | 121 | 141 | CT321F | CACGAGGGAGACATGGATCT | CT321R | GGAAGAGAAGCGGGAAGAGT | 230 |
| EG990656.1 | 1 | p2 | (AG)6 | 12 | 138 | 149 | CT322F | CCATTCCAATTCCCAACATC | CT322R | GGGCTTTAACTTCAGCAACG | 192 |
| EG990674.1 | 1 | c | (A)26ctcatttgtgttatacaatattctctatgttcaagaggggtgaagtgaggttttgaaacttttga(T)11 | 102 | 65 | 166 | CT323F | CTTATTTTGGTATTTTAAAGAAGCACT | CT323R | AAAAATCAAAAGTTTCAAAACCTCA | 198 |
| EG990718.1 | 1 | p2 | (TC)6 | 12 | 25 | 36 | CT324F | CACGAGGTGTGTTCTCTCTATCTT | CT324R | GGGTCTTTCCAACAAGCAAC | 199 |
| EG990784.1 | 1 | p3 | (CTT)6 | 18 | 225 | 242 | CT325F | TTCCATCTTCAATCCCCAAA | CT325R | GCTCCGATTACCCAAATCAA | 215 |
| EG990863.1 | 2 | p3 | (GTA)6 | 18 | 578 | 595 | CT326F | GTCTTCCTCCGTCGACTTTG | CT326R | TGGATAAGCCGATCTCAAGAA | 203 |
| EG990871.1 | 1 | p2 | (TC)14 | 28 | 65 | 92 | CT327F | CATCGTCCTCTCTCACACCA | CT327R | CCAATGCTCGTCTCTGTGAA | 244 |
| EG990875.1 | 1 | c | (CAT)6gctgtgtccttttctcactctctttacatatatctctcagaattctacacttttctagtttgttt(TC)7 | 98 | 114 | 211 | CT328F | GAGACATGGGGAAGAAGCAA | CT328R | GTTGGAACAGCAAGCAACAA | 246 |

| ID | SSR nr. | SSR type | SSR | size | start | end | Forward Primer | Forward primer | Reverse Primer | Reverse primer | Product size |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EG990921.1 | 1 | p3 | (TAA)5 | 15 | 298 | 312 | CT329F | CTAATGCCCAGGCTCTTGAC | CT329R | CGAAGTTGTAAGTTTTCAGGTGTG | 201 |
| EG990939.1 | 1 | p2 | (TC)6 | 12 | 158 | 169 | CT330F | CCCATTTTTACGCCTGTGAG | CT330R | GGGATTTGAAAGGGAAGCTC | 229 |
| EG990957.1 | 1 | c | (TC)7tac(CT)7 | 31 | 431 | 461 | CT331F | TCCTTTGGTCTCAGGTTTCG | CT331R | GCCATTCATTCTGCTCTGGT | 227 |
| EG990969.1 | 1 | p3 | (CTC)5 | 15 | 152 | 166 | CT332F | GGTCGTCGAATCAGAATGGT | CT332R | AACCTCGTCTTGGCAACATC | 248 |
| EG991187.1 | 1 | c | (CA)6tctataataaataaaacccatcaaagaattcaccaccctttttatttttcctctctctctcctt(TTTA)5 | 95 | 120 | 214 | CT333F | ACCAGACAGATCCCACCAAA | CT333R | GCAATTACGGACGGAAGAAA | 196 |
| EG991187.1 | 2 | p2 | (AG)6 | 12 | 327 | 338 | CT334F | TCCGTCCGTAATTGCTTAGG | CT334R | TTCAGAGCTTCCTCCTCCTG | 151 |
| EG991220.1 | 1 | p4 | (CTAT)5 | 20 | 159 | 178 | CT335F | TCATTGAAAAGCACAGAAGTGAA | CT335R | GTTGATAGGGGATGCAATGG | 219 |
| AW219913.1 | 1 | p3 | (TAT)7 | 21 | 225 | 245 | CT336F | GGGAGACAACAACAGGGTTT | CT336R | CAAAATCCTTGGGGTTACCA | 204 |
| EG975253.1 | 1 | c(p6) | (TCAGGC)5tcaagctgagttggcagag(AAT)5 | 64 | 25 | 88 | IITR N1F | CGAGGCAACAATATCATCTTCA | IITR N1R | TTTGGGATCGATCAAAGTGA | 204 |
| EG990254.1 | 1 | p5 | (TGCTG)5 | 25 | 102 | 126 | IITR N2F | CTTTCCCTTCCTTGCTTCCT | IITR N2R | TGAACCCAAAACACCAAACC | 206 |
| EG987728.1 | 1 | p5 | (TCTTC)5 | 25 | 51 | 75 | IITR N3F | CCACCACTTAGGTCGTGCTT | IITR N3R | CAAAAGCAAGGGAAAGGTCA | 169 |
| EG984655.1 | 1 | p5 | (CCTTT)5 | 25 | 110 | 134 | IITR N4F | CTCCCCTTCCACTTTCCTTC | IITR N4R | AAGAACAAGATGTTGGGCATTT | 244 |
| Contig802 | 1 | p5 | (AAAAT)5 | 25 | 144 | 168 | IITR N5F | TGCAAGTGACAACAGATTTGC | IITR N5R | ACGGCATTATCCGCTACAAA | 186 |
| Contig315 | 1 | c | (ATTGA)5aggaaaagaaaagaaa(AAG)7 | 62 | 969 | 1030 | IITR N6F | AAGGAAATTGCTCAGGCTCA | IITR N6R | CCCAAAGGCCAACATAATTG | 155 |
| Contig406 | 1 | p4 | (TTAA)6 | 24 | 867 | 890 | IITR N7F | AGGAGAAGGCTTGGGATGAC | IITR N7R | TTCCACCATGTGAAAAAGCA | 209 |
| Contig933 | 1 | p4 | (AAAC)5 | 20 | 341 | 360 | IITR N8F | CGGAAATCAACATCAACGAA | IITR N8R | TGTGACTTGGAAGCCCTTTT | 239 |
| Contig1497 | 1 | p4 | (AGAA)5 | 20 | 143 | 162 | IITR N9F | TCAGCCACAATGACACCAGT | IITR N9R | CTTCGTCGTCCGACACTACA | 243 |
| Contig1542 | 1 | p4 | (CTGG)5 | 20 | 541 | 560 | IITR N10F | ACTATGAATGCCGTGGAAGC | IITR N10R | AAAACAAACGCACCACACAA | 215 |
| EG987216.1 | 1 | p4 | (TACA)8 | 32 | 167 | 198 | IITR N11F | AGGAGAAGATGCTCCCAACA | IITR N11R | TGAGCATCCATGAAAGCAAG | 223 |
| Contig219 | 1 | c | (TTC)6atcatgccactactaccccatcatgatgaggagattcttctcaacagcaaccaatgtcgaa(TCT)5 | 94 | 199 | 292 | IITR N12F | TACCGTGTTGGTGACCTTGA | IITR N12R | ATGTCCAAACCCATCTCCAA | 221 |
| Contig464 | 1 | c | (CAC)5caacaacaacatcatttgtattcact(TC)7 | 55 | 34 | 88 | IITR N13F | CACGAGGATTTCACTCACACA | IITR N13R | AGGAGGGAGTTCAGGTTGGT | 192 |
| Contig521 | 2 | c | (ATA)8n(TAA)8 | 49 | 962 | 1010 | IITR N14F | AAAGAAGAAAGATCCTTGTTTGGA | IITR N14R | AAACAAAAATAACCGCTTGATGA | 165 |
| Contig848 | 1 | c | (TTC)6tccttcgccg(CCA)5 | 43 | 141 | 183 | IITR N15F | CTTTGCTTCAAAGGGTTTCG | IITR N15R | GCTCTGTCATTGTGGGAAAAA | 181 |
| Contig113 | 1 | p3 | (TTC)11 | 33 | 136 | 168 | IITR N16F | TCGAATTGAAGGCAAAGAGC | IITR N16R | GTAGCTGAGTGCCCTGGAAG | 247 |
| Contig417 | 1 | c | (A)12(TA)9 | 30 | 108 | 137 | IITR N17F | CAACCAGAAACATGAAGCACA | IITR N17R | CCTGTTATGCCTGCTGAGTG | 216 |
| Contig126 | 1 | p2 | (GA)18 | 36 | 27 | 62 | IITR N18F | CGAGGGAAAGAGTTAAGAGAAAA | IITR N18R | CCATAGCCAAAAACCCAGAA | 233 |
| Contig1686 | 1 | p2 | (AT)13 | 26 | 755 | 780 | IITR N19F | CTGCCTCAACAACCTCAACA | IITR N19R | CGAGCAAACCAAGGAAGAAG | 246 |
| Contig1716 | 1 | p2 | (CT)10 | 20 | 38 | 57 | IITR N20F | CCAAAGATCAACAACAACCAGA | IITR N20R | CCAACAGTGGCTTGCTTGTA | 207 |
| EG986602.1 | 1 | p2 | (TC)17 | 34 | 242 | 275 | IITR N21F | TCACCGAGTTATGTGGGTTTC | IITR N21R | CCCACAACCAAATACCCAAC | 216 |
| Contig38 | 1 | c | (T)11(G)10 | 21 | 521 | 541 | IITR N22F | GCCGGAGTATCCATCAAAAC | IITR N22R | AATCTTGTCAACCCCCTGTG | 192 |
| EG978028.1 | 1 | p1 | (C)13 | 13 | 497 | 509 | IITR N23F | CATGATTGCGTCAAAACTCG | IITR N23R | CCCAGGGGAGGGTTACTAGA | 185 |
| EG988734.1 | 1 | p1 | (A)21 | 21 | 116 | 136 | IITR N24F | TACCACCCTAGGCCCTAACC | IITR N24R | GGCCCCATTTTTATTTTGTTT | 161 |
| EG982898.1 | 1 | c | (A)20g(T)19 | 40 | 432 | 471 | IITR N25F | TAAGTGCAATTGTGCGAAGC | IITR N25R | AAGCACTTAGAGTGTGCTTTGTTG | 192 |
| Contig1393 | 1 | c | (A)13ttgaagaaacttacaattacctaaaagttaagagga(T)10 | 59 | 219 | 277 | IITR N26F | TAATCCATGTGCCAACAAGC | IITR N26R | TCTCCTTGCTCACCCAAAAT | 223 |

**Publications**

1. *G S Randhawa, D P Panigrahi and **K A Nagesh** (2010) **Understanding life: by making and breaking the cell.** Indian Journal of Microbiology. **50:**247-248.
2. **K A Nagesh,** P Bhatele, S K Tripathi, K S Dhugga, K S Gill and G S Randhawa. **Genetic diversity study of cluster bean (*Cyamopsis tetragonoloba* (L.) Taub) landraces with RAPD and ISSR markers** (Communicated)
3. *P Bhatele, **K A Nagesh,** S D Sunita, S K Tripathi and G S Randhawa. **Optimization of AFLP primers and Assessment of Genetic Diversity in *Cyamopsis* varieties** (Communicated)
4. **K A Nagesh,** P Bhatele, S K Tripathi, K S Dhugga, K S Gill and G S Randhawa. **Identification and characterization of EST-SSR markers in *Cyamospsis sp.*** (Communicated)

**Conference Abstracts**

1. **Nagesh K A,** Pranita B, S K Tripathi, Dhugga K S, Gill K S and Randhawa G S. **Data mining for simple sequence repeats in expressed sequence tags of cluster bean (*Cyamopsis tetragonoloba* (L.) Taub.).** World Congress on Biotechnology-2011, Hyderabad, India. 21-23 March 2011.(Page 516)
2. **Kuravadi A N,** Bhatele P and Randhawa G S. **RAPD analysis in Cluster bean (*Cyamopsis tetragonoloba* (L.) Taub).** International Conference on Recent Development, Future Prospects and Entrepreneurial Trends in Biotechnology, IET Biotechnology Institute, Alwar, Rajasthan, India.19-21 December 2009.(Page 54)
3. *Swati V, **Nagesh K A,** Vikas P, Dhugga K S, Gill K S and Randhawa G S. ***In vitro* plant regeneration studies in cluster bean (*Cyamopsis tetragonoloba* (L.) Taub.).** National Symposium on Biodiversity and Food Security – Challenges & Devising Strategies, Indian Institute of Pulse Research, Kanpur, Uttar Pradesh, India, 10-11 December 2011. .(Page 100)
4. *Pranita B T, Nagesh K.A , Tripathi S K, Sunita S D and Randhawa G S. **Genetic analysis of *Cyamopsis tetragonoloba* (guar) genotypes: a gum producing industrial crop of India.** 80th Annual Meeting of Society of Biological Chemists, Lucknow, Uttar Pradesh, India, 12-15 November 2011.

* This work is not included in thesis.