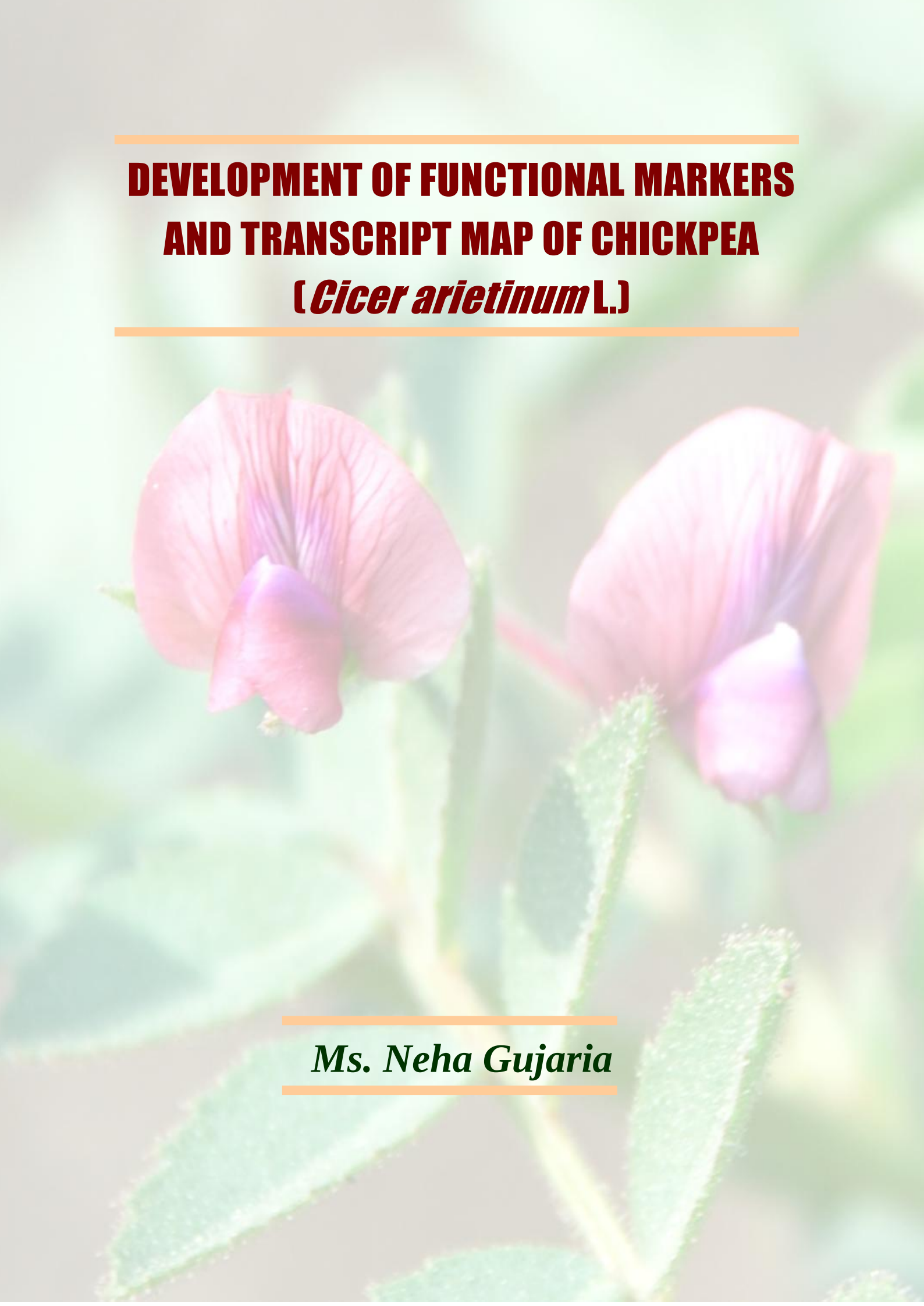
DEVELOPMENT OF FUNCTIONAL MARKERS AND TRANSCRIPT MAP OF CHICKPEA (*Cicer arietinum* L.)



Ms. Neha Gujaria

DEVELOPMENT OF FUNCTIONAL MARKERS AND TRANSCRIPT MAP OF CHICKPEA (*Cicer arietinum* L.)

A Thesis

submitted for the degree of

DOCTOR OF PHILOSOPHY

SUBMITTED BY

Ms. Neha Gujaria

CO- SUPERVISOR

Dr. Rajeev K. Varshney

SUPERVISOR

Prof. Mangla Bhide

DEPARTMENT OF ZOOLOGY

SCHOOL OF BIOLOGICAL AND CHEMICAL SCIENCES

DR. HARI SINGH GOUR UNIVERSITY SAGAR - 470 003 (M.P.) INDIA

2012

DECLARATION BY THE CANDIDATE (para 12b)

DECLARATION

I declare that the thesis entitled "DEVELOPMENT OF FUNCTIONAL MARKERS AND TRANSCRIPT MAP OF CHICKPEA (*Cicer arietinum* L.)" is my own work conducted under the supervision of Dr. Mangla Bhide, Prof. & Head, Chairman BOS, Department of Zoology, School of Biological and Chemical Sciences, Dr. Harisingh Gour Vishwavidyalaya, Sagar (M.P.) and co-supervision of Dr. Rajeev Kumar Varshney, Director, Centre of Excellence in Genomics (CEG), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad (A.P.) approved by Research Degree Committee. I have put in more than 200 days of attendance with the supervisor at the centre.

I further declare that to the best of my knowledge the thesis does not contain any part of work, which has been submitted for the award of any other degree either in this University or in any other Universities without proper citation.

Signature of Candidate

(Neha Gujaria)

Signature of the Co-supervisor

(Rajeev Kumar Varshney)

Signature of the Supervisor

(Mangla Bhide)

Signature of Head of the Department

CERTIFICATE OF THE SUPERVISOR (para 12 C)

CERTIFICATE

*This is to certify that the work entitled "DEVELOPMENT OF FUNCTIONAL MARKERS AND TRANSCRIPT MAP OF CHICKPEA (Cicer arietinum L.)" is a piece of research work done by Ms. Neha Gujaria under our guidance and supervision for the degree of **Doctor of Philosophy** of **Dr. HARISINGH GOUR VISHWAVIDYALAYA, SAGAR (M.P.), INDIA**. The candidate has put-in an attendance of more than 200 days with us.*

To the best of our knowledge and belief the thesis:

(1) embodies the work of the candidate herself
(2) has duly been completed
(3) fulfills the requirements of the Ordinance relating to the Ph.D. Degree of the University and
(4) is up to the standard both in respect of contents and language for being referred to the examiner.

Signature of the Co-supervisor

(Rajeev Kumar Varshney)

Signature of the Supervisor

(Mangla Bhide)

Forwarded

Signature of Head of the Department

ACKNOWLEDGEMENT

Endless compassion of Almighty turns my difficult task to a feeling of pleasant journey of my life. Emotions cannot be adequately expressed in words because emotions are transformed into a mere formality. Hence, my acknowledgements are many times more than what I am expressing here. I bow my head before my parents whose hardships, patience and perseverance today I stand has kept me going at all hard times.

I have received a lot of ad-hoc response from many generous grandees in many quarters to those pain stacking days encompassing the research work. I wish to vow my genuflexion with deep sense of gratitude to my co-supervisor **Dr. Rajeev Kumar Varshney**, Director-Centre of Excellence in Genomics, Applied genomics,ICRISAT and Leader Theme 1- Comparative and Applied Genomics, Generation Challenge Programme, for his conscientious guidance, gracious, cordial and meticulous explications and encapsilative remarks towards the representation of this dissertation. His valuable suggestions brought a panacea for me dealing with this research work. I can truly say that whatever he has done for me is not possible for every advisor to do for his advisee.

At this stage I feel pleasure to express my profound regards, indebtness and gratitude to my supervisor Dr (Prof) Mangla Bhide, , Prof. & Head, Chairperson BOS, Department of Zoology, School of Biological and Chemical Sciences, Dr. Harisingh Gour Vishwavidyalaya, Sagar (M.P.) for her expedient advice, debonair discussion, innovative ideas, abiding interest and invaluable support during this tenure of research work.

A depth of gratitude is owed to Dr. William Dollente Dar, Director General and Dr. Dave Hoisington, Deputy Director General of the ICRISAT for giving me opportunity to work in one of the best labs of International standards. With respect, regards and immense pleasure, I wish to acknowledge and express sincere thanks from my heart to several scientists including Dr. Trushar Shah, Dr. Abhishek Rathore, Dr. Vivek Thakur, Dr. T Mahendar for the valuable suggestions. I render special thanks Dr. RR Mir, Visiting Scientist and Dr. Himabindu Kudupa for their help in finalizing the thesis.

I feel immense pleasure to express my sincere thanks to Dr. Junichi Kashiwagi and Dr. L Krishnamurthy, Chcikpea Breeding for their help in conducting the research trials at Patancheru and Dr. H D Upadhyaya and Dr. P M Gaur for providing the genetic

material used in the present study. With immense pleasure, I express my cordial thanks to Mrs. Seetha Kannan, Mrs. Manjula B, Mrs B Poornima Reddy, Mr. Prasad KDV, Mr. Murali Krishna Y and Mr. B Anjaiah, for their administrative help. *It is my immense pleasure to render sincere regards to Mr. K Eshwar and Mr. G Somaraju for teaching me the laboratory techniques at the beginning of the research work.*Words are less to express my gratitude to Mr. A Gafoor, Mr. R Samantula, Mr. T Veerendar from Applied Genomics Laboratory, ICRISAT for providing technical help. I place on record my heartfelt thanks to all my labmates and friends especially Abhishek, Anuja, Ashish, Preeti, Sarwer, Spurthi, Manish Pandey, Pavana, Lekha, Kim, Deepa, Monisha, Ramesh, Mayank, Gautmi, Swathi, P. Swathi, Gnanesh, Rachit, Usha, Manish Roorkiwal, Prasad KVSS, Yogendra, Ashutosh, Bhanuprakash, Reetu, Prasad, Pradeep.

I also thank all the friends and staff at Department, School of Biological and Chemical Sciences, Dr. Harisingh Gour Vishwavidyalaya, Sagar (M.P.) for their kind help and co-operation. Assistance rendered by the members of Central Support Lab, Library and Learning System Unit at ICRISAT is gratefully acknowledged.

Words would fail to express the depth of my feelings for my strength, my husband Manish for his all time support, constant encouragement, affectionate love and for always being there for me. I feel indebted to my grandparent in-law Mr. Om Prakash Verma for his constant support and encouragement throughout the tenure of research work.

I seize the opportunity to express my heartiest sense of reverence, respect and affection to my parents Dr. S.C. Gujaria and Mrs.Vimla Gujaria, whose unending love, care, subdued spirit of co-operation, and timely encouragement throughout my research work and their support at every crucial moment of my life. I also express my heartiest sense of reverence, respect and affection to my brother Mr. Rajiv Gujaria, loving sister Mrs. Manvi and nephew Nishchay, who always supported and encouraged me during all the time.

ABBREVIATIONS

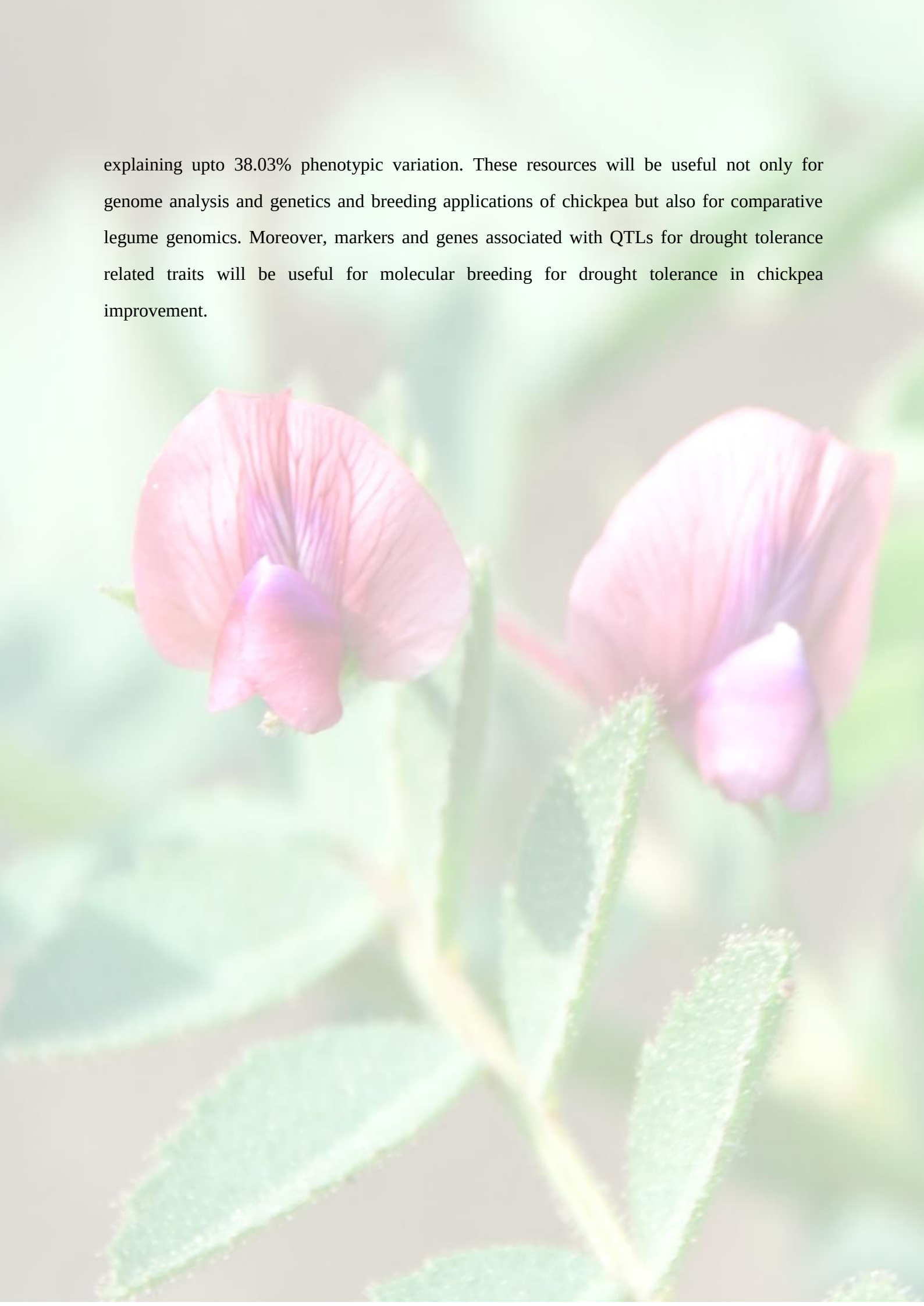
°C	:	degree Celsius
µl	:	Microliter
ABA	:	Abscisic acid
AFLP	:	Amplified Fragment Length Polymorphism
ANOVA	:	Analysis of Variance
ASR	:	Abscisic acid Stress and Ripening
BAC	:	Bacterial Artificial Chromosome
BES	:	BAC-end Sequences
BIBAC	:	Binary Bacterial Artificial Chromosome
BLAST	:	Basic Local Alignment Search Tool
bp	:	base pair
CAPS	:	Cleaved Amplified Polymorphic Sequences
C	:	complementary DNA
CIM	:	Composite Interval Mapping
cM	:	centiMorgan
CTAB	:	Cetyl Trimethyl Ammonium Bromide
DArT	:	Diversity Array Technology
DIVEST	:	Diversity Estimator
DNA	:	Deoxyribonucleic Acid
DREB	:	Drought Responsive Element Binding proteins
EMBL	:	European Molecular Biology Laboratory
EST	:	Expressed Sequence Tag
GCP	:	Generation Challenge Programme
ICCM	:	ICRISAT Chickpea Microsatellites
ICRISAT	:	International Crops research Institute for the Semi-Arid Tropics
Indel	:	Insertion deletion
ISSR	:	Inter Simple Sequence Repeats
JCVI	:	J. Craig Venter Institute
kbp	:	kilo base pairs
LDW	:	Leaf Dry Weight

LG	:	Linkage Group
LOD	:	Logarithm of odds (base 10)
MAS	:	Marker-Assisted Selection
Mb	:	Million bases
mM	:	milliMolar
MSA	:	Multiple Sequence Alignment
NCBI	:	National Center for Biotechnology Information
ng	:	nanograms
PAGE	:	Polyacrylamide Gel Electrophoresis
PCR	:	Polymerase Chain Reaction
PIC	:	Polymorphism Information Content
QTL	:	Quantitative Trait Loci
RAPD	:	Random Amplified Polymorphic DNA
RD	:	Rooting Depth
RDW	:	Root Dry Weight
RFLP	:	Restricted Fragment Length Polymorphism
RGA	:	Resistance Gene Analogues
RIL	:	Recombinant Inbred Line
RL	:	Root Length
RLD	:	Root Length Density
RSA	:	Root Surface Area
RT	:	Root Dry Weight/Total Dry Weight ratio
RV	:	Root Volume
SCAR	:	Sequence Characterized Amplified Region
SDW	:	Shoot Dry Weight
SIM	:	Simple Interval Mapping
SNP	:	Single Nucleotide Polymorphism
SSR	:	Simple Sequence Repeats
StDW	:	Stem Dry Weight
STMS	:	Sequence Tagged Microsatellite Sites

ABSTRACT

To develop repertoire of genic molecular markers (GMMs) including single nucleotide polymorphism (SNP) and intron spanning region (ISR) based markers, in chickpea, the third food legume crop of the world and the first food legume crop of India. In first approach, Solexa 1 Gb was carried out on pooled RNA from all drought challenged root tissues of genotypes ICC 4958 and ICC 1882. 15.6 and 22.1 million reads were generated and aligned against chickpea transcriptome assembly. A total of 26,082 SNPs were identified between these two genotypes. In second approach for SNP discovery through allele re-sequencing, primer pairs were designed for 970 genes/ expressed sequence tags (ESTs) of chickpea and 657 genes/ESTs of heterologous (closely related to chickpea) species and 2,046 SNPs were identified in 84,073 bp sequence data. In the third approach, ISR markers were designed by aligning chickpea unigenes to *Medicago truncatula*. In the forth approach, KASPar assay was designed for 96-plex SNPs and 56 polymorphic markers were identified on the parental genotypes and genotyping was done by designing Veracode assay for BeadXpress reader. From all the approaches: 87-EST-SNPs, 1627 allele-specific sequencing from chickpea and heterologous species, 121- intron spanning region (CISR) and 56 CKAM from KASPar assay were designed. SNP2CAPS analysis of 87 and 264 sequence alignments from *in silico* mining of ESTs and allele specific primers, as mentioned above, provided a total of 311 CAPS candidates. 311 CAPS candidates provided scorable amplification in 205 (65.92%) cases of which predicted assays were validated in 152 (74.15%) cases (CGMM). Screening of easily assayable 295 markers including 152 CGMMs, 87 CISRs and 56 CKAM on 5 parental genotypes of three mapping populations identified 75 polymorphic markers on the intra-specific mapping population. 73 of these GMMs together with 241 earlier developed markers could be integrated into the intra-specific genetic map. The transcript map developed here, therefore, has a total of 285 maker loci including 50 GMMs loci and spans 595.73 cM with an average inter marker distance of 2.09 cM. Identification of QTL for drought related root traits resulted in 12 significant QTLs. The QTL analysis revealed the presence of a "QTL hot-spot" region on LG04 that contained QTLs for several drought tolerance traits

explaining upto 38.03% phenotypic variation. These resources will be useful not only for genome analysis and genetics and breeding applications of chickpea but also for comparative legume genomics. Moreover, markers and genes associated with QTLs for drought tolerance related traits will be useful for molecular breeding for drought tolerance in chickpea improvement.



LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Chickpea (*Cicer arietinum* L.) is an economically important food legume crop in the arid and semi arid tropics. Chickpea ranks third among pulses (after common bean and pea), fifth among grain legumes, and 15th among grain crops of the world. The oldest report concerning this species was 5450 BC (Helbaek 1959) and it has been cultivated for at least 7000 years (van der Maesen 1972). It is grown over 11.08 million hectares and is a good source of nutrition especially to the vegetarians and poor farmers of developing countries (FAO, 2009). In Asia, chickpea contributes 86.73% of global production from 89.89% area. India leads the world production with 67.68% of this area (7.50 Mha), and 66.91% (6.54 Mt) of production followed by Pakistan (with 9.75% of area: 1.08 Mha and 0.741 Mt), Iran (0.56 Mha), Turkey (0.45 Mha) Myanmar (0.20 Mha), Australia (0.36 Mha), Ethiopia (0.23 Mha), Canada (0.04 Mha), Mexico (0.11 Mha), Syria (0.07 Mha), USA (0.04 Mha), Spain (0.02 Mha), and Eritrea (0.02 Mha) (FAO, 2009).

The seed of chickpea contains 20-30% protein, 40% carbohydrates and 3-6% oil and an extremely good source of calcium, magnesium, phosphorus, iron, zinc and manganese. Chickpea is also known for its use in herbal medicine and cosmetics. Chickpea is commercially divided into two types, kabuli and desi. The desi type of chickpea has small and colored seeds and the kabuli type have large and cream colored ones (see Millan et al. 2006).

Chickpea production is adversely affected by several biotic and abiotic stresses. Among the biotic stresses *Ascochyta* blight, *Fusarium* wilt and pod borer (*Helicoverpa armigiera*) are the main constraints whereas drought, heat, salinity and cold form a major part of the abiotic stresses. The estimated yield losses due to abiotic stress (6.4 million t) are much more than loss due to biotic stress (4.8 million t) (see Mantri et al. 2007). Among abiotic stresses, drought is the most important factor limiting chickpea productivity and the resulting yield losses vary between 30 and 60% depending on geographic location and climatic conditions during the crop season (Saxena et al. 1993). The value of annual global chickpea production losses caused by drought is

estimated at US$ 1.2 billion (Ryan 1997). Most of the desi varieties are found drought resistant as compared to the kabuli varieties (Yadav et al. 2004).

Drought is a meteorological event which can be described as absence of rainfall over a period of time long enough to cause water potential depletion in plant tissues and loss of moisture in the soil. From agricultural point of view, drought can be defined as the inadequacy of water availability, including precipitation and soil-moisture storage capacity, in quantity and distribution during the life cycle of a crop plant, which restricts the expression of full genetic potential of the plant. Drought can be categorized into two types intermittent and terminal associated to limited rainfall and can affect chickpea production. Intermittent drought is due to climatic patterns of sporadic rainfall that causes intervals of drought and can occur at any time during the growing season (Schneider et al. 1997) or when farmers have the option of irrigation, but the supply is occasionally limited. In contrast, terminal drought occurs when plants suffer lack of water during later stages of reproductive growth or when crops are planted at the beginning of a dry season (Frahm et al. 2004). Drought is accompanied by high temperatures with accelerates the effect of drought on crops (Sabaghpour et al. 2006). Autumn- or winter-sown crops in Mediterranean environments may experience intermittent drought stress during vegetative growth and/or terminal drought stress. Post rainy season crops, such as spring-sown crops in Mediterranean environments and winter-sown crops in the semi-arid tropics, are grown on residual moisture and experience progressively increasing terminal drought stress.

Terminal drought stress is more important to chickpea as the crop is grown globally as a post-rainy season crop under rainfed conditions. Development of short duration varieties has been the most successful approach in reducing chickpea yield losses due to terminal drought stress (Kumar et al. 1996). However, since yield in favorable crop growing conditions is generally correlated with the length of crop duration, any reduction of crop duration below the optimum would have a yield penalty (Saxena 1987). Deep and prolific attributes of the root system have been considered as main contributing factors for drought tolerance in chickpea (Serraj et al. 2004). Thus, there is a need to develop more drought tolerant chickpea varieties with deeper root

profiling. Due to the multigenic and quantitative nature of root profiling traits, it is difficult to breed drought tolerant varieties using conventional plant breeding.

Genomics tools and approaches possess great potential to develop the superior varieties with enhanced tolerance to different stresses (Tanksley et al. 1989a; Varshney et al. 2005a, 2005b, 2006). For instance, molecular marker based genetic maps have been useful in identifying and localizing important genes controlling both qualitatively and quantitatively inherited traits in a wide range of species (Varshney et al. 2006). Once the molecular markers identified is linked with agronomically desirable traits such as yield, quality, biotic and abiotic stress resistance, such markers can be used for marker-assisted selection (MAS) in breeding programs to transfer the desirable traits into the elite breeding lines. However, appropriate molecular markers and genetic maps integrated with molecular markers are prerequisites for MAS.

A diverse array of DNA-based marker technologies has been established to explore various DNA polymorphisms (Azhaguvel et al. 2006). Classically, the molecular markers can be grouped into three main categories (Gupta et al. 2002): (i) hybridization based markers- restriction fragment length polymorphisms (RFLPs), (ii) PCR-based markers: random amplification of polymorphic DNA (RAPDs), amplified fragment length polymorphism (AFLPs) and microsatellite or simple sequence repeat (SSR), and (iii) sequence or chip- based markers: single nucleotide polymorphisms (SNPs), diversity array technology (DArTs) and single feature polymorphisms (SFPs), which correlates variation in coding region affecting the function of genes and these molecular marker may be proved as candidate marker for a trait of interest.

Due to emphasis on functional genomics and advent of next generation sequencing (NGS) technologies, it has become possible now to develop the markers from genes or coding regions (Varshney 2010). As these markers are derived from genes and a putative function is known or can be deduced for the corresponding genes majority of times, these markers are popularly referred as ‘genic molecular markers (GMMs)’ (Varshney 2010) or ‘functional markers’ (Andersen and Lübberstedt 2003). A number of GMMs have several intrinsic advantages over

genomic DNA markers as they serve as a useful source for identification of 'perfect marker' for marker-assisted selection (MAS), estimating the functional genetic diversity present in germplasm collection, comparative mapping among related species and identification of chromosome duplication events. Genetic maps developed based on GMMs are popularly called as 'transcript maps' or 'functional maps'.

A number of methods have been used for developing GMMs in past for several crop species (Gupta and Rustgi 2004; Varshney 2010). Some of these methods include: (a) identification of SNPs (single nucleotide polymorphisms) by allele re-sequencing for candidate genes across different genotypes (Kota et al. 2008), (b) development of SSR markers from genes or expressed sequence tags (ESTs), called EST-SSR markers (see Varshney et al. 2002, 2005b), (c) identification of SNPs through *in silico* mining of ESTs coming from different genotypes and development of markers based on such SNPs, often referred as EST-SNP markers (Kota et al. 2008), (d) designing the primers from exonic regions to amplify the intronic region and detect either length or sequence polymorphism in introns (Feltus et al. 2006), referred here as intron spanning region (ISR) markers. A large number of GMMs as well as transcript maps have been developed in several crop species such as rice (Wu et al. 2002), wheat (Qi et al. 2004), barley (Stein et al. 2007; Kota et al. 2008; Sato et al. 2009), soybean (Choi et al. 2007), etc.

In the case of chickpea, there is no reference genome available. Therefore, there is a need to develop a robust collection of markers to identify genes responsible for drought tolerance. With an objective of developing transcriptomic resources, Varshney et al. (2009a) generated a set of 20,162 Sanger ESTs from four different genotypes. In recent year, new sequencing platforms have been invented to generate transcript reads. For instance, Illumina/Solexa 1Gb sequencing technology allows to sequence millions of short cDNA of average length of 36 bp per sample tag (read), reducing the library construction cost, runtime and also increasing the sensitivity. Their efficient in-depth sampling of the transcriptome compared to Sanger sequencing has also been demonstrated (Hanriot et al. 2008). These Illumina sequence reads can be used to identify SNPs. Other methods of SNP identification include mining of Sanger ESTs, allele specific sequencing, examining intronic regions for SNPs, etc. Because of their wide genic distribution,

co-dominant inheritance, technical simplicity, good genome coverage, bi-/multi-allelic nature, amenability to high-throughput automation, high reproducibility, chromosome-specific location and relative abundance, SNP markers are the preferred markers for plant genetics and breeding applications.

Once SNPs are identified, there is a need to have appropriate genotyping assays for SNP genotyping. At present, more than 30 SNP genotyping platforms have become available (Gupta et al. 2008; Varshney and Dubey 2009c). Genic markers are also developed by mining the EST sequences for SSR markers (Varshney et al. 2005b).

By using a variety of marker systems, genetic maps are developed on the populations segregating for traits of interest. In parallel, phenotyping is conducted on those populations. Analysis of genetic map and phenotyping data offers dissection of quantitative trait loci (QTL) that can be introgressed in molecular breeding approach. The application of this holistic approach, combining genomics with breeding and physiology, provides strategies for improving component traits of drought tolerance that should prove more effective and efficient than the conventional selection methods. Several studies were reported in past where genetic mapping is used for identification of the QTLs/genes for a trait of interest (Gupta and Varshney 2004). But in the case of chickpea, due to narrow genetic diversity in the primary gene pool, a very few reports are available on the mapping and QTL analysis on intra-specific mapping populations (Radhika et al. 2007; Kottapalli et al. 2009) and QTL identification (Singh et al. 2008).

In view of importance of chickpea and drought tolerance, this study was undertaken with following objectives:

Objectives:

1. Generation of Illumina/Solexa transcript reads
2. Large-scale identification of SNPs
3. Development of marker assays for SNP genotyping
4. Construction of a transcript map
5. Identification of QTLs responsible for drought toleranc

2. REVIEW OF LITERATURE

2.1 Leguminosae Family

Leguminosae or Fabaceae is a large and economically important family of flowering plants, which is commonly known as the legume family, pea family, bean family or pulse family. Leguminosae, is the third largest family among the angiosperms after Orchidaceae (orchid family) and Asteraceae (aster family), with 730 genera and over 19,400 species, according to the Royal Botanical Garden, UK. It is second only to Poaceae (grasses) in terms of agricultural and economic importance.

The species of this family are found throughout the world, growing in many different environments and climates. This family can be divided into three subfamilies: Mimosoideae, Caesalpinioideae, and Papilionoideae (Doyle and Luckow 2003). Of these, the Papilionoideae subfamily contains nearly all economically important crop legumes, including chickpea (*Cicer arietinum*), pigeonpea (*Cajanus cajan*), lentil (*Lens culinaris*), mungbean (*Vigna radiata*), cowpea (*Vigna unguiculata*), pea (*Pisum sativum*), common bean (*Phaseolus vulgaris*) and soybean (*Glycine max*). All these important crop legumes fall into two Papilionoid clades, namely Galegoid and Phaseoloid, which are often referred to as cool season and tropical season legumes, respectively. Despite their close phylogenetic relationships, crop legumes differ greatly in their genome size, base chromosome number, ploidy level, and self compatibility. Nevertheless, earlier studies indicated that members of the Papilionoideae subfamily exhibited extensive genome conservation based on comparative genetic mapping (Weeden et al.1992; Menancio-Hautea et al. 1993).

2.2 The Chickpeas

Chickpea (*Cicer arietinum* L.), the only cultivated species within the genus *Cicer*, is a self pollinated diploid (2n = 2x = 16) crop with a relatively small genome size of 740 Mbp (Arumuganathan and Earle 1991). It ranks third among food legumes in terms of production after common bean (*Phaseolus vulgaris*) and pea (*Pisum sativum*). Total annual world

production of chickpea is 9.8 million tones, and major producers India and Pakistan contribute 65% and 10% respectively, to the world harvest (FAO, 2009) (Fig. 1)

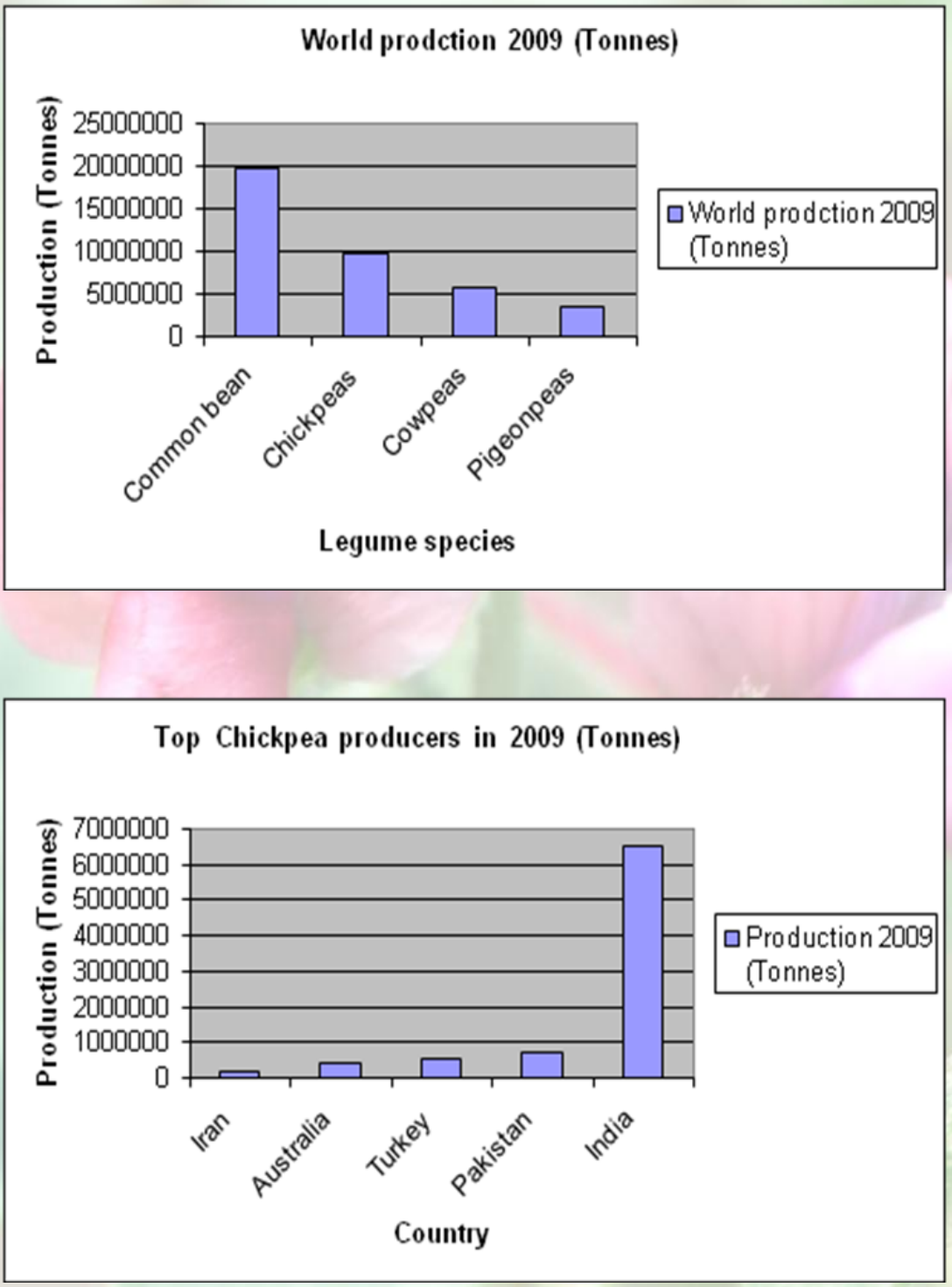


Figure 1 Graphical representation of production of - (a) legumes and (b) chickpea in the year 2009 (FAO, 2009)

2.2.1 Origin

The *Cicer* genus belongs to family Leguminoseae or Fabaceae, sub-family Papilionaceae and tribe Cicereae. It encompasses 9 annual and 34 perennial wild species. It is believed to be domesticated in the Old World about 7000 years ago (van der Maesen 1972) and most probably originated in South-eastern Turkey and adjoining Syria. *C. bijugum*, *C. echinospermum*, and *C. reticulatum*, the wild annual species of *Cicer*, closely related to chickpea are predominantly found in this region. South-west Asia and the Mediterranean are the two primary centres of origin, and Ethiopia the secondary centre of diversity (Vavilov, 1926; 1949-50). Wild annual *Cicer* originated mainly in the Mediterranean regions having a wide ecogeographic range, differing in habitat, topographic and climatic conditions (Abbo et al. 2003; Berger et al. 2003). Of the 9 annual species, chickpea (*Cicer arietinum* L.) is the only cultivated species. The eight other annual species of chickpea are wild and include: *C. reticulatum*, *C. echinospermum*, *C. pinnatifidum*, *C. judaicum*, *C. bijugum*, *C. cuneatum*, *C. chorassanicum* and *C. yamashitae*. According to van der Maesen (1987) the Cicer species was classified into four sections based on their morphological characteristics, life cycle and geographical distribution. Eight annual species namely *C. arietinum*, *C. reticulatum*, *C. echinospermum*, *C. pinnatifidum*, *C. bijugum*, *C. judaicum*, *C. yamashitae* and *C. cuneatum* were placed in section 'Monocicer', *C. chorassanicum* and *C. incisum* (perennial species) in section 'Chamaecicer', 23 perennial species in section 'Polycicer' and seven woody perennial species in section 'Acanthocicer'. The distribution of *Cicer* among different genepools is given in Fig. 2.

2.2.2 Chickpea – cytotaxonomy

Chickpea (*Cicer arietinum* L.) is an annual, highly self-pollinating, diploid (2n = 2x = 16) pulse crop. It is a herbaceous plant which branches from the base. It is almost a small bush with diffused, spreading branches. The plant is mostly covered with glandular or non glandular hairs but some genotypes do not possess hair. It is cultivated as a rainfed cool season crop or as a dry climate crop in semi arid regions. The crops optimum growth requires 18-26°C day and 21-29°C night temperature and annual rainfall of 600-1000mm (Duke 1981; Smithson et al. 1985). Cultivated chickpea (*C. arietinum*) is composed of two genetically distinct sub-types that are

readily distinguished based on seed size and colour. The desi types are shorter plants bearing white, pink, or blue colored flowers while kabuli types are taller plants bearing white flowers. *Desi* type, 'Microsperma', composed of small, angular, brown seeded varieties, with rough coat and *Kabuli* type, 'Macrosperma' composed of large, smooth, cream seeded varieties with smooth coat. According to Singh et al. (1997) desi types originated first than kabuli type. The production of these types also differs in the ratio of 3:1 for desi and kabuli types respectively. These two types also differ in their use, kabuli is used as whole grains and desi are processed into flour (Millan et al. 2006).

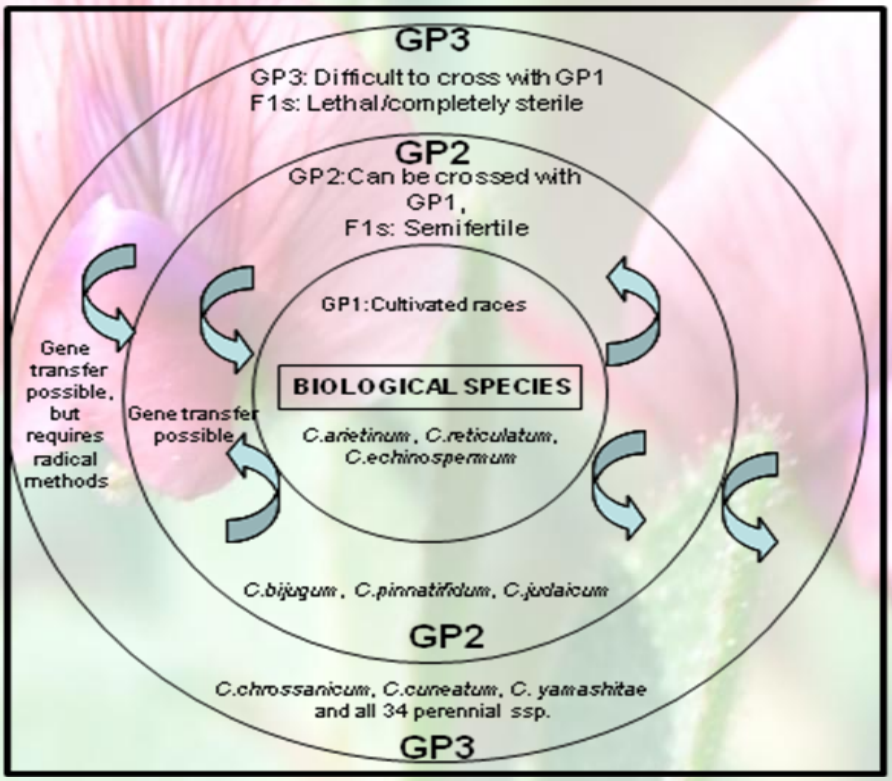


Figure 2 Distribution of species of *Cicer* in different gene pools

In kabuli seeds, the epidermis develops into a uniseriate palisade layer without thickening of the cell wall. In desi seeds, it develops into a multiseriate palisade layer which later develops into thick-walled sclereids which are heavily stainable with toluidine blue, indicating the presence of phenolic compounds contributing to testa colour. The walls of the subepidermal cells do not thicken in kabuli seeds, even though their size is considerably reduced. In desi seeds, these cells become thick walled as the seed matures. It is concluded that permeability and germinability of

the seeds may be considerably influenced by the observed differences in the structure of the palisade layers of the testas of desi and kabuli types. Desi and kabuli types differ in their dietary fiber components by a study at ICRISAT kabuli types contain higher amount of dietary fiber, particularly cellulose and hemicellulose.

2.2.3 Chickpea importance

Chickpeas are an excellent source of high quality protein, with a wide range of essential amino acids. Its potential as both a source of human food as well as animal feed, coupled with its ability to fix atmospheric nitrogen, is attracting an increasing number of farmers from semi-arid tropic (SAT) regions. They are also very high in dietary fiber and hence a healthy source of carbohydrates for persons with insulin sensitivity or diabetes. On an average, chickpea seed contains 23% of highly digestible protein, 64% crude fiber, 6% soluble sugar and 3% ash. The mineral component is high in phosphorous (343mg/100g), calcium (186mg/100g), magnesium (141mg/100g), iron (7mg/100g) and zinc (3mg/100g). Due to presence of high nutritional value components and near absence of anti-nutritive components, chickpea is considered as *nutraceutical* (or health benefiting food) (Williams and Singh 1987; McIntosh and Topping 2000; Charles et al. 2002; Millan et al. 2006). Besides, it has a traditional medicinal value with germinated chickpea reported as hypocholesteremic (Geervani 1991). *Desi* chickpea have a very low 'glycemic index' making them a healthy food source for people with diabetes (Walker and Walker, 1984). Furthermore, chickpea is an additional benefit to the farmers as it fixes a substantial amount of nitrogen for the subsequent crops and adds much needed organic matter that improves soil health, long-term fertility and sustainability of the ecosystems (Ahmad et al. 2005).

2.2.4 Constraints in chickpea production

Chickpea is a hardy, deep-rooted dryland crop and can grow to full maturity despite conditions that would prove fatal for most crops. It is grown on marginal land and rarely receives fertilizers or protection from diseases and insect pests (Singh and Reddy 1991).

Despite growing demand and high yield potential, chickpea yield is unstable and productivity is stagnant at unacceptably low levels. The chief constraints in chickpea production that reduce the yield and yield stability are a series of biotic and abiotic stresses. Biotic stresses like necrotrophic foliar fungal disease, *Ascochyta blight* (*Ascochyta rabiei* Labr.) and soil-borne necrotrophic fungal disease, *Fusarium wilt* (*Fusarium oxysporum* f. sp. ciceris) are considered the most serious biotic stresses (Muehlbauer and Kaiser 1998). Other fungi known to attack chickpea include leaf spot (*Alternaria* sp.), Ascochyta pisi, rust (*Uromyces ciceris-arientini*), gray mould (*Botrytis cinera*), powdery mildew (*Leviellula taurica*), *Pythium debar-yanum*, *P. ultimum*, dry root rot (*Rhizoctonia bataticola*), *R. solani*, foot rot (*Sclerotium rolfsii*), *Sclerotinia sclerotiorum*, wilt (*Verticillium albo-atrum*). Viruses isolated from chickpea include alfalfa mosaic, pea enation mosaic, pea leaf roll, pea streak, bean yellow mosaic, and cucumber mosaic. Pod borer (*Helicoverpa armigera*), is the most important pest, and feeds on leaves and developing seeds. Groundnut aphid (*Aphis craccivora*), pea aphid (*Acyrthsosiphon pisum*), cowpea bean seed beetle (*Callosobruchus maculatus*), and Adzuki bean seed beetle (*Callosobruchus chinensis*) are minor. Many storage insects specifically Bruchid sp. are a serious pest of stored chickpea (Duke 1981). In general, estimates of yield losses by individual pests, diseases or weeds range from 5-10 % in temperate regions and 50-100 % in tropical regions. Other important diseases of chickpea include pod borer (*Helicoverpa armigera*), botrytis grey mould, root rots, and rust (Millan et al. 2006; Mantri et al. 2007). In order of importance, drought, cold and salinity are the three main abiotic stresses that affect chickpea growth and productivity worldwide (Croser et al. 2003). In fact, the estimated collective yield losses due to abiotic stresses (6.4 mt) have been significantly higher than for biotic stresses (4.8 mt) (Ryan, 1997).

In West Asia and North African countries, low temperature causing freezing injury or death or delayed onset of podding reduces yield tremendously. Heat and salinity are relatively important problems following cold stresses (Muehlbauer and Singh, 1987; Singh, 1997). However, among the abiotic factors, drought stands to be the major problem in chickpea growing regions because the crop is grown on residual moisture and the crop is eventually exposed to terminal drought (Serraj et al. 2004).There is a big gap between the potential yield and the actual yield. Drought is

of major concern as 90% of chickpea production is under rain-fed conditions (Coram et al. 2007) Drought stress causes a 40-50% reduction in chickpea yield globally (Ahmad et al. 2005). It is estimated that if the yield loss due to drought stress is alleviated, chickpea production could be improved up to 50% that is equivalent to approximately US$ 900 millions (Ryan, 1997).

2.2.5 Drought stress in chickpea

From agricultural point of view, drought is the inadequacy of water availability, including precipitation and soil-moisture storage capacity, in quantity and distribution during the life cycle of a crop plant, which restricts the expression of full genetic potential of the plant (Mitra, 2001).

Soil water shortage has been identified as a major constraint to increasing chickpea production, adaptation and stability of crop performance throughout the world. Due to this single factor the annual yield losses are globally very high, ranging from 30-40% that depend on geographical region and length of crop season (Anwar et al. 2003; Yadav et al. 2004; Sabaghpour et al. 2006). It is estimated that if the yield loss due to drought stress is alleviated, chickpea production could be improved up to 50% that is equivalent to approximately US$ 900 millions (Ryan, 1997).

Intermittent and terminal droughts are the two distinct kinds of drought associated with limited rainfall in the semi-arid tropics. Intermittent drought is due to climatic patterns of sporadic rainfall that cause intervals of drought at varying intensities. Terminal (end-of-season) drought occurs in lowland tropical environments when crops are planted at the beginning of a dry season. As 90% of chickpea crops are grown on conserved soil moisture in the post-rainy season, the crop invariably suffers from terminal drought if there is little or no winter rainfall. Chickpea crop relies on stored soil moisture for growth during the later stages of reproductive growth i.e. during the critical flowering, pod-filling and seed development periods as the terminal drought stress intensifies (Frahm et al. 2004; Kumar and Abbo, 2001). The crop, in particular, is affected due to drought stress because of late sowings (Toker and Cagirgan, 1998). Estimates of yield losses due to terminal drought range from 35 to 50% across the semi-arid tropics (Schneider et al. 1997; Sabaghpour et al. 2006). Rahangdale et al. (1994) reported that water stress decreases seed yield by 15.2%. Terminal drought stress is normally accompanied by increasing temperature towards maturity, often to levels, more than 30°C, those which may

affect pod filling (Toker and Cagirgan, 2006). Extensive research efforts have been made to reduce the yield loss of chickpea under the drought environments. However, many issues related to drought are yet to be resolved since drought is a highly complex phenomenon.

Drought resistance is defined based on the relative yield of a genotype compared with other genotypes subjected to the same drought (Hall, 1993). Various morphological, physiological and biochemical characters confer drought resistance. Physiologically, drought tolerance is a complex phenomenon involving drought escape, dehydration avoidance, dehydration tolerance mechanisms (Blum, 1988). However, crop plants use more than one mechanism at a time to resist drought.

Drought escape: Drought escape is a particularly important strategy of matching phenological development with the period of soil moisture availability to minimize the impact of drought stress on crop production in environments where the growing season is short and terminal drought stress predominates (Turner, 1986). This mechanism involves rapid phenological development (early flowering, early podding and early maturity), developmental plasticity (variation in duration of growth period depending on the extent of water-deficit) and remobilization of preanthesis assimilates to grain (Mitra, 2001). Early phenology is the most important mechanism to escape terminal drought stress. This strategy has been shown to be associated with high initial growth vigor and high yield in chickpea (Sabaghpour et al. 2006; Berger et al. 2003). One drawback of these varieties are that they are not able to give high yield even under favorable conditions due to the reduced plant biomass as the photosynthetic period of the plant is reduced (Gaur et al. 2008).

Drought avoidance: Avoidance is related to the maintenance of high tissue water potential or turgor pressure and consists of mechanisms that reduce water loss from plants or maintain water uptake (Stoddard et al. 2006). The process of maintenance of water uptake is achieved through increased rooting depth, efficient root system and increased hydraulic conductance. Deep and prolific root systems have been associated with enhanced avoidance of terminal drought stress in chickpea (Serraj et al. 2004). The water loss can be reduced through epidermal (stomatal and lenticular) conductance, reduced absorption of radiation by leaf rolling or folding, reduced

evaporation surface (leaf area) due to leaf shedding and change in leaf morphology (e.g. few leaflets, tiny leaves). Reduction in leaf area is reported to reduce water loss in some chickpea accessions (Saxena, 2003).

Drought tolerance: Drought tolerance is defined as the plant capacity to sustain high plant water status or cellular hydration under the effect of drought (Blum, 2004). Plants withstand low tissue water potential by maintenance of turgor through osmotic adjustment (a process which induces solute accumulation in cell), increase in elasticity in cell and decrease in cell size and desiccation tolerance by protoplasmic resistance (Turner, 1986; Mitra, 2001). Earlier research works conducted to identify mechanisms of drought resistance in grain legumes under soil water stress of -0.6 MPa showed that drought resistance in chickpea was due to a significant decrease in osmotic potential (Amede and Schubert, 2003). Additionally, a positive relationship between Osmotic adjustment and grain yield in water-deficit has been shown. Osmotic adjustment leads to better extraction of water from the soil, stimulates root growth and facilitates a better translocation of preanthesis carbohydrate reserves to the grain during the grain filling period (Moinuddin and Chopra, 2004). Many substances play important roles in plant osmoregulation for drought resistance, including proline, glycine betaine, LEA proteins and soluble sugars such as levan, trehalose, sucrose, etc. The osmoregulation mechanism and the genetic engineering of plant drought-tolerance were studied by Du et al. (2004).

Unfortunately, most of these adaptations to drought have disadvantages. A genotype of short duration usually yields less compared to that of normal duration. The mechanisms that confer drought resistance by reducing water loss (such as stomatal closure and reduced leaf area) usually result in reduced assimilation of carbon dioxide. Osmotic adjustment increases drought resistance by maintaining plant turgor, but the increased solute concentration responsible for osmotic adjustment may have detrimental effect in addition to energy requirement for osmotic adjustment. Consequently, crop adaptation must reflect a balance among escape, avoidance and tolerance while maintaining adequate productivity (Mitra, 2001).

2.3 Physiology and Genetics of Drought

In the past decade, excellent progress has been made in unraveling abiotic stress pathways especially drought stress at the molecular level in plants. Plants express a number of genes in response to water deficit. Hundreds of genes that are induced under drought have been identified. However, because plant responses to stress are complex, the functions of many of the genes are still unknown (Chaves et al. 2003; Boominathan et al. 2004). Drought stress response in plants involves an array of different pathways associated with perception, signal transduction, gene expression and the synthesis of numerous novel compounds (e.g., proteins that scavenge oxygen radicals, chaperone proteins, or osmotically active compounds) (Xiong and Zhu, 2003; Du et al. 2004, Gong et al. 2008) (Fig 3). Knowledge of these processes is essential for a holistic understanding of plant resistance to stress, which is needed to improve crop management and breeding techniques.

The phytohormone, abscisic acid (ABA) plays a key role in mediating responses to abiotic stress and promotes characteristic developmental changes that help plants cope with water deficit (Lokko et al. 2007). Nayyar et al. (2005) found higher ABA contents in the wild Cicer species (*Cicer reticulatum*) than in the cultivated species under water stress.

Many of the traits that explain plant adaptation to drought such as phenology, root size and depth, hydraulic conductivity and the storage of reserves are those associated with plant development and structure, and are constitutive rather than stress induced (Chaves et al. 2003).

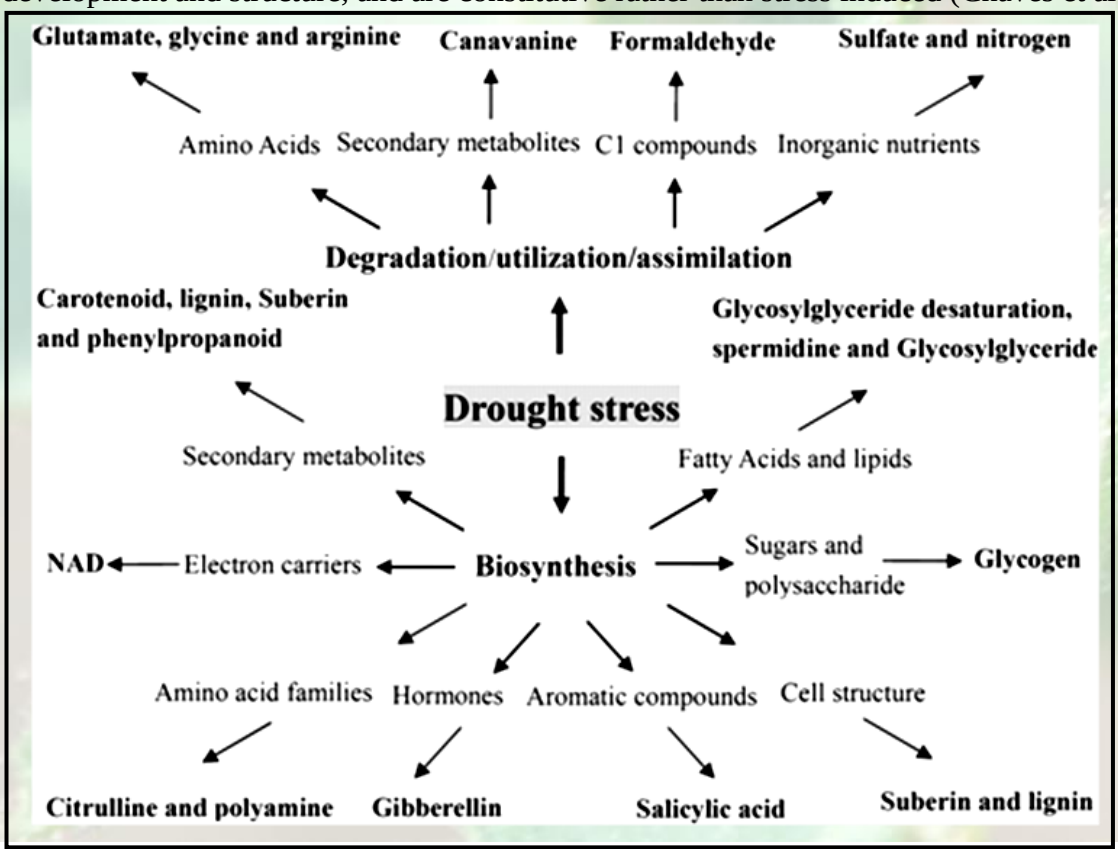


Figure 3

Biochemical pathways significantly affected by drought stress (Gong et al. 2010)

Roots play a primordial role in sensing soil water deficits as root development is fundamentally involved particularly in drought and mineral deficiency (Price et al. 2002; Matsui and Singh, 2003). They are able to measure decreasing soil water availability during a period of drought, which results in an increased release of ABA from stellar tissues of root (Sauter et al. 2001; Wilkinson and Davis, 2002). It is clear that the root tip is an important component in the sensing and signaling environment cues to the whole plant (Aiken and Smucker, 1996). The proportion of root length density distributed at deeper soil layers (115-120 cm) was shown to be higher under receding soil moisture conditions in chickpea (Ali et al. 2002; Ali et al. 2005). Root depth, length to weight ratio was also evaluated under progressively receding soil moisture conditions (Krishnamurthy et al. 2007; Price et al. 2002; Serraj et al. 2004; Kashiwagi et al. 2005). These parameters were useful as screening criteria for drought-tolerance in the past (Matsui and Singh, 2003). The effect of terminal drought on the dry matter production, seed yield and its components including pod production and pod abortion was investigated (Leport et al. 2006). Root development is known as a constitutive drought tolerance mechanism (Sinclair and Muchow, 1989). Therefore, research on drought in chickpea is primarily focused on root tissue.

2.4 Methods of Enhancing Drought Tolerance

A range of tools, from conventional breeding to the use of transgenic plants, from the use of molecular markers for selective breeding to the application of structural and functional genomics and proteomics play a very vital role in selection and improvement of plants for drought tolerance. However, because plant responses to drought stress are complex, the functions of many of the genes are still unknown. However in the last few years methodological advances in robotics and miniaturization are revolutionizing the way plant responses to stress are studied and understood.

2.4.1 Conventional breeding

Conventional breeding for drought tolerance is based on selection for yield and its components, screening of germplasm for new secondary traits, creating new crosses to recombine sources of variation in new genotypes under a given water-limited environment. This is the earliest method used in crop improvement.

Since about 90% of the chickpea crop is grown under rain-fed conditions, mechanism of drought escape is critical in the selection for drought resistant genotypes. Developing short-duration varieties have been the most effective strategy to date (Kumar and Abbo, 2001). The world's shortest-duration chickpea developed, ICCV 96029 (flowers about 23 to 27 days after sowing), led to substantial increase in chickpea area and productivity (Kumar and Rao, 2001). Time of flowering is a major trait of a crop's environmental adaptation, particularly when terminal drought and high temperatures restrict the growing season (Millan et al. 2006). Agronomic studies have concluded that early podding and greater biomass at podding contribute to high yields in cool-season pulses (Turner et al. 2001). The early-maturing varieties are preferred by the farmers because of a stable yield than the late maturing varieties (Gaur et al. 2008). The chickpea-breeding program at ICRISAT has placed high emphasis on development of early maturing varieties. Several varieties (e.g. ICCV 2, ICCC 37, JG 11, and KAK 2) that mature in 85 to 100 days and some breeding lines that mature in 75 to 80 days at Patancheru have been developed that greatly contributed to enhancement of productivity of chickpea in terminal drought-prone areas of peninsular India (Kumar and Rao, 1996). A study conducted with a common set of 73 genotypes showed that high-yielding genotypes flowered early, podded early and had a relatively long flowering period (Berger et al. 2003).

Because the large environmental variation necessitates evaluation of material at several locations and/or over years, trait-based selection could have an advantage. Efforts have been made to identify morphological traits that could contribute to drought tolerance/avoidance in chickpea (Turner et al. 2001). Two important drought avoidance traits have been suggested: a large root system is apparently more efficient for extraction of available soil moisture, and a smaller leaf

area helps to reduce transpirational water losses (ICRISAT, 1992). Research at ICRISAT led to the identification of the drought avoidance chickpea variety ICC 4958 with a large root system (Saxena et al. 1993; Serraj et al. 2004; Kashiwagi et al. 2005). Promising drought-tolerant, Fusarium-wilt-resistant lines with high-yields, ICCV 94916–4, ICCV 94916–8, ICCV 94920–3, ICCV 94924–2 and ICCV 94924–3 were obtained from a three-way cross involving ICC 4958, Annigeri and ICC 12237, a Fusarium wilt resistant accession. Efforts were also made to combine large roots trait of ICC 4958 and the few pinnules (smaller leaf area) trait of ICC 5680. Seven varieties were developed (ICCV 98901 to ICCV 98907) which proved to be more drought tolerant and yielded similar to the high yielding parent (Saxena, 2003; Serraj et al. 2004).

Reduction in leaf area is expected to reduce water loss. Saxena (2003) reported two chickpea accessions, ICC 5680 and ICC 10448, with a smaller leaf area. In another study, Toker and Canci (2003) did not observe any advantage of multipinnate or tiny leaf types in drought tolerance.

While such an approach has been partly successful, huge investments in land, labour and capital are required to screen a large number of progenies. In addition, there is evidence of increasingly marginal returns from conventional breeding, suggesting a need to seek more efficient methods for genetic enhancement of drought resistance. On the other hand, associating drought responses with the expression of specific physiological mechanisms can help greatly in establishing screening protocols, which allow better management of genotype *vs* environment (G × E) interactions (Serraj et al. 2003).

2.4.2 Genetic engineering approach

Multigenic and quantitative nature of drought makes it difficult to breed for abiotic stress tolerance using conventional plant breeding. Genetic engineering promises to provide a precise approach for developing drought tolerant crops that can augment traditional breeding methods. Crop improvement through genetic engineering has become a reality (Dunwell, 2000).

In recent years, introduction of drought-induced genes involved in different biochemical pathways from different sources to sensitive plants has evolved as one of the promising methods. Many different genes responsible for biosynthesis of different solutes and osmolytes conferring drought resistance are considered for transfer in some plants especially tobacco (Pilon-Smits et al. 1995; Bohnert et al. 1995).

Transgenic plants with resistance to major biotic constraints are being developed and tested by ICRISAT and its research partners, especially for legume crops (Sharma and Ortiz, 2000). In chickpea, transgenic technology is being exploited primarily for insect and pest resistance, drought tolerance, and quality enhancement. Development of transgenic chickpea plants was reported by Kar et al. (1997), where they introduced Bt Cry IA gene from the bacterium *Bacillus thuringiensis* (Bt) for the development of *Helicoverpa* pod borer resistant chickpeas. At ICRISAT, PGIP (Poly Galactourinase Inhibiting Protein) gene and other antifungal genes such as chitinases and glucanases are being introduced into chickpea for resistance to fungal disease Botrytis gray mold (Sharma and Ortiz, 2000; Singh and Jauhar, 2006). Insecticidal genes as those derived from Bt and plant genes such as protease inhibitor from soybean and pigeonpea are being considered by ICRISAT researchers for transformation of chickpeas. Alpha-amylase inhibitor gene (αAI1) isolated from the seeds of *Phaseolus vulgaris* L. (common bean) has been introduced into chickpea through *Agrobacterium* mediated transformation to analyse the ability of the gene to inhibit growth of bruchid weevil *C. maculatus*, which causes severe damage to chickpea seeds during storage (Ignacimuthu and Prakash, 2006). Research developments were also made in transferring drought responsive elements and osmoregulation genes for the tolerance to drought, salinity and cold (Ali et al. 2003). Efforts on genetic engineering of chickpea for enhanced tolerance to water stress are being carried out at ICRISAT by using two different approaches. The first one employs a single gene approach using the P5CSF129A (pyrroline-5-carboxylate synthetase) gene driven by a CaMV 35S promoter for proline accumulation. The second approach involves using a transcription factor, DREB1A, a drought responsive element (DRE), driven by the stress-inducible promoter from the rd29A gene that acts as a major “switch”, triggering a cascade of genes in response to a given stress. Plants

expressing the P5CSF129A and DREB1A (Nayak et al. 2009) genes demonstrated substantial resistance to water stress in comparison with wild type variety under experimental greenhouse conditions (ICRISAT, 2006).

The applicability of this technology will however depend on the identification of key genes, number of genes conferring a particular trait and public acceptance of cultivars. Lack of multidisciplinary approach and precise screening techniques, incomplete knowledge about genetic basis of drought resistance, negative correlation of drought resistance traits with productivity and unavailability of appropriate genes to obtain transgenic plants are the main constraints for genetic improvement of drought resistance.

2.4.3 Molecular breeding approach

The use of molecular markers to help identify DNA regions tightly linked to agronomic traits in crops can facilitate and speed up the breeding strategies for crop improvement (Mackill et al. 1999; Varshney et al. 2005a,b; Varshney et al. 2009a,b; Nayak et al. 2009). A DNA marker is typically derived from a small region of DNA that shows sequence polymorphism between individuals within a species. Functional marker development requires allele sequences of functionally characterized genes from which polymorphic, functional motifs affecting plant phenotype can be identified (Andersen and Lubberstedt, 2003). DNA markers have been utilized to study root traits or to explore their relationship to abiotic stress tolerance in rice (Champoux et al. 1995; Price et al. 2002; Price and Tomos, 1997) and maize (Zhu et al. 2008). Most types of molecular markers have been tested in chickpea including restriction fragment length polymorphism (RFLP) (Tanksley et al. 1989b; Udupa et al. 1993; Simon and Muehlbauer, 1997), random amplified polymorphic DNA (RAPD) markers (Simon and Muehlbauer, 1997), simple sequence repeat (SSR) or sequence tagged microsatellite site (STMS) markers (Hüttel et al. 1999; Winter et al. 1999; Sethy et al. 2003, 2006a,b; Lichtenzveig et al. 2005; Choudhary et al. 2006; Varshney et al. 2009a,b; Nayak et al 2009), single nucleotide polymorphism (SNP) markers (Kota et al. 2008; Varshney et al. 2009a; Nayak et al. 2009, Rajesh and Muehlbauer 2008). Based on these marker resources mapping of quantitative trait loci (QTL) that relate performance and yield to drought seems to be an interesting genetic avenue. Thus, regions of

chromosomes can be identified that carry genes that improve stress tolerance (Bohnert et al. 1995).

However, there is a low level of polymorphism detected in cultivated chickpea using RFLP markers. In contrast, SSR markers have been shown to be highly polymorphic and are so the marker of choice for marker assisted selection (MAS) (Gupta and Varshney, 2000, Varshney 2010). Although the use of MAS may be helpful for crop improvement (Serraj et al. 2005), its practical application in legumes for the genetic improvement of resistance or tolerance to stress has been limited, being mainly hampered by lack of investment and the genetic complexity of most stress-related traits (Dita et al. 2006). Moreover, the density of the intraspecific genetic map is still very low (Kota et al. 2001). Hence, there is a need for the development and utilization of EST-based markers useful for studying important abiotic stresses such as drought. These markers serve the most important criteria of high reproducibility, detection of co-dominance polymorphism and suitability for rapid large-scale low cost screening molecular breeding applications (Buhariwalla et al. 2005). EST-SSRs and EST-SNPs can be possibly detected from EST dataset of a particular organism reported after a thoroughput single pass sequencing of ESTs from stressed tissues. These markers have been earlier developed and reported in rice (Cho et al. 2000), wheat (Eujayl et al. 2001), chickpea (Huettel et al. 1999; Winter et al. 1999; Sethy et al. 2003, 2006; Choudhary et al. 2006, Nayak et al. 2009, Varshney et al. 2009a, Rajesh and Muehlbauer 2008). These markers are employed in designing locus-specific primers and will be useful for the evaluation of genetic diversity and molecular mapping (Kota et al. 2001; Sethy et al. 2003). RFLP markers developed from ESTs (EST-RFLP) have been extensively used for the construction of high-density genetic linkage maps (Harushima et al. 1998; Davis et al. 1999) and physical maps (Kurata et al. 1997; Semagn et al. 2006).

2.4.3.1 *Functional markers*

Due to emphasis on functional genomics and advent of next generation sequencing (NGS) technologies, it has become possible to develop the markers from genes or coding regions (Varshney 2010). As these markers are derived from genes and a putative function is known or

can be deduced for the corresponding genes majority of times, these markers are popularly referred as ‘genic molecular markers (GMMs)’ (Varshney 2010) or ‘functional markers’ (Andersen and Lübberstedt 2003). GMMs have a number of intrinsic advantages over genomic DNA markers as they serve as a useful source for identification of ‘perfect marker’ for marker-assisted selection (MAS), estimating the functional genetic diversity present in germplasm collection, comparative mapping among related species and identification of chromosome duplication events. The two broadly used approach to develop GMMs is EST approach and SNP approach, where as to genotype these GMMs high throughput genotyping platforms like GoldenGate assay, KASPar assay, etc and cost effective SNP2CAPS programme are being used by molecular biologists (Varshney 2010)

2.4.3.1.1 ***EST Approach:*** High-throughput sequencing technology has provided a mechanism to gain insight into genomes at the RNA level by large-scale single pass sequencing of randomly picked clones from cDNA libraries constructed from mRNA isolated at a particular development stage and/or tissue (Adams et al. 1993; Luo et al. 2005). ESTs were originally intended as a way to identify gene transcripts, but have since been instrumental in gene discovery, for obtaining data on gene expression and regulation, sequence determination, and for developing highly valuable molecular markers, such as EST-based RFLPs, SSRs, SNPs, and CAPS (cleaved amplified polymorphic sequences). ESTs have been used for designing probes for DNA microarrays that is used to determine gene expression (Sreenivasulu et al. 2004, Semagn et al. 2006).

There has been increasing initiatives on EST sequencing and EST based applications contributing to functional genomics in many plant species, such as Arabidopsis (White et al. 2000; Weber et al. 2006), barley (Kota et al. 2008), rice (Yamamoto and Sasaki, 1997; Chandra Babu et al. 2003), *Medicago truncatula* (Covitz et al. 1998; Journet et al. 2002, Cheung et al 2006), maize (Lawrence et al. 2004; Quyen et al. 2005), sugarcane (Carson and Botha, 2000; Vettore et al. 2003), grape (Silva et al. 2005), Lotus (Endo et al. 2000; Nelson et al. 2006), oil palm (Ho et al. 2007), citrus (Forment et al. 2005) and mint (Sterky et al. 1998). In case of

chickpea, a few reports are available for the EST development like Choudhary et al. 2009, Buhariwalla et al. 2005, Varshney et al. 2009a, Ashraf et al. 2009.

Efforts to identify genes underlying drought tolerance are mostly focused on model species and major cereal crops such as rice and maize (Bruce et al. 2002; Nguyen et al. 2004) with lesser attention being given to legumes. However, chickpea is taxonomically one of the closest crops to *Medicago* and both are well adapted to dry environments. Moreover, root system being a primary sensor of drought stress, generation of ESTs from chickpea roots is encouraged (Davies and Zhang, 1991). Targeted EST development has begun in chickpea, focusing on ABA-related mechanisms of water-deficit tolerance in epicotyl tissue (Romo et al. 2004). ESTs have been generated from stems and leaves of Ascochyta-blight-resistant chickpea genotype (Coram et al. 2007). A study to explore the mechanisms of drought tolerance in chickpea resulted in generation of ESTs from a subtractive suppressive hybridization (SSH) library, using the accessions ICC 4958 and Annigeri, which are both considered as different sources of drought avoidance and tolerance (Buhariwalla et al. 2005). The chickpea root EST database was developed by Jayashree et al. (2005) which is useful for chickpea genomics. An extensive collection of plant-derived ESTs are available in the database NCBI dbEST, a division of GenBank: http://www.ncbi.nlm.nih.gov/dbEST/ and in the relational database, ICRISAT chickpea EST database: http://www.icrisat.org/ gt1/cpest/home asp. Till date, a total of 7,907 ESTs generated from chickpea are recorded which is very low when compared to model species such as *Medicago truncatula* (249,450) and *Lotus japonicas* (157,951). Very recently, Varshney et al. 2009a (20,162 Ca-ESTs) and Ashraf et al. 2009 (6,272 Ca-ESTs) added a huge number of Ca-ESTs to the existing repertoire of chickpea ESTs. There is a need to generate higher number of ESTs in chickpea in the near future that can assist in rapid progress of drought genomics research.

2.4.3.1.2 ***SNP approach:*** Recent progress in genome resource development for model and major crop plants has energized genetic research. However, this activity has largely bypassed ‘‘orphan crops’’ such as chickpea which are crops of relevance to food security and income for

subsistence farmers in developing countries. Despite the limited genome resources, access to most of the genes in these organisms can be gained through cDNA sequences, which represent expressed genes. Partial cDNA sequences, known as ESTs, when determined from multiple genotypes of a species, facilitate the identification of SNPs in protein encoding genes and can be used in conjunction with mapping populations to generate genetic linkage maps that represent. Advances in genome analysis have made it possible to utilize SNPs in *Arabidopsis thaliana* (Weber et al. 2006), *Glycine max* (Choi et al. 2007), *Zea mays* (Barbazuk et al. 2007) and *Vigna unguiculata* (Hyten et al. 2010).

In chickpea, though few reports are available on identification and development of SNPs based on allele-specific re-sequencing of some genes (Rajesh and Muehlbauer 2008, Nayak et al, 2010, Singh et al 2008). The number of available SNP markers till date are comparatively very low in chickpea, that are required for high-throughput assay development. Hence, development of large-scale informative marker system such as SNPs is necessary for utilization of these resources in molecular breeding programmes.

Further hopes regarding the use of SNPs in routine analyses have been raised through the identification of haplotypes. Haplotypes are closely linked SNPs which occur along a chromosome in clearly defined structures or patterns (alleles) that extend over hundreds of base pairs or even several kilobases. The haplotype structure of SNPs alleviates the problem of scoring an extremely large number of SNPs by not requiring the analysis of each individual SNP in a genome but only of a limited number for genome coverage. Considerable efforts have been put into the identification of the haplotype pattern of the human genome (http://hapmap.org) but at present it is still not clear how efficient the analysis of haplotyopes is for the identification of quantitative traits (Johnson et al. 2001; Foster and Sharp, 2004). In plants, large scale SNP development and analysis project have been performed predominantly in diploid crop plants such as maize (Ching and Rafalski, 2002), barley and soybean (Zhu et al. 2003) where meanwhile more than 1.000 genes with SNPs were identified each. These data have demonstrated that SNPs are present in large numbers in crop plants and that they share similar

features (e.g. presence as haplotypes) as in other eukaryotic species (Rafalski 2002; Kahl et al. 2005).

2.5 Next Generation Sequencing (NGS) Platform- Example of Illumina/ Solexa Sequencing Technology

Next generation sequencing (NGS) technologies have opened fascinating opportunities for the analysis of plants with and without a sequenced genome on a genomic scale. During the last few years, NGS methods have become widely available and cost effective. They can be applied to a wide variety of biological questions, from the sequencing of complete eukaryotic genomes and transcriptomes, to the genome-scale analysis of DNA-protein interactions**.** These NGS technologies hold great potential to impact plant genetics and breeding in addition to impact human health and microbial biology (Varshney et al. 2009g). Three major sequencing platforms that are currently being used in plant species include Genome sequencer FLX (Roche/454 Life Sciences, http://www.454.com/), Applied Biosystems SOLiD (http://www3.appliedbiosystems.com) and Illumina Genome Analyze (http://www.illumina.com/). Details about mechanism and chemistry of these platforms have already been discussed in details in several reviews (Mardis 2008, Shendure & Ji 2008). These three platforms provide thousands of million sequence reads in a single run in reduced time and less costs as compared to conventional Sanger sequencing technology (Lister et al. 2009). Among these three approaches, FLX/454 platform is superior in terms of read length (about 400 bp) but is rather expensive in terms of cost when compared with the Solexa and AB SOLiD (Varshney et al. 2009g). Yet another approach based on single molecule synthesis is gaining attention and is termed as 3rd generation sequencing. Apart from this many new sequencing technologies are emerging and/or are at their infant stages to facilitate genome wide marker discovery in both model/major and orphan crop species. A number of laboratories and companies like Biotage, Helicos, Li-Cor, Microchip Biotechnologies, Nanofluidics, Nanogen, Network Biosystems and Visigen are working on development of 3rd generation sequencing platforms (Hudson 2008, Gupta 2009).

Sequence data generated for parental genotypes of the mapping populations by using NGS technologies can be used for mining the SNPs at large scale. While in the case of model plant species or major crop species, it is easier to align the NGS data from individuals to the reference genome sequence data, if available or the transcript sequence data available through EST sequencing projects. In case of under-resourced crop species where appropriate or adequate sequence data are not available, the best possible strategy is to sequence the cDNAs with NGS technologies and then align with the transcript data of the species, if available or of the related major/ model crop species. These approaches have been discussed in a separate review article (Varshney, 2009).

The advent of NGS technologies such as Roche FLX/454, Solexa and ABI-SOLiD has created the potential for generating considerably increased amounts of information for many organisms including orphan legume crop like chickpea. Roche FLX/454 technology provide inexpensive, genome-wide information producing approximately 100Mb sequence data in a single run (Mardis, 2008), while Illumina/Solexa 1Gb sequencing technology allows to sequence millions of short cDNA of average length of 36 bp per sample tag (read), reducing the library construction cost, runtime and also increasing the sensitivity. At present, Illumina/Solexa technology 1G has improved and generates 75+ bp reads for a total of > 33 Gb of paired-end data per run. Their efficient in-depth sampling of the transcriptome compared to Sanger sequencing has also been demonstrated (Hanriot et al. 2008). The only drawback with NGS technologies is the generation of short reads and for the proper alignment of these reads requires a sophisticated software programme. However, with the availability of various *denovo* assembly software programs such as CAP3 (Huang and Madan 1999), PCAP (Huang et al. 2003), RePS (Wang et al. 2002), and Phusion (Mullikin et al. 2003), MAQ, SOAP, ELAND, MOSAIK, VALVET, EULER, SSAKE, SHARCGS can effectively assemble the shorter reads. Previously, combinatorial strategy involving cDNA normalization and FLX-454 deep sequencing platform has been employed in transcriptome characterization studies in *Medicago* (Cheung et al. 2006), Coral (Meyer et al. 2009), pigeonpea (Dubey et al. 2011), *Melitaea cinxia* (Glanville fritillary butterfly) (Vera et al. 2008) and many other non-model organisms.

In summary, it is possible now to mine large scale SNPs in major as well as under resourced crop species and to undertake molecular breeding (Varshney et al. 2009e). Apart from developing SNP markers, NGS technologies can be and are being used for other applications such as de novo sequencing, association mapping, alien introgression, transcriptome expression and polymorphism, population genetics, evolutionary biology and genome-wide assembly in several crop species (Varshney 2009b,c,g,h; Fig 4).

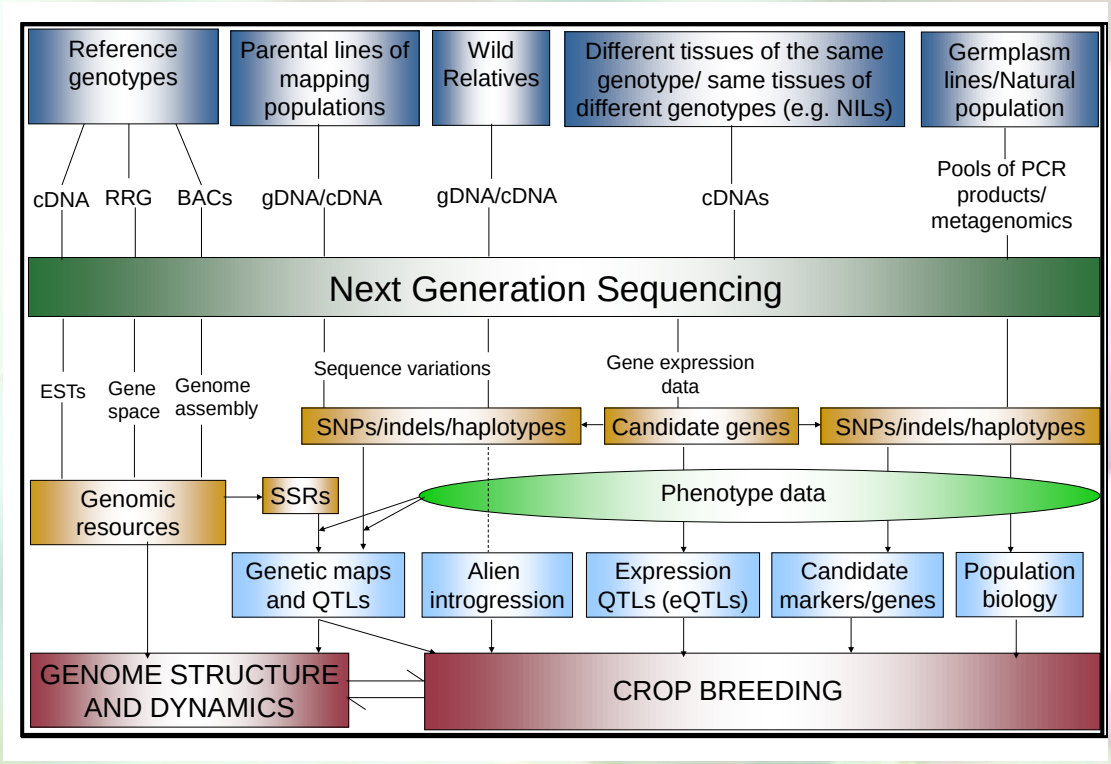


Figure 4 An overview of NGS pipeline (Varshney et al. 2009, Trends in Biotechnology)

As NGS technologies can provide a larger number of SNPs, development of high-throughput and cost effective genotyping platforms for these SNPs is yet another important task. Although there are several high-throughput SNP genotyping platforms are available, each of them has its own merits and demerits.

Illumina/Solexa sequencing is one of the commonly used sequencing technology. Illumina/Solexa Genome Analyzer I performs sequencing-by-synthesis of a random array of clonal DNA colonies attached to the surface of a flow cell by a unique "bridged" amplification reaction. The flow cell surface is coated with single stranded oligonucleotides that correspond to

the sequences of the adapters ligated during the sample preparation stage. Single-stranded, adapter-ligated fragments are bound to the surface of the flow cell exposed to reagents for polyermase-based extension. There are about 8 million such colonies on each of the 8 lanes of the cell. At each cycle of synthesis all four nucleotides, labelled with four different fluorescent dyes and blocked at the 3'-ends, are introduced in the flow cell (Fig. 5).

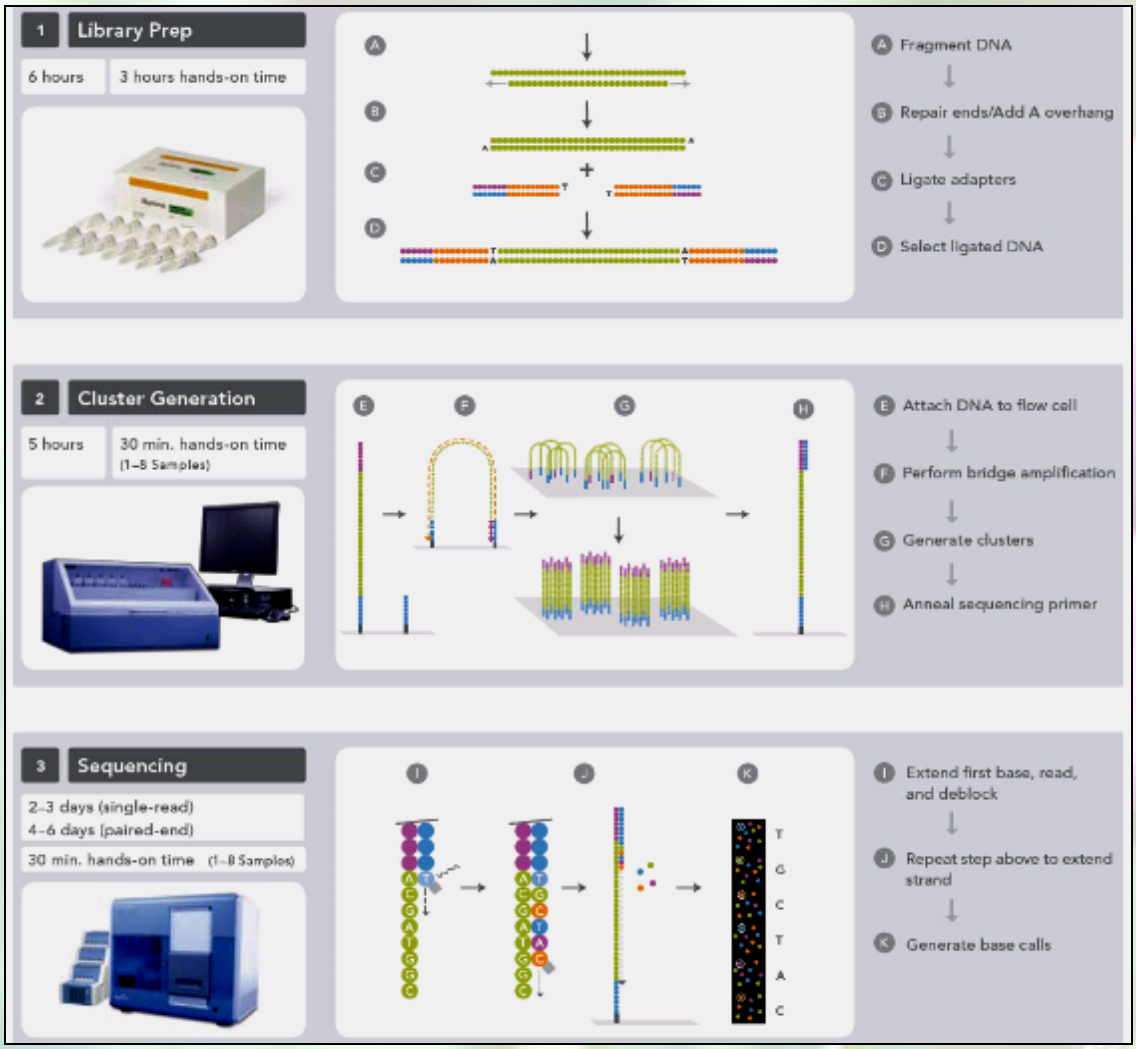


Figure 5 An overview of Illumina/Solexa sequencing (www.illumina.com)

Illumina/Solexa technology has been used for generation of millions of reads in several plant species. For instance, 3,948,871 reads from two separate short RNA libraries developed from total RNA extracted from *M. truncatula* leaves (Szittya et al. 2008), more than 6 million raw reads ranging from 18 to 30 nucleotides in length in soybean seeds (Song et al. 2011), ~3.8 billion nucleotides of high quality sequence in Arabidopsis (Cokus et al. 2008) etc.

Illumina/Solexa 1Gb sequencing yielded 120 million–173 million reads that were aligned to a Arabidopsis reference genome sequence and used for identification of 8, 23,325 unique SNPs in Arabidopsis (Ossowski et al. 2008). Illumina/Solexa 1G sequencing generated 574 Mbp data which was used to identify and mark repetitive regions and define putative gene space in barley (Wicker et al. 2008). In case of polyploidy crop, *Brassica napus*, 20 million reads were generated from each of two cultivars: Tapidor and Ningyou 7, and 23,330–41,593 putatative SNPs were identified in the two cultivars (Trick et al. 2009). In the case of chickpea, in addition to this study, Garg et al. (2011) generated ~107 million high-quality trimmed reads from root and shoot tissues of ICC 4958.

2.6 SNP Genotyping Platforms

A number of SNP genotyping platforms e.g. pyrosequencing (Alderborn et al. 2000; Ching and Rafalski 2002; Varshney et al. 2008), mass spectrometry (Rodi et al. 2002), Affymetrix chips (Borevitz et al. 2003) GoldenGate assays (Fan et al. 2003; Rostoks et al. 2006) are available. However, in view of cost-effective and high-throughput SNP genotyping very recently, KASPar assay from KBiosciences and Illumina's BeadXpress array has been developed.

2.6.1 Illumina's goldengate assay

This assay involves activation of genomic DNA using paramagnetic particles and PCR based amplification of activated DNA using three oligos and a universal PCR primer pair for each SNP. Two of the oligos used are allele specific oligos which on ligation to DNA containing target allele extends and ligates to the third locus specific oligo (LSO) which contains SNP specific tag and sequence complementary to the universal primer. The universal primer carries allele specific fluorescent label and contains an address sequences which helps in binding of the amplified product to the beads of fiber optic array. Data analysis is done using scatter plots. These beads are present in micro-titer plate which facilitates the genotyping in multiple of 96. GoldenGate assays have been developed for several crop species such as barley (Rostocks et al. 2006), wheat (Akhunov et al. 2009), soybean (Hyten et al. 2010), cowpea (Muchero et al. 2009) and chickpea (Doug Cook, unpublished) etc. SNP genotyping based on GoldenGate assay has

been found very successful in constructing genetic map, undertaking trait mapping and association mapping in several crop species like rice (Arai-Kichise et al. 2011), maize (Yan et al. 2009), wheat (Akhunov et al. 2009), barley (Close et al. 2009), oil seed rape (Durstewitz et al. 2010) and legumes such as soybean (Hyten et al. 2008), cowpea (Muchero et al. 2009) and pea (Deulvot et al. 2010).

2.6.2 VeraCode assays for BeadXpress system

With an objective of developing a cost-effective genotyping platform with <768 SNPs, Illumina has introduced SNP genotyping platform that is flexible in assay content and multiplexing (up to 384 analytes), and can serve medium- to high-throughput applications. The Illumina BeadXpress platform supports the GoldenGate Genotyping Assay on digitally inscribed VeraCode microbeads to allow streamlined workflow, rapid detection, unparalleled data reproducibility and consistency. Thus, it is a highly valuable tool for biomarker research and validation, pharmaceutical development, as well as the development of molecular diagnostic tests. Deulvot et al. (2010) used the Illumina GoldenGate and the Veracode technologies on a BeadXpress Platform, and genotyped a mapping population as well as a germplasm collection in pea and validated 384-SNP dataset. Also in rice one 96-plex and three 384-plex OPA sets were designed for genotyping using VeraCode on the Illumina BeadXpress Reader (Thomson et al. 2010).

2.6.3 KASPar assay

The KASPar assay is a PCR based novel homogeneous fluorescent SNP genotyping system. This technology utilizes a unique form of allele-specific PCR that is distinct and different to conventional amplification refractory mutation system (ARMS) using a Fluorescence Resonance Energy Transfer (VIC/FAM) based homogeneous format. The chemistry involves two competitive allele-specific tailed forward primers and one common reverse primer. The KASPar assay system relies on the discrimination power of a novel form of competitive allele-specific PCR to determine the alleles at a specific locus within genomic DNA for SNP typing (Chen et al. 2010; http://www.kbioscience.co.uk/)

2.6.4 CAPS assays

With a need to have an assay that is robust, yet cost effective, and could be performed using standard gel-based procedures, SNP2CAPS was considered. In this context, CAPS markers have been shown to meet these criteria. However, converting SNPs to CAPS markers can be a difficult process if done manually. In order to address this problem, program, SNP2CAPS, that facilitates the computational conversion of SNP markers into CAPS markers SNPs by relating the SNP position to the presence / absence of a restriction site in amplicon using 'SNP2CAPS' programme (Thiel et al.2004). Such kind of studies was done on barley (Kota et al. 2008), chickpea (Choudhary et al. 2009; Varshney et al. 2007b), arabidopsis (Hou et al. 2010), brassica (Möhring et al. 2005) etc.

2.6.5 Intron spanning region (ISR) markers

ISR markers are the current markers of choice to explore poorly characterized genomes for DNA polymorphism. These gene based markers are used to scan introns for suitably variable markers (Feltus et al. 2006). The CISP markers are designed from conserved exon sequences that flank introns in order to maximize (intronic) polymorphism discovery rates within a taxon while maintaining cross taxa applicability via DNA conservation in the priming sites. This technique has been effective in both plants and animals (Aitken et al. 2005; Feltus et al. 2006; Fredslund et al. 2006a,b). For instance, a total of 19,719 Allium ESTs, 15,661 Musa ESTs, and 2,074 Oryza BACs were used to identify CISPs and evaluate their suitability as pan-taxon genomic resources for both well studied models and resource poor taxa in legumes. The intron based polymorphism is exploited in legumes, pearl millet and other grasses to develop functional markers (Feltus et al. 2006; Fredslund et al. 2006a; Yadav et al. 2008).

2.7 Linkage Mapping in Chickpea

Linkage maps provide the platform for mapping and tagging of useful traits using the information of tightly linked markers. There has been considerable progress on genetic mapping in chickpea and other legume species in recent years. A large number of linkage maps are published on inter-specific crosses (*C. arietinum* x *C. reticulatum*, *C. arietinum* x *C.*

echiniospermum) (Gaur and Slinkard 1990a, 1990b; Kazan et al. 1993; Simon and Muehlbauer 1997; Winter et al. 1999, 2000; Tekeoglu et al. 2002; Pfaff and Kahl 2003; Cobos et al. 2006) as compared to intra-specific crosses (Cho et al. 2002; Flandez- Galvez et al. 2003a ; Radhika et al. 2007; Kottapalli et al. 2009). The most probable reason for very few studies are available on the mapping of intra-specific mapping population could be due to scarcity of molecular markers and extremely low level of genetic polymorphism detected within the cultivated gene pool (Udupa et al. 1993; Labdi et al. 1996).

In chickpea, the first inter-specific map was reported by Gaur and Slinkard in 1990s (1990a, 1990b). This map was developed by using morphological markers i.e. isoenzyme markers. They have integrated 29 morphological and isozyme marker loci on the F2 mapping population of the C. *arietinum* x *C. reticulatum*. Similar report came after a year from Kazan et al. 1993, reporting mapping of 28 new isozyme markers onto the F2 lines of *C. arietinum* × *C. reticulatum* and *C. arietinum* × *C. echinospermum*. 91 morphological, isozyme, RFLPs and RAPDs markers were mapped on *C. arietinum* × *C. reticulatum* and *C. arietinum* × *C. echinospermum* F2 lines by Simon and Muehlbauer, 1997. Also several other reports like integration of STMS, amplified fragment length polymorphism (AFLP) by Winter et al. 1999, morphological isozyme, inter simple sequence repeat (ISSR) and RAPD loci by Santra et al. 2000, STMS markers by Tekeoglu et al. 2002; Flandez-Galvez et al. 2003; Udupa and Baum, 2003; Cho et al. 2004; Tar'an et al. 2007). Few reports of linkage mapping in chickpea is listed in are listed in the Table 1.

Table 1: Molecular genetic maps developed for chickpea (modified from Upadhyaya et al. 2011)

Mapping population	Markers mapped	References
Interspecific		
C. arietinum × *C. reticulatum* (F_2)	29 marker loci (morphological and isozyme)	Gaur and Slinkard 1990a, 1990b
C. arietinum × *C. reticulatum* (F_2) *C. arietinum* × *C. echinospermum* (F_2)	28 marker loci (morphological and isozyme)	Kazan et al. 1993
C. arietinum × *C. reticulatum* (F_2) ; *C. arietinum* × *C. echinospermum* (F_2)	91 marker loci (morphological, isozyme, RFLPs and RAPDs)	Simon and Muehlbauer 1997
C. arietinum × *C. reticulatum* (F_2)	120 STMS loci	Winter et al. 1999
C. arietinum × *C. reticulatum* (F_2)	354 marker loci (SSRs, DAF, AFLPs, ISSRs, RAPDs, isozyme, cDNA, SCAR and morphological)	Winter et al. 2000
C. arietinum × *C. reticulatum* (F_2)	116 marker loci (RAPDs, ISSRs, isozyme, and morphological)	Santra et al. 2000
FLIP 84-92C × PI 599072 (RIL)	144 marker loci (111 RAPDs, 21 ISSRs, 1 morphological, 11 isozyme)	Santra et al. 2000
C. arietinum × *C. reticulatum* (F_2)	117 marker loci (Addition of RGA Potkin 1-2 171 to linkage group 5 of Santra et al. 2000)	Rajesh et al. 2002
C. arietinum ICC4958 × *C. reticulatum* PI 489777 (RIL)	56 marker loci (55 SSRs and 1 RGA)	Tekeoglu et al. 2002
C. arietinum × *C. echinospermum* (F_2)	83 marker loci (SSRs, RAPDs, ISSRs and RGA)	Collard et al. 2003
C. arietinum × *C. reticulatum* (F_2)	296 marker loci (47 defense response gene markers to the map of Winter et al. 2000)	Pfaff and Kahl 2003
C. arietinum 'Lasseter' × *C. echinospermum* 'PI 527930' (F_2)	83 marker loci (54 RAPDs, 14 SSRs, 9 ISSRs, 6 RGA)	Collard et al. 2003

C. arietinum 'Hadas' × *C. reticulatum* 'Cr205' (RIL)	93 marker loci (91 SSRs, 2 CytP450 markers)	Abbo et al. 2005
C. arietinum (ILC72) × *C. reticulatum* (Cr5-10)	89 marker loci (RAPDs, ISSRs, STS)	Cobos et al. 2006
C. arietinum ICC 4958× *C. reticulatum* (RIL) PI 489777	521 marker loci (SSR, RAPD, AFLP, RGA)	Nayak et al. 2010
C. arietinum ICC 4958× *C. reticulatum* (RIL) PI 489777	1,921 marker loci (BEC-SSR, DArT, SNP)	Thudi et al. 2011
Intraspecific		
ICCV 2 × JG 62 (RIL)	103 marker loci (68 SSRs, 34 RAPDs, 4 ISSRs, and 5 morphological)	Cho et al. 2002
ILC 1272 × ILC 3279	55 (SSRs and Ascochyta blight resistance loci)	Udupa and Baum 2003
ICC 12004 × Lasseter (F_2)	69 marker loci (54 SSRs, 12 RGAs, 3 ISSRs)	Flandez- Galvez et al. 2003a
CA 2139 × JG 62 (RIL), CA 2156 × JG 62 (RIL)	138 marker loci (118 RAPDs, 13 SSRs 4 morphological and 3 ISSRs)	
JG 62 × Vijay (RIL), Vijay × ICC 4958 (RIL)	273 marker loci (RAPDs and ISSRs)	Radhika et al. 2007
ICC 4991 × ICCV 04516 (F2)	84 marker loci (82 SSRs and 2 ESTs)	Kottapalli et al. 2009
WR315 × C104	102 marker loci (20 ISSR, 75 STMS, 6 RAPD, and 1 STS)	Sharma et al. 2004
ICCV-2 × JG-62	138 (STMS)	Gour et al. 2011
Consensus map		
Five narrow crosses (Desi × Kabuli types)	229 (STMSs, SCARs, ASAP, fusarium resistance gene, morphological, RAPD and ISSR)	Millan et al. 2010
Five wide crosses (*C. arietinum* × *C. reticulatum*)	555 marker loci (STMS, RAPD, cross-genome markers)	Millan et al. 2010

With the advancement of the genomics, a number of SSR markers were developed in chickpea by using SSR enriched library (Huettel et al. 1999; Winter et al. 1999; Sethy et al. 2003, 2006a, 2006b; Choudhary et al. 2006; Nayak et al. 2009), EST- sequences (Choudhary et al. 2006,2009) and BAC-end sequences (Lichtenzveig et al. 2005, Thudi et al. 2011). The first SSR inter-specific genetic map was developed by Tekeoglu et al. 2002. The group had integrated 55 SSR and 1 RGA marker on 142 RILs derived from ICC 4958 (C. *arietinum*) and PI 489777 (*C. reticulatum*). This map spanned 1,174.5 cM with 167 marker loci. Pfaff and Kahl 2003 had integrated 296 marker loci (47 defense response gene markers to the map of Winter et al. 2000) on *C. arietinum* × *C. reticulatum*- F2 lines. Nayak et al. (2010) constructed a genetic map on 132 RILs of ICC 4958 × PI 489777 contains 521 loci organized into eight linkage groups that span 2,602 cM, with an average inter-marker distance of 4.99 cM. Very recently, Thudi et al. 2011 has developed a comprehensive genetic map comprising 1,291 markers on eight linkage groups (LGs) spanning a total of 845.56 cM distance with an average inter-marker distance of 0.65 cM on the same inter-specific mapping population as of Nayak et al. 2011.

For intra-specific mapping population, the first map was reported by Cho et al. 2002. The map had 103 marker loci (68 SSRs, 34 RAPDs, 4 ISSRs, and 5 morphological markers) on ICCV 2 × JG 62, recombinant inbred lines. Udupa and Baum (2003) developed another map based on 97 RILs of ILC 1272 × ILC 3279, by integrating 55 SSRs and Ascochyta blight resistance loci. Flandez-Galvez et al. reported the genetic map in same year, 2003. It was developed on population of 85 F2 plants from an intraspecific cross between desi cultivars ICC12004 and Lasseter and was based on 66 markers including 51 SSRs. Recently using consensus map approach, Radhika et al. (2007) had developed an integrated intraspecific map of *C. arietinum* using two recombinant inbred line populations Vijay × ICC 4958 and JG 62 × Vijay. Another consensus map was very recently developed by Millan et al. 2010. It was developed from wide crosses comprised 555 loci, including 135 STMSs and 33 cross-genome markers in addition to AFLPs, ASAPs, chickpea gene-specific markers, isozymes, ISSRs, RAPDs, RGAs, *Fusarium* resistance genes and two morphological traits (flower color and growth habit) covering 652.67 cM while, the consensus map derived from narrow crosses include 229 loci, including 99

STMSs, 3 SCARs, 1 ASAP, *Fusarium* resistance, genes, 5 morphological traits as well as RAPD and ISSR markers distributed on eight linkage groups covering 426.99 cM. Gaur et al. (2011) has developed an intraspecific genetic map on 126 RILs of mapping population ICCV-2 × JG-62 on which 138 markers were mapped into eight LGs that spanned 630.9 cM with an average marker density of 4.57 cM.

A transcript map is an important resource for geneticist to identify quantitative trait loci and also for map based cloning, as well as to a breeder who depends on the marker assisted selection in cultivar developement. Due to scarcity of genomic resources and a low polymorphism in cultivated chickpea genotypes, initial genetic mapping studies were restricted to inter-specific mapping populations. While several research groups used the *C. arietinum* × *C. reticulatum* mapping population for developing genetic map by deploying a variety of molecular markers (Winter et al. 1999, 2000; Nayak et al 2009). But with the advancement of next generation sequencing and less cost comparative to sanger sequencing, a large number of GMMs as well as transcript maps have been developed in several crop species such as rice (Wu et al. 2002), wheat (Qi et al. 2004), barley (Stein et al. 2007; Kota et al. 2008; Sato et al. 2009), soybean (Choi et al. 2007), common bean (Galeano et al. 2009) by using above mentioned methodologies and several others.

2.8 Trait Mapping and QTL Analysis in Chickpea

A large number of genes/QTLs have been tagged by using different kinds of molecular markers in several crops, but very few markers have been validated and deployed in breeding programmes of legume crops. The conventional phenotypic approach for selection of specific traits requires the evaluation of the trait from multiple environments over several years; this is often very expensive, time consuming, and labor intensive (Yuan et al. 2010). On the other hand, genomics offers great opportunities for dissecting quantitative traits into their single genetic determinants (Dudley, 1993; Tanksley, 1993; Lee, 1995; Young, 1996,2005; Beavis and Kein, 1996; Quarrie, 1996; Prioul et al. 1997; Tuberosa et al. 2002). Recently, a number of studies on quantitative trait loci (QTL) analysis relating root morphology and physiology in plants were reported (Beebe et al. 2006; Li et al. 2005, 2007; Chen et al.2002; Cichy et al.

2009). The first QTL descriptive analysis began by relating to root weight in field-grown maize (Reiter et al. 1991) and subsequently to root architecture traits such as root hair length, and lateral root branching and length (Zhu et al. 2005a, b). Meanwhile, QTL analysis of root traits and efficiency in legumes started much later.

Identification of QTLs is a roadmap to the way to MAS (Ribaut et al. 2002; Morgante and Salamini, 2003) and gene pyramiding. The MAS in breeding has revolutionized the improvement of temperate field crops (Varshney et al. 2006; Koebner 2004) and will have similar impacts on breeding of legume crops, particularly for traits where phenotyping is only possible late in the season, and where screening of traits is difficult

Due to the availability of large number of molecular markers and advancement in genotyping facilities, there is a increase in number of studies reporting QTLs for drought related traits in different crops under drought stress indicates a growing interest in this approach (Tuberosa et al. 2002; Varshney et al. 2005a,b; Tuberosa et al. 2005). These markers, after validation in desired genetic background can be used for marker-assisted selection. Although progress in molecular mapping has been slow, some efforts have been made to identify the QTLs/ markers associated with resistance to some diseases.

In chickpea majority of gene/QTL mapping studies have been reported on biotic stresses like *Ascochyta blight* and *Fusarium* wilt resistance. As a result several QTLs for *Ascochyta* blight resistance are available (Millan et al. 2006). The major constraint of chickpea production among biotic stresses is *Fusarium* wilt which causes great yield losses. For the first time, Fusarium resistant loci (Foc1) was tagged with two RAPD markers (UBC170550 and CS27700) at 7% recombination from the resistant loci by Mayer et al. (1997). Since this was a RAPD marker, more precise ASAP primer was designed out of it. This marker was mapped in the LG-6 of the map of Simon and Muehlbauer (1997). A large number of studies involving inter- and intraspecific RIL populations revealed the organisation of resistance genes for fusarium wilt races 1, 3, 4 and 5 (*foc*1 and *foc*3, *foc*4 and *foc*5; Mayer et al. 1997; Ratnaparkhe et al. 1998; Tullu et al. 1998, Winter et al. 2000; Sharma et al. 2004) in two adjacent resistance gene clusters

on LG 2 flanked by STMS markers, TA96 (*foc1–foc4* cluster), TA27 (*foc3– foc5* cluster) (Millan et al. 2005). Moreover, several potential resistance and pathogenesis related genes were localized on the LG 2 (Hüttel et al. 2002; Pfaff & Kahl 2003). Milan et al. (2005) therefore speculated that LG 2 is a hot spot for pathogen defense. It is mentioned to note that not all fusarium wilt resistance genes are located on LG 2 of the genetic map of Winter et al. (2000). This can be supported by the study done by Rubio et al. (2003). He showed that two different genes can confer resistance race 0 (*foc0*1 and *foc0*2) but also demonstrated linkage of the *foc0*1 to RAPD marker *OPJ20*600 (resistance derived from line JG62). Subsequently, this locus was mapped on LG5 tightly flanked by a RAPD (3 cM apart) and an STMS (2 cM apart) marker (Cobos et al. 2005). Udupa and Baum (2003) identified and mapped a major locus (*ar1*, mapped on linkage group 2), which confers resistance to pathotype I, and two independent recessive major loci (*ar2a*, mapped on linkage group 2 and *ar2b*, mapped on linkage group 4), with complementary gene action conferring resistance to pathotype II. Out of two pathotype II-specific resistance loci, one (*ar2a*) linked very closely with the pathotype I specific resistance locus, indicating a clustering of resistance genes in that region of the chickpea genome. Apart from these, there also reports for resistance gene localization of various races of Fusarium wilt pathogen (races 1, 2 and 3by Gowda et al. 2009; Sharma et al. 2004b; races 4 and 5 by Winter et al. 2000).

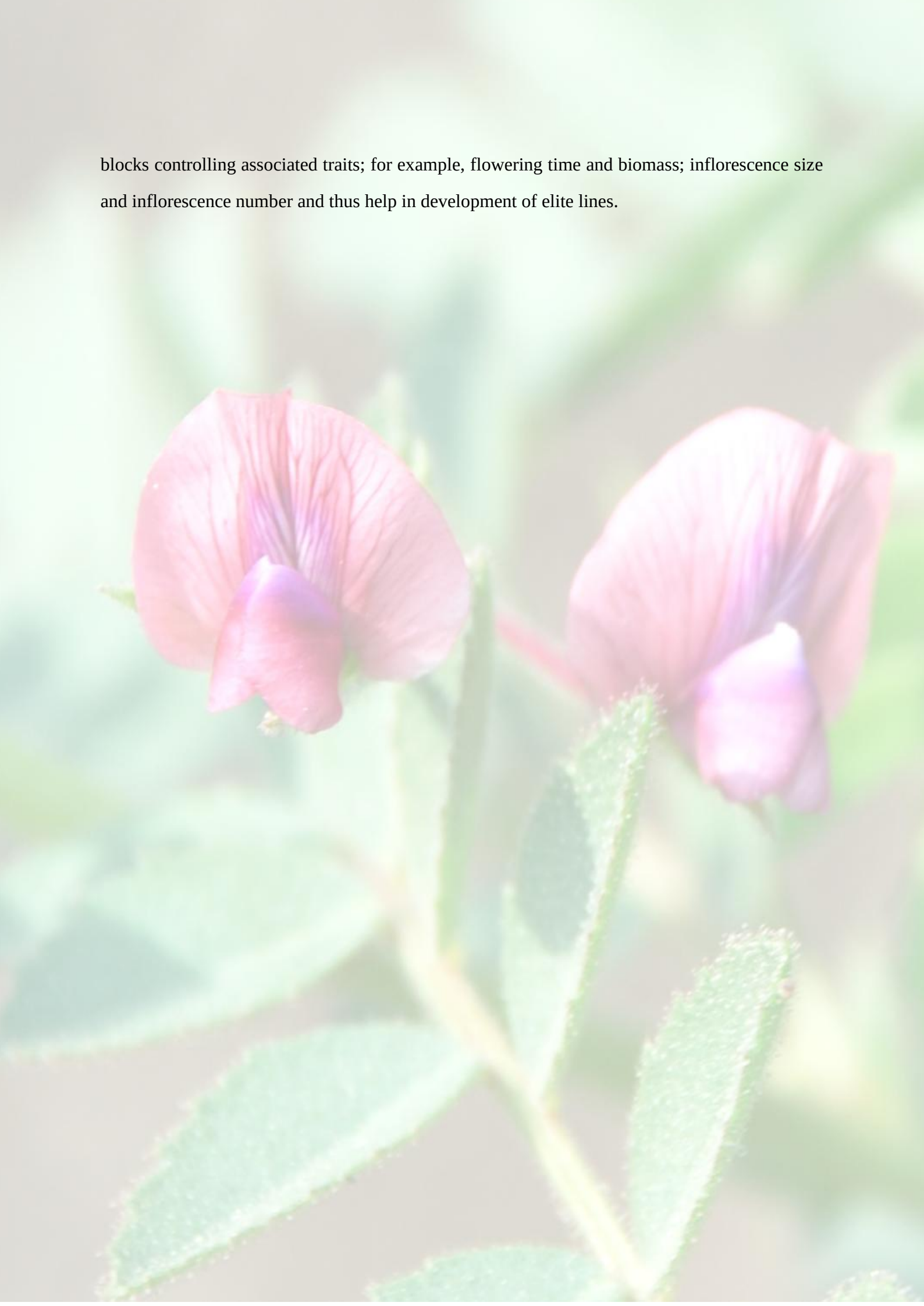
Although considerable progress has been made in identifying QTLs in chickpea related to fusarium wilt and ascochyta blight disease resistance, but information on the genetic basis of traits related to drought tolerance in chickpea is limited. A total of 160 morphological markers including flower colour, double podding, seed coat thickness and resistance to Fusarium wilt race 0 (foc 0) and 134 molecular markers (3 ISSRs, 13 STMS and 118 RAPD) were used to identify QTL associated with flower colour, seed coat thickness and single/double podding locus (Cobos 2005). LG3 included a gene for resistance to foc 0 (foc01/foc01) flanked by RAPD marker OPJ20600 and STMS marker TR59.

Rajesh et al. 2002 and Cobos et al. 2005 tagged the gene responsible for single pod/double pod in chickpea with the marker TA80. Or et al. (1999) reported a major photoperiod response gene (*Ppd*) affecting time to flowering. Cobos et al. 2005 discovered that maker GAA47 was associated with the QTL responsible for the trait flower colour. Cho et al. (2002) identified a single QTL for days to 50% flowering on LG3 with a LOD score of 3.03 and a QTL for seed weight on LG4 accounting for 52% of the total phenotypic variation. Abbo et al. (2005) identified QTLs responsible for Beta- carotene concentration, leutin and seed weight. The QTLs responsible for first flower time was detected by Lichtenzveig et al. (2006) on LG1 and LG2. For improving cold tolerance, AFLP markers have been identified in an F2 progeny of a cross between the chilling sensitive cultivar Amethyst and the chilling tolerant ICCV-88516 (Clarke and Siddique,2003). In the case of flowering, a major gene (efl-1) for time of flowering was reported by Kumar and van Rheenen (2000). QTL for flowering was located on the LG 03 in an inter-specific RIL population and explained 28% of the total phenotypic variation (Cobos et al. 2006).

For a successful MAS, direct QTL-allele-specific markers (such as STS markers derived from cloned QTL alleles) are needed (Sorrells and Wilson, 1997). Although success in terms of cloning of QTL alleles is highly limited, map-based cloning and candidate gene approaches would increasingly facilitate isolation and characterization of agronomically important QTLs (Lynch and Walsh, 1998). Only few reports are available for successful map-based positional cloning of QTLs like for glucose and fructose content in tomato (Brix 9-2-5) (Fridman et al. 2000), fruit weight in tomato (fw 2.2) (Frary et al. 2000), heading date in rice (Hd1) and (Hd6) (Takahashi et al. 2001), etc. Cloning the genetic determinants of QTLs is expected to bridge the missing link in understanding the relationship between genotype and phenotype.

In summary, QTL mapping of yield, quality components, biotic and abiotic stress components can provide crop breeders with a better understanding of the basis for genetic correlations between economically important traits like linkage and/or pleiotropic relationships between gene

blocks controlling associated traits; for example, flowering time and biomass; inflorescence size and inflorescence number and thus help in development of elite lines.



3. MATERIALS AND METHODS

3.1 Plant Material

3.1.1 Genotypes used for generating Illumina/Solexa sequence reads

Two chickpea genotypes, *Cicer arietinum* - ICC 4958 and ICC 1882, parents of a mapping population were selected for drought experiments. ICC 4958 is an accession available from the ICRISAT germplasm collection. It is a drought-tolerant donor parent that produces high yields under drought prone environments. Whereas ICC 1882 is a drought sensitive genotype selected for imposing drought stress and was collected from chickpea breeding division of ICRISAT.

3.1.2 Genotypes used for sequence diversity analysis

For allele specific resequencing, a set of 22 diverse chickpea genotypes originating from 7 countries was collected from Genebank and Chickpea Breeding Division of ICRISAT for the development of SNP markers through allele-specific sequencing. These chickpea genotypes represent 9 *Cicer* species including 11 cultivated and 11 wild chickpea genotypes (Table 2). All the newly developed markers were optimized for amplification initially on two genotypes ICC 4958 and ICC 1882.

Apart from these two chickpea genotypes, parental polymorphism was also detected on the parents of one intra-specific mapping population- ICC 283 and ICC 8261 and one inter-specific mapping population- PI 489777. All the genotypes were collected from Chickpea Breeding Division of ICRISAT.

3.1.3 Mapping population used for developing transcript map

Recombinant inbred lines (RILs, $F_{10:11}$) derived from two genotypes ICC 4958 and ICC 1882 of *Cicer arietinum* were used to construct the gene based transcript map. This population was derived from a single F_1 plant and advanced by single seed descent from the F_2 to F_{10}.

Table 2 : List of chickpea genotypes used for allele re-sequencing for identification of SNPs

S.No.	Accession	Species	Geographical origin	Type

1	ICC 1882	*C. arietianum*	India	Desi
2	ICC 283	*C. arietinum*	India	Desi
3	ICC 3137	*C. arietinum*	Iran	Desi
4	ICC 4958	*C. arietinum*	India	Desi
5	ICC 506EB	*C. arietinum*	India	Desi
6	ICC 8261	*C. arietinum*	Turkey	Kabuli
7	ICCV2	*C. arietinum*	India	Kabuli
8	Annigeri	*C. arietinum*	India	Desi
9	ICCC 37	*C. arietinum*	India	Desi
10	JG 62	*C. arietinum*	India	Desi
11	Vijay	*C. arietinum*	India	Desi
12	ICC 17162	*C. cuneatum*	Ethiopia	Wild
13	ICC 17148	*C. judaicum*	Lebanon	Wild
14	ICC 17248	*C. microphyllum*	Pakistan	Wild
15	ICC 17152	*C. pinnatifidum*	Turkey	Wild
16	ICC 17123	*C. reticulatum*	Turkey	Wild
17	IG 72933	*C. reticulatum*	Turkey	Wild
18	IG 72953	*C. reticulatum*	Turkey	Wild
19	PI 489777	*C. reticulatum*	Turkey	Wild
20	ICC 17116	*C. yamashitae*	Afghanistan	Wild
21	ICC 17122	*C. bijugum*	Turkey	Wild
22	ICC 17138	*C. pungens*	Afghanistan	Wild

3.2 DNA Isolation

Total genomic DNA from above mentioned genotypes and each RIL was extracted from leaves of 15 days old seedlings using high throughput mini DNA extraction protocol as mentioned in Cuc et al. (2008). The steps of high throughput DNA extraction are mentioned below:

Sample preparation

- *Leaves were harvested from 15 days old seedlings.*

- *Leaf tissue of 70-100mg was placed in 12 x 8-well strip tube with strip cap (Marsh Biomarket, USA) in a 96 deep-well plate together with two 4mm stainless steel grinding balls (Spex CertiPrep, USA)*

CTAB extraction

- *Each sample was mixed with 450µl of preheated (65ºC) extraction buffer (100 mM Tris-HCl (pH-8, 1.4 M NaCl, 20mM EDTA, CTAB (2-3%w/v), β- mercaptoethanol) was added to each sample and secured with eight strip caps.*
- *Samples were processed in a Geno Grinder 2000 (Spex CertiPrep, USA), following the manufacturers instructions, at 500 strokes/min for 5 times at 2 min interval.*
- *Plate was fitted into locking device and incubated at 65ºC for 10 min with shaking at periodical intervals.*

Solvent extraction

- *Each of the samples was mixed with 450µl of chloroform-isoamylalcohol (24:1) by inverting twice.*
- *Plate was centrifuged at 5500 rpm for 10 min. The aqueous layer (300µl) is transferred to fresh strip tubes (Marsh Biomarket, USA)*

Initial DNA precipitation

- 0.7 vol (210µl) of isopropanol (stored at –20ºC) was added to each sample and inverted once to mix.
- Plate was centrifuged at 5000 rpm for 15 min.
- Supernatant was decanted from each sample and pellet was air dried for 20 min.

RNAase treatment

- 200µl low salt TE (10 mM Tris EDTA (pH-8)) and 3µl RNAase was added to each sample and incubated at 37ºC for 30 min.

Solvent extraction

- 200µl of phenol-chloroform-isoamylalcohol (25:24:1) was added to each sample and inverted twice to mix.
- Plate was centrifuged at 5000 rpm for 5 min.
- Aqueous layer was transferred to a fresh 96 deep-well plate (Marsh Biomarket, USA).
- 200µl chloroform-isoamylalcohol (24:1) was added to each sample and inverted twice to mix.
- Plate was centrifuged at 5000 rpm for 5 min.
- Aqueous layer was transferred to a fresh 96 deep-well plate.
- 315µl ethanol-acetate solution (30ml ethanol, 1.5ml 3M NaOAc (pH-5.2)) was then added to each sample and placed in –20ºC for 5 min.
- Plate was again centrifuged at 5000 rpm for 5 min.
- Supernatant was decanted from each sample and pellet was washed with 70% ethanol.
- Plate was centrifuged at 6000 rpm for 10 min.
- Supernatant was again decanted from each sample and samples were air dried for approximately 1 hour.
- Pellet was re-suspended in 100µl low-salt TE and stored at 4ºC.

Quantification of DNA

DNA quantification was carried on 0.8% agarose gel containing 0.5 µl/10 ml Ethidium bromide (EtBr, 10mg/ml). The DNA was normalized to 5 ng/µl working stock by visualizing the standard λ DNA molecular weight markers (2.5 ng/µl, 5 ng/µl, 10 ng/µl) on 0.8% agarose gel.

3.3 Drought Stresses Imposition on Selected Chickpea Genotypes

Selected chickpea genotyes, ICC 4958 and ICC 1882 (*Cicer arietinum*), were exposed to different drought stresses.

Types of drought stresses

Four different types of drought stresses were imposed on ICC 4958 and ICC 1882 at different conditions in ICRISAT, Patancheru (17° 30' N; 78° 16' E; altitude 549 m). The different types of drought stresses are described below:

3.3.1 PEG induction

Polyethylene glycol (PEG) is a polymer which is produced in different molecular weights (4000 to 8000). It is known to induce controlled drought stress in nutrient cultures. Four different concentrations of PEG (molecular weight 8000) were estimated for slow drought induction onto ICC 4958 and ICC 1882 in greenhouse maintained at 30±1°C.

Germination of seeds: 15 seeds of each genotype ICC 4958 and ICC 1882 were surface sterilized with the 20% chloride solution. First all the seeds were placed in 20% chloride solution followed by three washes with de-ionized water/Milli Q water. Thereafter the sterilized seeds were germinated for about 3-4 days on the wet tissue paper in a petri plate.

Growing healthy plants: Germinated seeds (4 no.s) of each genotype were transferred to different containers filled with Yoshida Solution (Annexure 2). Nylon mesh was used to keep distance between the germinated seeds to avoid intermingling of roots among different seedlings. The solution was changed after every 7 days to retain the nutrient balance in the container. Plants were kept for two weeks in these containers for proper growth.

PEG stress treatment: Four different concentrations of PEG (Mol wt: 8000 m) was used to impose slow drought on healthy grown plants- 50mM, 10mM, 5mM and 1mM. Two healthy grown plants of each genotype (ICC 4958 and ICC 1882) were selected as a control and a stressed plant. These seedlings were transplanted into conical flasks covered with aluminum foils. In one flask, control plant will be grown only in nutrition solution (Yoshida solution),

while the plant to which different PEG concentration stresses (50mM, 10mM, 5mM and 1mM) are to be imposed is transplanted to medium containing nutrition solution and respective concentration of PEG. The conical flasks were named clearly to avoid any misreading or confusion during harvesting of tissues (Fig. 6a).

Stress intensity optimization: The control and the stressed plants were grown for a week and than relative water content (RWC) was recorded on a daily basis upto two more weeks. The PEG concentration of 50mM and 10mM were found to be very high and resulted in sudden death of plants. Therefore, experiment was carried out with 5mM and 1mM concentration of PEG and the relative water content (RWC) was measured at different intervals of time in both ICC 4958 and ICC 1882. Comparatively, 1mM PEG concentration gave a slow drought stress effect than 5mM PEG, hence the experiment was optimized by prolonging slow stress period to 7 days with 1mM PEG treatment, and allowing an adaptive chance for survival (Fig. 6a). The stress imposed on the plant might not be purely stress due to unavailability of water but also might be due to the some toxic effect of the PEG chemical (Jacomini et al. 1988).

Figure 6(a) Plants grown under PEG induction treatment in glass house at ICRISAT

Collection of tissues: On a daily basis RWC was recorded from stressed and controlled genotypes of ICC 4958 and ICC 1882. The result of this experiment is shown in Fig. 6b, where relative water content (RWC) of the both the genotype is measured. When RWC reaches 50-60 % that indicates drought condition, the leaf, root and flower/early-pod tissues were harvested individually. The tissues from the control plants were also collected at the same time. The tissues were snap frozen in liquid nitrogen and preserved at –80° C until RNA extraction.

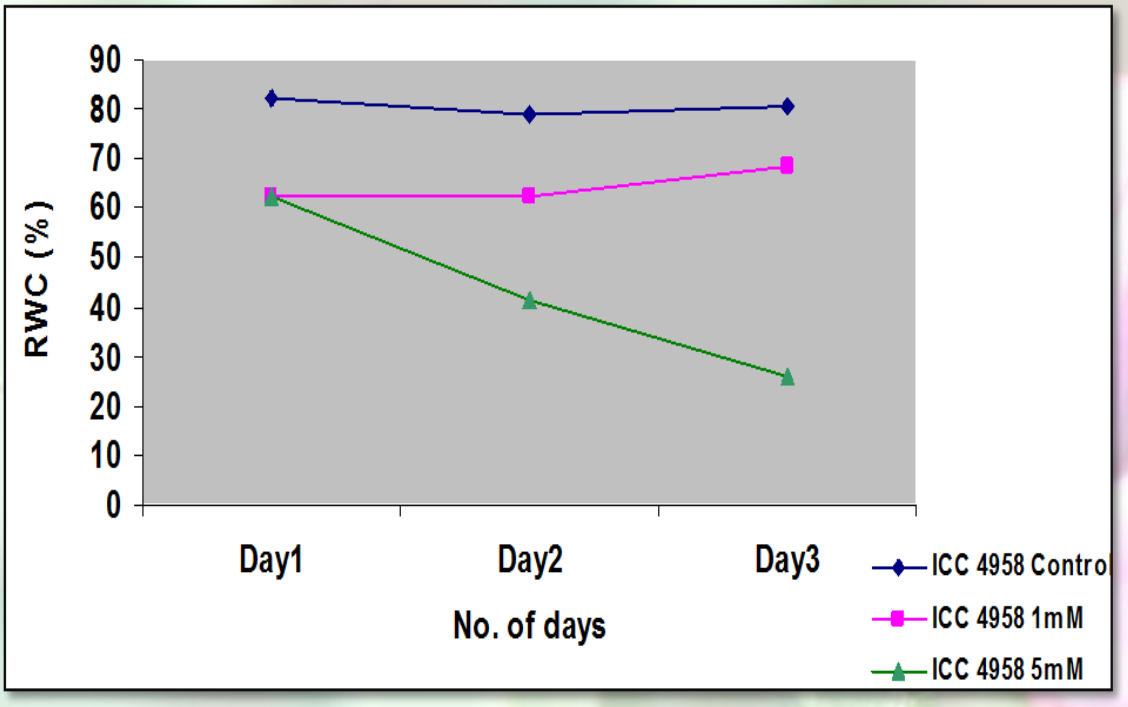


Figure 6 (b) Evaluation of chemically induced stress using 1mM PEG

3.3.2 Sudden dehydration shock

Germination of seeds: 10 seeds of each genotype ICC 4958 and ICC 1882 were surface sterilized with the 20% chloride solution. First all the seeds were placed in 20% chloride solution followed by three washes with de-ionized water/Milli Q water. Thereafter the sterilized seeds were pre-germinated for about 3-4 days on the wet tissue paper in a petri plate.

Growing healthy plants: 5 germinated seeds of each genotype were transferred to different containers filled with Yoshida Solution. Nylon mesh was used to keep distance between the germinated seeds to avoid intermingling of roots among different seedlings. The solution was

changed after every 7 days to retain the nutrient balance in the container. Plants were kept for two weeks in these containers for proper growth.

Sudden dehydration treatment: All the 10 well grown plants of ICC 4958 and ICC 1882 were imposed to sudden dehydration by removing the hydroponics solution (Fig.7). Plants were transferred to different trays and maintained under greenhouse conditions for 7 days. Relative water content was recorded on a daily basis until it reaches 50-60%.

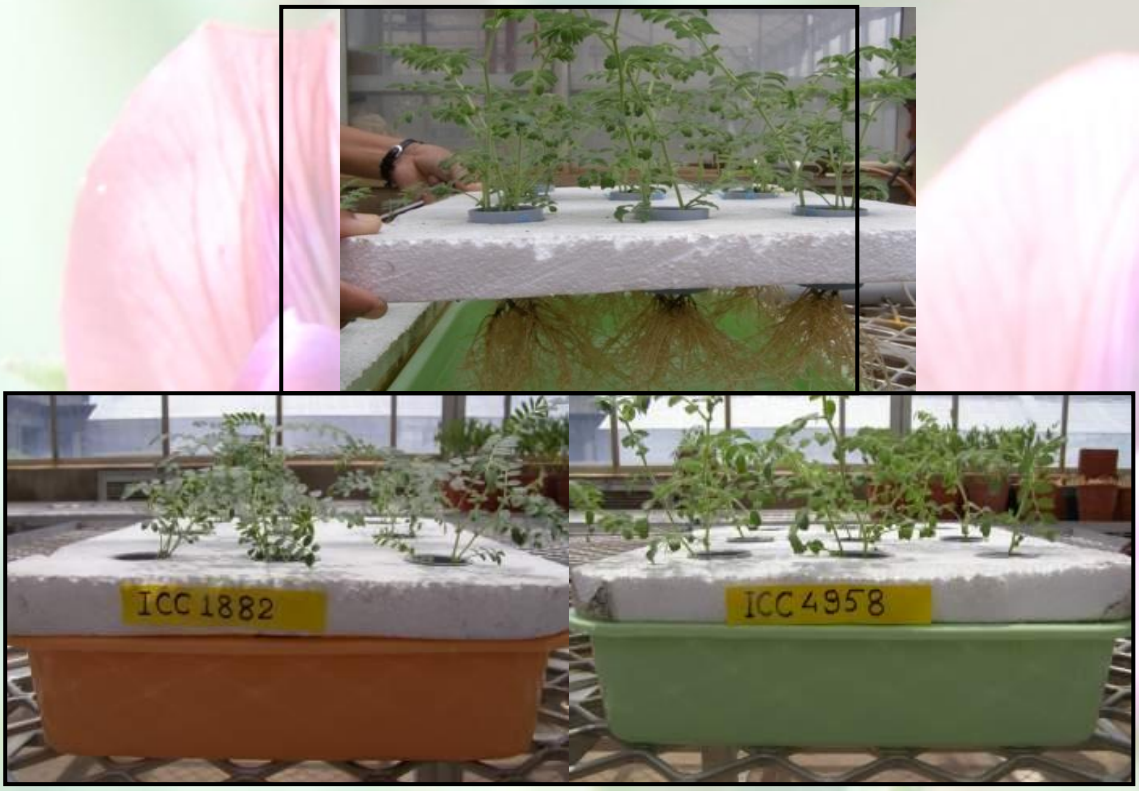


Figure 7 Sudden dehydration: Healthy plants transferred to trays in glasshouse at ICRISAT

3.3.3 Dry down experiment under field conditions

ICC 4958 and ICC 1882 were grown in lysimeters, consisting of PVC cylinders (20 cm diameter, 120 cm height) filled with a vertisol at ICRISAT (max:32.1°C -27.1°C, min: 16.8 °C-13.8°C, max : min relative humidity 29.8–42.8% : 87.4–94.3%). Temperatures and relative humidity of the air were collected from a temperature and relative humidity recorder (Gemini Tinytag Ultra 2 TGU-4500 Datalogger), which was located in the crop canopy. A PVC end plate was placed on top of four screws at the bottom of the cylinders, 3cm from the very bottom, to

prevent the soil from seeping through. The endplate did not fit the cylinder tightly and allowed water drainage. To allow a rigorous control of the bulk density of the soil profile, the vertisol used to fill the tubes was sieved in particles smaller than 1 cm. The cylinders were filled with 42 kg of dry soil leaving the top 15 cm empty and irrigated more than the field capacity requirement then allowed to drain. A top up using dry soil was done to ensure that all cylinders would be filled at the same level, about 5 cm from the brim. This top up varied little between cylinders so that the bulk density was similar in all tubes, at a value of approximately 1.2. Weighing of the cylinders indicated that all saturated cylinders weighed between 58 and 59 kg. The soil that was used to fill up the lysimeters had been fertilized with di-ammonium phosphate at a rate of 100 mg kg^{-1} soil (Fig. 8a).

Figure 8(a) Plants grown under slow drought stress treatment under field conditions in PVC cylinders at ICRISAT

Space arrangement of the lysimeters and weighing: The experimental design was a complete randomized block design (RBD) with treatment as the main factor (three blocks) and genotype as a sub-factor replicated five times within each block. In each block, planted in adjacent but separated trenches, the lysimeters were arranged next to one another and therefore the crop of chickpea was planted at a density of approximately 25 plants m^{-2}, a plant population close to the field planting (row-to-row distance of 30 cm and plant-to-plant spacing of 10cm). This was an important characteristic of the lysimetric approach to be able to accurately assess water extraction pattern of a crop cultivated in conditions that are quite similar to the field. The three

trenches were 1.2 m deep and 1.8 m wide, separated to the adjacent trench by a 20 cm concrete wall.

The top of the cylinders was equipped with a metal collar and rings that allowed the cylinder to be lifted. Weighing of the cylinders was done by lifting the cylinders with a block-chained poulie, and a S-type load cell (Mettler-Toledo, Geneva, Switzerland) was inserted between the rings of the cylinder and the poulie. The scale, of 100 kg capacity allowed repeated measurements and gave an accuracy of 10 g on each weighing.

Sowing and crop management: Ten seeds of each genotype were sown in the soil. The cylinders were then irrigated with 500 ml of water immediately after sowing and twice on alternate days with 250 ml until the seedlings uniformly emerged. The plants were thinned to two individuals per cylinder at 7 days after sowing (DAS) and then to a single plant per cylinder 14 DAS. One block was assigned to a well-watered treatment (WW) and two blocks to a water-stressed treatment (WS). One of these blocks was kept until maturity while the second block of WS treatment was prepared for the purpose of root sampling at five weeks after stress imposition. The WS treatment was imposed by cessation of watering from 23 DAS . WW plants were watered every 5 days to maintain the soil above 70% field capacity until maturity. Before initiating the weighing, the top of the cylinders were covered with a round and slit plastic sheet, on top of which 2 cm of low density polyethylene granules were laid. These layers prevented more than 90% of soil evaporation, so that cylinder weightings allowed to measure plant transpiration. The cylinders were weighed every five days from 23 DAS until 48 DAS, then once a week until 61 DAS, and then every two weeks until maturity. The transpiration data calculated from each pair of consecutive weighing were assigned to the date of the latest weighing so that, for simplicity, transpiration at 28 DAS meant the transpiration during the 23-28 DAS interval. Plant water uptake was estimated from the losses in weight of each cylinder. Then a normalized transpiration ratio (NTR) was obtained by dividing each TR value by the average of the TR values obtained in the second, third, fourth days, fifth and six days of the experiment. The experiment was terminated for each plant subjected to water deficit when the NTR was less than 0.1. The NTR for these 6 days is plotted in the Fig. 8b. The blue line represents the drop in the transpiration ratio from 1.10 to 0.10 in ICC 1882, whereas the red line

shows the drop in the transpiration ratio from 1.20 to 0.10 in ICC 4958. This is the time point at which the samples have been harvested.

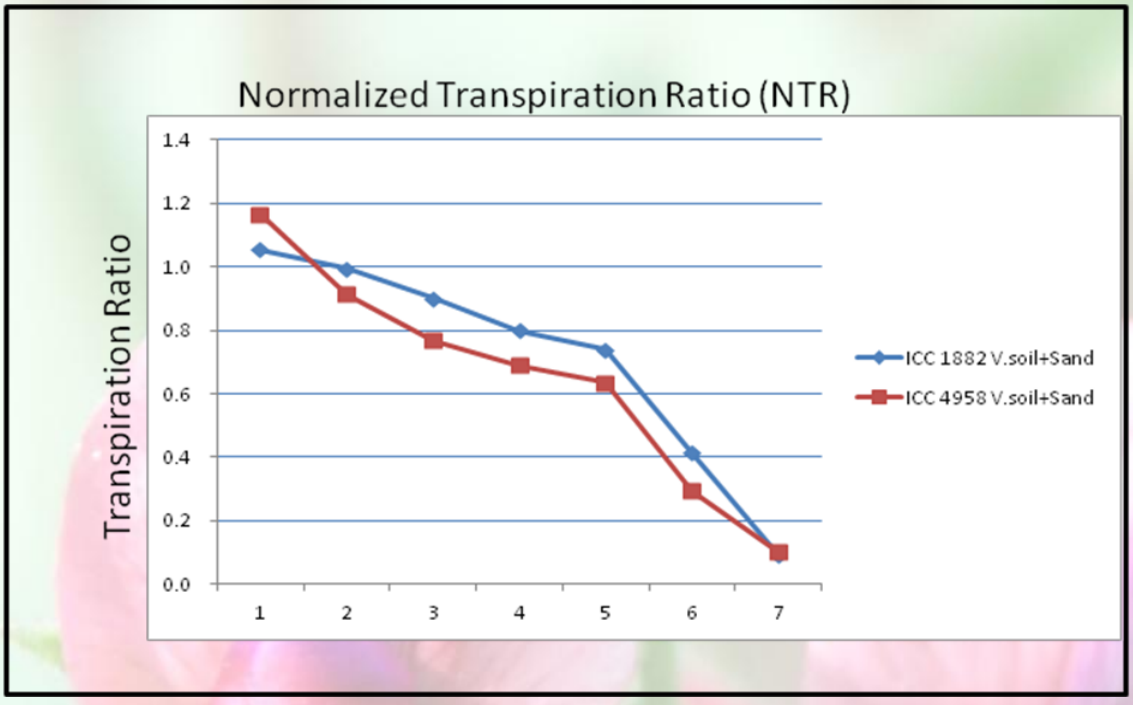


Figure 8(b) Normalized Transpiration ratio (NTR) of ICC 4958 and ICC 1882 during Dry down stress study under field conditions

3.3.4 Dry down experiment under greenhouse conditions

ICC 4958 and ICC 1882 were grown under greenhouse (day: night temperature: 32°C: 25°C, relative humidity: 40-80%). Seeds were sown in three replications for each genotype (Fig. 9 a,b and c).

Space arrangement of the pots and sowing: The plants were grown in pots of 18 cm height and 20 cm diameter filled with 4 kg of a vertisol collected from the ICRISAT farm under glasshouse conditions. The pots were set on benches and each row of pots is maintained at 30 cm of distance from another row and each plant is maintained at spacing of 10 cm. A total of 6 pots were prepared (3 for each genotype). Three seeds were sown per pot and 10–15 days after sowing, each pot was thinned to a single plant. Pots were kept well-watered for 6 weeks (42 DAS).

Measurement of Transpiration rate and harvesting of plants: Assessments of leaf transpiration rate under different VPD (Vapour pressure deficit) conditions were performed.

Figure 9(a) Comparison of stressed *vs* control plants of ICC 4958 and ICC 1882

Figure 9(b) Plants grown under slow drought stress treatment in glass house

Figure 9(c) Comparison of roots of plants of ICC4958 and ICC 1882

A measurement of leaf transpiration rate (g cm^{-2} h^{-1}) was done at 42 days after sowing (DAS) when the plants were at the late vegetative stage, by sequentially weighing potted plants at regular time intervals. Pots were saturated for 1 day before starting the experiment and allowed to drain overnight. They were bagged the following morning with a white plastic bag wrapped around the stem to avoid soil evaporation (Fig. 9a). Pots were subsequently weighed. The experimental design was a randomized complete block design with two water treatments (well-

watered and water stressed). Each morning, the pots were weighed. Three pots of each genotype were maintained in a well-watered condition by watering the soil daily to return the soil to ~80% of pot capacity. Three pots of each genotype were allowed to dry progressively over approximately a 2-week period. Water was added to the drying pots if needed so that there was only a maximum of 70 g net loss of water each day. The transpiration values were normalized as described previously (Kholova et al. 2010) to facilitate comparison. A transpiration ratio (TR) was obtained by dividing each individual transpiration value by the mean of the transpiration of the well-watered control, and this was done for each genotype. Then a normalized TR (NTR) was obtained by dividing each TR value by the average of the TR values obtained in the second, third and fourth days of the experiment, before plants were stressed (the first day of transpiration was quite erratic, probably because of recent pot saturation, and was not used). The experiment was terminated for each plant subjected to water deficit when the NTR was less than 0.1 (Fig. 9d), where blue line represents the drop in the transpiration ratio from 1.12 to 0.10 in ICC 1882, whereas the red line shows the drop in the transpiration ratio from 0.91 to 0.10 in ICC 4958. Root tissues were harvested at the same time and stored in -80°C till the RNA isolation.

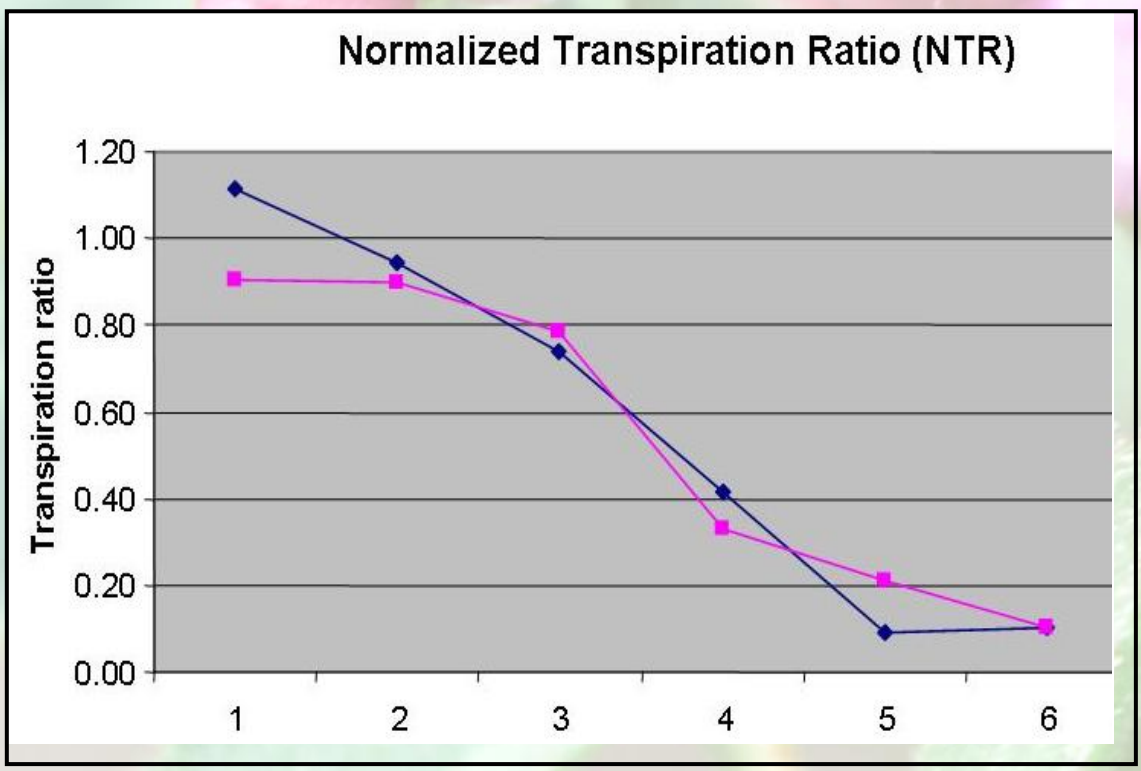


Figure 9 (d) Normalized Transpiration ratio (NTR) of ICC 4958 and ICC 1882 during Dry down stress study under glasshouse

3.4 RNA-Sequencing using Illumina/Solexa Technology

3.4.1 RNA isolation and sequencing

Stressed tissues as well as the well watered (WW) tissues were immediately dipped in liquid nitrogen after harvesting and then stored at -80°C. RNA was isolated from these tissues by using modified Hot Acid Phenol method described by Schmitt et al. 1990.The protocol of RNA isolation is given below:

Protocol for RNA isolation

- About 3g root tissue was taken and grinded to fine powder using liquid nitrogen in autoclaved pestle and mortar
- This powder was homogenized by adding 3ml of RNA extraction buffer (10 mM Tris, pH 8.0, 10 mM EDTA, 0.5% SDS) and 3ml of Acid phenol before thawing and vortex for 5 minutes in 50 ml centrifuge tubes
- After proper vortexing, of the sample, 3 ml of chloroform-isoamyl alcohol (CIA) (24:1) was added and again vortexed thoroughly for 1 minutes
- Samples were centrifuged at 5000rpm for 20 min at 4ºC
- Transfer the clear supernatant to a fresh tube and add equal volumes of CIA and vortexed thoroughly for 1 min
- Centrifuge at 5000 rpm for 20 min at 4ºC
- This step was repeated till clear layer of supernatant was obtained
- The clear upper aqueous phase was transferred into a fresh autovlaved 5ml centrifuge tube and one third volume of 8 molar lithium chloride (LiCl) was added and incubated overnight at 4^{o}C
- Next day the tubes were centrifuged at 5000 rpm for 20min at 4ºC to obtain RNA pellet
- Supernatant was discarded and pellet was re-suspended in 1ml of autoclaved 0.1% diethyl pyrocarbonate (DEPC) water
- To the re-suspended RNA, $1/3^{rd}$ of 8M LiCl was added and incubated overnight at 4ºC

- Next day the tubes were centrifuged at 12000 rpm for 20 min and the supernatant was discarded
- The RNA pellet was air dried for 15 min and dissolved in 20-30µl of DEPC water
- The concentration was checked on the equilibrated formaldehyde agarose gel

Formaldehyde agarose gel preparation (Sambrook et al. 1989)

- 1.2% of Formaldehyde agarose gel was prepared of the size 10 × 14 × 0.7 by mixing 0.6 gm of agarose and 5 ml of 10 × FA gel buffer and volume was made upto 50 ml with autoclaved DEPC water
- The solution was heated to melt the agarose and cooled to 65°C in water bath
- To this add 1.8 ml of 37% formaldehyde and 0.5 µl of a 10 mg/ml of Ethidium bromide (EtBr) solution
- The solution was mixed thoroughly and poured in gel casting tray
- Formaldehyde agarose gel was first equilibrated in 1 X FA gel running buffer for 30 min

The image taken by gel documentation system (G:Box, SYNGENE), is presented in Fig. 10.

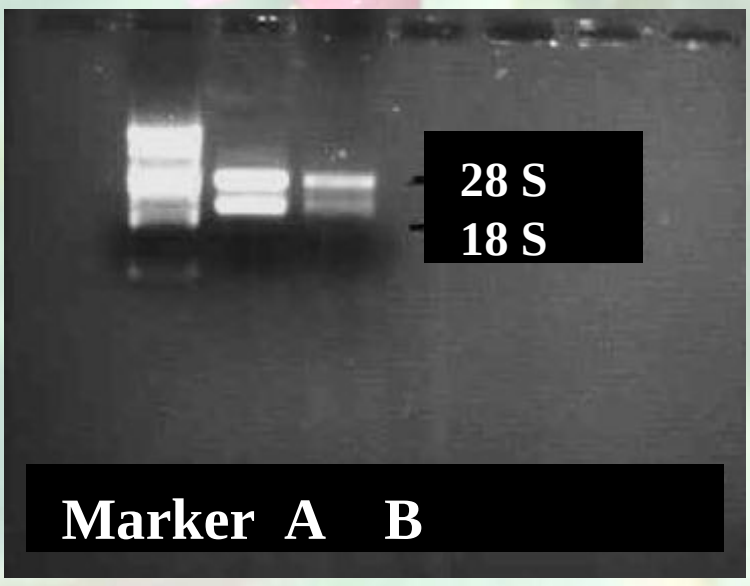


Figure 10 Formaldehyde agarose gel showing total RNA samples (A, B) along with RNA ladder

Sample preparation

- 1 µl of 5X RNA loading buffer was added to 4 µl of extracted RNA and mixed.

- The mixture was heated for 3 min at 65°C
- 5 µl of sample was loaded on the equilibrated FA gel.
- The gel was run at 55 V/1 hr in 1 X FA running buffer

Quantification of RNA

The Quantification of RNA was performed by UV spectrophotometer (UV-160A; Shimadzu). Absorbance was measured at 260 nm (A_{260}) and calculated as follows:

Concentration of RNA= 40 µg/µl × A_{260} × dilution factor

RNA sequencing and data processing

Total RNA isolated from all the stressed root samples from all the stages for each genotype pooled separately. Pooled RNA for both genotypes was used for Illumina/Solexa sequencing on Illumina's Genome Analyzer I at National Centre for Genome Resource (NCGR), Santa Fe (NM), USA.

Generated Illumina/Solexa sequence dataset was processed by using Alpheus pipeline developed at NCGR following standard criteria (Neil et al. 2008).

3.5 Identification of SNPs

3.5.1 SNP identification based on Illumina

Generated Illumina sequence reads for both genotypes (ICC 4958 and ICC 1882) were aligned to the chickpea transcriptome assembly developed by Hiremath et al. (2011). A range of criteria and thresholds were used for identification of SNPs between two genotypes as mentioned in Azam et al. 2012.

3.5.2 SNP mining from ESTs dataset

All available Sanger ESTs in public domain at the time of analyses were assembled by CAP3 and then scanned the multiple sequence alignments (MSA) of the contig to find sequence differences that correlate with the source of the EST. Subsequently, these contigs (CAP3 output) were run into a perl script based inhouse module, *Divest* (Jayashree et al. 2009). It detects the

redundancy in the sequence alignment files. Details about criteria used for SNP mining are given in the Results section.

3.5.3 Primer pairs used for allele-specific sequencing

For SNP discovery based on allele re-sequencing approach, ESTs/ genes were selected from a range of species as mentioned below:

970 chickpea ESTs/genes and 657 heterologous genes or transcription factors (identified in *Medicago truncatula, Medicago sativa, Lotus japonicus, Lupinus* spp., *Arachis hypogaea, Pisum sativum, Crotolaria tenuifolia, Phaseolus vulgaris, Phaseolus coccineus, Glycine max, Glycine soja, Robinia pseudoacacia* and *Trifolium pratense*) primer pairs were designed using PRIMER3 in such a way that the SNP position should come in between the expected amplicon size and were custom synthesized by MWG (MWG-Biotech AG, India).

(a) Based on Illumina sequencing, 691 primer pairs were designed using context sequence of chickpea transcriptome assembly to which reads of ICC 4958 and ICC 1882 aligned and variants were identified. In the second approach, on the basis of SNPs identified in Sanger ESTs, 279 number of primer pairs were designed from chickpea ESTs. In addition additional set of 217 primer pairs Buhariwalla et al. 2005 and unpublished work at ICRISAT, 12 primer pairs from Rajesh and Muehlbauer 2008 and 50 primer pairs from Singh et al. (2008) were designed. Details of these primer pairs are given in Annexure 1.

(b) For developing the cross species genic markers from close relative of chickpea like *Medicago truncatula, Medicago sativa, Lotus japonicus, Lupinus* spp., *Arachis hypogaea, Pisum sativum, Crotolaria tenuifolia, Phaseolus vulgaris, Phaseolus coccineus, Glycine max, Glycine soja, Robinia pseudoacacia* and *Trifolium pratense* were used. A total of 657 primers were developed based on transcription factor/gene sequences of chickpea-related species. ESTs from heterologous species were aligned with chickpea transcript assembly to identify SNPs. Highly conserved coding sequences

were used for the development of primers (Choi et al 2004 a,b; Nayak et al. 2010). Details of the primer sequences are given in Annexure 1.

3.5.4 Primer pairs used for CAPS assays

From a separate study (Varshney et al. 2009a), primer pairs were designed for multiple sequence alignment of 742 ESTs generated from ICC 4958, ICC 1882, ICCV 2 and JG 11, which could be converted into CAPS using SNP2CAPS programme. For 87 EST contigs containing SNPs that could be optimized for CAPS assay and primer pairs was designed using PRIMER3 and were custom synthesized by MWG (MWG-Biotech AG, India). Details of the primer sequences are given in Annexure 2.

3.5.5 Primer pairs used for ISR assays

121 intronic regions identified based on comparison of chickpea ESTs with *Medicago* genome. CISRs were designed based on comparative analysis of 4,558 unigenes generated in this study along with unigenes developed in a separate study (2, 595 from salinity stressed genotypes ICCV2 and JG11 and 7097 public domain ESTs). A total of 9,569 unigenes and *Medicago* genome sequences were used to design CISR primers using CISPRIMERTOOL software (www.icrisat.org/gt-bt/CISPTool.htm). A total of 121 primers pairs were developed for intron length polymorphism. The details of primer sequences in given in Annxure 3.

3.5.6 Amplification profile

PCR for chickpea and heterologous species gene/transcript factors markers and EST-SNPs was carried out in 20 µl reaction in a GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA) containing 5 ng template DNA, 0.2 mM dNTPs, 2mM $MgCl_2$, 2 pmol of forward and reverse primer, 1U of *Taq* DNA polymerase (Sib enzyme), and 1X PCR buffer. The amplification cycles were: initial denaturation of 5 min at 94°C followed by 10 cycles of denaturation for 15 sec at 94°C, touchdown from 61°C to 51°C with 1°C decrease in each cycle for 20 sec followed by extension at 72°C for 30 seconds. The next 40 cycles included denaturation at 94°C for 15 sec, annealing at 54°C for 30 sec and extension at 72°C for

30 sec followed by final extension of 30 min at 72°C and left at 4°C until further use. For testing the amplification, PCR products were separated on 1.2 % agarose gel electrophoresis.

For CISR markers and ICCeM (from separate study), PCR was carried out in a reaction volume of 5 µl in a GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA) containing 5 ng template DNA, 0.2 mM dNTPs, 10mM $MgCl_2$, 2 pmol of forward and reverse primer (M13 tailed), 1U of *Taq* DNA polymerase (Sib enzyme), 1X PCR buffer along with 2 pM (FAM, Vic, Ned and Pet) to enable detection of the fragments in the ABI-3700 automated sequencing system. The amplification cycles were: initial denaturation of 3 min at 94°C followed by 5 cycles of denaturation for 20 sec at 94°C, touchdown from 60°C to 56°C with 1°C decrease in each cycle for 20 sec followed by extension at 72°C for 30 secs. The next 40 cycles included denaturation at 94°C for 20 sec, annealing at 56°C for 20 sec and extension at 72°C for 30 sec followed by final extension of 20 min at 72°C and left at 4°C until further use. For testing the amplification, PCR products were separated on 1.2 % agarose gel electrophoresis.

3.5.7 Sequencing

For allele re-sequencing based SNP discovery, PCR products obtained by using 970 chickpea and 657 heterologous genes/ transcription factors were sequenced in both directions using Sanger sequencing methodology. Initially, PCR amplicons producing solitary bands were purified by treating with 1 U each of Exonuclease I (Exo) and shrimp alkaline phosphatase (SAP). Subsequently, the EXO/SAP treated PCR products were incubated at 37°C for 45 min followed by denaturation at 80°C for 15 min to deactivate remaining EXO activity. All the Exo/SAP treated products were sequenced from both ends using respective forward and reverse primers at Macrogen Inc., Seoul, South Korea (http://www.macrogen.com/).

3.5.8 SNP identification and diversity analysis from allele specific re-sequencing

Sequencing data were inspected manually for possible sequencing error. The forward and reverse sequences for the given gene/EST were merged to prepare the contig for each genotype by using DNA Baser v 2.9 software (http://dnabaser.com) under the following parameters: minimum match percentage, 80%: minimum of overlap, 25 bases; quality value of 25.

Subsequently, contigs for all genotypes were aligned using Clustal W (Thompson et al. 1994; http://www.ebi.ac.uk/Tools/clustalw2/index.html) and aligned files are saved with .aln extension. The .aln files for each gene/marker were subsequently opened in BioEdit version 7.0.5.3. SNP identification among different genotypes was undertaken manually after considering the base quality.

The multiple sequence alignment files for each gene/marker after manual confirmation of identified SNPs in BioEdit programme, were saved as FASTA file. Subsequently, FASTA files for different genes were put together in a single file. This file was used in SNP DIVersity ESTimator module (*divest.pm)* developed at ICRISAT (Jayashree et al. 2009) for calculating the polymorphism information content (PIC) value of individual SNPs as well as gene/marker, nucleotide diversity (π), number and PIC value of haplotypes for each gene/marker. The formula used to calculate above parameters are given below:

1. SNP/Indel frequency is calculated as:
   the total length of the sequence in base pairs / number of SNPs or indels.

2. The nucleotide diversity π is calculated using the formula:

$$\pi = k/aL$$

where k is the number of SNPs identified in an alignment of 'n' genotypes,

L is the number of basepairs and $\sum_{i-2}^{n} 1/(i-1)$

3. The PIC of SNP is calculated using the formula:

$$PIC = 1 - \sum_{i}^{k} p_i^2$$

where p_i is the frequency of the *i*th allele at a given SNP locus.

4. The haplotype diversity is calculated using the formula:

$$H = [n/(n-1)] \cdot \left[1 - \sum_{i}^{k} p_i^2\right]$$

where K is the number of haplotypes, p_i is the frequency of haplotypes and n is the total number of reads.

3.6 Genotyping Assays

Four genotyping assays were employed for generating marker genotyping data: (a) CAPS assays, (b) polyacrylamide gel electrophoresis using MDE (mutation detection enhancement) gel, (c) KASPar assay, and (d) VeraCode assay.

3.6.1 CAPS assays

All the predicted CAPS candidates were amplified using the same PCR conditions as mentioned above. Amplicons were then subjected to restriction digestion with the corresponding restriction enzyme followed by electrophoretic separation on agarose gel electrophoresis (3% agarose, 1X TBE buffer, 1 hr, 120V) and visualized by means of ethidium bromide staining (Varshney et al. 2007c).

3.6.2 Polyacrylamide gel electrophoresis using MDE gel

PCR products, generated by using ISR markers, were denatured and separated on single-strand confirmation polymorphism (SSCP) gels containing a mutation detection enhancement (MDE) solution. The (MDE) gel electrophoresis is done on 310 x 380 x 0.4 mm gels containing 24% v/v MDE monomer (Cambrex Bio-science Rockland, Rockland, ME, USA) for 16 h at 200 V at room temperature, before being silver stained as described by Tegelstrom 1992.

3.6.3 KASPar assay

To genotype some polymorphic markers in parents of intra-specific mapping population (ICC 4958 × ICC 1882), KBiosciences' SNP genotyping platform 'KASPar assay' was used. For developing the KASPar assay primer/markers, Illumina sequence information, allele re-sequence information, Sanger based sequences and tentative orthologous gene (TOG) based sequences were selected. A SNP variant was specified around a polymorphic locus in the selected sequences and the two alleles were separated by a forward slash. Sequence information is provided along with SNP ID, Forward sequence (50 bp), SNP and 50 bp continuing sequence (reverse sequence, but in the

same direction as the forward sequence), such that a stretch of 101 bp length sequence was considered. Marker genotyping assays were used at KBioscience company, UK.

3.6.4 VeraCode genotyping assay for BeadXpress reader

Illumina's BeadXpress array is based on VeraCFode technology. It targets specific SNPs in the genomic DNA. The genotyping application is based on sequence-specific extension and ligation of correctly hybridized query oligos, which are distinguished by their shared primer landing sites.

For designing the primers for VeraCode assays, ADT score was calculated for the targeted polymorphic marker. Based on high ADT score, sequence containing forward sequence (50 bp), SNP (reference allele/variant allele) and reverse sequence of 50 bp i.e. a stretch of 101 bp length sequence, source, sequence orientation was submitted to Illumina Inc. Subsequently, 48-plex SNP assays were developed at ICRISAT using BeadXpress Reader System (http://www.illumina.com).

DNA concentration was determined using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific). Ninety-six well plates were prepared with DNA at uniform concentration (50 ng/µL). The assay started with the activation of DNA through a chemical reaction with biotin. The biotinylated DNA was then purified from excess biotin. Predesigned assay oligonucleotides (oligos) were added and hybridized to DNA. The mixture was bound to streptavidin-conjugated paramagnetic particles (SA-PMPs). After the oligo hybridization, mis- and non-hybridized oligos were washed away, and allele-specific extension and ligation of the hybridized oligos was performed. The extended and ligated products form a synthetic template that is transferred to a PCR reaction and amplified. The strand containing the fluorescent signal in the PCR products was isolated and hybridized to the VeraCode beads via the address sequence. After the hybridization, the VeraCode beads were washed and scanned on the BeadXpress Reader. The BeadXpress Reader uses lasers to excite the Cy3 and Cy5 fluors of the single-stranded PCR products bound to the VeraCode beads. Light emissions from these fluors are then recorded in a data file. Fluorescence data were analyzed to derive genotyping results using Illumina's GenomeStudio software package.

3.7 Genetic Mapping and QTL Analysis

3.7.1 Generation of marker data

Genotyping data for 5 CAPS polymorphic markers, 3 CISR polymorphic markers, 9 ICCeM markers (from another study) were generated on 232 recombinant inbred lines (RILs) of intra-specific mapping population (ICC 4958 ×ICC 1882) derived from ICC 4958, a drought tolerant parent and ICC 1882, a drought sensitive parent. From KASPar assay parental screening, 56 polymorphic markers were identified on the parents of mapping populationICC 4958 × ICC 1882. These markers were genotyped through BeadXpress array in accordance with manufacturer's protocols (http://www.illumina.com).

3.7.2 Scoring and processing of marker data

Based on the amplicon sizes in the parents, data were scored for ICCeM (from separate study) and CISR primers (this study), where as for CAPS markers the difference in the pattern of restriction digestion shown by both the parents is scored. For SNP markers genotyping data delivered from BeadXpress array scored using Illumina's BeadStudio data analysis software. The allele of the female parent was always scored as 'A' irrespective of the size of the amplicon and the allele of the male parent was always scored as 'B' and the recombinant inbred lines having alleles from both the parents were designated as 'H' and missing data was scored as '-'.

Allelic scores in the mapping population (ICC 4958 × ICC 1882) were recorded as following:

'A' – allele of female parent (ICC 4958)

'B' – allele of male parent (ICC 1882)

'H '– heterozygous (presence of both parental alleles)

'-' – missing data (failed amplification)

To test the segregation of the mapped markers Chi-square analysis ($P<0.05$) was applied against the expected Mendelian segregation ratio of 1:1 for RILs

3.7.3 Construction of genetic map

All the genotyping data generated by the different genotyping assays given above on 232 RILS for intra-specific mapping population was used to construct the transcript map. Genotyping data was analyzed using the $\chi 2$ test to assess the goodness-of-fit to the expected 1: 1 segregation ratio for each marker. In addition to these, marker genotyping data for 235 marker loci were compiled from 21 DArT (Thudi et at. 2011), 78 BEC-SSR (Thudi et al. 2011), 52 SSR (Winter et al. 2000), 36 SSR (Nayak et al. 2010) and 48 SSR (Gaur et al. 2011; Sethy et al. 2007).

With an objective of developing the high-quality genetic map, JoinMap® 4 programme (Van Ooijen 2006; http://www.kyazma.nl) was used to construct high quality genetic map using LOD 3-10 and Kosambi mapfunction. Marker order was fixed for anchor marker of the framework map developed by Thudi et al. (Unpublished). The intermarker distances calculated from JoinMap® 4 programme were used to construct linkage map by using MAPCHART version 2.2 (Voorips, 2006).

3.7.4 Phenotyping data on the mapping population

Phenotyping data was collected by Crop Physiology Team on 10 different traits that are associated with drought. These data were recorded in two different seasons 2007 and 2008 in three replication by the Physiology Division at ICRISAT, Patancheru. For taking the root related observations, chickpea plants were grown in PVC cylinders as described earlier under drydown experiment in field conditions. These traits are given below:

1. Shoot dry weight (SDW),
2. Stem dry weight (StDW),
3. Leaf dry weight (LDW),
4. Root dry weight (RDW),
5. Rooting depth (RD),
6. Ratio between RDW and total dry weight (R/T),
7. Root length (RL),
8. Root length density (RLD),

9. Root surface area (RSA) and
10. Root volume (RV)

The methods employed for recording the observations of the root traits are explained in brief in the following sections:

*Root dry weight (RDW):*The root dry weight (RDW) was recorded after drying the root in a hot air oven at 80ºC for 72 h.

Rooting depth (RD): Rooting depth is measured after removing the soil particles from the root system under running water. The roots were straightened and depth to which the roots had been penetrated was measured.

Ratio of RDW and total dry weight (R/T): It is ratio between the RDW and the total dry weight. Total dry weight is the collective dry weights of shoot and root system of each plant.

Root length (RL): The soil from the roots is removed by rinsing the roots in running water. The root length was measured by using a digital image analysis system (WinRhizo, Regent Instruments INC., Quebec, Canada). The image of the root is taken and the total length is obtained by measuring the length of tap and branch roots of the chickpea root system.

Root length density (RLD): The root length density (RLD) was calculated by dividing the total root length per cylinder by the cylinder volume at the maximum rooting depth.

Root surface area (RSA): The root surface area was measured by using a digital image analysis system (WinRhizo, Regent Instruments INC., Quebec, Canada).

Root volume (RV): The root volume is the volume of the soil in the PVC cylinder, at the maximum rooting depth.

Shoot dry weight (SDW): Total shoot which include stem and leaves of each plant were dried in a hot air oven at 80ºC for 72 h. The weight of this dry matter was measured as SDW. SDW is the sum of StDW and LDW.

Stem dry weight (StDW): The weight of dried stem (without leaves) is measured after drying the stem in hot air oven at 80ºC for 72 h.

Leaf dry weight (LDW): The weight of dried leaves (without stem) is measured after drying the leaves in hot air oven at 80ºC for 72 h.

3.7.5 QTL analysis for drought related root traits

Both the genotyping and phenotyping data for root traits were analyzed for mapping QTLs by using the method composite interval mapping (CIM) proposed by Zeng (1994) in the Windows QTL Cartographer, version 2.5 (Wang et al. 2007). CIM analysis was performed using the Model 6, scanning the transcript map and estimating the likelihood of a QTL and its corresponding effects at every 2 cM, while using significant marker cofactors to adjust the phenotypic effects associated with other positions in the genetic map. The number of marker cofactors for the background control was set by forward–backward stepwise regression. A window size of 10 cM was used, and therefore cofactors within 10 cM on either side of the QTL test site were not included in the QTL model. Thresholds were determined by permutation tests (Churchill and Doerge, 1994; Doerge and Churchill, 1996) using 1,000 permutations and a significance level of 0.05. QTLs were determined significant if the corresponding likelihood ratio (LR) score was greater than 11.5 (equal to a LOD score of 3).The per cent phenotypic variance (PV) explained by a QTL was estimated at the highest probability peaks. QTL analysis was done on the mean data of year 2005 for ten traits and for year 2007 for eight traits. For declaring the major QTL, a PV (%) greater than 20, explained by the QTL was taken into consideration.

4. RESULTS

4.1 Generation of Illumina/Solexa Transcript Reads

Drought stress was imposed on ICC 4958 and ICC 1882 as described earlier in Materials and Methods section. The total RNA isolated from stressed root tissues of ICC 4958 and ICC 1882 was used for RNA-seq using Illumina's Genome Analyzer I at National Centre for Genome Resources (NCGR), Santa Fe (NM), USA.

In summary, 15.7×10^6 transcript reads for ICC 4958 and 22.1×10^6 reads from ICC 1882 were generated. These reads were short (36 basepair) and single ends. For further analysis of these reads, *Alpheus* pipeline (Neil et al. 2008) was used. As the reference genome of chickpea was not available, these Illumina/Solexa sequences were aligned to the chickpea trancript assembly (TA), containing 98,534 tentative unique sequences (TUSs). This transcript assembly was generated from 435,018 454/FLX sequencing reads from (pooled and normalized) RNA of 22 tissues at different developmental stages of ICC 4958 and 21,491 Sanger ESTs (Hiremath et al. 2011). Upon mapping of these Illumina/Solexa tags on TA, a total of 11.4×10^6 (73.09%) tags from ICC 4958 could be mapped onto 24,572 contigs, while in the case of ICC 1882, 13.9×10^6 (22.1%) tags could be mapped onto 25,896 contigs of TA respectively.

4.2 Large Scale Identification of SNPs

4.2.1 SNP mining based on Sanger EST dataset

At the time of analysis, a resource of 27,259 Sanger ESTs including 11,904 drought responsive ESTs, 8,258 salinity responsive ESTs and 7,097 public domain chickpea ESTs was available. CAP3 assembly was performed on this dataset and a total of 9,569 unigenes (2,431 contigs and 7,138 singletons) were identified. Analysis of 742 contigs that are comprised of 5 or more ESTs using *in silico* mining approach identified 21,405 SNPs.

4.2.2 SNP mining based on Illumina/Solexa reads

Though for SNP identification, the 18,559 consensus contigs were taken for consideration where both the genotypes tags were aligned. Details of the alignment of solexa reads with the chickpea transcriptome assembly is given in Table 3.

Table 3: Summary of Illumina/Solexa sequencing analysis

Parameter	Number of Illumina/Solexa Reads	
	ICC 1882	ICC 4958
Number of reads	2,20,90,582	1,56,60,209
Number of transcripts	25,896	24,572
Number of gene matches	25,896	24,572
Average read length	36	36
Average read quality	27	23

Illumina/Solexa sequences of ICC 4958 and ICC 1882 were aligned against chickpea transcript assembly using *Alpheus* pipeline as mentioned earlier (Fig. 11). Detailed analysis provided a total of 26,082 nucleotide variants (transitions, transversions and *InDels*) between two genotypes. SNPs were detected in 18,559 contigs out of which only 9,237 were of high confidence (>5 read-depth) from 6,908 contigs (Fig. 12). Of these 26,082 SNPs, a total of 2,405 *indel*, where as 11,386 transition and 12,291 transversion were recorded.

On the basis of allele frequency difference range (<0.1-1.0) and read depth (3- >500), 21,620 SNPs were classified as depicted in Table 4. Maximum number of SNPs (3,856) had frequency difference range 0.20-0.29 and read depth 11-100. An additional feature of SNP analysis based on Illumina/Solexa dataset is prediction of expression values for different contigs that have SNPs. In this context, data normalization for more precise quantification was done by considering per million reads for discerning the expression values, since the numbers of tags mapped were slightly higher. Data analysis showed that 10,333 contigs (genes) had fold difference of <2X, 4,020 contigs showed a fold difference between 2 to 5, while 538 contigs had fold difference varying from 6 to >10X between the genotypes, 2,920 contigs were identified to be exclusively expressed only in ICC 1882 genotypes and 4,031 contigs only in ICC 4958 (Fig. 13).

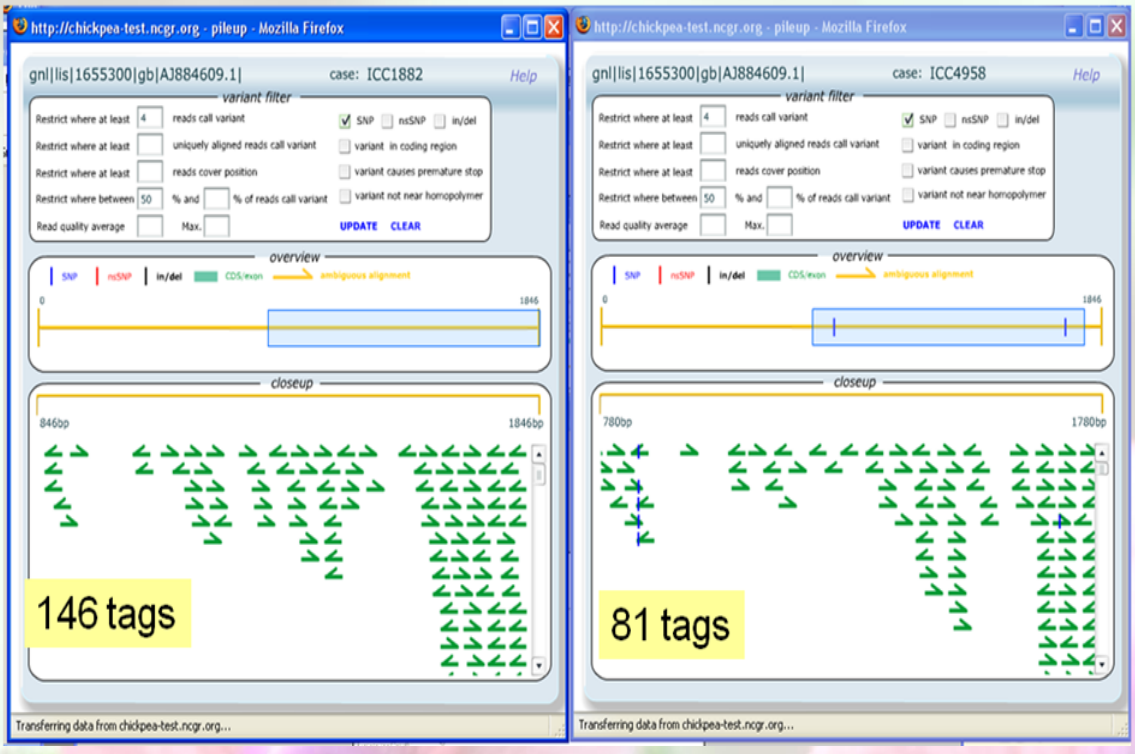


Figure 11 Methodology for analyzing Illumina/Solexa data set using *Alpheus* pipeline

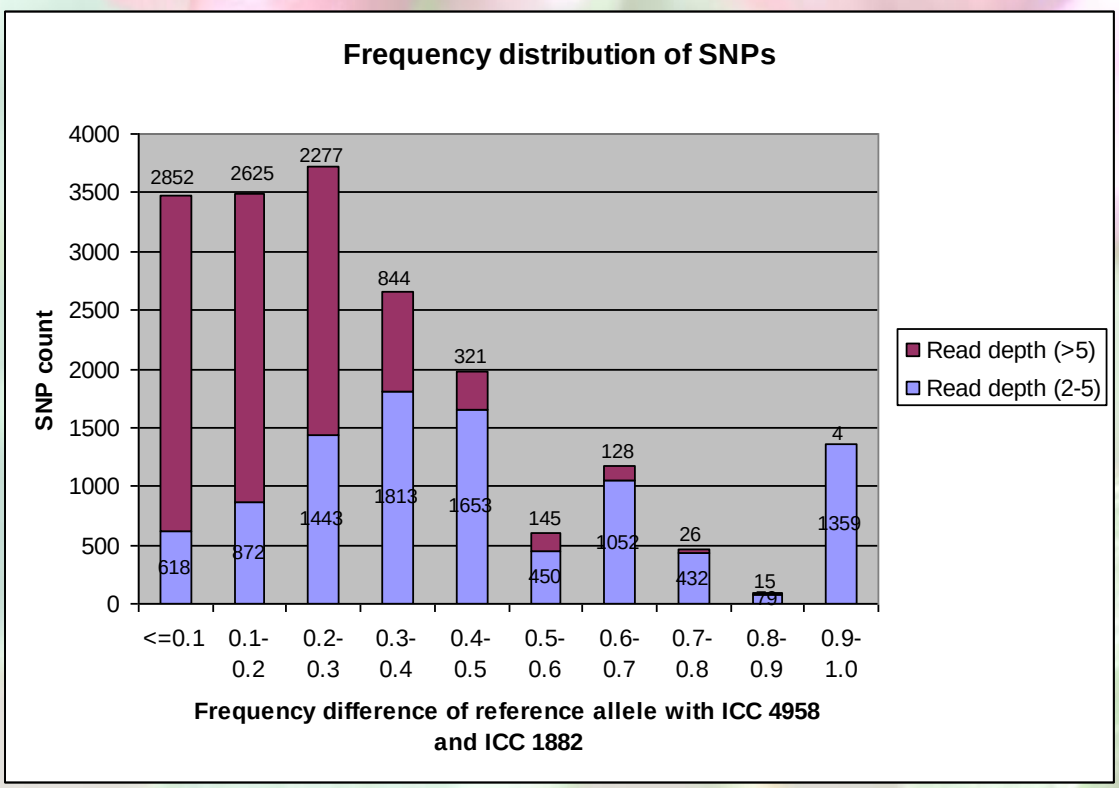


Figure 12 Distribution of SNPs in ICC 4958 and ICC 1882 on the basis of difference in frequency of reference allele

Table 4 : Number of SNPs classified based on frequency difference and read depth

Frequency different range	Number of reads/tentative contigs			
	>500	101-500	11-100	03-Oct
<0.1	389	751	2,109	158
0.10-0.19	107	414	2,431	500
0.20-0.29	17	123	3,856	827
0.30-0.39	4	47	1,478	992
0.40-0.49	1	13	746	828
0.50-0.59	8	18	502	1,442
0.60-0.69	-	17	297	1,361
0.70-0.79	-	1	85	374
0.80-0.89	-	-	55	166
0.90-1.0	-	-	40	1,463

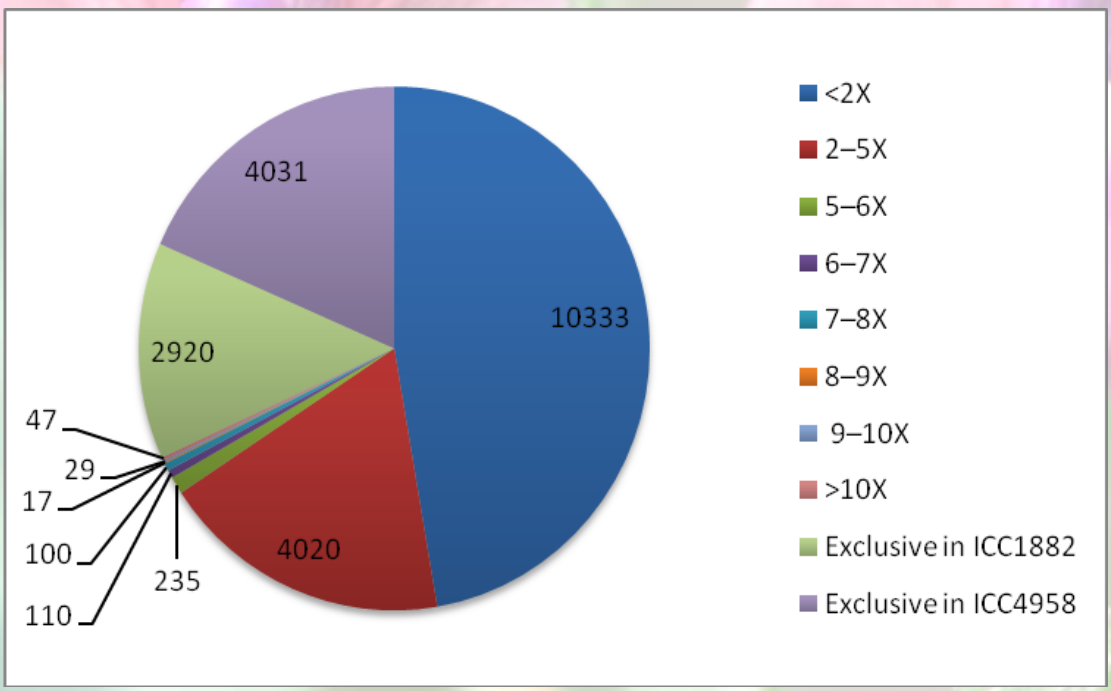


Figure 13: Expression of genes based on Illumina/Solexa sequencing

4.2.3 SNP identification based on allele specific re-sequencing

In the third approach of identification of SNPs, a total of 1,627 primer pairs were designed for genes or transcription factors (TFs) identified in chickpea or other legume species e.g. *Medicago truncatula*, *Medicago sativa*, *Lotus japonicus*, *Lupinus* spp., *Arachis hypogaea*, *Pisum sativum*,

Crotolaria tenuifolia, *Phaseolus vulgaris*, *Phaseolus coccineus*, *Glycine max*, *Glycine soja*, *Robinia pseudoacacia* and *Trifolium pratense*. Subsequently these primer pairs were used to amplify 2-20 genotypes and primer sequence information is given in Annexure 1. All successful 1,139 amplicons generated were used for sequencing in both directions. Sequence data obtained on all the genotypes for a given gene were aligned and compared to identify the SNPs.

4.2.3.1 *Use of chickpea ESTs/genes-derived primer pairs*

A total of 970 primer pairs were designed as following: (a) 691 for SNPs identified based on Illumina sequencing as mentioned above (Ca2C, Ca2S, CaESTs and Ca series), (b) 217 ESTs (AGLC series; Buhariwalla et al. 2005), and (c) 62 candidate genes (CaHa series; Singh et al. 2008; Rajesh and Muehlbauer 2008). Primer sequences for all above mentioned genes/ESTs are given in Annexure 1. After screening these primer pairs on two chickpea genotypes namely ICC 4958 and ICC 1882, only 738 primer pairs showed scorable amplification. Subsequently all 738 primer pairs were used on 2-19 genotypes. Sequencing of amplicons in both directions provided high quality sequences for 398 markers/genic regions on 2-19 genotypes. The genic sequences generated which were generated here were deposited in dbEST divion of GenBank (HO214215-HO214331)

4.2.3.2 *Use of heterologous genes-/transcription factors (TFs)-derived primer pairs*

A total of 657 primer pairs were designed based on transcription factor/gene sequences of chickpea related legume species. For instance, 479 primer pairs were designed for genes/transcription factors of *Medicago truncatula* and *Medicago sativa* (297), *Lotus japonicus* (144) and *Lupinus albus and L. luteus* (38). Some of these markers were taken from Nelson et al. (2006). Another set of 178 primer pairs were designed based on gene sequences of *Glycine max* and *Glycine soja* (97), *Trifolium pratense* (38) and other legume species (43) e.g. *Arachis hypogaea, Pisum sativum, Crotolaria tenuifolia, Phaseolus vulgaris, Phaseolus coccineus* and *Robinia pseudoacacia* that had significant similarity with Illumina sequence reads and SNPs between chickpea genotypes ICC 4958 and ICC 1882 identified in another study (unpublished results). Primer sequences for all these genes/markers are given in Annexure1. Screening of these 657 primer pairs on two genotypes showed scorable amplification with 401 primer pairs.

Subsequently sequencing of these amplicons provided high quality sequences for 134 markers/ genes on 3- 20 genotypes (Table 5).

In summary, a total of 1,139 markers showed scorable amplification on the panel of 5 genotypes and high quality sequences were generated for 532 markers on 2-20 genotypes (Table 5).

4.2.3.3 *SNPs and sequence diversity*

Analysis of the multiple sequence alignments for 532 markers developed through allele specific re-sequencing showed occurrence of SNPs only in case of 264 genes. Among these genes, 166 have been originated from chickpea, while 98 have come from heterologous species (34 from *Medicago truncatula*, 13 from *Lupinus* sp., 13 from *Lotus japonicus*, 19 from soybean and 19 genes from other legumes). DIVEST analysis of above mentioned sequence alignment has shown a total of 2,046 SNPs in 84,073 bp sequence data generated for 2- 20 genotypes for 264 genes. Among 2,046 SNPs identified, 964 accounted for transversion where as 1,167 for transition. Apart from this, 71 *InDels* were also identified. Partial multiple sequence alignment of a gene for SNP discovery is presented in Fig. 14.

As the genes surveyed for sequence diversity were originally derived from gene / transcription factor sequences and examined on cultivated and wild species, the sequence diversity was estimated as per origin of the gene/TF sequence as well as in terms of the cultivated vs wild species.

Although almost equal number of genes were derived from chickpea (970) and heterologous species (688), a slightly higher proportion of chickpea genes (17.11%) as compared to heterologous species (14.91%) showed SNPs. As shown in Table 6, 166 genes coming from chickpea showed 569 SNPs with a frequency of 1/93 bp while 98 genes coming from heterologous species provided 1,477 SNPs with a frequency of 1/21.09 bp. The nucleotide diversity index (π) for th+e chickpea genes ranged from 0.5×10^{-3} to 25.9×10^{-3} (mean = 5.3×10^{-3}), while genes derived from heterologous species showed a higher nucleotide diversity index of 1.0×10^{-3} to 74.0×10^{-3}(Mean= 14.6×10^{-3}) (Table 6).

Table 5 : Development, amplification and sequencing status of gene sequences based primers derived from chickpea and heterologous species

Marker Series	Marker type	Primer Id series/ source of genes	Markers designed	Markers amplified	Number of genotypes surveyed		High quality sequences
					Cultivated	Wild	
Chickpea	Illumina/ 454 chickpea sequences (691)	Ca2C, Ca2S	554	365	1-6	1-2	197
		CaESTs, Ca	137	115	2	1	45
		Total	691	480			242
	Chickpea ESTs (279)	AGLC	217	214	5-9	1-11	145
		Rajesh and Muehlbauer, 2008	12	7	2	1	5
		Singh et al. 2008	50	37	4	2	6
		Total	279	258			156
	Total Chickpea (970)		970	738			398
Heterologous species	*Medicago truncatula*		297	155	2-9	1-11	57
	Lupinus spp.		38	14	5-9	2-11	14
	Lotus spp.		144	54	8-9	6-11	15
	Glycine max		97	97	2	1	24
	Trifolium pratense		38	38	2	1	9
	Others		43	43	2	1	15
	Total Heterologous (657)	Total	657	401			134
Total		1,627	1,627	1139			532

Figure 14 Partial multiple sequence alignment (MSA) of a gene showing SNP position in 5 genotypes along with chromatogram; SNP position is depicted in box;(-)- is referred as *InDel* in MSA

Table 6: Sequence diversity analysis based on source of genes and species types

	Number of genotypes surveyed (Average)	Total length of sequence surveyed (Average in bp)	Number of genes showing SNPs	Number of SNPs identified (Average)	PIC value of individual SNP (Average)	Nucleotide diversity range (Average)	Number of haplotypes (Average)	Haplotype diversity range (Average)	PIC value of haplotypes (Average)
Source of genes									
Chickpea (970)	2-19 (6.01)	52,917 (318.78)	166	569 (3.42)	0.24-0.50 (0.37)	0.0005-0.0259 (0.0051)	1-9 (2.3)	0.2500-1.1670 (0.5303)	0.25-0.92 (0.45)
Heterologous species (657)	3-20 (9.68)	31,156 (317.92)	98	1,477 (15.07)	0.11-0.46 (0.34)	0.0010-0.0740 (0.0146)	1-10 (3.7)	0.2230-1.1670 (0.7704)	0.20-1.00 (0.61)
Within species type									
Cultivated (1-9)	264	84,073 (318.46)	85	197 (2.3178)	0.20-0.50 (0.41)	0.0000-0.0270 (0.0051)	1-3 (1.4)	0.0000-1.1670 (0.6809)	0.00-0.91 (0.46)
Wild species (0-11)	264	84,073 (318.46)	66	1,445 (21.894)	0.19-0.50 (0.36)	0.0000-0.0884 (0.0235)	1-10 (2.0)	0.0000-1.1670 (0.9759)	0.00-1.00 (0.76)
Across/ total									
264 genes across 2-20 genotypes	264	84,073 (318.46)	264	2,046 (7.750)	0.11-0.50 (0.34)	0.0000-0.0634 (0.0080)	1-10 (2.9)	0.2230-1.1670 (0.6201)	0.20-1.00 (0.50)

The PIC value ranged from 0.24- 0.50 (average 0.37) in chickpea genes, while 0.11- 0.46 (average 0.34) in heterologous species genes. As PIC values of the bi-allelic SNPs can not exceed than 0.50, sequence data for these candidate genes region were analyzed in terms of haplotypes as well. Number of haplotypes observed varied from one to nine with an average 0.45 (range 0.25 to 0.92) haplotype PIC value in chickpea genes, while heterologous species genes showed one to ten haplotypes with an average PIC of 0.61 (range 0.20 to 1.0). Haplotype diversity calculated was higher with heterologous species genes (0.7704) as compared to chickpea genes (0.5303).

It is interesting to note that at least 8 genic regions coming from heterologous species (LG80, LG99, LG101, LG103, LG104, LG111, LUP120, TC77515) and 1 candidate gene region (AGLC212) from chickpea showed > 50 SNPs across the genotypes examined. However, higher PIC values (>0.50) were observed with 3 chickpea derived candidate genes region (Ca2C21276, Ca2S126415_1648_0587, Ca2C3599).

While analyzing the sequence data as per cultivated species and wild species, a higher level of sequence diversity was observed in wild species as compared to cultivated species. On surveying 264 genes in cultivated and wild species, the number of SNPs present in wild species (1,445 SNPs) were more than 10 times higher as compared to cultivated species (197 SNPs). Similarly, nucleotide diversity was higher (43.6×10^{-3}) in case of wild species as compared to cultivated species (1.6×10^{-3}). In terms of PIC value of SNPs, cultivated species depicted average PIC value 0.41 (ranging from 0.20-0.50) and wild species showed average PIC value 0.36 (ranging from 0.19-0.50).

As the number of SNPs are much higher in case of wild species, the number of haplotypes ranged from 1-10 (average 2.0) in wild species in contrast to 1-3 (average 1.4) in cultivated species. Haplotype diversity was also observed high in case of wild species (0.9759) as compared to cultivated species (0.6809). Similarly, a higher PIC value of haplotype was observed in wild species (average 0.76) as compared to cultivated species (0.46).

In total 2,046 SNPs, with an average SNP frequency of 1 SNP per 41.09 bp were observed on surveying 84,073 bp sequence data across all 264 candidate genes region. PIC values of SNP ranged from 0.11-0.50 with an average of 0.34 across all the 220 genic regions. Number of haplotypes across 264 genic regions ranged from 1-10 with PIC value ranging from 0.20-1.00 (avg. 0.50). On an average 2.9 haplotypes (ranged from 1-10) were present in 264 genic regions with an average haplotype diversity of 0.6201 (ranged from 0.2230-1.1670).

4.3 Optimization of Marker Assays for SNP Genotyping

4.3.1 CAPS assays

An attempt was made to develop CAPS assays for SNPs identified through allele specific re-sequencing and *in silico* mining of database,were converted into CAPS. Out of 264 SNP marker, a total of 224 (85%) showed the presence of a putative restriction site which were further considered for CAPS assay. Finally a total of 311 candidate genes (224 from allele specific re-sequencing and 87 from *in silico* mining of database) were selected for wet lab verification on five chickpea genotypes (ICC 4958, ICC 1882, ICC 283, ICC 8261, PI 489777). Details about these 311 CAPS candidates are given in Table 7. While scorable amplification was observed in 182 (58.52%) cases out of 311 CAPS candidate, CAPS assays were observed succeeded in 152 cases (83.51%) out of 182 scorable amplified CAPS candidate. These CAPS markers have been referred as Chickpea Genic Molecular Marker (CGMM). While details of the validated 152 CGMMs are presented in Table 8, a representative CAPS profile for 7 CGMMs on a panel of 5 chickpea genotypes has been shown in Fig. 15. Validated CAPS include 124 (55.35%) out of 224 candidates identified based on allele re-sequencing data and 28 (32.18%) out of 87 candidates identified through mining of ESTs. This clearly indicates that allele re-sequencing approach is more effective than EST mining for conversion of SNPs into CAPS.

Table 7: Details on identification of CAPS candidates

S. no.	Primer ID	Marker name*	Forward primer	Reverse primer	Expected product size (bp)	Restriction enzyme for assay	Source of gene/ markers*
1	AGLC100	CGMM001	TCCAGGTGGAGGAGTCAGAT	TCAAACGTCTTCCACCTTCA	300	*Bfa*I	A
2	AGLC111	CGMM002	ACTAGTCCTGCAGGTTTAAACGA	CCTCTTCCCTCAATTTTCCTCACA	400	*Hpy*CH4IV	A
3	AGLC136	CGMM003	TCGATCGCAGTTTGAATCAC	AGGACAGAGCCACGAAGAAG	500	*Hph*I	A
4	AGLC163	CGMM004	AGTCCTGCAGGTTTAAACGAAT	ATAAGGGGCAAACTCCATGA	160	*Dde*I	A
5	AGLC168	CGMM005	TGAGAGGCTTAGGGATGAGC	CCTCCCCTTCATTATCACCA	580	*Bfa*I	A
6	AGLC178	CGMM006	GCCACGGTTGGTCTGTTCGA	CCGTTAAGGTTGCCGGACGA	300	*Tsp*509I	A
7	AGLC179	CGMM007	TCAAAATCCTGGTGGAGGTC	TGCCACTGCTGGTAAAGAGA	200	*Ava*I	A
8	AGLC196	CGMM008	TGATAAATTGCAGGTGGAGAGAGA	TAAGTTAGACTCCCAGGCAAGGTA	200	*Ban*II	A
9	AGLC198	CGMM009	GTGAGGGCGTAGCTTCACAT	CGAGCAGTTGAGGACCAAAT	500	*Hin*fI	A
10	AGLC202	CGMM010	TCAGAATCCCCATTAGTGCAG	GCTTTGGGATAGGATTTCCAG	650	Xba*I*	A
11	AGLC28	CGMM011	TGCACCATGTTGGAGAAAGA	CGTCCTGGATTCATCGAAGT	350	*Aci*I	A
12	AGLC3	CGMM012	GTCCAGTTCGCCCAATTCTA	ATGGAAGGCCGTGTCAATAA	500	*Apo*I	A
13	AGLC52	CGMM013	CTTTACCAAAACCACCTTCACCA	CAGGTCGCGTTGTTGCA	350	*Dde*I	A
14	AGLC74	CGMM014	CAGGTCCGCGTTGTTGCAA	GGAAGAGTGAGATTGTTGCGTGA	390	*Mbo*I	A
15	AGLC77	CGMM015	GCAGGTCGCGTTGTTAGCA	ATTACTATGCTTCCTTCTCCTCCA	690	*Hpy*CH4III	A
16	AGLC17	CGMM016	AGTGAGTTGGTTCGGAAACG	AACATGCGCTCAAGTTCAGA	260	*Tsp*509I	A
17	AGLC8	CGMM017	TGTTGTCTCGCCAATTCAAAGCA	CGTTTGGTGGCATTCCTGCA	480	*Hpy*188I	A
18	AGLC82	CGMM018	CCGAGGTCTTGCCATTGGTA	CAGATTCGTTATTGCCTTCCCGTA	250	*Hin*dIII	A
19	Ah6928256	CGMM019	CGCAAGAAGCAGTCAGATGT	GTAACACCCTGGTTGGTTGG	800	*Hin*cII	A
20	Ca10273CD20F12	CGMM020	TGAGCAGAAATTGCAACAGG	GCGCTTTCCAAGAACTATGC	570	*Hin*cII	A
21	Ca10556169746056	CGMM021	AACGTCAACTGGAAAAACGG	GGTATCGGAGGGTTTGGAAT	200	*Tsp*509I	A
22	Ca12411169747340	CGMM022	ACTGCCAATCCTACCCTCCT	ACAGCTGGAGCATATTGCCT	950	*Aci*I	A
23	Ca14402150174189	CGMM023	CTTGTGTTCCTTTGGTGGCT	GGAGCTTGATGCTGAGGTTC	350	*Hpy*CH4IV	A

24	Ca1448960219076	CGMM024	GAGGTTCCCGATCCTTCTTC	AGCCCTCATCAAACCATCAG	650	*Mse*I	A
25	Ca15775169744160	CGMM025	ACCTCAACCACCCTCCTTCT	CTCTTCCCAGTCGCTGAATC	200	*Aci*I	A
26	Ca1773847832705	CGMM026	ATGGGTGCAGGTTTCATAGC	TGAAGCAGAGCACTTTGTGG	850	*Hpy*CH4III	A
27	Ca21249	CGMM027	CAACCGTTGAAACCTTCGAT	GTCTCCTCCGTTTTCCCTTC	220	*Aci*I	A
28	Ca21567	CGMM028	GGATTCCACCGACAAGAGAA	CTGATTCACGCTCAGGAACA	120	*Tsp*509I	A
29	Ca9351169747446	CGMM029	ATCGCCAAAATGACCAAAAC	CCGAAAGCAATCTTCACCAT	150	*Hpy*188I	A
30	Ca93773928149	CGMM030	TTGTCATGGTGGTTGCATCT	GACATCGCCTTCTCAAGCTC	850	*Hpy*CH4IV	A
31	CaESTCg3	CGMM031	GTGGACTGGCTCACTGACAA	AGGCAACTCCTCTACGGTCA	200	*Mbo*II	A
32	CaESTSn16	CGMM032	GGCCTAAACTTGGAACATGG	CACCAAACAAACAAGCTGGA	690	*Hpy*CH4III	A
33	CaESTSn30	CGMM033	TGTGAATACATGACGTGAACCT	AGTGCTGGGATTGAGATGGT	190	*Hpy*CH4IV	A
34	Ct6874464	CGMM034	TGATGCATACATAAATGGACATGA	GGTTTTGGTGATTTGCACAG	280	*Hin*cII	A
35	Ct6875390	CGMM035	TCCATTCCATGTTCTCAGGA	GGAACAATCCTGTTCTCCGA	450	*Hpy*CH4IV	A
36	Ct6875951	CGMM036	GTGGTGCAGTGGTGTTATGG	CGAAGGCATGAGTTGAGGAT	300	*Nla*III	A
37	ge13bg	ge13bg	GTGCTGGTTTCCATTTAACG	CGCCATTAGAGAAAGATGGA	590	*Bcc*I	A
38	Gm103068	CGMM037	GAAACCTGATCCAATGGTGG	AAGCCAGTGGGAATTTCCTT	400	*Rsa*I	A
39	Gm2091985	CGMM038	AAGCAAAGGATGTTGTTAAAGG	CCACAGTTCTTTATGATTGTCTCC	550	*Nla*III	A
40	Gm2113547	CGMM039	TTGTGAAACGAGGATGGGAT	CATTCCAGCATAGCGGATTT	280	*Hin*cII	A
41	Gm2117188	CGMM040	GTGGTATGAGCACCTTGGCT	ACTTTTTCGCCAAGAACACG	200	*Nla*III	A
42	Gm2118543	CGMM041	TGGGCAGTATGGGGAGTTAG	GAAAGCCGAACATTTTGTCC	280	*Hha*I	A
43	Gm2120918	CGMM042	CTCAATCTGGTGGTGTTAGGC	TACTGCATCATCAAGCTCCA	500	*Tsp*509I	A
44	Gm2124498	CGMM043	TGATGAACCTGATGCCAAAA	TCCTTGGATTTGGATTTCCC	300	*Ava*II	A
45	Gm2125123	CGMM044	AAATGCGATCTGCCTGAGTT	CAAAGGTTTTCTCGGTGCTA	590	*Hpy*188I	A
46	Gm2129282	CGMM045	GTACCCGTTATGGTGCCAGT	CTCAACCTCCTGATGGTGCT	300	*Aci*I	A
47	Gm2132124	CGMM046	GGGTTGGATTGGAGAGTCTG	GGCAACAATGTCCAACTGTG	420	*Hph*I	A
48	Ms6943512	CGMM047	AGATGGACAAATCACCGAGC	TAAACACCGCTCTCATCCAA	240	*Hpy*CH4IV	A
49	Mt123479	CGMM048	AGAGGGGATACGGCAAAGTT	GTTTCCAGCATCACCACCTT	290	*Hpy*188I	A
50	Mt127721	CGMM049	TCGTAACTTCGGTGTTGGTG	TCAGGAACAATCTTCCCAGC	420	*Hin*fI	A

51	Mt133126	CGMM050	GCCGAGATTGGTGAGTTTGT	TGGAGGAACACTTGACAACA	190	*Hae*III	A
52	Mt6799803	CGMM051	AAGGCAAGCCTAAGCTGTTG	TGATCCATTTGGCATTAGCA	520	*MboI*	A
53	Mt6803180	CGMM052	GGTCATGAAAATTGATGGTGAA	TTTGGAAGCTCCACAAGAAAA	220	*Hae*III	A
54	Mt6817377	CGMM053	TGACGTGAGGTGTGGTGTTT	TGTATGACCTCTCCCTTGGC	500	*Hpy*188I	A
55	Ps1768451	CGMM054	TAAAGGAGAGAAAATGTACATCCA	TACTCGATCCAATCGATGCC	280	*Mbo*I	A
56	Ps1770858	CGMM055	CAATCCCCATAATTGGCAAC	CGTAGGTTACGCCAATGGTC	290	*Hin*fI	A
57	Pv4540	CGMM056	TTGGAGGACTTCCTCAATGG	ATGAACAAAGCAGCAGAGGG	600	*Ava*I	B
58	rgr4	rgr4	GCCAAACGGAGATTAGATGA	ATTCCCCTATTTGGGTTCAC	290	*Tsp*509I	B
59	SHMT	SHMT	TGAAATTGAATGAGGGGAAA	TGGCTCCTCCACATAGTCAT	500	*Hpy*CH4IV	B
60	tk	tk	TTTGGCCTGAAGTGAAAAAG	GTGTTTGGTTCTACGTTGGG	700	*Nde*I	B
61	Tp6849647	CGMM057	CTTGCGATCCCTGCTTCTAC	GACAGGGACAGGGACAGAGA	200	*Eco*RV	A
62	Tp6849720	CGMM058	CTGGGGATGGTGCTTACCTA	AACGGTTGATCAAATGGAGC	800	*Mbo*II	A
63	Tp6850763	CGMM059	CATCACGAGCTTGAAACTGC	CCATTCCAACAATATGGGCT	500	*Dde*I	A
64	Tp6854083	CGMM060	GCAAGGAAAGTTGCAAAAGC	TTTTCTCCTCTGGCTGTGGT	720	*Hpy*CH4V	A
65	Tp6857294	CGMM061	GCTGATATGGCACATGCACT	AGTTCGAATACCAACGCCAC	500	*Hin*fI	A
66	Ca2S124718_0933_2144	CGMM062	GTCGTGGCCGTTAAATTATG	TACCAATGATTTGCGATTCC	290	*Aci*I	A
67	Ca2C8663	CGMM063	TCAAACCCTCCCATCACTAA	TCACCTGGGTTGTTGACTTT	270	*Mbo*II	A
68	Ca2C41871	CGMM064	CACATGCATTCAACCAAAAA	TGGATAAGGGACCAAGAACA	150	*Nla*III	A
69	Ca2S290507_2844_3854	CGMM065	GCAAAGCGTGAATCAATTTT	CCTGAACCGTGAATCAAATC	390	*Aci*I	A
70	Ca2C43617	CGMM066	GATGGTTATCTTCAGCGTGG	AAACCTCCACAATCCTGACA	300	*Bfa*I	A
71	Ca2S125676_2925_1604	CGMM067	TGTCGCTAATCCTCTTGGAC	GGAAGGAGTGATTGTGATGG	220	*Hpy*CH4IV	A
72	Ca2C33173	CGMM068	TGGCAACTACCGTAACCATT	ATTGACTTCACTGTCAGGGC	180	*Aci*I	A
73	Ca2S232393_3220_2305	CGMM069	AGTGACTTGGTCAAATGGGA	AATATCAGCATGGTTTCCGA	400	*Rsa*I	A
74	Ca2S143601_3004_0194	CGMM070	GCCAAGTGGAAAAGAAAGAA	ATCCCATTGAAGAACAGCAG	520	*Eco*RI	A
75	Ca2S295752_0698_0195	CGMM071	AAGTTGCCAATTCCAGACAG	TTTATCGGGTTTGTGATGCT	190	*Eco*RI	A
76	Ca2S126415_1648_0587	CGMM072	TGCATCCATAGATACAACAAACA	CAAGAAGCATTTTGTCCTTACA	350	*Tsp*509I	A
77	AGLC212	CGMM073	ATGTTTGAAGAAGGATGCCA	ATCATCATCCCACTCGTCAT	781	*Bcl*I	A

78	AGLC213	CGMM074	AAGTTGCCAATTCCAGACAG	TTTATCGGGTTTGTGATGCT	580	*Bsa*MI	A
79	AGLC216	CGMM075	AGAGTAATAGTGGGGGTGGC	GTATGAAAGGCCAGCAGCTA	469	*Nla*III	A
80	AGLC217	CGMM076	GGAAGCAGCCATCTAAGGAT	CCAAACCTCAAATAGGGGTC	189	*Msp*I	A
81	AGLC214	CGMM077	AGAAAATGTCCAAGCACAGG	ATTGGTTTTGGTGGTTGTTG	242	*Xmn*I	A
82	AGLC122	CGMM078	GCCACGTTAATACAACAGGG	CGGGGGAATGTCTTTTAACT	605	*Hpy*8I	A
83	LG73	CGMM079	CACAACCCCTGTTGAGCCCTG	AATCACAGTGTTCATGTCTTGACACG	409	*Aci*I	A
84	LG83	CGMM080	GGAGAACTGGCTCGGTATGCTGC	TGGGAATGTTGTGATGCTTCAACC	287	*Hpy*CH4V	A
85	LG87	CGMM081	CCAATGGCTGAGGAATCTAGCAC	AACAACTGATTCAGGTGCAGGGAG	562	*Bsu*RI	A
86	LG90	CGMM082	TTGGAATGCCTTCTCTCTCACCTAC	CCTTCCTCCATACATTCCATTTCAAC	333	*Nla*III	A
87	LG91	CGMM083	GTTGAAATGGAATGTATGGAGGAAGG	GACAATGAGCCCTTCAGGTCCTG	699	*Alw*26I	A
88	LG95	CGMM084	CATGGGCAAATGATTATATTCGCC	CCTCTTCCCCTATTGTAGAGAGGCC	339	*Bcl*I	A
89	LG99	CGMM085	CAGCATCAATAGGAACATGGTTCTC	CAACTGTTGTCCGTGTTTTGGC	612	*Fin*I	A
90	LG101	CGMM086	GTGAGTCATACCAGCCTTGTAACCC	GAAGCGTGCTGCTCGTCACAG	250	*Bsa*BI	A
91	LG105	CGMM087	TCCACGGTCGCATGAGGCC	CGAAATACAGCCTCTTGCACAGG	338	*Ava*I	A
92	LG111	CGMM088	GAGCATGAAATTCTACCTCCTCTTCC	TAATTCTCTTTGGGGGAGAAGGC	661	*Acc*B1I	A
93	LUP51	CGMM089	TAGGTCTAGATACTGGTGTCCC	GGTTTGTAGTTCTTGTTTGGAG	500	*Taq*I	A
94	LUP94	CGMM090	ACTACCATTGAGGAAGTCAGAG	GATGTCGTTGAAGTTTGAGTAG	229	*Nmu*CI	A
95	LUP120	CGMM091	GTTTGACTCCATCACAGATTG	CTGCTCTCAACAAGAATCAAC	549	*Acc*16I	A
96	LUP235	CGMM092	AGAGAACATGAAGCCATGAA	CCATCTTCAGGAATAGCAAG	319	*Fin*I	A
97	LUP240	CGMM093	GTCTGTTCATTCCTTCAGAATC	CACCATACTTCACTCCTTCAC	434	*Bst*UI	A
98	LUP255	CGMM094	GTGTCTACACCGATGTTAATGT	TTCTTGACAGGCTTGAGAAT	771	*Xba*I	A
99	LUP302	CGMM095	TTAGTTGGGAATACAGCACC	GGGACAAACAAATGTGAAGT	261	*Nla*III	A
100	LUP318	CGMM096	CCTAGTGATTCTGGGATTACAT	GTTGCTTCAATGTCACCTGTAT	270	*Hin*fI	A
101	LUP326	CGMM097	CTCACAGGAAATTCGAACAC	CTCTCTAACAATGTGAGTCATACC	544	*Acc*I	A
102	TC76606	CGMM098	AAATGTCGGACGAGGAACAC	ACCATGCTTCCCAGTTTTTG	378	*Mbo*I	A
103	TC77515	CGMM099	GACAACTGTGCAGGGATTGA	CAGCTGCTCAGAAACCATGA	740	*Bse*JI	A
104	TC77624	CGMM100	CACCACAAGAAATGAAAAGGAA	GGCTCAAGGAAAACGTTGAT	170	*Alu*I	A

105	TC77707	CGMM101	TTCTGTTCTTCCACCCCAAC	CGAGCAAGAATCGAACACAA	372	*Hin*P1I	A
106	TC81224	CGMM102	CCGGGGAAGTTGTAGCATTA	GCCAAGCCAAAATCACAAAT	404	*Mnl*I	A
107	TC85414	CGMM103	GTCAGGGTGTGCTTTTGGTT	ATCACTTGGGTCAAGATCGG	221	*Bsu*RI	A
108	TC86258	CGMM104	GCACAGGACAATTGCCTACA	GCATAGGAAGGGGAATCACA	185	*Fnu*4HI	A
109	TC87719	CGMM105	GGTTGCAGTTAAACGCCATT	CATCGTTCGGAAACAACCTT	398	*Hin*fI	A
110	TC92821	CGMM106	TTATTGCTGGGCCTTTCAAC	GAGCAAACATCCAACAGCAA	166	*Bsa*JI	A
111	TC94373	CGMM107	TGGGTTGATTGGGATGATTT	GGTCTGTGCTGCAAGATTCA	274	*Eco*RI	A
112	TC96130	CGMM108	TGAACTCCCTCCTGGCTTTA	TGCCTTGTCTGTGCCAGTAG	596	*Bse*3DI	A
113	TC103928	CGMM109	GCGGACGATGGTAAAAGTGT	CATCAGCTTCCTCTTCCTCG	275	*Mse*I	A
114	Ca2C10102	CGMM138	AAAGGTTAACTGCTTGGCCT	AGCCAACTCACCAACATCAT	150	*MseI*	A
115	Ca157594586605	-	TTGGCTCTTTCCATTCCAAC	CCAAAGCCTGGGAAAATGTA	319	*Tsp*DTI	A
116	Ca20156	-	ACCCAATCGTGATTGTCCAT	CCCACGGTGACTACTGTCCT	221	*Hin*cII	A
117	Ca21249	CGMM027	CAACCGTTGAAACCTTCGAT	GTCTCCTCCGTTTTCCCTTC	149	*Tse*I	A
118	Ca22434	-	GCCAGAGCATGCTTCACATA	CTTGGTGGCTTGCTTTCTTC	175	*Bfa*I	A
119	Ca12863169747399	-	CATGCAAAACAATTCCATCG	CTTCTGAAACATAAGCCGCC	736	*Ear*I	A
120	Ca1683047832555	-	TCGAGTTCACAGTCTCCGTG	TCACCAAATGCACCATCAGT	226	*SspI*	A
121	Ca10880169748548	-	TTTTTGCCCAAGACAAGACA	TTCTCTTTTTGGTGGCTGCT	101	*Mse*I	A
122	Ca13382169746951	-	GGAGCGAGCAGAATAAGGTG	TGCATAGCATATGGGGATGA	219	*Hpy*CH4V	A
123	Ca15994169746275	-	ATGGAATGAGCAAGGGTGAC	TCACCTCCGCAACATGATTA	195	*Hpy*AV	A
124	Ca2C42668	-	AGGCACTTGTCATCCCTGTA	CATCTGAGACATCTTCCGCT	97	*NIa*III	A
125	Ca2S117228_1922_1005	-	TTTGGTAGCAGAAACCTTGC	GCAGTAAGATGTTGGAAACACA	189	*Mbo*II	A
126	Ca2S110494_3324_2020	-	TCTGACATCAATGAACAAGCA	TTTCCGTTTGTAGTGGTGGT	166	*Mbo*I	A
127	Ca2S242431_3223_1815	-	AGAGCACATGACCCACTCAT	GAACAAAACGTGACAAGGCT	65	*Eco*RV	A
128	Ca2C3599	-	TCATAACCTTAGCGCATTCC	ATGACTGGACGAAACCAAGA	215	*Nla*III	A
129	Ca2C42103	-	AAAATGGGTTGGAAAGAAGG	GATTGGGTCTGTGACGGTAG	166	*NIa*III	A
130	Ca2C4224	-	GCACACCACGCATGTATAAG	TTGAATGTTTGGTGCCTTCT	139	*Mbo*II	A
131	Ca2C6533	-	TAGCACCAAATTCCCCATTA	TTGTGTTGACAATTGGTTCG	196	*Mse*I	A

132	Ca2C28092	-	AAAGGTACCTCCACCCAAAG	ACATAATTTTCCCCCTGCTC	159	*Tsp*DTI	A
133	Ca2S289800_2555_0786	-	AGAGGGAGGGGATAGGAGTT	GAAACCAGTTCCTCACTGGA	193	*Hpy*CH4IV	A
134	Ca2SFE670434	-	AAAATAGCCATCCTTGTAGTTCA	CAATACCCTGCATCCTTGTT	296	*Aci*I	A
135	Ca2C12526	-	TTGAGGAGGAGAAACACAGC	GACAGGCTTCCAGAAACTGA	153	*Aci*I	A
136	Ca2C25794	CGMM139	AGAGGCGGACAAGAGATTGA	TGCAAGACGCTCATAACCAG	288	*Bfa*I	A
137	Ca2C32639	-	TGCCACTATGGCACTCTCAG	GAGCTCCGCAGTGTGGTAAT	263	*Mly*I	A
138	Ca2C33338	CGMM140	TTTCTCCGGAAACACTACCG	CATGTCAGTGGAATTGCTGC	372	*Aci*I	A
139	Ca2C3559	CGMM141	ATGTTTGGTGGACTTGGAGC	TACCCAACATGGCACTTGAA	285	*Aci*I	A
140	Ca2C38039	-	CGGTAGGAACAGGGAAAGTG	GCCCTGAAGCTGTTGATCTC	153	*Mbo*II	A
141	Ca2C6285	-	GGTCCCACATTTCATCCAAC	CGTTGTTCCGGTTTCTGTTT	253	*Mbo*I	A
142	Ca2C6628	-	GTACTTGGGTTCCAAGCAGG	TTTACGAATCGATGGCACAA	174	*Mnl*I	A
143	Ca2C11129	CGMM142	GCAATCCTGAATGGAAGGAA	ATGTCCGTCAGGGAGATCAG	182	*Hsp*92II	A
144	Ca2C161	CGMM143	GGACAACCGAAGAGGATCAA	GTTGTCCGTCCTTCCTGGTA	257	*Bfa*I	A
145	Ca2C1614	-	CCGGTAAGAAGCAGTCGAAG	TTTCCATTCTTTACCGCCAG	167	*Mse*I	A
146	Ca2C17163	-	CTTGTTGGTGATGATCCGTG	TACAAGCCCAATATTCCCCC	371	*Hpa*I	A
147	Ca2C23568	CGMM144	GGCCAAAGAATGTTACCAGC	ACAAGCTAGATAGGCCGTGG	288	*Mnl*I	A
148	Ca2C26976	-	AAGTCCCACTGGCAATGAAG	GGCAGGAAAATGAGAAGCTG	320	*Mbo*I	A
149	Ca2c31438	-	AAACCAGTCTCCCACCACAG	AGGATAATTCCACGAACCCC	270	*Hin*dIII	A
150	Ca2c32031	-	ACACCGTTGGAAGACGAAAC	CCGGACTTTCACGATTCACT	293	*Aci*I	A
151	Ca2c34206	-	TGGCTGAGATGATTCCTGGT	TCAGTCCAAGGAACACCTTTC	229	*Mse*I	A
152	Ca2C34413	-	CAATCCTCGCCTAAAAACCA	CAAGCTAGTGCCATGGTTCA	245	*Hpa*II	A
153	Ca2C34583	-	ATGCAGCAGAAGCAAGTGAA	CCGACGATACTCGACGTCTT	281	*Nla*III	A
154	Ca2c36391	-	TGACACGTACCAAGAGCAGG	TCGAGGAAATGCCCAGTATC	320	*Mnl*I	A
155	Ca2c37368	-	AAGACTCCGCAATGGAGAGA	CCTAGAAGGGGCCCTAACAC	215	*Hpy*CH4V	A
156	Ca2C37998	CGMM145	GCGAGGAAACAAGTGGAGAG	TTCCCCCATTGGTAAAAACA	276	*Bfa*I	A
157	Ca2C38128	CGMM146	TAAAAACGCGTACCGAATCC	GTCGGCTTTTCAAGCAACTC	588	*Nde*I	A
158	Ca2C3892	-	TCATGCCAAATGAAAGTGGA	TGGATCAGTGATGGGAACAA	239	*Bfa*I	A

159	Ca2C41582	CGMM147	CATCTTCGCACTCTTCTCCC	GATCAGCAGCAACCACAAGA	341	*Taq*I	A
160	Ca2C42261	-	TACGCTGAAGGTACGCCTCT	CTGCGTCTGTCAGATTGCTC	203	*Acu*I	A
161	Ca2S032873_1646_3001	-	GAAACTCAAGCTCAGCCGAC	TGCAAAGTCATCTGCTTCTGA	158	*Mbo*II	A
162	Ca2C36636	-	CATTGAGAGGGTGCTTGGAT	GAATGAGAGGTCGACCGAAA	243	*Taq*I	A
163	Ca2C10894	-	TTTTTGCAACACAAACCCAA	CAACCGTTGATGTCATGAGG	290	*Mnl*I	A
164	Ca2C12857	-	GCCATCTTTTGGTGTGGAAG	CACACACATATAAACAACCCCAA	285	*Tsp*509I	A
165	Ca2C41121	-	CACGCTTCCAACAACAGAGA	CCTCCTCTTCCCAATTTTCC	348	*Tsp*509I	A
166	Ca2S082575_1212_1008	-	ACCATGTTTCCCAGAACTCG	CGCACTATGGAAATGGGAAG	150	*Hph*I	A
167	Ca2C6324	-	AACGGATTCGATATGGACGA	TCTCTTGCGACCCTTTCTGT	364	*Acu*I	A
168	chs	-	GAATTCCCTCGTTTTCCTTC	AAATGAATTGTGGAAAGGCA	127	*Alu*I	A
169	AGLC93	-	CTTCAAGTTCTTCGTTTGACGCAA	CCTTTCTCCCACAACCTCTCCA	306	*Hpy*CH4IV	A
170	AGLC57	-	CTCCTCTTCTCCGTCGTAGCA	CTGGTCCTTGACGGGAGTGA	596	*Tsp*DTI	A
171	AGLC2	-	GCAGCAGCCAGCCTAAGTAT	GGTCTGTGTTGGGCTTGTTT	415	*Eco*RV	A
172	AGLC7	-	TAATCATCGGTCATGAGTCTGTCA	CAAAATCGAAGATCTGCATCTGCA	262	*Mae*III	A
173	AGLC14	-	CTCGCTGGACCTCTCATCTT	AGTGCAAGCCACGAGAAAGT	392	*Hpy*99I	A
174	AGLC15	-	CCCTCTTCCCTCCGTTCTAA	CTTGACCACCGGACAAAACT	465	*Hpy*99I	A
175	AGLC19	-	CATCCCACACTACTTTTTACCTCA	CCTCCTCCGACAAATTCATCACA	271	*Hpy*CH4IV	A
176	AGLC22	-	TGTCAGACTGAGCTGTGTATGAGA	TTGCCCGTATGGTTATGTTAGGAA	815	*Tsp*GWI	A
177	AGLC23	-	AATGGTGATTCGTCAGTCGCCTA	CTGTCTGAAGAAAGTGAACGAA	295	*Ava*II	A
178	AGLC30	-	TCGTGCAGCTCATCCATATT	TTGTCCACGCATCATCTCAT	245	*Dde*I	A
179	AGLC44	-	CTTTACCAAAACCACCTTCACCAA	GCAGGTCGCGTTGTTGCAA	686	*Fai*I	A
180	AGLC72	-	CATGTTTTCTACCCTCACAATGCA	TACTCACTTGTTGTTCCAGACA	110	*Tsp*509I	A
181	AGLC84	-	TTTAATTACGCGGTTTCCACGA	GAAGACTTGAGACATGGGCACA	361	*Acu*I	A
182	AGLC94	-	TTTGTGATGGTCCTGCTCTCTCA	ACCGCTTCAGGATCAACTCGA	370	*Hin*fI	A
183	AGLC112	-	CGACTCCCTCATCACCTCCA	CTTTGGGTCTCTGTTGTTGCTGA	155	*Tsp*509I	A
184	AGLC126	-	TTCACAACAATGGCTGAACC	ATGCCTTCTTTGCTGCAGAT	592	*EarI*	A
185	AGLC131	-	ATGGCTTCAGCTAGGGAAGAA	ACCCATACCACCCAAATTGA	137	*Bsp*HI	A

186	AGLC137	-	GATGCAGGGTCGTTTCAAAT	AGCGATCAACACCGAGAGAT	340	*Bst*NI	A
187	AGLC193	-	GGAACATATGTATTGCGTGCAT	TGAATGTGTGTCTGAAAATTGATG	689	*Hpy*CH4III	A
188	AGLC171	-	GAGTACTTGCCAACTAGCTTAGGA	TTGGATATAACAGATGACGGGGAA	140	*Sfa*NI	A
189	Mt106141	-	GAGACTGCTGTTAGGGACCG	CCCTTGAGCCAGCAAGATAC	212	*Mse*I	A
190	Mt106628	-	TGTCCTGGTTCTTCACTTGCT	TTGCACTTCAATGGTTTAGCC	173	*Hpy*188I	A
191	Mt123479	CGMM048	AGAGGGGATACGGCAAAGTT	GTTTCCAGCATCACCACCTT	287	*Taq*I	A
192	Mt124331	-	GGGTTCTCTTATGGCAGGGT	GGGCATGGTCATTGACTTCT	315	*Ssp*I	A
193	Mt125375	-	TATTGCTGCTGCACTGAAGG	TGTTCTCCCTTGAATGGTCC	190	*Bcc*I	A
194	Mt127721	CGMM049	TCGTAACTTCGGTGTTGGTG	TCAGGAACAATCTTCCCAGC	420	*Hin*fI	A
195	Mt133126	CGMM050	GCCGAGATTGGTGAGTTTGT	TGGAGGAACACTTGACAACA	174	*Hae*III	A
196	Mt6799803	CGMM051	AAGGCAAGCCTAAGCTGTTG	TGATCCATTTGGCATTAGCA	354	*Mbo*I	A
197	Mt6803180	CGMM052	GGTCATGAAAATTGATGGTGAA	TTTGGAAGCTCCACAAGAAAA	92	*Hae*III	A
198	Mt6811198	-	CCAAAGAACCACCTCGTCAT	TCTTCAACGATGGGAATAGGA	91	*Acu*I	A
199	Mt6815341	-	CGATGGACGAGGAAATTGTT	TTTCCCTTCTTTCTGTCCCC	149	*Hpy*CH4V	A
200	Mt6817377	CGMM053	TGACGTGAGGTGTGGTGTTT	TGTATGACCTCTCCCTTGGC	295	*Dde*I	A
201	Mt6836854	-	GAGAATCTGGGTCTCCTCCC	ACAGTGAACTCGACCCCATC	244	*Hpy*CH4V	A
202	LG80	-	TGTGGAGAGGAGACAATTTCAAACC	GGTGGAAGCATTGGATTGGTGC	476	*Rsa*I	A
203	LG103	-	GTGCATTCCATTGCAATAGCATCC	TGGACACTGTTGACCTGCACCC	658	*Mse*I	A
204	LUP276	-	GGAAACTTAACACCAAAGAGAG	TTGCTCCATACACTTCTTGTAG	224	*FaeI*	A
205	Gm2077934	-	CCCATCAATGAGGCTCAAAT	GCTCCCACTCTGCTTCTTTG	253	*Tso*I	A
206	Gm2084815	-	CAGCGAAGGTGGAGGTTAGT	TGAACATCTGTGGAGGAGGA	123	*Tsp*509I	A
207	Gm2096212	-	GGAAGAAGAATATCTTTAAGTTGGC	CATTATGCCTTCTGAAGCAGC	446	*Fai*I	A
208	Gm2099239	-	TCAACAATGCTATTGACGGC	CAATTATTGGGTTTGTTTGGG	166	*Tse*I	A
209	Gm2123242	-	GAAAGAGAAGCTTCAAGAAAATGC	TTCCAATCAGCACTCCCAAT	328	*Hin*P1I	A
210	Gm2127540	-	AGCTGCGCTTTCTTGCAAAT	AAAGTCTTTGAAACCACAACTACAA	143	*Rsa*I	A
211	Gm2131909	-	GGGAGAACAAGATTGGCAAG	GGTCACCCTTCACGAACACT	113	*Mnl*I	A
212	Gm2132124	CGMM046	GGGTTGGATTGGAGAGTCTG	GGCAACAATGTCCAACTGTG	367	*Tsp*DTI	A

213	Gm2076345	-	AGGTTCTTGGCAGGGGTACT	TCCTCCACCTTACCCTCCTC	74	*Bcc*I	A
214	Ct6876038	-	CCTCGAGGTCGAATCTTCTG	CTTCTGGATGTTGTAATCGGC	185	*Mfe*I	A
215	Ms6943512	CGMM047	AGATGGACAAATCACCGAGC	TAAACACCGCTCTCATCCAA	220	*Hpy*CH4IV	A
216	Pv4675	-	TTTCAATGGCATCCTTTCTCT	TCCTAGACCCACCAGAAGGA	613	*Hpy*CH4IV	B
217	Rp1788051	-	TGGTGAACGTTCCAAAGACA	CCCTTCTTGTCACCACCAAT	350	*Fai*I	A
218	CaHa61	CaHa61	CACAACAACAACCCACAAGA	GCCTTAACTTGCTTGGCATA	280	*Hha*I	A
219	Ca2C20537	-	GATAGAGTCATCCTCCTATCCA	CTTTCCCACTTGGAACTTTG	725	*Mbo*I	A
220	Ca2C24067	-	TATACCCAGTGAGGTTGTGAG	GGTAATAGGGTTTAGGGTGAAC	487	*Bfa*I	A
221	Ca2C9868	-	CTGTATCCACTTTGTGCATC	CCTGACCCTGATGCTCATTT	468	*Eco*RI	A
222	kpil	-	GTAGGGATACCCGCTGAACT	CCTATGTGACGTCCTGTTCC	259	*Hpa*II	A
223	Mt6831918	-	GGCGGCTTCTTAAGTCCTCT	CCGGTGCAATTTGGATTAAA	191	*Hin*dIII	A
224	Rp17788514	-	TCCAAATCAAACCAGTACCCA	TACCGTGAGGCCTTGTAACC	166	*Nae*I	A
225	Ca2C42878	-	AAACAAGGGCAGGGAACTTT	ACATTCCTCGTCCCAAGATG	246	*Bsi*SI	A
226	Ca2C42611	-	GGAGGTGCTATCATTTCCGA	CCAGTCAGCAGCATCAGGTA	672	*Acl*WI	A
227	Ca2C43782	-	GGGAAAGCAAAAGGAGAAGG	AGGTGTTTGTAGCCCTGGTG	236	*Bfa*I	A
228	Ca2C19195	-	CAATGCATGGTGACACACAA	TCCCAGAATTCCTCAACCTG	498	*Apa*I	A
229	Ca2SGR399431	-	GGGTTTTTGGGTTTTAGGGA	GATAAAAAGGGGCTGGAAGC	679	*Fau*I	A
230	Ca2C718	-	CGCTGGAGACAAGCTAAACC	CTGCTGTACCAAGCCACTGA	379	*Smi*I	A
231	Ca2C585	CGMM110	CCATCAACAACCATGACTCG	TCCACCTGCATTATTCCCTC	201	*Age*I	A
232	Ca2C11361	CGMM112	CTCATTGCTTGAAGATGGCA	CAATGGCATCATCTGGTACG	207	*Ace*III	A
233	Ca2C13051	-	AATCCTTCACCATTTCAGCG	AGATTTCCGTTGAATCACGC	461	*Acc*II	A
234	Ca2C42782	CGMM111	TGGAGACATGGACCAAAACA	AGGCAACTCCTCTACGGTCA	289	*Alw*26I	A
235	Ca2SGR399333	-	GCGATGTAAAACGAGGGTGT	CCACCAAAAATTCCATACCG	404	*Asp*LEI	A
236	Ca2C43336	CGMM113	GCTGCTGTTGGGATTTCATT	TTATTGCCACAGTTGGTGGA	214	*Bsp*HI	A
237	Ca2C42692	-	TCATCTCGATATGCTCGCTG	ACATGGACTGACATGGACGA	168	*Aci*I	A
238	Ca2C43936	-	CCGGAGGAATCTACGAATGA	TTTTTCTCATCTTCGGTCGG	272	*Acc*II	A
239	Ca2C21271	CGMM114	ATGACGTGCAGACAAAGCAG	ATTTGGGCAATTAGCCCTTT	275	*Aan*I	A

240	Ca2SGR409670	-	TGAAAATCAATGAAAGGGCA	CGGGACATAACTGAACACCC	189	*Asu*HPI	A
241	Ca2C39414	-	CATGCCAAGTCCAATCACTG	GCTGCTCAAACAACACCAAA	356	*Bsa*MI	A
242	Ca2SGR399716	-	GGGGGTTTTCCCTTTTGTTA	CATATGCTCAGCTTGCCAGA	150	*Ahl*I	A
243	Ca2SGR399815	CGMM115	CATGGTCGTGTTACAGTGGC	CAACAAGCTTCCAGCAATCA	328	*Acu*I	A
244	Ca2SGR409100	CGMM116	CTGAGCATGCTGTCTTGGAA	GCGCTCTGAACAAACATGAA	261	*Alu*BI	A
245	Ca2SGR399970	-	CTCCTCAAACGAAATCGACA	TCAGAACTGTGTCTGGCTCG	215	*Bfi*I	A
246	Ca2C42582	-	GTTATTGGCTCTCTGCAGGC	CTTCGATTCTTCCCCGTGTA	220	*Bsm*BI	A
247	Ca2C43484	-	GTAGCCTGCGAAAAGCTACG	TCGCCACTACTTTGGGAATC	188	*Alw*26I	A
248	Ca2C43499	-	GGCCCAAGGTTCAACTACAA	TTACAGCGAAACTGCGAATG	568	*Hin*dIII	A
249	Ca2C43467	-	AGACATCAACCGAGATTCCG	GTTCTTTTCGCCTTTCCCTC	268	*Bau*I	A
250	Ca2C43375	-	GGGTCAAGATATAAAGGGCG	ATGGATACATGGGGTTTCCC	154	*Acc*II	A
251	Ca2C8893	-	CCTTTGTGACACGATGTTGG	CCGTTGGTGAAAGTGGAGAT	405	*Aat*II	A
252	Ca2C42817	-	TCCGTGGTTGGTTTTACCTC	AGGTTCAGGGCATTCTCCTT	163	*Acc*II	A
253	Ca2C27483	CGMM117	TTTGATTTCCAGTGCTGCTG	TAAAGCCCATGTCATCGTCA	155	*Bbv*CI	A
254	Ca2C43919	-	GTTCTGGAGACGCCACTAGC	TGCGACATTCATCACTCTCC	153	*Asp*LEI	A
255	Ca2C43916	-	AGCCGACGTTGAAATACCAC	CAGAACCACCGGATCACTTT	248	*Acl*I	A
256	Ca2C42666	-	TCGCTTCTCCCACTTATGCT	CGAGATTCCCACTGTCCCTA	154	*Aha*III	A
257	Ca2C42621	-	CCCCATACGCAAACAACTCT	TCGGTGTTCAACCATCAAAA	799	*Aji*I	A
258	Ca2SGR408900	CGMM118	AACATCGGGAATGTTTCGAG	CGATGAACACCCGAAGATTT	530	*Acc*II	A
259	Ca2SGR408994	CGMM119	TCCGTTTTCAACAACCATGA	GGTAGGACAAACGCGAAAAA	239	*Acl*WI	A
260	Ca2C42867	CGMM120	TAGGTCCATTTCACAAGCCC	TGGAGCTGAATTGCATGAAG	234	*Aci*I	A
261	Ca2C35363	-	GCTCCATTAGCATTGCCTTC	AAAGGGTCTTGTTGGTGTGC	201	*Bal*I	A
262	Ca2C43662	-	AAAAGATGGACGCCACATTC	CTAGGGTTTTGAGGCTGCTG	490	*Alw*I	A
263	Ca2C16048	CGMM121	CAACCGGAACTACGGTGTCT	TCGTCCCCATCCTCATACAT	412	*Hin*dIII	A
264	Ca2C43545	CGMM122	AAGGAAAGGAGGAGGACCAA	GCAAGCAATTTTCCAAGCTC	282	*Aci*I	A
265	Ca2C30733	-	TGGTTTCATTCTCCAATCACA	CGATCGATGCATAAATCACAA	484	*Uba*F11I	A
266	Ca2C20880	-	CGTTCAGTGGGAGCTTTCTC	CGTTGCAACTCCATCATACG	420	*Alw*26I	A

267	Ca2C25517	-	ACTGGCCAACAATCACACAA	GGAGGCAGCACAAGAGAAAC	217	*Ase*I	A
268	Ca2C33267	-	ACAACAGAATTGCCCTGACC	AGGCCCGAGAAATTTCCTTA	696	*Acl*WI	A
269	Ca2C43828	-	TCTCTCTTACGATGGGTGGG	GAACATGAAATTGCCTCCTCA	456	*Bsg*I	A
270	Ca2C12943	CGMM123	AAACGCGTACCCTTGACATC	GAAGAAGACCCTGAACTGCG	237	*Bst*6I	A
271	Ca2C34423	-	AGGAAGCAGAGGGTGGAAAT	GCCTTCTCGTTCTCATACGC	232	*Bse*NI	A
272	Ca2C24803	-	ATCGGAGCCTTACTGGGTTT	TCGTTTGCAGCCTTAGCTTT	194	*Bse*GI	A
273	Ca2C31127	CGMM124	TACCAGAACAAGACCTGCCA	TGTGATTTGGCATCCTTTGA	289	*Bce*83I	A
274	Ca2C43744	-	CGGTGAAGAGGAAAAAGCTG	ATGTACCACCCTAAAGCCCC	249	*Bsh*FI	A
275	Ca2C43650	-	GGGGGTTGATTATTCATTTGC	TAATTGGCATGTTGGAACGA	337	*Bse*MII	A
276	Ca2C43539	-	CCCACAGGTATGGGATAACG	ACCCAACATGAGATGCAACA	163	*Asp*LEI	A
277	Ca2C43058	CGMM125	TGGTGAGCAAAAGTGGAGTG	GCAATCCTATGAGCGGTGTT	246	*Acc*BSI	A
278	Ca2C33657	-	CACATACCATGAATGTGTGCC	GCCAATTTTGATCCAACCAC	421	*Cst*MI	A
279	Ca2C23395	-	CAAGTGGAGAACGGTGACAA	GGGCATGCAAAGAAAGTGAA	370	*Hin*dIII	A
280	Ca2C42642	CGMM126	TTACTTTCCGTCACCCGTTC	TTATCGGCCTTAGCAGCAGT	595	*Asu*II	A
281	Ca2C22854	CGMM127	AGCGTGCTCAGATAAAGGGA	GAACATCATGGACAAACCCC	461	*Afa*I	A
282	Ca2C43505	-	GCCTTCTTCAAACCAAATTCA	AATTGGCTCTGTAACGGCTG	236	*Acl*WI	A
283	Ca2C43222	-	GTTTCAATGAGAAGCGGAGG	TTTCCTGTTGTGGTTTGCTG	155	*Mnl*I	A
284	Ca2C34533	-	CAGGCTTCCAGAAGAAGGTG	CTTTGCAGCAGGTTTCTTCC	206	*Sth*132I	A
285	Ca2C42710	CGMM128	TAGGCTTCACCCAGCAACTT	AGGCAAGTCCCAATGAGCTA	245	*Asi*256I	A
286	Ca2C5199	-	AACCGACCTCCACTCTCCTT	AATGTTAAGCATTCACCGCC	194	*Acu*I	A
287	Ca2C8942	-	ACCAGTTGATGACCGAAAGG	CTTTCCTTTTTGCAGCCTTG	290	*Fau*NDI	A
288	Ca2C42587	-	GCACAAACAATTGCTCCAGA	AATAACACGCCCCATTCAAG	386	*Sth*132I	A
289	Ca2C20777	CGMM129	TTAACGGTGCCGTACACAAA	TAATGACCTCACCCTTTGCC	249	*Eco*57I	A
290	Ca2C5572	-	AGTTGGATCTGACTGGGGTG	TCCCAGCTTAAATCCATTGC	426	*Bse*3DI	A
291	Ca2C23573	CGMM130	CAATTTACGCTTTTGCCCAT	GGGTACGGTACGGAGAGTCA	282	*Bcc*I	A
292	Ca2C44220	CGMM131	GTTGTGGGACAAAAATGGCT	CAAGCTTGGTTGGGAAGGTA	307	*Bpm*I	A
293	Ca2C17306	-	TCTGCATCAAGGCACTGAAC	GGGGTTCCCTACGGAATAAA	489	*Alu*I	A

294	Ca2C44110	-	CTTCAACAGCAAGAGGCACA	AGAAGAGCAAGGCTGATCCA	237	*Ass*I	A
295	Ca2C3917	CGMM132	TCGAGGGAGCTGAAATGAGT	CAAAGCAACCATTGCTCTCA	295	*Afa*I	A
296	Ca2C9564	-	GCCCTGCTGGCTACTATCTG	ATTGAAGCAAGGAAGCCTGA	503	*Bsp*LU11I	A
297	Ca2C19564	CGMM133	TCAATCACATGGCTTTGGAA	TTGGCGGTGGAGATAGTTTC	244	*Bst*6I	A
298	Ca2C28874	CGMM134	CGCCTCCATATTATCGTCGT	TCCAAAACACGTGGAACAAA	283	*Bsl*FI	A
299	Ca2C14402	-	AATGACACGCCGTTACATGA	GAAGAGGTGAGCTGGTTTGC	220	*Bsc*GI	A
300	Ca2C22017	CGMM135	ATACATACGGGCTACGGCTG	TGGCTTCTTGCAAATCAGTG	408	*Aha*III	A
301	Ca2C44166	-	CGGGGAGTTGAAAATAAGCA	CTTCATGATGGATCAGCCCT	658	*Ahl*I	A
302	Ca2C27784	CGMM136	GCTTGCCAGCAAAAATAAGC	GAAGGCGAAAATTCAGGACA	296	*Alu*BI	A
303	Ca2C2842	-	GACCCAAGCAAAGACCACAT	AAGGGAGAGAAGAAGCCAGC	473	*Bsb*I	A
304	Ca2C2529	-	TGGCATATAATCGGGTCCAT	CCTCCCTCGAAGAATCATCA	437	*Sim*I	A
305	Ca2C44338	-	CAATGGCTTCCTACACCGTT	GAGATTCATCAAGGGACGGA	509	*Bse*YI	A
306	Ca2C33309	-	TTGCTCAAACAGATGGTGCT	ACTGTCACTACGGGTCCCAC	238	*Bse*GI	A
307	Ca2C18912	-	TTGTTATAACGAGGGCCCAG	ACGACTCTTTGCTCCTTCCA	261	*Hin*4II	A
308	Ca2C44635	-	AAGGAGAGTCCAGTGCCAGA	ATAATGCTCTGCCAAATGCC	298	*Bsi*SI	A
309	Ca2C44705	-	GCTGAGGAAGTGAATGAGGC	CAGCCAAATCAACAACCCTT	223	*Bsr*GI	A
310	Ca2SCK148696	CGMM137	TGCTTCTCGAATTCCTTCGT	TTCCAGTGATGCATCTTCCA	185	*Asp*S9I	A
311	Ca2C16596	-	TGGGAGGTTGGAATCTTGAG	CCAAGGGTTATCAATGGGTG	351	*Asu*II	A

[c]Source of gene/markers- A: Present study; B: Rajesh and Muehlbauer, 2008

Table 8: Polymorphism assessment of validated CAPS markers

S. no.	Marker name	Primer ID	Restriction enzymes	Amplicon length (bp)	Number of alleles	Status of polymorphism		
						ICC 4958 × PI 489777	ICC 4958 × ICC 1882	ICC 283 × ICC 8261
1	CGMM001	AGLC100	*Bfa*I	300	1	Monomorphic	Monomorphic	Monomorphic
2	CGMM002	AGLC111	*Hpy*CH4IV	400	2	Polymorphic	Polymorphic	Polymorphic
3	CGMM003	AGLC136	*Hph*I	500	2	Monomorphic	Polymorphic	Polymorphic
4	CGMM004	AGLC163	*Dde*I	160	2	Polymorphic	Monomorphic	Monomorphic
5	CGMM005	AGLC168	*Bfa*I	580	2	Polymorphic	Monomorphic	Monomorphic
6	CGMM006	AGLC178	*Tsp*509I	300	2	Polymorphic	Monomorphic	Monomorphic
7	CGMM007	AGLC179	*Ava*I	200	2	Polymorphic	Polymorphic	Monomorphic
8	CGMM008	AGLC196	*Ban*II	200	2	Polymorphic	Monomorphic	Monomorphic
9	CGMM009	AGLC198	*Hin*fI	500	2	Polymorphic	Monomorphic	Monomorphic
10	CGMM010	AGLC202	*Xba*I	650	2	Polymorphic	Monomorphic	Monomorphic
11	CGMM011	AGLC28	*Aci*I	350	2	Polymorphic	Monomorphic	Monomorphic
12	CGMM012	AGLC3	*Apo*I	500	2	Polymorphic	Monomorphic	Monomorphic
13	CGMM013	AGLC52	*Dde*I	350	2	Polymorphic	Monomorphic	Monomorphic
14	CGMM014	AGLC74	*Mbo*I	390	2	Polymorphic	Monomorphic	Monomorphic
15	CGMM015	AGLC77	*Hpy*CH4III	690	2	Polymorphic	Monomorphic	Monomorphic
16	CGMM016	AGLC17	*Tsp*509I	260	1	Monomorphic	Monomorphic	Monomorphic
17	CGMM017	AGLC8	*Hpy*188I	480	2	Polymorphic	Monomorphic	Monomorphic
18	CGMM018	AGLC82	*Hin*dIII	250	2	Polymorphic	Monomorphic	Monomorphic
19	CGMM019	Ah6928256	*Hin*cII	800	1	Monomorphic	Monomorphic	Monomorphic

20	CGMM020	Ca10273CD20F12	*Hin*cII	570	2	Polymorphic	Monomorphic	Monomorphic
21	CGMM021	Ca10556169746056	*Tsp*509I	200	2	Polymorphic	Monomorphic	Monomorphic
22	CGMM022	Ca12411169747340	*Aci*I	950	1	Monomorphic	Monomorphic	Monomorphic
23	CGMM023	Ca14402150174189	*Hpy*CH4IV	350	2	Polymorphic	Monomorphic	Monomorphic
24	CGMM024	Ca1448960219076	*Mse*I	650	2	Polymorphic	Monomorphic	Monomorphic
25	CGMM025	Ca15775169744160	*Aci*I	200	1	Monomorphic	Monomorphic	Monomorphic
26	CGMM026	Ca1773847832705	*Hpy*CH4III	850	2	Polymorphic	Monomorphic	Monomorphic
27	CGMM027	Ca21249	*Aci*I	220	1	Monomorphic	Monomorphic	Monomorphic
28	CGMM028	Ca21567	*Tsp*509I	120	2	Polymorphic	Monomorphic	Monomorphic
29	CGMM029	Ca9351169747446	*Hpy*188I	150	1	Monomorphic	Monomorphic	Monomorphic
30	CGMM030	Ca93773928149	*Hpy*CH4IV	850	2	Polymorphic	Monomorphic	Monomorphic
31	CGMM031	CaESTCg3	*Mbo*II	200	2	Polymorphic	Monomorphic	Monomorphic
32	CGMM032	CaESTSn16	*Hpy*CH4III	690	2	Polymorphic	Monomorphic	Monomorphic
33	CGMM033	CaESTSn30	*Hpy*CH4IV	190	1	Monomorphic	Monomorphic	Monomorphic
34	CGMM034	Ct6874464	*Hin*cII	280	2	Polymorphic	Monomorphic	Monomorphic
35	CGMM035	Ct6875390	*Hpy*CH4IV	450	1	Monomorphic	Monomorphic	Monomorphic
36	CGMM036	Ct6875951	*Nla*III	300	2	Polymorphic	Monomorphic	Monomorphic
37	ge13bg	ge13bg	*Bcc*I	590	1	Monomorphic	Monomorphic	Monomorphic
38	CGMM037	Gm103068	*Rsa*I	400	2	Polymorphic	Monomorphic	Monomorphic
39	CGMM038	Gm2091985	*Nla*III	550	2	Polymorphic	Monomorphic	Monomorphic
40	CGMM039	Gm2113547	*Hin*cII	280	2	Polymorphic	Monomorphic	Monomorphic
41	CGMM040	Gm2117188	*Nla*III	200	1	Monomorphic	Monomorphic	Monomorphic
42	CGMM041	Gm2118543	*Hha*I	280	2	Polymorphic	Monomorphic	Monomorphic

43	CGMM042	Gm2120918	*Tsp*509I	500	2	Polymorphic	Monomorphic	Monomorphic
44	CGMM043	Gm2124498	*Ava*II	300	2	Polymorphic	Monomorphic	Monomorphic
45	CGMM044	Gm2125123	*Hpy*188I	590	2	Polymorphic	Monomorphic	Monomorphic
46	CGMM045	Gm2129282	*Aci*I	300	2	Polymorphic	Monomorphic	Monomorphic
47	CGMM046	Gm2132124	*Hph*I	420	1	Monomorphic	Monomorphic	Monomorphic
48	CGMM047	Ms6943512	*Hpy*CH4IV	240	1	Monomorphic	Monomorphic	Monomorphic
49	CGMM048	Mt123479	*Hpy*188I	290	2	Polymorphic	Monomorphic	Monomorphic
50	CGMM049	Mt127721	*Hin*fI	420	2	Polymorphic	Monomorphic	Monomorphic
51	CGMM050	Mt133126	*Hae*III	190	2	Polymorphic	Monomorphic	Monomorphic
52	CGMM051	Mt6799803	*Mbo*I	520	2	Polymorphic	Monomorphic	Monomorphic
53	CGMM052	Mt6803180	*Hae*III	220	1	Monomorphic	Monomorphic	Monomorphic
54	CGMM053	Mt6817377	*Hpy*188I	500	2	Polymorphic	Monomorphic	Monomorphic
55	CGMM054	Ps1768451	*Mbo*I	280	1	Monomorphic	Monomorphic	Monomorphic
56	CGMM055	Ps1770858	*Hin*fI	290	2	Polymorphic	Monomorphic	Monomorphic
57	CGMM056	Pv4540	*Ava*I	600	1	Monomorphic	Monomorphic	Monomorphic
58	rgr4	rgr4	*Tsp*509I	290	1	Monomorphic	Monomorphic	Monomorphic
59	SHMT	SHMT	*Hpy*CH4IV	500	1	Monomorphic	Monomorphic	Monomorphic
60	tk	tk	*Nde*I	700	2	Polymorphic	Monomorphic	Monomorphic
61	CGMM057	Tp6849647	*Eco*RV	200	2	Monomorphic	Polymorphic	Monomorphic
62	CGMM058	Tp6849720	*Mbo*II	800	2	Polymorphic	Monomorphic	Monomorphic
63	CGMM059	Tp6850763	*Dde*I	500	1	Monomorphic	Monomorphic	Monomorphic
64	CGMM060	Tp6854083	*Hpy*CH4V	720	1	Monomorphic	Monomorphic	Monomorphic
65	CGMM061	Tp6857294	*Hin*fI	500	1	Monomorphic	Monomorphic	Monomorphic

66	CGMM062	Ca2S124718_0933_2144	*Aci*I	290	1	Monomorphic	Monomorphic	Monomorphic
67	CGMM063	Ca2C8663	*Mbo*II	270	1	Monomorphic	Monomorphic	Monomorphic
68	CGMM064	Ca2C41871	*Nla*III	150	2	Polymorphic	Monomorphic	Monomorphic
69	CaHa61	CaHa61	*Hha*I	280	2	Polymorphic	Monomorphic	Monomorphic
70	CGMM065	Ca2S290507_2844_3854	*Aci*I	390	1	Monomorphic	Monomorphic	Monomorphic
71	CGMM066	Ca2C43617	*BfaI*	300	1	Monomorphic	Monomorphic	Monomorphic
72	CGMM067	Ca2S125676_2925_1604	*Hpy*CH4IV	220	1	Monomorphic	Monomorphic	Monomorphic
73	CGMM068	Ca2C33173	*Aci*I	180	2	Polymorphic	Monomorphic	Monomorphic
74	CGMM069	Ca2S232393_3220_2305	*Rsa*I	400	2	Polymorphic	Monomorphic	Monomorphic
75	CGMM070	Ca2S143601_3004_0194	*Eco*RI	520	2	Polymorphic	Monomorphic	Monomorphic
76	CGMM071	Ca2S295752_0698_0195	*Eco*RI	190	2	Polymorphic	Monomorphic	Monomorphic
77	CGMM072	Ca2S126415_1648_0587	*Tsp*509I	350	2	Polymorphic	Monomorphic	Monomorphic
78	CGMM073	AGLC212	*Bcl*I	781	1	Monomorphic	Monomorphic	Monomorphic
79	CGMM074	AGLC213	*Bsa*MI	580	2	Polymorphic	Monomorphic	Monomorphic
80	CGMM075	AGLC216	*Nla*III	469	1	Monomorphic	Monomorphic	Monomorphic
81	CGMM076	AGLC217	*Msp*I	189	1	Monomorphic	Monomorphic	Monomorphic
82	CGMM077	AGLC214	*Xmn*I	242	1	Monomorphic	Monomorphic	Monomorphic
83	CGMM078	AGLC122	*Hpy*8I	605	1	Monomorphic	Monomorphic	Monomorphic
84	CGMM079	LG73	*Aci*I	409	2	Polymorphic	Monomorphic	Monomorphic
85	CGMM080	LG83	*Hpy*CH4V	287	1	Monomorphic	Monomorphic	Monomorphic
86	CGMM081	LG87	*Bsu*RI	562	2	Monomorphic	Polymorphic	Monomorphic
87	CGMM082	LG90	*Nla*III	333	1	Monomorphic	Monomorphic	Monomorphic
88	CGMM083	LG91	*Alw*26I	699	1	Monomorphic	Monomorphic	Monomorphic

89	CGMM084	LG95	*Bc*lI	339	1	Monomorphic	Monomorphic	Monomorphic
90	CGMM085	LG99	*Fin*I	612	1	Monomorphic	Monomorphic	Monomorphic
91	CGMM086	LG101	*Bsa*BI	250	2	Monomorphic	Polymorphic	Monomorphic
92	CGMM087	LG105	*Ava*I	338	1	Monomorphic	Monomorphic	Monomorphic
93	CGMM088	LG111	*Acc*B1I	661	1	Monomorphic	Monomorphic	Monomorphic
94	CGMM089	LUP51	*Taq*I	500	2	Polymorphic	Monomorphic	Monomorphic
95	CGMM090	LUP94	*Nmu*CI	229	1	Monomorphic	Monomorphic	Monomorphic
96	CGMM091	LUP120	*Acc*16I	549	1	Monomorphic	Monomorphic	Monomorphic
97	CGMM092	LUP235	*Fin*I	319	1	Monomorphic	Monomorphic	Monomorphic
98	CGMM093	LUP240	*Bst*UI	434	1	Monomorphic	Monomorphic	Monomorphic
99	CGMM094	LUP255	*Xba*I	771	1	Monomorphic	Monomorphic	Monomorphic
100	CGMM095	LUP302	*Nla*III	261	1	Monomorphic	Monomorphic	Monomorphic
101	CGMM096	LUP318	*Hin*fI	270	1	Monomorphic	Monomorphic	Monomorphic
102	CGMM097	LUP326	*Acc*I	544	1	Monomorphic	Monomorphic	Monomorphic
103	CGMM098	TC76606	*Mbo*I	378	2	Polymorphic	Monomorphic	Monomorphic
104	CGMM099	TC77515	*Bse*JI	740	1	Monomorphic	Monomorphic	Monomorphic
105	CGMM100	TC77624	*Alu*I	170	1	Monomorphic	Monomorphic	Monomorphic
106	CGMM101	TC77707	*Hin*P1I	372	1	Monomorphic	Monomorphic	Monomorphic
107	CGMM102	TC81224	*Mn*lI	404	1	Monomorphic	Monomorphic	Monomorphic
108	CGMM103	TC85414	*Bsu*RI	221	1	Monomorphic	Monomorphic	Monomorphic
109	CGMM104	TC86258	*Fnu*4HI	185	1	Monomorphic	Monomorphic	Monomorphic
110	CGMM105	TC87719	*Hin*fI	398	1	Monomorphic	Monomorphic	Monomorphic
111	CGMM106	TC92821	*Bsa*JI	166	1	Monomorphic	Monomorphic	Monomorphic

112	CGMM107	TC94373	*Eco*RI	274	1	Monomorphic	Monomorphic	Monomorphic
113	CGMM108	TC96130	*Bse*3DI	596	1	Monomorphic	Monomorphic	Monomorphic
114	CGMM109	TC103928	*Mse*I	275	1	Monomorphic	Monomorphic	Monomorphic
115	CGMM110	Ca2C585	*Dra*I	900	1	Monomorphic	Monomorphic	Monomorphic
116	CGMM111	Ca2C42782	*Nco*I	300	1	Monomorphic	Monomorphic	Monomorphic
117	CGMM112	Ca2C11361	*Mbo*II	250	1	Monomorphic	Monomorphic	Monomorphic
118	CGMM113	Ca2C43336	*Bsp*HI	600	1	Monomorphic	Monomorphic	Monomorphic
119	CGMM114	Ca2C21271	*Mbo*I	300	1	Monomorphic	Monomorphic	Monomorphic
120	CGMM115	Ca2SGR399815	*Acu*I	500	1	Monomorphic	Monomorphic	Monomorphic
121	CGMM116	Ca2SGR409100	*Alu*I	300	1	Monomorphic	Monomorphic	Monomorphic
122	CGMM117	Ca2C27483	*Pvu*II	250	1	Monomorphic	Monomorphic	Monomorphic
123	CGMM118	Ca2SGR408900	*Bst*UI	900	1	Monomorphic	Monomorphic	Monomorphic
124	CGMM119	Ca2SGR408994	*Hin*6I	600	1	Monomorphic	Monomorphic	Monomorphic
125	CGMM120	Ca2C42867	*Aci*I	550	1	Monomorphic	Monomorphic	Monomorphic
126	CGMM121	Ca2C16048	*Hae*III	900	1	Monomorphic	Monomorphic	Monomorphic
127	CGMM122	Ca2C43545	*Hpa*II	600	1	Monomorphic	Monomorphic	Monomorphic
128	CGMM123	Ca2C12943	*Ear*I	250	1	Monomorphic	Monomorphic	Monomorphic
129	CGMM124	Ca2C31127	*Mbo*II	300	1	Monomorphic	Monomorphic	Monomorphic
130	CGMM125	Ca2C43058	*Bsm*FI	350	1	Monomorphic	Monomorphic	Monomorphic
131	CGMM126	Ca2C42642	*Bst*BI	850	1	Monomorphic	Monomorphic	Monomorphic
132	CGMM127	Ca2C22854	*Nde*I	800	1	Monomorphic	Monomorphic	Monomorphic
133	CGMM128	Ca2C42710	*Mbo*I	300	1	Monomorphic	Monomorphic	Monomorphic
134	CGMM129	Ca2C20777	*Acu*I	850	1	Monomorphic	Monomorphic	Monomorphic

135	CGMM130	Ca2C23573	*Bcc*I	300	1	Monomorphic	Monomorphic	Monomorphic
136	CGMM131	Ca2C44220	*Bpm*I	300	1	Monomorphic	Monomorphic	Monomorphic
137	CGMM132	Ca2C3917	*Rsa*I	450	1	Monomorphic	Monomorphic	Monomorphic
138	CGMM133	Ca2C19564	*Ear*I	400	1	Monomorphic	Monomorphic	Monomorphic
139	CGMM134	Ca2C28874	*Bsm*FI	700	1	Monomorphic	Monomorphic	Monomorphic
140	CGMM135	Ca2C22017	*Bst*BI	600	1	Monomorphic	Monomorphic	Monomorphic
141	CGMM136	Ca2C27784	*Acu*I	300	1	Monomorphic	Monomorphic	Monomorphic
142	CGMM137	Ca2SCK148696	*Hin*dIII	800	1	Monomorphic	Monomorphic	Monomorphic
143	CGMM138	Ca2C10102	*Mse*I	150	1	Monomorphic	Monomorphic	Monomorphic
144	CGMM139	Ca2C25794	*Bfa*I	288	2	Polymorphic	Monomorphic	Monomorphic
145	CGMM140	Ca2C33338	*Aci*I	372	2	Polymorphic	Monomorphic	Monomorphic
146	CGMM141	Ca2C3559	*Aci*I	285	2	Polymorphic	Monomorphic	Monomorphic
147	CGMM142	Ca2C11129	*Hsp*92II	182	2	Polymorphic	Monomorphic	Monomorphic
148	CGMM143	Ca2C161	*Bfa*I	257	2	Polymorphic	Monomorphic	Monomorphic
149	CGMM144	Ca2C23568	*Mnl*I	288	2	Polymorphic	Monomorphic	Monomorphic
150	CGMM145	Ca2C37998	*Bfa*I	276	2	Polymorphic	Polymorphic	Monomorphic
151	CGMM146	Ca2C38128	*Nde*I	588	2	Polymorphic	Monomorphic	Monomorphic
152	CGMM147	Ca2C41582	*Taq*I	341	2	Polymorphic	Monomorphic	Monomorphic

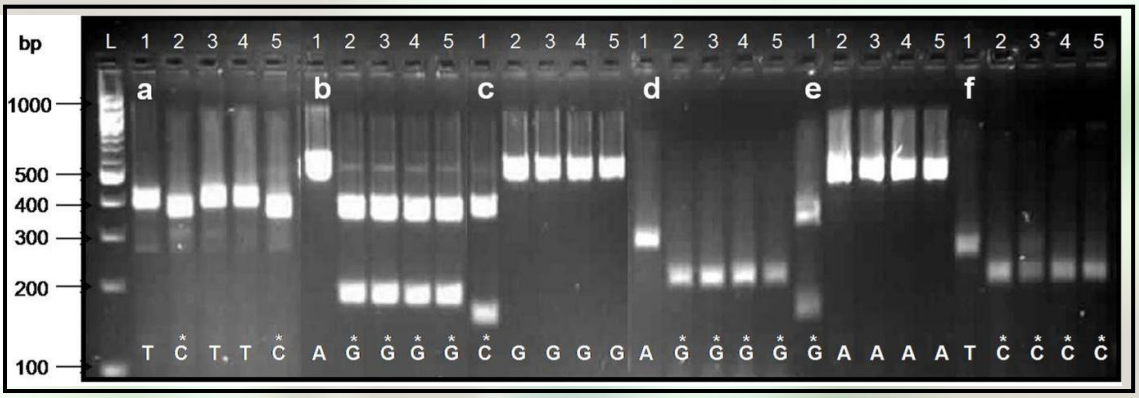


Figure 15: Some selected examples of assaying SNPs via CAPS markers

Restricted digested products for 6 CAPS candidate markers have been shown on 1.2% agarose gel as following: CGMM002-HpyCH4IV (recognition site- A/CGT), b) CGMM020-HincII (recognition site- GTY/RAC), c) CGMM009- HinfI (recognition site- G/ANTC), d) CGMM023-Hpych4IV (recognition site- A/CGT), e) CGMM051-MboI (recognition site- /GATC), and f) CGMM041-HhaI (recognition site- GCG/C). Nucleotide variation has been depicted in the picture and the asterisk (*) marked nucleotide identified in the recognition site. Names of genotypes for DNA samples in each panel (a to f) are as following: Lane 1: PI 489777; Lane 2: ICC 4958; Lane 3: ICC 1882; Lane 4: ICC 8261; Lane 5: ICC 283 and Lane L: 100 bp ladder

4.3.2 Intron spanning region (CISR) markers

Another approach for development of functional markers used was based on intronic regions. Based on alignment of 9,569 unigenes of chickpea (in a separate study), and genome sequences of *Medicago*, primer pairs were developed in such a way that primers are located in conserved exonic regions and regions amplified contain introns to span introns (ISRs) and be located in highly conserved exons. Alignment of 9,569 unigenes of chickpea with available genome sequence of *Medicago* (8 chromosomes data and 0 chromosome – unaligned sequence data) was performed using BLASTN ($E \leq 1 \times 10^{-10}$ and >90% identity) and primer design criteria was fixed for amplicon ranging between 200 to 2000 bp. A total of 267 chickpea unigene sequences showed the alignment with 784 intronic regions after alignment. Further 144 unigenes were selected for primer designing based on the criteria of amplicon length ranging from 200-2000 bp where single amplicon could be able to amplify more than one intronic region. Submission of these 144 unigenes to Primer 3 software yielded 121 primer pairs coming from 92 chickpea unigenes. As these markers were developed for chickpea, these have been referred as *C*hickpea ISR (CISR) markers.

The detailed distribution of CISR markers based on 9,569 unigenes of chickpea on chromosome wise *Medicago* genome sequence is shown in Table 9. PCR profile used for CISR markers was similar as earlier described and the amplicons were resolved on MDE gels (Thudi et al. 2010). Schematic representation of CISR design process is shown in Fig. 16. In total, out of 121 CISRs 71% (87) of the primer pairs showed amplification on two chickpea genotypes representing 74 unigenes.

Table 9: Details on development, amplification and polymorphism assessment of CISR markers

***Medicago truncatula* (Mt) chromosome number**	**Number of corresponding chickpea unigenes having intronic region**	**Number of primer pair designed for amplification of intronic regions**	**Number of primer pairs yielding scorable amplification**	**Number of primer pairs showing polymorphism across 5 chickpea genotypes**
Chromosome (0) +strand	13	3	2	1
Chromosome (1) +strand	9	3	1	0
Chromosome (2) +strand	13	2	2	1
Chromosome (3) +strand	36	20	18	4
Chromosome (4) +strand	40	15	10	5
Chromosome (5) +strand	78	43	30	5
Chromosome (6) +strand	15	8	5	1
Chromosome (7) +strand	38	18	12	3
Chromosome (8) +strand	25	9	7	1
Total	267	121	87	21

To identify the polymorphic markers all the amplified primer pairs were screened on the 5 chickpea genotypes. Details of polymorphic markers is given under Table 10.

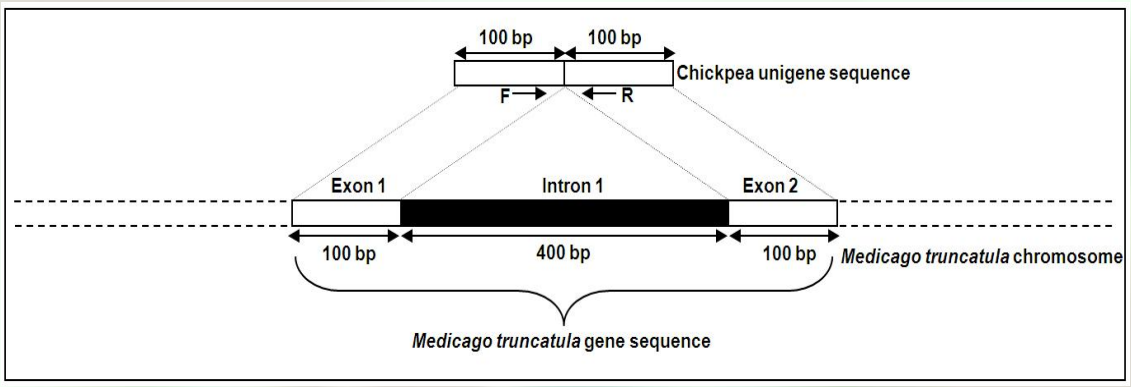


Figure 16 A schematic representation of the strategy for the development of CISR markers. An example showing alignment of a chickpea unigene to a gene on one *Medicago truncatula* chromosome. The chickpea unigene shows similarity with two exons (Exon 1 and Exon 2) separated by one intron (Intron 1). To develop the CISR marker, primer pairs (F: forward primer, R: reverse primer) have been designed by using Exon 1 and Exon 2 sequences in such a way that they amplify the intronic region.

Table 10 : Polymorphism status of easily assayable GMMs in intra-specifc mapping populations

Marker series	Primer pairs designed	Marker assays optimized	Number of polymorphic markers	
			ICC 4958 × ICC 1882	ICC 283 × ICC 8261
CGMM	311	152	6	2
CISR	121	87	3	3
CKAM	56	56	56	-
Total	488	295	65	5

4.3.3 VeraCode assays for BeadXpress system

From KASPar assay, a set of 56 polymorphic markers were identified on the parents of the intra-specific mapping population (ICC 4958 × ICC 1882) under a separate study. A custom assay of 96-plex SNPs OPA (Oligos Pool All) was designed using Illumina's Assay Design Tool and manufactured by Illumina. The 96-plex also includes 56 polymorphic KASPar assay markers. The genotyping was carried out with this 96-plex on the 232 RILs of ICC 4958 × ICC 1882 at Centre of Genomics-ICRISAT. Genotyping data was successfully generated for all the 56 markers and these SNP markers were referred as *C*hickpea *K*aspar *A*ssay *M*arker (CKAM).

4.4 Construction of a Transcript Map

4.4.1 Marker polymorphism and genotyping

As mentioned above, successful assays were developed for 152 CGMM, 87 CISR and 56 CKAM. In addition, 51 ICCeM markers for EST-SSRs were also available from another study. All these markers were screened on the parents of two intra-specific mapping populations (ICC 4958 × ICC 1882; ICC 283 × ICC 8261).

In case of CGMMs, 6 (4.2%) and 2 (1.4%) CGMMs were polymorphic in ICC 4958 × ICC 1882 and ICC 283 × ICC 8261 populations, respectively (Table 10). Out of 87 CISR markers tested, 3 (3.45%) CISR markers showed polymorphisms for each of two intra-specific mapping populations. In case of ICCeMs, 9 (17.65%) markers showed polymorphism for each of two intra-specific mapping populations. From KASPar assay, 56 CKAM markers were also reported to be polymorphic between ICC 4958 × ICC 1882.

In summary, out of 346 markers screened (8 CGMM, 6 CISR, 18 ICCeM and 56 CKAM) markers showed polymorphism in 4 genotypes. Only <10% markers showed polymorphism between the parents of two intra-specific populations. It is also important to mention that 8 markers (CGMM002, CGMM003, CISR117, ICCeM033, ICCeM035, ICCeM050, ICCeM051 and ICCeM058) were polymorphic for both intra-specific mapping populations.

Based on markers developed in this study, a total of 75 markers were polymorphic in the ICC 4958 × ICC 1882 population. By using the respective markers assays, genotyping data could were generated for 5 CAPS, 3 CISR, 9 ICCeM and 56 CKAM based markers on 232 recombinant inbred lines (RILs). The snapshots of CAPS genotyping, CISR genotyping, ICCeM genotyping and Veracode assays BeadXpress genotyping are given in Figs. 17, 18, 19 and 20.

4.4.2 Genetic map construction

As mentioned above, genotyping data was obtained for 73 markers in this study. In an earlier study in the ICRISAT, a genetic map comprising of 235 marker loci was developed on ICC 4958 × ICC 1882 population. The JoinMap® 4 analysis on these marker genotyping data,

however, could integrate only 6 (35.29%) GMM (1 CGMM, 2 CISRs and 3 ICCeM) loci and only 44 (78.57%) of 56 polymorphic CKAM markers in the genetic map of chickpea.

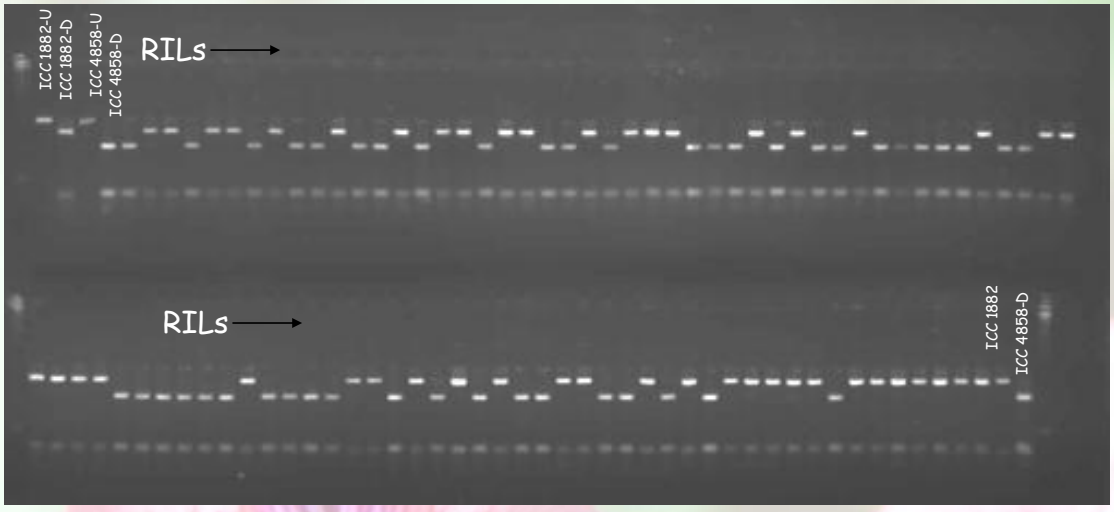


Figure 17 A snapshot showing SNP genotyping through CAPS assay on the mapping population ICC 4958 × ICC 1882

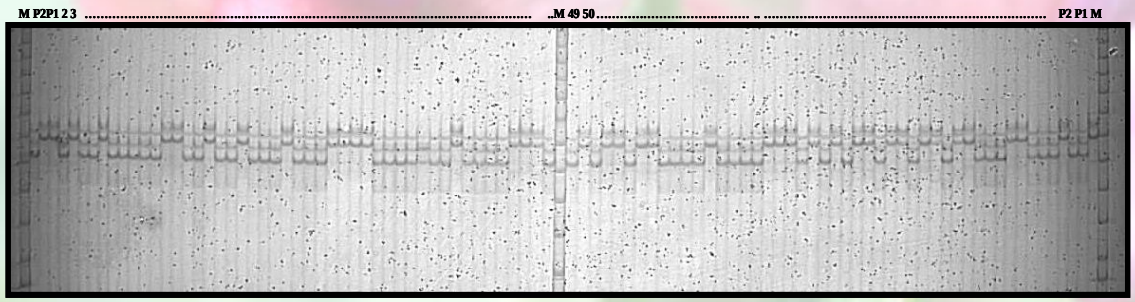


Figure 18 A snapshot showing genotyping of CISR marker on the mapping population ICC 4958 × ICC 1882

In summary, the developed genetic map has a total of 285 marker loci integrated onto 8 different linkage groups spanning 595.73 cM with an average inter marker distance of 2.09 cM. The average number of marker loci per linkage group was 35.63, with a maximum of 69 marker loci on linkage group 4 and a minimum of 19 marker loci in linkage group 2 (Fig. 21; Table 11). This genetic map has a total of 50 gene based loci, therefore, can also be referred as a 'transcript map'. The marker density for the gene-based markers on this transcript map is 6.38 GMM loci per linkage group.

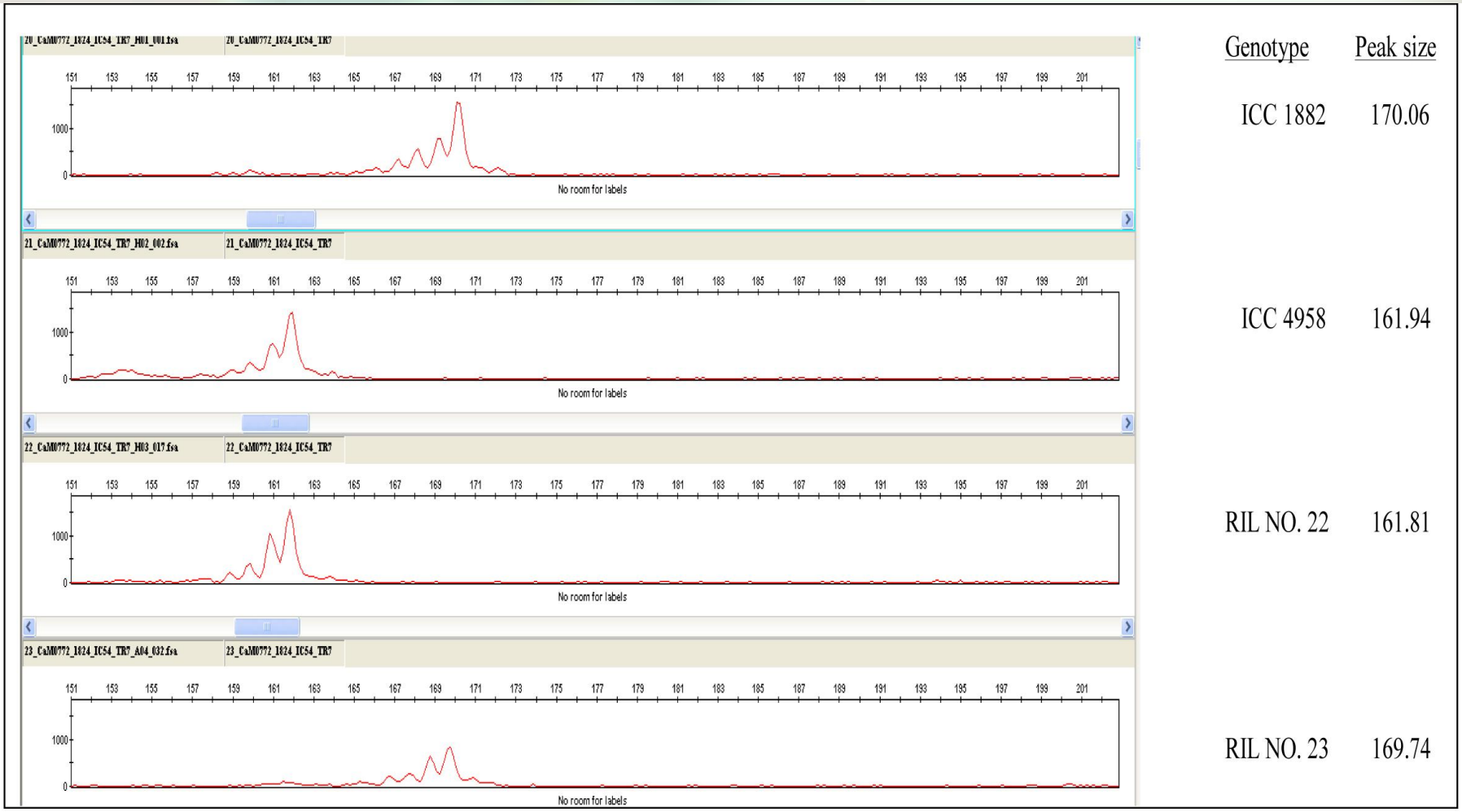


Figure 19 A snapshot showing genotypingt of ICCeM marker on the mapping population ICC 4958 × ICC 1882using ABI -3730

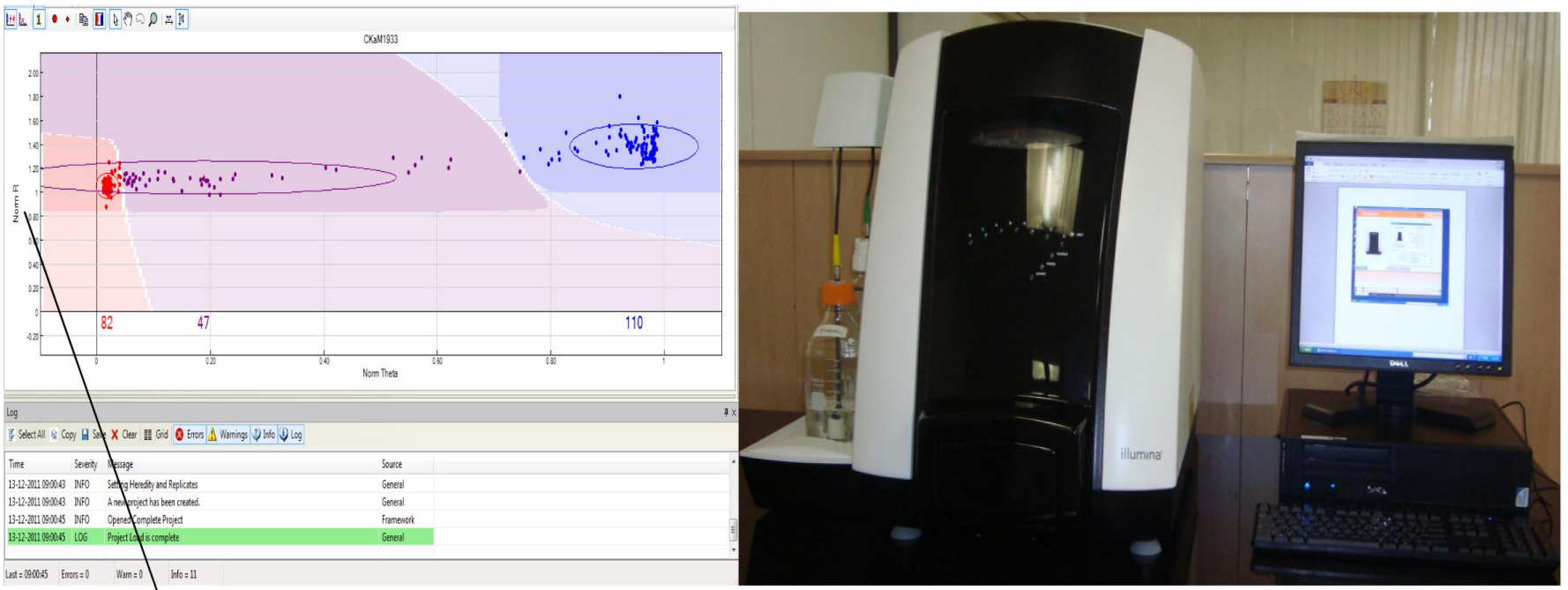


Figure 20 (a) A BeadXpress array analyzer installed in ICRISAT

(b) A snapshot of allele calling through SNP Genotyping by BeadXpress array of marker CKAM1933 on ICC 4958 × ICC 1882

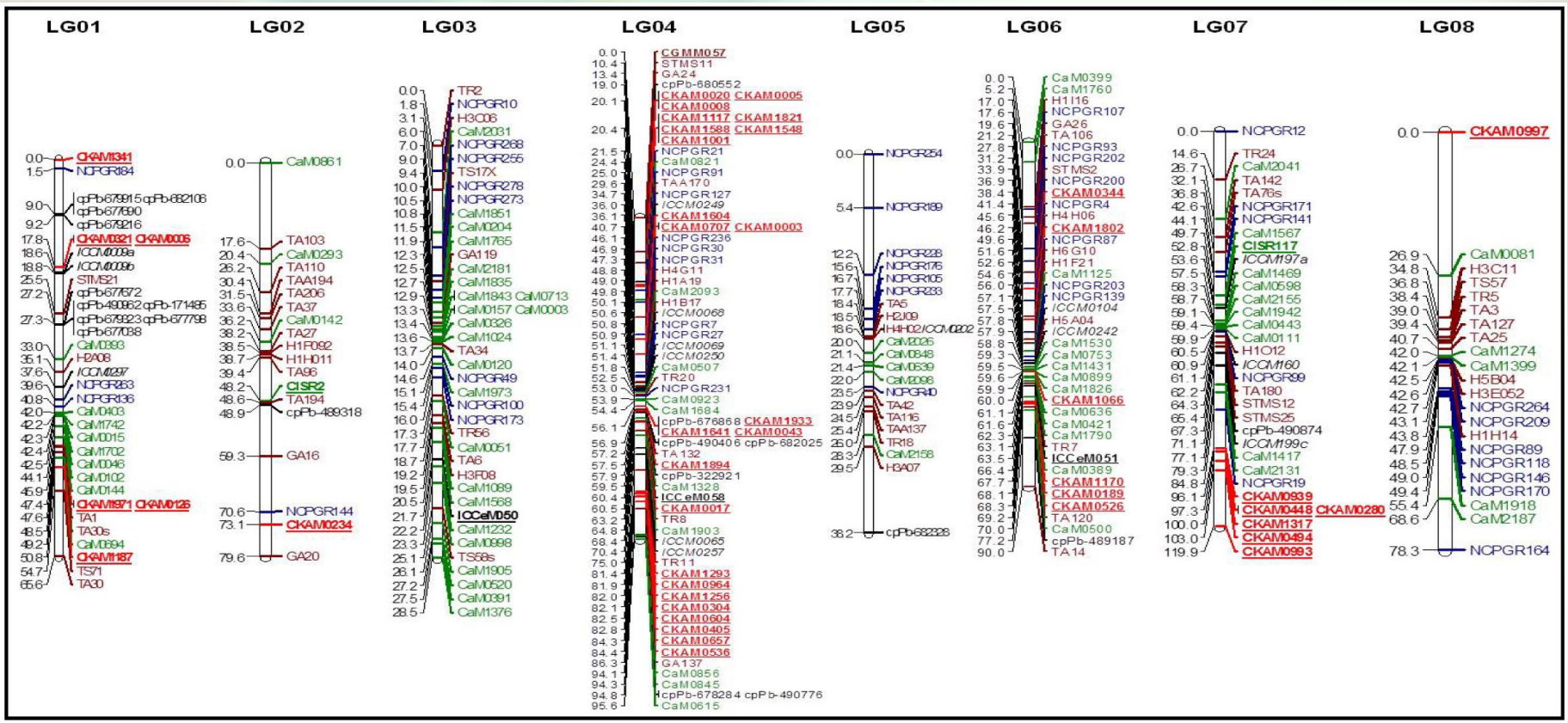


Figure 21 A transcript map of chickpea based on recombinant inbred lines of ICC 4958 × ICC 1882

Distances between the mapped loci (in cM) are shown to the left of the linkage group and all the loci are the right side of the map. Newly developed and integrated marker loci have been shown in bold, underlined and colored fonts: CGMM loci: brown colour, CISR loci: green colour, ICCeM loci: black colour., CKAM: red colour. In addition to these new loci, ; DArT loci-black colour, ICCM loci-italics, black colour; CaM loci-green colour; NCPGR loci- blue colour; Winter et al. 2000 loci-brown colour

Table 11: Distribution of mapped marker loci on different linkage groups of the intra-specific map of chickpea

Linkage group	Markers integrated	Map distance (cM)
LG1	37	65.649
LG2	19	79.599
LG3	41	28.536
LG4	69	95.618
LG5	22	38.156
LG6	42	89.988
LG7	33	119.862
LG8	22	78.32
TOTAL	**285**	**595.728**

The linkage groups are written as 'LG'.
The map distances and inter-marker distances were calculated in centi-Morgan (cM) map units.

4.5 Identification of QTLs for Drought Tolerance

4.5.1 Phenotyping data analyses

Phenotyping data for drought tolerance related root traits was obtained from Crop Physiology Division of ICRISAT. The data was collected for year 2005 and 2007 with three replications. The analysis of variance (ANOVA) on pooled data showed highly significant differences among genotypes and environments. The mean square values varied significantly for all the traits. Highly significant differences were found in genotypes (RILs) for each trait except for the trait RD. For more details of ANOVA please see Nayak 2011.

4.5.2 QTL analysis for drought tolerance related root traits

Above mentioned genotyping and phenotyping data were used for QTL analysis for drought tolerance traits. In summary, QTL analysis identified of a total of 12 significant QTLs (LOD ≥ 3) for ten root traits. Two QTLs each were identified for root dry weight (RDW) and root surface area (RSA), where as for leaf dry weight (LDW), root density (RD), root length density (RLD), root volume (RV) and stem dry weight (StDW) only one QTL was identified. For root length (RL), ratio of RDW to total dry weight (RT)

Table 12: Main effect QTLs for drought tolerance related root traits using single locus analysis using QTL cartographer

S. No.	Trait/QTL	Environment	LG	Marker interval	Position	Additive effect	LOD	R2 (%)
1	**Leaf dry weight** *QTLLDW1*	V	4	NCPGR91-TAA170	27.61	0.11904	15.4920865	28.2048
2	**Root Depth** *QTLRDp1*	V	1	ICCM0009b-STMS21	19.67	2.21304	3.34888783	7.118
3	**Root dry weight**	V	1	STMS21-cpPb-677672	25.7	0.02814	3.25195978	5.5939
	QTLRDW1	V	4	TAA170-NCPGR127	30.07	0.04055	5.23857	11.5656
	QTLRDW2							
4	**Root length** *QTLRL1*	V, VII	4	NCPGR91-TAA170	25.61	466.766	4.39-12.65	13.15-29.71
5	**Root length density** *QTLRLD1*	V	4	NCPGR91-TAA170	25.61	0.02571	12.600653	29.7704
6	**Root Surface Area**	V	1	CaM0144-CKaM1971	47.07	-30.7747	3.03028065	4.8414
	QTLRSA1	V	4	NCPGR91-TAA170	25.61	67.6993	9.62792391	24.0773
	QTLRSA2							
7	**Ratio of Root dry weight/total dry weight** *QTLRT1*	V, VII	4	NCPGR91-TAA170	27.61	-1.16789	9.35-9.75	18.14-18.86
8	**Root Volume** *QTLRV1*	V	4	NCPGR91-TAA170	27.61	0.67427	5.21014739	10.1761
9	**Stem Dry Weight** *QTLStDW1*	V	4	NCPGR91-TAA170	25.61	0.06495	9.15303478	25.1416
10	**Shoot dry weight** *QTLSDW1*	V, VII	4	NCPGR91-TAA170	25.61	0.20146	11.59-14.31	30.94-38.03

[c]Environment- I: pat= Patancheru 2008-09, II: dug=Durgapura 2008-09, III: seh=Sehore 2008-09, IV: Ndl=Nandyal 2008-09 YEAR, V: Patancheru 2005,VI: Patancheru 2006,VII: Patancheru 2007, VIIIa: Patancheru 2009-10, Irrigated, VIIIb: Patancheru 2009-10, Rainfed, IXa: Durgapura 2009-10, Irrigated, IXb: Durgapura 2009-10, Rainfed, Xa: Sehore 2009-10, Irrigated, Xb: Sehore 2009-10, Rainfed, Xia: Nandyal 2009-10, Irrigated, XIb: Nandyal 2009-10, Rainfed; Position= distance (cM) between the QTL and the first marker of the relevant chromosome; $\mathbf{R^2}$ (%) =Percentage of phenotypic variation explained by the QTL; SDW-shoot dry weight; StDW-stem dry weight; LDW-leaf dry weight; RT- RDW/total dry weight ration; RLD-root length density; RSA- root surface area; LOD- logarithm of odds

and shoot dry weight (SDW) one consistent QTL were detected in both the years (2005 and 2007). Most of the QTLs are identified in the same genomic region on LG04 spanning NCPGR91-TAA170. All the 12 QTLs identified here are either present on LG1 or LG04. The details for these QTLs are given below as well as summarized in Table 12 and presented in Fig. 22.

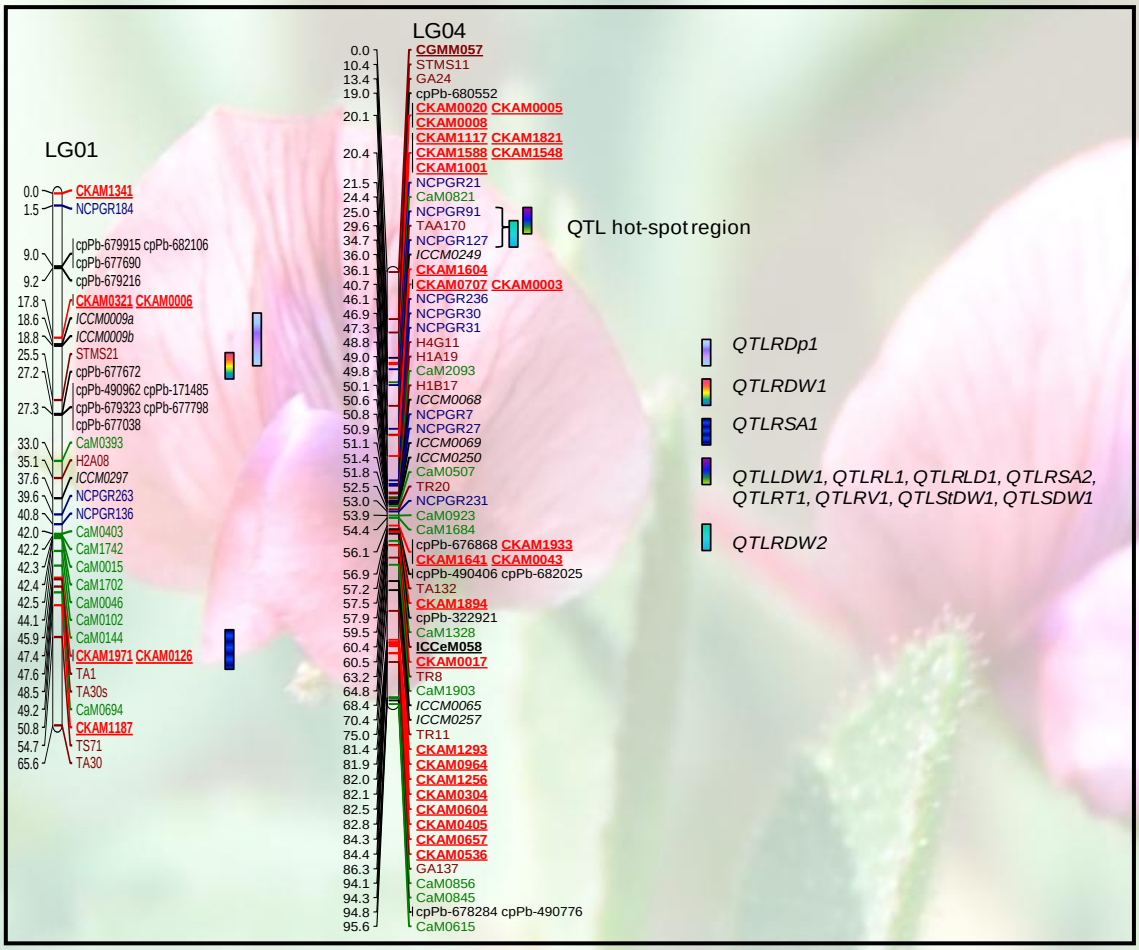


Figure 22 QTL map of LG01 and LG04 for drought tolerance related traits based on intra-specific mapping population- ICC 4958 ICC 1882

4.5.2.1 *QTLs for shoot dry weight (SDW)*

Only one major QTL, '*QTLSDW1*' was detected for shoot dry weight. This QTL was detected in the year 2005 and 2007 data at the same position at 25.61cM on LG04. It was flanked by NCPGR91-TAA170. In both the years, LOD score varies from 11.59-14.31 and phenotypic variation (R^2%) for this QTL varied from 30.94-38.03 respectively, with positive additive effect

of 0.201. The phenotypic variation observed for this trait was the highest among all the traits studied across years.

4.5.2.2 *QTLs for stem dry weight (StDW)*

For stem dry weight, only one QTL was identified in 2005 year data on LG04 flanked by markers NCPGR91 and TAA 170 at threshold LOD of 9.15 and explained 25.14% of phenotypic variation. This QTL (*QTLStDW1)* is a major QTL explaining >20% of phenotyping variation (Table 13).

4.5.2.3 *QTLs for leaf dry weight (LDW)*

The phenotypic data of leaf dry weight was available only for the year 2005 and was used to identify QTLs for LDW. For LDW also one major QTL- '*QTLLDW1'* was observed on LG04 flanked by markers NCPGR91 and TAA 170 at threshold LOD of 15.49 and contributing 28.20% of phenotypic variation with additive effect of 0.119.

4.5.2.4 *QTLs for root dry weight (RDW)*

For root dry weight (RDW), two putative QTLs were identified in 2005 year data. These are named as *QTLRDW1* and *QTLRDW2*. *QTLRDW1* was detected on LG01 with flanking markers STMS21 and cpPb-677672 at threshold LOD of 3.25 and *QTLRDW2* was detected on LG04 with left flanking marker TAA170 and right flanking marker NCPGR127 with LOD of 5.24. The QTL-'*QTLRDW2*' explaining 11.57 % phenotypic variation as compared to the '*QTLRD1*' contributing only 5.59% phenotypic variation for RDW.

4.5.2.5 *QTLs for rooting depth (RD)*

For RD, QTL (*TLRD1*) was identified in the data of year 2005. This QTL was identified on LG01 flanking by markers ICCM009b and STMS21. This QTL was observed at LOD score of 3.35, contributing 7.12 % phenotypic variance with positive additive effect.

4.5.2.6 *QTLs for ratio of RDW to total dry weight (RT)*

Only one QTL was detected for trait RT and is called as '*QTLRT1'*. The QTL "*QTLRT1"* is flanked by markers NCPGR91 and TAA170 identified in the year 2005 data as well as in 2007 data on LG04. The range of LOD for this QTL varies from 9.35-9.75 and this QTL contributes 18.14 to 18.86 % of phenotypic variation.

Table13: List of major QTLs explaining >20% phenotypic variation

Trait name	QTL	Environment	LG	Marker interval	Position	Additive effect	LOD	R^2 (%)
RSA2005M	*QTLRSA2*	V	4	NCPGR91-TAA170	25.61	67.70	9.63	24.08
StDW2005M	*QTLStDW1*	V	4	NCPGR91-TAA170	25.61	0.06	9.15	25.14
LDW2005M	*QTLLDW1*	V	4	NCPGR91-TAA170	27.61	0.12	15.49	28.20
RLD2005M	*QTLRLD1*	V	4	NCPGR91-TAA170	25.61	0.03	12.60	29.77
RL2005M	*QTLRL1*	VII	4	NCPGR91-TAA170	25.61	466.77	12.65	29.71
SDW2005,07M	*QTLSDW1*	V, VII	4	NCPGR91-TAA170	25.61	0.20	11.59-14.31	30.94-38.03

LG-linkage group; cM- centi Morgan genetic distance; **R^2** %- phenotypic variation explained in percent; LOD- logarithm of odds; A- additive effect;
The trait names are suffixed by '05'- data from year 2005 and '07'-trait data from year 2007

4.5.2.7 *QTLs for root length (RL)*

For root length, one major QTL was detected on LG04 for the year 2005 and 2007 data. This was named as *´QTLRL1*' and is flanked by markers NCPGR91 and TAA170. The LOD for this QTL varies from 4.39 to 12.65 and explained 13.15 to 29.71% phenotypic variation.

4.5.2.8 *QTLs for root length density (RLD)*

For the year 2005, only one major QTL '*QTLRLD1*' was detected on LG 04. This QTL is present on the hot-spot region and flanked by markers NCPGR91 and TAA170. The threshold LOD was 12.60 and this QTL explains 29.77 % phenotypic variation.

4.5.2.9 *QTLs for root surface area (RSA)*

Maximum of two QTLs were detected in case of RSA. QTL '*QTLRSA1'* was identified in the data of year 2005, defined by marker interval 'CaM0144-CKAM1971' on LG1 and a major QTL '*QTLRSA2'* was identified on LG04 flanked by markers 'NCPGR91-TAA170' in 2005 data. The threshold LODs were 3.03 and 9.63 respectively. These two QTLs were contributing 4.84% and 24.08% of phenotypic variation with negative and positive additive effect respectively.

4.5.2.10 *QTLs for root volume (RV)*

For root volume also, only one QTL '*QTLRV1*' was detected in year 2005 data on LG04 and is defined by marker interval NCPGR91-TAA170. This QTL is contributing 10.18 % phenotypic variation at LOD 5.21 and positive additive effect.

As most of the QTLs of the trait (except for RD) are observed in the genomic region defined by marker interval "NCPGR91-TAA170-NCPGR127" was considered as the "QTL hot-spot region for root traits", this genomic region/QTL explained highest phenotypic variation (Fig. 23). A total of six major QTLs were observed pertaining to RSA, StDW, LDW, RLD, RL and SDW. Only one stable major QTL was observed in case of SDW (Table 13)

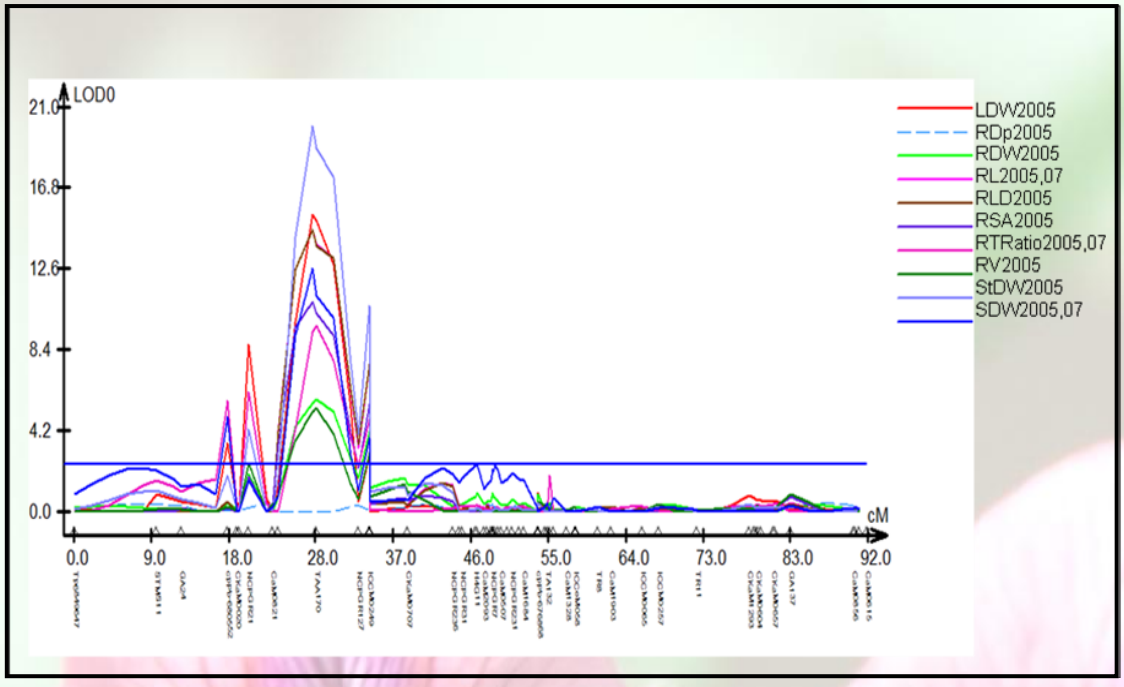


Figure 23 A snapshot of "QTL hot spot region" located on LG04 of intra-specific mapping population- ICC 4958 ICC 1882

5. DISCUSSION

The present study reports generation of Illumina sequence reads from transcriptome, identification of large scale SNPs, optimization of a range of SNP genotyping assays including CAPS, CISR, KASPar and VeraCode assays and construction of transcript map. Furthermore, QTLs for several drought tolerance root traits have also been identified based on the transcript map that may be used for molecular breeding application in chickpea improvement. A review on use of genomic resources to harness genetic diversity, co-authored by PhD candidate (Neha Gujaria) has recently been published in **Plant Genetic Resources: Characterization and Utilization** (Upadhyaya et al. 2011).

5.1 Effect of Drought Stresses on Chickpea Genotypes

Physiological studies suggested that plant roots are the primary site of perception and injury for several types of water stress, including salinity and drought. In many circumstances, it is the stress-sensitivity of the root that limits the productivity of the entire plant (Atkin et al. 1973; Steppuhn H. 2005). Plants are known to use more than one mechanism to resist unfavourable environmental conditions such as drought, which include 'drought escape' mechanism, in which the plants undergo rapid phenological development, completing their lifecycles before serious water deficit conditions prevail in their environment (Chaves et al. 2003), and by 'drought avoidance' mechanism by maintaining relatively high tissue water potential even at low soil-moisture content, balancing between water loss and turgor pressure allowing to survive drought (Turner, 1986).

A large number of studies were carried out on drought responsive genes that were cloned and characterized from an array of plant species and notably in model plant species such as *Medicago truncatula* (Bell et al. 2001), *Arabidopsis thaliana* (Asamizu et al. 2000a), *Oryza sativa* (Yamamoto et al. 1997), *Glycine max* (Shoemaker et al. 2002), *Lotus japonicus* (Asamizu et al. 2000b), *Populus alba* (Sterky et al. 1998) etc. Despite extensive studies, very little information about drought tolerance mechanisms is known in legume species. A recent study

carried out at ICRISAT has contributed 20,162 Sanger ESTs from drought- and salinity –stress challenged tissues to the EST collection.

For generating the Illumina sequencing based transcript reads, field and soil conditions were mimicked, where the water is lost at a slow rate and thus, gives the plant a chance of adaptive response to sustain drought stress by drought avoidance mechanisms. For this, prolonged slow drought stresses (sudden dehydration, PEG treatment, slow drought stress under glasshouse condition and slow drought under field environment) were administered to the plants of each of the two contrasting chickpea genotypes-ICC 4958-*a tolerant genotype* with higher root biomass and ICC 1882- *a sensitive genotype* with lower root biomass under this study (Fig. 9c). To understand the molecular basis of drought tolerance mechanisms, transcriptional changes and relative abundance in response to the same in chickpea genotypes, Illumina /Solexa 1G sequencing approach was carried out on the RNAs isolated from root tissues imposed by above mentioned stresses.

Figure 9(c) Comparison of roots of plants of ICC4958 and ICC 1882

As described earlier, artificial drought stress, by using polyethylene glycol (PEG) treatment, was imposed on the plants of two different chickpea genotypes. PEG, a high molecular legitimate osmoticum, modifies the osmotic potential of nutrient solution medium and thus induces water deficit in a relatively controlled manner simulating the natural gradual drought process similar to soil drying thus constituting a convenient way to study the effects of drought on the

morphogenic responses (Larher et al. 1993). Higher concentration of PEG accumulation results in differential hydraulic behavior of plants and has been observed to physically damage the roots (Jacomini et al. 1988). In this study, slow drought stress was induced to allow plants to have an adaptive niche to cope drought. The PEG concentration of 50mM and 10mM were found to be very high and resulted in sudden death of plants, therefore, experiment was carried out with 1mM concentration of PEG. 1mM PEG concentration gave a slow drought stress effect than 5mM PEG, and allowing an adaptive chance for survival. The stress imposed on the plants might not be purely stress due to unavailability of water but also might be due to the some toxic effect of the PEG chemical (Jacomini et al. 1988).

The second stress imposed during the present study in glass house was "dehydration stress". The sudden removal of plant roots from hydroponic medium could lead to some shock stress and such water deprivation induced was not easily possible with PEG-induction, as plants were stressed partially to toxic effects of PEG and not necessarily to water deficit. Under sudden dehydration stress, plants undergo drought avoidance mechanism for sustaining and express shock proteins in an elevated level. The third and forth stresses imposed under glasshouse and field conditions was slow drought on chickpea genotypes. The slow drought imposition on both the genotypes could lead to some shock stress genes in stressed genotypes as compared to well watered plants.

5.2 Transcriptomic Resources for Development of Functional Markers

5.2.1 Illumina/solexa sequencing- based transcript reads

The recent development of next-generation sequencing technology has brought new opportunities and challenges to the field of molecular biology. As millions of sequencing reads can be generated in a very short amount of time, efficient processing and analyzing of sequencing data has become critical. Furthermore, combining sequencing with other experimental techniques, a number of approaches have been developed to help explore the complex biological systems, among which Solexa sequencing is considered to be a very powerful tool for transcriptomics. The efficiency of Illumina / Solexa sequencing for the identification of differentially expressed genes has been well evidenced in a study by Hoen et al.

(2008) in which the results obtained by Illumina / Solexa were compared with different microarray platforms. The sequence-based analysis does not require background correction as reported in microarray, artifacts due to cross-hybridization are avoided, rare genes with low-abundance may also be detected, and hence, the number of transcripts / genes analyzed is comparatively greater than other technologies.

Illumina/Solexa 1Gb sequencing was performed on pooled RNAs from different stresses challenged tissues of both drought-tolerant as well as sensitive genotypes. In summary, 15.6 million reads from ICC 4958 and 22.1 millions reads from ICC 1882 genotypes, were generated. These reads were single-end and their average size was 36 bp. Since the draft genome of chickpea is not available, we used a chickpea transcript assembly comprising of 98,534 contigs/TUSs developed in a separate study (Hiremath et al. 2011), as a reference genome for aligning and analyzing Solexa datasets using *Alpheus* pipeline (Neil et al. 2008). In majority of earlier studies, short reads were analyzed by using the reference genomes. For instance, 3.9 million short reads isolated from *Medicago truncatula* were aligned to *Medicago truncatula* genome (Szittya et al. 2008). In the same way, 173 million short reads of *Arabidopsis thaliana* were aligned to *Arabidopsis thaliana* reference genome (Ossowski et al. 2008). Detailed analysis provided alignment of ICC 4958 genotype's reads to 25,896 TUSs and with 24,572 TUSs in the case of ICC 1882. In summary, this study generates a large number of transcript reads in a faster manner. These results have been recently published by the PhD candidate (Neha Gujaria) as a co-author in **Plant Biotechnology Journal (Hiremath et al. 2011)**.

5.2.2 SNP resource based on Illumina/Solexa sequencing

Analysis of transcript reads of two genotypes identified a total of 26,082 nucleotide variants (transitions, transversions and Indels). Of these 26,082 SNPs, a total of 2,405 *indel*, whereas 11,386 (43.65%) transition and 12,291 (47.12%) transversion were recorded. Maximum number of SNPs (3,856) had frequency difference range 0.20-0.29 and read depth 11-100. Altogether, SNPs were detected in 18,559 contigs. However using stringent criteria (read depth >100), high confidence 9,237 SNPs were identified in 6,908 contigs.

Alpheus pipeline predicted expression values for different TUSs having SNPs in two genotypes. For instance, a comparison of expression values revealed 4,554 contigs with more than two fold expression between ICC 4958 and ICC 1882. 4,031 contigs showed exclusive expression in ICC4958 and 2,920 contigs were exclusively expressed in ICC 1882. 201 SNPs were identified in contigs which show high differential expression between ICC 4958 and ICC 1882. Similar kinds of studies on identification of SNPs by NGS technology were conducted in *Eucalyptus grandis* -23,742 SNPs (Novaes et al. 2008), 8,23,325 unique SNPs in *Arabidopsis thaliana* (Ossowski et al. 2008), >7000 SNPs in Maize (Barbazuk et a., 2007), ~1000 SNPs were found in wheat (Akhunova et al. 2009)) and 39,022 putative SNPs were identified in *Glycine max* (Xiaolei et al. 2010).

5.2.3 SNP resource and sequence diversity based on allele specific re-sequencing

In the present study, a total of 1,627 primer pairs were designed for genes or transcription factors (TFs) identified in chickpea or other legume species e.g. *Medicago truncatula*, *Medicago sativa*, *Lotus japonicus*, *Lupinus* spp., *Arachis hypogaea*, *Pisum sativum*, *Crotolaria tenuifolia*, *Phaseolus vulgaris*, *Phaseolus coccineus*, *Glycine max*, *Glycine soja*, *Robinia pseudoacacia* and *Trifolium pretense* using allele-specific re-sequencing approach (Nelson et al.2006) and 1,139 markers showed scorable amplification on the panel of 5 genotypes and high quality sequences were generated for 532 markers on 2-20 genotypes and SNP identification was carried out in on high quality sequences. Similar kind of re-sequencing approach was done by Buhariwalla et al. (2005) but with limited number of ESTs in chickpea.

In summary, total primer pairs were designed for 1,912 genes, however, 1,139 primer pairs showed scorable amplification in the two genotypes examined. High-quality sequences were obtained for 532 candidate genic regions. Analysis of sequence alignment for these genic regions showed 2,046 SNPs in 264 genes and 71 *InDel* across a set of 2-20 genotypes. SNP and *InDel* markers have been proven very useful for estimating the linkage-disequilibrium and association mapping for crop improvement (Rafalski 2002).

Scanning of 84,073 bp sequence data (264 candidate genic regions) led to the identification of 2,046 SNPs with an average SNP frequency of 1 SNP per 41.09 bp in the chickpea germplasm

surveyed. However, earlier SNP report in chickpea presented SNP frequency as 1/61 bp in coding region while 1/71 bp in genomic region (Rajesh and Muehlbauer 2008). The SNP frequency in the present study is even higher than the earlier report and it can be attributed to use of higher number of genotypes and especially more wild species. Interestingly SNP frequency in chickpea seems to be higher as compared to other crop species like: wheat (1/1000 bp, Bryan et al. 1999); rice (1/89 bp, Nasu et al. 2002); barley (1/300 bp, Kota et al. 2008); common bean (1/76 bp, Gaitán-Solís et al. 2008) and soybean (1 in 277 bp, Zhu et al. 2003). However, this is an outcome of using 11 wild species representing the secondary as well as tertiary genepool in the present study. In case SNP frequency is considered only in germplasm of only cultivated species, the SNP frequency is 1/ 427 bp. In this scenario, it is evident that the cultivated genepool of chickpea is narrow as compared to other legume or cereal species mentioned above.

In general, SNPs are bi-allelic, 79 SNPs for 26 candidate genic regions showed three alleles that enhance the value of SNPs identified in the present study. For instance, the marker CGMM101 (primer pair TC77707) is having two base pair substitution at 128 and 136 base pair length from G→T→A and C→A→G respectively, showing haplotype diversity of 0.70 across all accessions, but 0.95 across wild chickpea accessions (Fig. 24)

The chickpea lines coming from wild species had >10X higher SNPs as compared to lines coming from the cultivated species. Similarly, nucleotide diversity (π) was about 5X higher in wild species (43.6×10^{-3}) as compared to the cultivated species (1.6×10^{-3}). While comparing this π with the earlier report of 14×10^{-3} by Rajesh and Muehlbauer (2008), it is clear that π in cultivated species is lower and is higher in wild species in the present study. It is interesting to note that the LG101 marker derived from *Lotus japonicus* showed the highest π as 74.0×10^{-3} (cultivated species - 4.8×10^{-3}, wild species- 88.4×10^{-3}) where as genic region for AGLC57, derived from chickpea EST, showed minimum π as 0.6×10^{-3}.

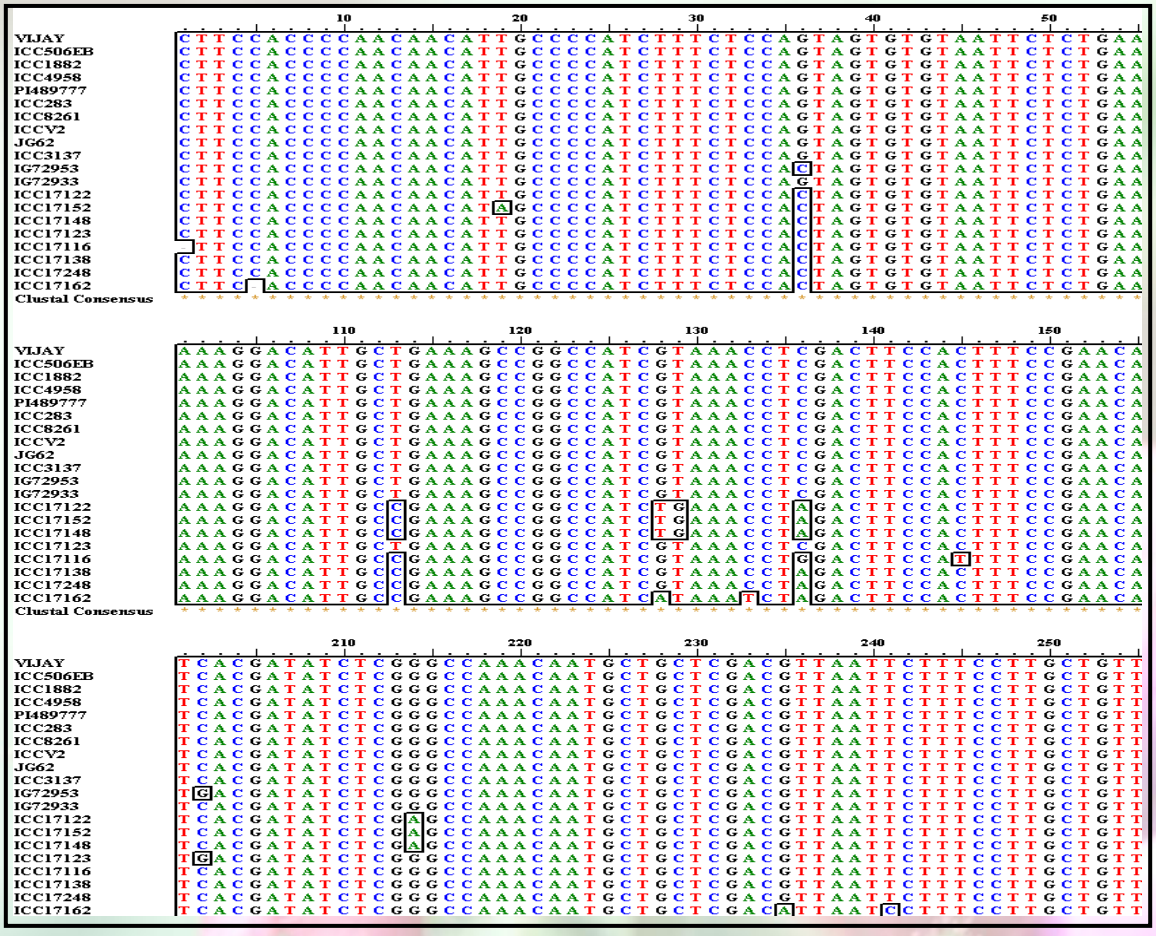


Figure 24 Partial sequence alignment of CGMM101 marker (Primer ID TC77707) genic region across 20 chickpea accessions, showing bi-allelic and tri-allelic SNPs. The asterisks represent similar sequences

However, unlike number and frequency of SNPs and π, the PIC value of SNPs in the wild species is lower (average 0.36) than the cultivated species (average 0.41). As PIC value is a direct function of allelic frequency for the given SNP in the germplasm collection and as compared to few genotype per wild species (4 for *C. reticulatum* and 1 each for 7 wild species), 11 lines were examined for cultivated species in the present study, a higher PIC value in cultivated species is not unexpected. Identification of SNPs and sequence diversity analysis has been published the PhD applicant as the first author in **Theoretical and Applied Genetics** (Gujaria et al. 2011).

5.3 Comprehensive Set of Genic Molecular Markers (GMMs)

Molecular markers are important genetic tools for understanding genome dynamics and facilitating molecular breeding. In case of chickpea, the progress in the area of development of molecular markers and genetic map however has remained slow. Nevertheless, in the past few years, significant progress has been made in the area of development of molecular markers (Hüttel et al. 1999; Sethy et al. 2006; Lichtenzveig et al. 2005; Buhariwalla et al. 2005; Choudhary et al. 2006, 2009; Hyten et al. 2010) and genetic maps (Winter et al. 2000; Pfaff and Kahl 2003; Choi et al. 2007; Radhika et al. 2007; Millan et al. 2010; Nayak et al. 2010; Thudi et al. 2011). However majority of these markers have derived from genomic DNA library and therefore developed markers or integrated marker loci onto the genetic map do not essentially represent genes. While gene-based molecular markers, popularly called GMMs (Varshney et al. 2007a), have been developed in large number in several crop species including some legumes like soybean (Choi et al. 2007), concerted efforts were not undertaken in case of chickpea. The present study, therefore, reports development of a comprehensive set of GMMs based on a range of genotyping assays. These markers will be useful for chickpea genetics research and breeding applications.

5.3.1 Cleaved amplified polymorphic sequence (CAPS) markers

The CAPS marker technique is also known as PCR-AFLP markers (Konieczny and Ausubel 1993). It identifies the restriction fragment length polymorphism resulting from the single base substitution like SNPs and insertions/deletions, which modify the recognition sites of the restriction enzymes (Chelkowski and Stepien 2001). The primers are designed based on prior sequence information available in the GenBank of genomic and cDNA sequences and cloned RAPD amplicons. It is robust and cost-effective assay that can be implemented in laboratories lacking sufficient infrastructural facilities. CAPS markers are characterized by their co-dominant inheritance and locus specific nature and useful for genotyping applications (Parsons and Heflich 1997; Weiland and Yu 2003).

In this study for the development of CAPS markers, a total of 311 CAPS candidates including 224 identified through allele re-sequencing approach and 87 identified by *in silico* mining of

ESTs were assayed on five chickpea genotypes. While scorable amplification was observed in 181 (58.12%) cases, CAPS assays were observed in 152 cases (83.98%). A higher success rate was observed for conversion of SNPs identified through allele-specific sequencing (59.90%) as compared to those that were identified through EST mining (32.18%) approach. This can be attributed for the possible sequencing error instead of presence of true SNPs. Also CAPS identified based on *in silico* mining of ESTs did not show polymorphism in these five genotypes. This is possible as the genotypes deployed for CAPS validation in the present study are different than those from which ESTs that were used for mining (Raju et al. 2011). In case of CGMMs, only 6 (4.2%) CGMMs and 2 (1.4%) CGMMs were polymorphic in two intra-specific mapping populations namely ICC 4958 × ICC 1882 and ICC 283 × ICC 8261, respectively. These polymorphic markers were genotyped subsequently on their respective populations.

5.3.2 Chickpea intron spanning region markers

CISR markers are an effective means to explore poorly characterized genomes for both DNA polymorphism and noncoding sequence conservation on a genome-wide or candidate gene basis, and also provide anchor points for comparative genomics across a diverse range of species. A very few reports are available for development of CISR markers in legumes. In CISRs, relatively conserved exons located near exon–intron boundaries are used to scan introns for suitably variable markers (Feltus et al. 2006). These gene-based markers are used to identify candidate genes inexpensively and have been used successfully in identifying polymorphic markers in legumes, pearl millet and other grasses (Feltus et al. 2006; Fredslund et al. 2006a; Yadav et al. 2008).

In the present study, 121 CISR markers were developed from 92 chickpea unigenes. Probably so far, very few CISR markers were developed in chickpea. From these 121 CISRs markers, 87 were easily assayable, and only 3 (3.45%) markers showed polymorphism in the cultivated *Cicer arietinum* genotypes i.e. ICC 283 and ICC 8261; ICC 4958 and ICC 1882, respectively. This reason could be attributed to the occurrence of very less polymorphism between the cultivated genepool. These 87 CISR markers could be used for genetic improvement of chickpea. Also reports are available for cross-species amplification of CISR in bean and peanut,

probably these marker dataset could be used to test genetic transferability between closely related legume species (Fredslund et al. 2006a, b).

5.3.3 VeraCode assay- based SNP markers

Next generation sequencing platforms can be used for to rapid generation of polymorphisms and genotype data for genetic mapping (Varshney 2010). Therefore, high-throughput BeadXpress assay (which uses GoldenGate genotyping chemistry) was used to generate the genotyping data for 56 polymorphic SNP markers identified through KASPar assay. The genotyping data generated here was more accurate and offers considerable quality-control advantages as compared to other genotyping assays (Mefford et al. 2009).

5.4 First Generation 'Transcript Map' of Chickpea

Chickpea is an important grain legume crop of rainfed agriculture in the semi-arid tropics. The crop suffers from terminal drought which causes considerable decreases in yield nearly upto 50% loss. Development of drought tolerant cultivars is the best strategy to minimize the yield loss. Majority of the wild species harbor resistance to these diseases but the introgression is thwarted due to cross compatibility barrier and linkage drag. Hence limited success has been achieved in chickpea breeding.

The development of genetic linkage map will greatly expedite the ability of breeders to tag and follow the introgression of specific chromosome segments linked to desirable traits from wild species into breeding lines of cultivated chickpea. Without the availability of a genetic map, it is difficult to utilize molecular markers or to combine molecular and conventional genetic techniques in chickpea for genomics-enabled crop improvement (Varshney et al. 2005b). Genic molecular markers are the markers of choice because they are ubiquitous throughout the genome, bi-allelic to multi-allelic, co-dominant and breeder friendly (Varshney 2009b).

Development of molecular markers in legumes has started sometimes in 1990s (Keim et al. 1990; Fatokun et al. 1992; Nadimpalli et al. 1994), but the progress has been very slow in developing their respective genetic maps in majority of these legume species. As mentioned earlier, narrow genetic diversity in gene pools of these legume species in general and scarcity of

polymorphic markers in particular were the main constraints. Due to availability of larger number of molecular markers in these legume species during last 2-3 years, progress has been accelerated in developing genetic maps for several legume species such as chickpea (Nayak et al. 2010), pigeonpea (Bohra et al. 2011) and groundnut (Varshney et al. 2008).

Due to low level of polymorphism between *Cicer arietinum* species, inter-specific cross has become the choice for analysing QTLs for simple traits like resistance to *Fusarium* wilt (Winter et al. 2000), *Ascochyta* blight (Rakshit et al. 2003) and carotenoid concentration (Abbo et al. 2005). Very few reports are available on intra-specific cross for identifying QTL(s) responsible for important agronomic traits like double podding (Cho et al. 2002; Radhika et al. 2007), *Ascochyta* blight resistance (Udupa and Baum 2003; Cho et al. 2004; Tar'ran et al. 2007; Kottapalli et al. 2009), *Fusarium* wilt resistance (Cobos et al. 2005) and some yield related traits like seed weight (Radhika et al. 2007). Since, there is no comprehensive gene based genetic map was developed in intra-specific mapping population so far, the present investigation emphasizes genetic map construction based on GMMs

Based on polymorphism obtained in the ICC 4958 × ICC 1882 cross, genotyping data were generated for 3 CISR, 5 CAPS, 56 CKAM and 9 ICCeM markers. These data were assembled with the genotyping data already available for 235 marker loci on this population from another study. By using entire dataset, a genetic map comprising of 285 markers loci spanning on 8 different linkage groups (LGs) with a total of 595.73 cM genetic distance with an average inter marker distance of 2.09 cM was developed. LG7 of the intra-specific map spanned the highest genetic map distance 119.9 cM. LG5 with 38.2 cM distance covered the least map distance among all others. The average number of marker loci per linkage group was 35.63, with a maximum of 69 marker loci on linkage group 4 and a minimum of 19 marker loci in linkage group 2. As compared to earlier intra-specific genetic maps (Radhika et al. 2007; Nayak 2011), this map is the most dense genetic map for chickpea.

The newly developed genetic map has a total of 50 gene-based marker loci. Therefore we will call this genetic map as a 'transcript map'. The transcript map developed, has a total of 50 GMM loci with an average of 6.38 GMM loci per linkage group. There were no earlier reports on

construction on the transcript map based on GMMs in intra-specific mapping population of chickpea. By using the resource of GMMs developed in this study, the first generation transcript map based on interspecific mapping population has recently been published in **Theoretical and Applied Genetics (Guajria et al. 2011)**. This thesis, however, presents development of first transcript map based on an intraspecific mapping population. In the other legume species, transcript maps have been reported in closely related legumes species like cowpea (Muchero et al. 2009) and soybean (Choi et al 2007).

It is evident that integration of more GMM loci is required to enhance the density of GMM loci on a transcript map. This first generation transcript map will be useful for trait mapping, comparative mapping with other legume species as well as linking genetic map with physical map of chickpea as the GMM loci integrated in this map are sequence based and represent the genes. It is also anticipated that the integrated GMM loci will serve as anchor markers for other chickpea maps so that those maps can be aligned with the reference genetic map.

5.5 Genetic Dissection of QTLs for Drought Related Root Traits

A very few reports were available for QTL analysis in intra-specific crosses. The reason behind this could be the low level of polymorphism observed in cultivated chickpea as compared to crosses derived from wild progenitor. Majority of the studies focused on disease resistance like *Fusarium* wilt (Cobos et al. 2005; Sharma et al. 2004a,b; Singh et al. 2008) and *Ascochyta* blight (Tar'ran et al. 2007; Iruela et al. 2006; Lichtenzveig et al. 2006). However, information on genetic dissection of traits related to drought tolerance is limited.

In present study, to identify QTLs responsible for draught tolerance in chickpea, phenotyping data was collected for two years-2005 and 2007 in three replications. For the year 2005, data was available for ten root traits, but for 2007 data was available for only eight root traits. By analysisng the phenotyping data mentioned above together with the genetic map presented in this study, a total of 12 QTLs were identified for 10 root traits. For eight traits only one QTL was detected but in case of RDW and RSA two QTLs were observed. One interesting thing that has come out from QTL analysis, is the identification of the 'Hot-Spot' region on the LG04. This region was common for all the traits except for RD. The hot-spot region was flanked by

NCPGR91-TAA170-NCPGR127. In fact, similar kind of observations were made by Nayak (2011) while carrying out QTL analysis based on the genetic map with normal SSR marker loci. The marker 'TAA170' was also detected in previous study related to QTL for days to first flowering (DFF) under drought condition by Rehman (2009) and resistance to *Ascochyta* blight in chickpea by Aryamanesh et al. (2009).

For RT, SDW and RL, QTLs were observed in both the years' environment 2005 and 2007. *QTLRT1*, *QTLSDW1*, *QTLRL1* were observed between NCPGR91-TAA170, contributing 18.5%, 34.48%, and 21.43% of phenotypic variation. *QTLSDW1* and *QTLRL1* are considered as major QTL contributing more than 20% of phenotypic variation (Ravi et al. 2011). For breeding purposes, QTL with large additive effect which are stable across environments and which do not depend on epistatic interactions, for instance, *QTLSDW1*, *QTLRL1* are most desirable for chickpea breeding through marker assisted selection.

For RDW and RSA, two QTLs were detected, referred as *QTLRDW1*, *QTLRDW2* and *QTLRSA1*, *QTLRSA2* respectively. *QTLRDW1* and *QTLRSA1* were observed on LG01. *QTLRDW1* was flanked by markers STMS21-cpPb-677672 and *QTLRSA1* was flanked by markers CaM0144-CKAM1971. These two QTLs were explaining 5.59% and 4.84% phenotypic variance and were recoreded as minor QTLs. For RSA one major QTL- *QTLRSA2* was detected on LG04 spanning 4.6 cM and flanked by marker NCPGR91-TAA170 with 24.08% PV. Though, *QTLRDW2* was also located on the LG04 between TAA170-NCPGR127, spanning 5.1cM, considered as minor QTL contributing 11.57% phenotypic variation.

The highest phenotypic variation was explained by *QTLSDW1* for the trait SDW (R^2= 34.48%), where as the lowest phenotypic variation was contributed by *QTLRSA1* for the trait RSA (R^2=4.84%). Six of the 12 QTLs- *QTLLDW1*, *QTLRL1*, *QTLRLD1*, *QTLRSA2*, *QTLStDW1*, *QTLSDW1* are reported as major QTLs which were contributing more >20% phenotypic variance. Based on QTL mapping studies in other species, it can be generalized that higher phenotypic variation for the given trait in the mapping population and high/reasonable marker density genotyping data are the pre-requisites to identify the major QTLs explaining higher phenotypic variation.

In this study, a hot-spot region was detected on LG04, flanked by markers NCPGR91-TAA170, is the same region reported by Nayak 2011. Therefore, the region spanned by 9.7 cM and flanked by NCPGR91-TAA170-NCPGR127 present on LG04, is supported the evidence of presence of QTL hot-spot region in chickpea genome, which is associated with several drought tolerance related traits. This can also be supported by the study done by Xu et al. (2006), who discovered the hot-spot QTL region of sub-mergence tolerance QTL in rice on chromosome 9. This major named *Sub1*, with a LOD score of 36 and an R^2 value of 69 % (Xu and Mackill, 1996), may also contain the presence of several transcription factors. QTLs for complex traits like flowering time, plant architecture, sugar content, fruit weight etc. as identified to be controlled by transcription factors in plant species like Arabidopsis, rice, maize and tomato (Paran and Zamir 2003; Salvi and Tuberosa 2005) may be present in the this region.

5.6 Future Needs and Directions

From last few decades, to increase the yield of chickpea, the breeders around the world were focusing on conventional breeding approaches, but it had not helped them much. In fact the use of molecular markers in improving the breeding efficiency in plant breeding was suggested as early as in 1989 (Tanskley et al. 1989a,b; Melchinger 1990). In this regard, once linkage between a gene for the agronomic trait of interest and marker locus is established, then DNA diagnostic tests can be used to guide plant breeding (Morgante and Salamini 2003; Gupta and Varshney, 2004). Due to the selection of useful lines for breeding with the help of linked molecular markers, thus increasing yield by pyramiding genes for resistance/ tolerance into elite germplasm. Molecular markers are reported to play a crucial role in crop improvement programmes (Varshney et al. 2009d,f). Molecular marker technologies help in improving the efficiency of breeding several fold since selection is not directly on the trait of interest but on the molecular marker tightly linked to the trait, thereby accelerating the generation of new varieties, especially when the characters are difficult to score. Marker assisted selection (MAS) was successfully used for the breeding of resistant soybean to cyst nematode (Diers, 2004), resistant pinto bean to common bacterial blight (Mutlu et al. 2005), and of resistant narrow-leafed lupin (*Lupinus angustifolius* L.) to phomopsis stem blight (Yang et al. 2002) and anthracnose (Yang et

al. 2004). In recent years, the power of functional molecular markers has been demonstrated over normal markers for MAS. Developed transcriptomic resource and functional molecular markers is going to be very useful for chickpea genetics research and breeding applications.

Genetic mapping and trait mapping are pre-requisites for undertaking MAS in crop breeding. Though a few genetic maps were available on intra-specific mapping populations, this study presents probably the most dense intra-specific genetic map for chickpea and that too based on gene-based markers. Molecular mapping, in some earlier studies, has resulted in identification of molecular markers for several important traits like flower colour associated with the SSR marker GAA47 by Cobos et al. (2005), single gene for double pod associated with another SSR marker TA80 by Rajesh et al. (2002) and Cobos et al. (2005). Also some other disease resistance traits molecular mapping has been extensively used in the resistance gene localization of various races of *Fusarium* wilt pathogen (races 1, 2 and 3 by Gowda et al. 2009; Sharma et al. 2004b; races 4 and 5 by Winter et al. 2000). In the the present study, a 'Hot-Spot' region on LG04 was confirmed which is associated with nine root traits responsible for drought tolerance in chickpea. This region was flanked by tightly linked marker interval 'NCPGR91-TAA170-NCPGR127' spanning 9.7 cM on the chromosome and explaining the highest phenotypic variation (34.48%) for root traits. Therefore, this study provides the candidate genomic region for introgression in elite breeding lines for developing drought tolerant lines. Another important area that needs to be followed up is the fine mapping in the hot-spot region and cloning of QTL region to identify the genes responsible for conferring drought tolerance so that the similar set of genes can be used across the legume crop species for improving drought tolerance.

SUMMARY

Chickpea is a third most important food legume in the worldwide. It is generally planted after the main rainy season and grown on stored soil moisture making terminal drought stress a primary constraint to productivity. It is a self pollinating, annual diploid plant (2n=2x=16) with a small genome size of 740 Mbp. Chickpea seeds contain 23% protein, 64% total carbohydrates (47% starch, 6% soluble sugar) and only 3-6% oil and an extremely good source of calcium (190mg/100g), phosphorus (343mg/100g), iron (7mg/100g), zinc (3mg/100g), magnesium and manganese.

Despite of its economical importance chickpea productivity is low because of yield loses due to abiotic stresses such as drought, cold and salinity and biotic stress which includes foliar and soil borne fungal diseases (*ascochyta blight*, *fusarium wilt*, and *botrytis grey mold*), insect pests (*Helicoverpa borer*). The estimated yield losses due to abiotic stress (6.4 million ton) are much more than loss due to biotic stress (4.8 million ton). Of all, drought is a major constraint causing 40-50% reduction in chickpea yield globally. Considering the constraining issues and relative affect on the global yield, it is very crucial to improve drought tolerant genotypes in chickpea for stabilization of the yield. Therefore, improving resistance to biotic and tolerance to abiotic stresses as well as a general increase in dry matter are major aims of chickpea breeders around the world.

Drought is a complex phenomenon, and identifying genes for drought tolerance is more challenging. It is essential to identify genes that confer drought tolerance which can be deployed in breeding programme. The application of a holistic approach “genomics-assisted breeding” which combines genomics with breeding and physiology provides strategies for improving component traits of drought tolerance that should prove more effective and efficient than the

conventional selection methods. It is essential to identify QTLs (quantitative trait loci) or genes that confer drought tolerance and can be deployed in breeding programmes. The importance of root traits in regards to drought tolerance was extensively studied in several crops, and the role of root traits especially root length and root length density were proven to be the traits related to drought tolerance in chickpea from earlier physiological studies held at ICRISAT.

In past, several crops genetic mapping based approaches were used to identify the genes/QTLs for a trait. The objective of genetic mapping is to identify the simply inherited markers in close proximity to genetic factors affecting QTL. This localization relies on the processes that create a statistical association between marker and QTL alleles and processes that selectively reduce that association as a function of the marker distance from the QTL. The marker systems like RFLP, RAPD, AFLP, SSR, ISSR, CAPS, DArT and SNPs brought revolution in applied crop breeding programs. Despite considerable effort in developing molecular markers in chickpea cultivated gene pool especially SSRs and SNPs, low rates polymorphism have limited the number of markers that have been integrated into chickpea genetic maps. The aim of the study is the development of functional markers that are associated with genes that regulates or controls (Transcription factors) the drought tolerance in chickpea and identification of the QTLs responsible for drought related traits.

In view of importance of chickpea and drought tolerance, this study has been undertaken with following objectives in chickpea: (a) Generation of Illumina/Solexa transcript reads (b)Large-scale identification of SNPs (c) Development of marker assays for SNP genotyping (d) Construction of a transcript map (e) Identification of QTLs responsible for drought tolerance.

The conclusions from the research work are summarized as:

(a) A total of 26, 082 SNPs were identified from Illumina/Solexa sequencing of drought stresses genotypes of ICC 4958 and ICC 1882.

(b) In second approach, allele specific re-sequencing of 2-20 genotypes of chickpea, a total of 2,046 SNPs were identified. A novel set of 1,627 genic molecular markers were for genes or transcription factors (TFs) identified in chickpea or other legume species e.g. *Medicago truncatula*, *Medicago sativa*, *Lotus japonicus*, *Lupinus* spp., *Arachis hypogaea*, *Pisum sativum*, *Crotolaria tenuifolia*, *Phaseolus vulgaris*, *Phaseolus coccineus*, *Glycine max*, *Glycine soja*, *Robinia pseudoacacia* and *Trifolium pretense*.

(c) In third approach, 181 *C*hickpea *I*ntron *S*panning *R*egion (CISR) markers were developed by aligning the chickpea unigenes to *Medicago* genome.

(d) In the fourth approach, a set of 96-plex KASPar assay was designed and 56 polymorphic markers were obtained on the parents of mapping population ICC 4958 × ICC 1882. These 56 CKAM markers were successfully genotyped on the 232 RILs of the mapping population using VeraCode assays for BeadXpress system at ICRISAT.

(e) From allele-resquencing 1,139 markers showed scorable amplification on the panel of 5 genotypes and high quality sequences were generated for 532 markers on 2-20 genotypes. Along with these, 87 CISR and 58 EST-SNPs showed scorable amplification.

(f) Diversity analysis of 532 multiple sequence alignment showed a total of 2,046 SNPs in 84,073 bp sequence data generated for 2- 20 genotypes for 264 genes with SNP frequency of 1 SNP per 41.09 bp. Among 2,046 SNPs identified, 964 accounted for transversion where as 1,167 for transition. Apart from this, 71 *InDels* were also identified. The nucleotide diversity index for the chickpea genes ranged from 0.5×10^{-3} to 25.9×10^{-3} (mean = 5.3×10^{-3}), while genes derived from heterologous species showed a higher nucleotide diversity index of 1.0×10^{-3} to 74.0×10^{-3}(mean= 14.6×10^{-3}). Haplotype diversity was higher with heterologous species genes (0.7704) as compared

to chickpea genes (0.5303). PIC values of SNP ranged from 0.11-0.50 with an average of 0.34 across all the 264 genic regions.

(g) SNP2CAPS analysis showed 311 putative CAPS candidates (224 from allele specific re-sequencing and 87 from chickpeaEST database). Out of which, CAPS assay was successful in 152 cases (124 from from allele specific re-sequencing and 28 from chickpeaEST database). All the 152 CAPS candidates, referred as *C*hickpea *G*enic *M*olecular *M*arkers (CGMM), were checked for polymorphism on the panel of 5 genotypes (ICC 4958, ICC 1882, ICC 283, ICC 8261, PI 489777).

(h) 6 CGMMs and 2 CGMMs showed polymorphism for the intra-specific mapping populations namely ICC 4958 × ICC 1882 and ICC 283 × ICC 8261, respectively. Out of 87 CISR markers tested, 21 markers showed polymorphism across 5 genotype, while 3 CISR markers showed polymorphisms for each of two intra-specific mapping populations.

(i) 6 CGMMs and 3 CISRs could be genotyped on 232 RILs of ICC 4958 × ICC 1882. For constructing the transcript map on intra-specific mapping population, 6 CGMMs, 3 CISRs, 56 CKAM and 9 ICCeM (from separate study) markers genotyping data was utilized along with 235 marker loci data from previous study.

(j) The comprehensive transcript map developed here consists of 285 marker loci spanning 595.73 cM with an average inter marker distance of 2.09 cM. This is probably the first transcript map developed so far in intra-specific mapping population.

(k) Genotyping data for the intra-specific mapping population was analyzed together with phenotyping data for drought tolerance traits obtained for two years i.e. 2005 (ten traits) and 2007 (eight traits). The QTL analysis detected 12 significant QTLs (LOD $\geq$ 3) for

the ten root traits using single-locus analysis, of which six were major QTLs and showed more than 20% phenotypic variation.

(l) The QTL analysis revealed the presence of a “QTL hot-spot” region on LG04 that contained QTLs for several drought tolerance traits including shoot dry weight (SDW) explaining upto 38.03% phenotypic variation. This genomic region was found to contain the QTLs for all except for rooting depth (RD) traits.

Due to large scale transcriptomic resource development using next-generation sequencing approaches and high-throughput genotypic platforms, there will surge of molecular markers especially SNP markers in resource-poor crops like chickpea and the genes of interest can be directly tagged and manipulated in near future. The genomic resources in terms of huge sequence data will definitely improve the research abilities in chickpea and can expect the physical map for chickpea in near future. This will enable the linking of known genetic maps with physical maps, and pave the way for development of comparative genome maps. Therefore, this is the first report on development of large-scale genic markers including development of easily assayable markers and transcript map of chickpea. These resources should be useful not only for genome analysis and genetics and breeding applications of chickpea but also for comparative legume genomics.

BIBLIOGRAPHY

- Abbo S, Molina C, Jungmann R, Grusak MA, Berkovitch Z, Reifen R, Kahl G and Winter P (2005) Quantitative trait loci governing carotenoid concentration and weight in seeds of chickpea (*Cicer arietinum* L). Theoretical and Applied Genetics 111:185–195
- Adams MD, Bento Soares M, Kerlavage AR, Fields C and Venter JC (1993) Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. Nature Genetics 4:373–380
- Ahmad F, Gaur P and Croser J (2005) Chickpea (*Cicer arietinum* L) In: Genetic resources, chromosome engineering and crop improvement – grain legumes [eds Singh R, Jauhar P] CRC Press: Boca Raton, FL pp 185–214
- Aiken RM and Smucker AJM (1996) Root system regulation of whole plant growth. Annual Review of Phytopathology 34: 325–346
- Aitken KS, Jackson PA and McIntyre CL (2005) A combination of AFLP and SSR markers provides extensive map coverage and identification of homo(eo)logous linkage groups in a sugarcane cultivar. Theoretical and Applied Genetics 110: 789–801
- Akhunov E, Nicolet C and Dvorak J (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina Golden-Gate assay. Theoretical and Applied Genetics doi:101007/s00122-009-1059-5
- Akhunova AR , Macmil SL , Qu C , Wang P, Wiley GB, Kenton S , Roe AB and Akhunov ED (2009) SNP discovery in polyploid wheat using 454 sequencing technology. http://www.intl-pag.org/17/abstracts/P03e_PAGXVII_144.html
- Alderborn A, Kristofferson A and Hammerling U (2000) Determination of single-nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. Genome Research 10:1249–1258

- Alderborn A, Kristofferson A and Hammerling U (2000) Determination of single-nucleotide polymorphisms by real-time pyrophosphate DNA sequencing. Genome Research 10:1249–1258

- Ali MY, Johansen C, Krishnamurthy L and Hamid A (2005) Genotypic variation in root system of chickpea (*Cicer arietinum* L). Journal of Agronomy and Crop Science 191:464–472
- Ali MY, Krishnamurthy L, Saxena NP, Rupela OP, Kumar J and Johansen C (2002) Scope for genetic manipulation of mineral acquisition in chickpea. Plant and Soil 245: 123–134
- Amede T and Schubert S (2003) Mechanisms of drought resistance in grain legumes. I.Osmotic adjustment. Ethiopia Journal of Science 26: 37–46
- Andersen JR and Lübberstedt T (2003) Functional markers in plants. Trends in Plant Science 8:554–560
- Anwar MR, McKenzie BA and Hill GD (2003) Water-use efficiency and the effect of water deficits on crop growth and yield of Kabuli chickpea (*Cicer arietinum* L) in a cool-temperate sub-humid climate. Journal of Agricultural Science 141:285–301
- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana.* Nature 408:796–815
- Arai-Kichise Y, Shiwa Y, Nagasaki H, Ebana K, Yoshikawa H, Yano M and Wakasa K (2011) Discovery of Genome-Wide DNA Polymorphisms in a Landrace Cultivar of Japonica Rice by Whole-Genome Sequencing. Plant Cell Physiology 52: 274–282
- Arumuganathan K and Earle ED (1991) Estimation of nuclear DNA content of plants by some important plant species. Plant Molecular Biology Reporter 9:208–218
- Aryamanesh N, Nelson MN, Yan G, Clarke HJ and Siddique KHM (2009) Mapping a major gene for growth habit and QTLs for ascochyta blight resistance and flowering time in a population between chickpea and *Cicer reticulatum*. Euphytica 173:307–319
- Asamizu E, Nakamura Y, Sato S and Tabata S (2000a) A large scale analysis of cDNA in *Arabidopsis thaliana*: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries. DNA Research 7:175–180
- Asamizu E, Nakamura Y, Sato S and Tabata S (2000b) Generation of 7137 non-redundant expressed sequence tags from a legume, *Lotus japonicas*. DNA Research 7:127–130

- Ashraf N, Ghai D, Barman P, Basu S, Nagaraju G, Mondol MK, Chakraborty N, Datta A and Chakraborty S (2009) Comparative analyses of genotype-dependent expressed sequence tags and stress-responsive transcriptome of chickpea wilt illustrates predicted and unexpected genes and novel regulators of plant immunity. BMC Genomics 10: 415
- Atkin RK, Barton GE and Robinson DK (1973) Effect of root-growing temperature on growth-substances in xylem exudate of *Zea mays*. Journal of Experimental Botany 24: 475–487
- Azam S, Thakur V, Pradeep R, Shah Trushar, Jayashree B, BhanuPraksh B, Farmer AD, Studholme DJ, May GD, Edwards D, Jones DGJ and Varshney RK (2012) Coverage based consensus calling (CBCC) of short sequence reads and comparison of CBCC results to identify SNPs in chickpea (*Cicer arietinum*; Fabacea) a crop science without a genome. American Journal of Botany 99: 1–7
- Azhaguvel P, Vidya SD, Sharma A and Varshney RK (2006) Methodological advancement in molecular markers to delimit the gene(s) for crop improvement. In: Texiera da Silva J (ed) Floriculture, Ornamental and Plant Biotechnology Advances and Tropical Issues, Global Science Books, London, pp 460–99
- Barbazuk WB, Emrich SJ, Chen HD, Li L and Schnable PS (2007) SNP discovery via 454 transcriptome sequencing. Plant Journal 51:910–918
- Beavis WD and Kein P (1996). In: Genotype-by-Environment Interaction (eds) Kang MS, Gauch HJ, CRC Press, Boca Raton, USA, pp 123–149
- Beebe SE, Rojas-Pierce M, Yan XL, Blair MW, Pedraza F, Munoz F, Tohme J and Lynch JP (2006) Quantitative trait loci for root architecture traits correlated with phosphorus acquisition in common bean. Crop Science 46:413–423
- Bell CJ, Dixon RA, Farmer AD, Flores R, Inman J, Gonzales RA, Harrison MJ, Paiva NL, Scott AD, Weller JW and May GD (2001) The *Medicago* Genome Initiative: a model legume database. Nucleic Acids Research 29:114–7
- Berger J, Turner NC and French RJ (2003) The role of phenology in adaptation of chickpea to drought In: Solution for a better environment In: Proceeding of the 11th Australian Agronomy Conference, Geelong, Victoria, February 2003, pp 1–4

- Berger JD, Abbo S and Turner NC (2003) Ecogeography of annual wild *Cicer* species: the poor state of the world collection. Crop Science 43: 1076–1090
- Blum A (1988) Plant breeding for stress environment Boca Raton, FL, USA, CRC Press
- Blum A and Nguyen HT (2004) Physiology and biotechnology integration for plant breeding: Epilogue. In: Nguyen HT, Blum A (eds) Physiology and Biotechnology Integration for Plant Breeding, Marcel Dekker, Inc. New York, USA
- Bohnert HJ, Nelson DE and Jenson RG (1995) Adaptations to environmental stresses. The Plant Cell 7:1099–1111
- Bohra A, Dubey A, Saxena RK, Penmetsa RV, Poornima KN, Kumar N, Farmer AD, Srivani G, Upadhyaya HD, Gothalwal R, S Ramesh, Singh D, Saxena KB, PB Kishor PBK, Singh NK, Town CD, May GD, Cook DR and Varshney RK (2011) Analysis of BAC-end sequences (BESs) and development of BES-SSR markers for genetic mapping and hybrid purity assessment in pigeonpea (*Cajanus* spp.). BMC Plant Biology 11:56
- Boominathan P, Shukla R, Kumar A, Manna D, Negi D, Verma PK and Chattopadhyay D (2004) Long Term Transcript Accumulation during the Development of Dehydration Adaptation in Cicer arietinum. Plant Physiology 135: 1608–1620
- Borevitz JO, Liang D, Plouffe D, Chang HS, Zhu T, Weigel D, Berry CC, Winzeler E and Chory J (2003) Large-scale identification of single-feature polymorphisms in complex genomes. Genome Research 13:513–523
- Bruce WB, Edmeades GO and Barker TC (2002) Molecular and physiological approaches to maize improvement for drought tolerance. Journal of Experimental Botany 53:13–25
- Bryan GJ, Stephenson P, Collins A, Kirby J, Smith JB and Gale MD (1999) Low levels of DNA sequence variation among adapted genotypes of hexaploid wheat. Theoretical and Applied Genetics 99:192–198
- Buhariwalla HK, Jayashree B, Eshwar K and Crouch JH (2005) Development of ESTs from chickpea roots and their use in diversity analysis of the *Cicer* genus. BMC Plant Biology 5:16
- Carson DL and Botha FC (2000) Preliminary analysis of expressed sequence tags for sugarcane. Crop Science 40:1769–1779

- Champoux MC, Wang G, Sarkarung S, Mackill DJ, O'Toole JC, Huang N and McCouch S R (1995) Locating genes associated with root morphology and drought avoidance in rice via linkage to molecular markers. Theoretical and Applied Genetics 90:969–981
- Chandra Babu R, Nyugen BD, Chamarerk V, Shanugasundaram P, Chezian P, Jeyaprakash P, Ganesh SK, Palchamy A, Sadasivam S, Sarkarung S, Wade LJ and Nyugen HT (2003) Genetic analysis of drought resistance in rice by molecular markers. Crop Science 43:1457–1469
- Charles MT, Dominique R, Kumar J and Dangi OP (2002) A preliminary study of the functional properties of chickpea leaves In: Annual Meeting of the Canadian Society of Food and Nutrition, May 2002, Edmonton, Alberta, Canada
- Chaves MM, Maroco JP and Pereira JS (2003) Understanding plant responses to drought- from genes to the whole plant. Functional Plant Biology 30: 239–264
- Chelkowski J and Stepien L (2001) Molecular markers for leaf rust resistance genes in wheat. Journal of Applied Genetics 42:117–126
- Chen M, Presting G, Barbazuk W, Goicoechea J, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Higingbottom S, Phimphilai J, Phimphilai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Sun S, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Bancroft I, Salse J, Regad F, Mohapatra T, Singh, N, Tyagi A, Soderlund C, Dean R and Wing R (2002) An integrated physical and genetic map of the rice genome. Plant Cell 14:537–545
- Chen W , Mingus J , Mammadov J, Backlund JE, Greene T, Thompson S and Kumpatla S (2010) KASPar: A simple and cost-effective system for SNP genotyping. Plant & Animal Genomes XVIII Conference, January 9-13, San Diego, CA (http://www.intl-pag.org/18/abstracts/P03e_PAGXVIII_194.html)
- Chen W and Singh KB (1999) The auxin, hydrogen peroxide and salicylic acid induced expression of the Arabidopsis GST6 promoter is mediated in part by an ocs element. Plant Journal 19: 667–677
- Cheung F, Haas BJ, Goldberg MD, May GD, Xiao Y and Town CD (2006) Sequencing *Medicago truncatula* expressed sequenced tags using 454 life sciences technology. BMC Genomics 7: 272

- Ching A and Rafalski A (2002) Rapid genetic mapping of ESTs using SNP pyrosequencing and indel analysis. Cellular and Molecular Biology Letters 7:803–810
- Cho S, Chen W and Muehlbauer FJ (2004) Pathotype specific genetic factors in chickpea (*Cicer arietinum* L.) for quantitative resistance to *Ascochyta* blight. Theoretical and Applied Genetics 109:733–739
- Cho S, Kumar J, Schultz JL, Anupama K, Tefera F and Muehlbauer FJ (2002) Mapping genes for double podding and other morphological traits in chickpea. Euphytica 128:285–292
- Cho YG, Ishii T, Temmykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N and Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and genbank sequences in rice. Theoretical and Applied Genetics 100:713–722
- Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P,Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB and Cook DR (2004a) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M sativa*. Genetics 166:1463–1502
- Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, EllisN, Doyle J, Kiss GB, Young ND and Cook DR (2004b) Estimating genome conservation between crop and model legume species. Proceedings of National Academy of Sciences, United States of America 101:15289–15294
- Choi IY, Hyten DL, Matukumalli LK, Song Q, Chaky JM, Quigley CV, Chase K, Lark KG, Reiter RS, Yoon MS, Hwang EY, Yi SI, Young ND, Shoemaker RC, van Tassel CP, Specht JE and Cregan PB (2007) A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. Genetics 176:685–696
- Choudhary S, Sethy NK, Shokeen B and Bhatia S (2006) Development of sequence tagged microsatellite site markers for chickpea (*Cicer arietinum* L). Molecular Ecology Notes 6:93–95

- Choudhary S, Sethy NK, Shokeen BA and Bhatia S (2009) Development of chickpea EST-SSR markers and analysis of allelic variation across related species. Theoretical and Applied Genetics118:591–608
- Churchill GA and Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:963–971
- Cichy KA, Snapp SS and BlairMW (2009) Plant growth habit, root architecture traits and tolerance to low soil phosphorus in an Andean population. Euphytica 165: 257–268
- Clarke H and Siddique KHM (2003) Chilling tolerance in chickpea novel methods for crop improvement. In: Sharma RN, Yasin M, Swami SL, Khan MA, William AJ (ed), International Chickpea Conference, Indira Ghandhi Agricultural University, India, Raipur, India, pp 5–12
- Close TJ, Bhat PR, Lonardi S, Wu Y, Rostoks N, Ramsay L, Druka A, Stein N, Svensson JT, Wanamaker S, Bozdag S, Roose ML, Moscou MJ, Chao S, Varshney RK, Szűcs P,Sato S, Hayes PM, Matthews DE, Kleinhofs A, Muehlbauer GJ, Young JD, Marshall DF, Madishetty K, Fenton RD, Condamine P, Graner A and Waugh R (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10:582
- Cobos MJ, Fernández MJ, Rubio J, Kharrat M, Moreno MT, Gil J and Millán T (2005) A linkage map of chickpea (*Cicer arietinum* L.) based on populations from Kabuli × Desi crosses: Location of a resistance gene for Fusarium wilt race 0. Theoretical and Applied Genetics 110:1347–1353
- Cobos MJ, Rubio J, Strange RN, Moreno MT, Gil J and Millan T (2006) A new QTL for *Ascochyta blight* resistance in an RIL population derived from an interspecific cross in chickpea. Euphytica 149:105–111
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M and Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452: 215–219
- Collard BCY, Pang ECK, Ades PK and Taylor PWJ (2003) Preliminary investigation of QTLsassociated with seedling resistance to Ascochyta blight from *Cicer*

echinospermum, a wild relative of chickpea. Theoretical and Applied Genetics 107:719–729

- Coram TE and Pang ECK (2005) Isolation and analysis of candidate *ascochyta* blight defence genes in chickpea Part I Generation and analysis of an expressed sequence tag (EST) library. Physiological and Molecular Plant Pathology 66:192–200
- Coram TE, Mantri NL, Ford R and Pang ECK (2007) Functional genomics in chickpea: an emerging frontier for molecular-assisted breeding. Functional Plant Biology 34: 861–873
- Covitz PA, Smith LS and Long SR (1998) Expressed sequence tags from a root-hair-enriched Medicago truncatula cDNA library. Plant Physiology 117:1325–1332
- Croser JS, Clarke HJ, Siddique KHM and Khan TN (2003) Low temperature stress: implications for chickpea (*Cicer arietinum* L) improvement. Critical Review in Plant Science 22:185–219
- Cuc LM, Mace E, Crouch J, Quang VD, Long TD and Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea*). BMC Plant Biology 8:55
- Davies WJ and Zhang J (1991) Root signals and the regulation of growth and development of plants in drying soil. Annual Review of Plant Physiology and Plant Molecular Biology 42:55–76
- Davis GL, Mcmullen MD, Baysdorfer C, Musket T, Grant D, Staebelld M, Xua G, Polaccoa M, Kosterd L, Melia-Hancockc S, Houchinsa K,Chaoc S, and Coe EH (1999) A maize map standard with sequenced core markers, grass genome reference points, and 932 expressed sequence tagged sites (ESTs) in a 1736 locus map. Genetics 152:1137–1172
- Deulvot C, Charrel H, Mart A, Jacquin F, Donnadieu C, Lejeune-Hénaut I , Burstin J and Aubert G (2010) Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. BMC Genomics 11: doi:101186/1471-2164-11-468

- Diers BW (2004) Marker assisted selection in breeding self-pollinated crop plants. In: Goodman RM (ed.) Encyclopedia of Plant and Crop Science. Marcel Dekker, Inc., New York, pp 202–204
- Dita MA, Rispial N, Prats E, Rubiales D and Singh KB (2006) Biotechnology approaches to overcome biotic and abiotic stress constraints in legumes. Euphytica 147:1–24
- Doerge RW and Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–294
- Doyle JJ and Luckow MA (2003) The rest of the iceberg legume diversity and evolution in a phylogenetic context. Plant Physiology 131: 900–910
- Du Jinyou, Xiaoyang C Wei L and Qiaong G (2004) Osmoregulation Mechanism of drought stress and Genetic Engineering Strategies for Improving Drought Resistance in Plants. Forestry studies in China 6: 56–62
- Dubey A, Farmer A, Schlueter J, Cannon S, Abernathy B, Tuteja R,Woodward J, Shah T, Mulasmanovic B, Kudapa H, Raju NL, Gothalwal R, Pande S, Xiao Y, CD Town, NK Singh, May GD, Jackson S and Varshney RK (2011) Defining the transcriptome assembly and its use for genome dynamics and transcriptome profiling studies in pigeonpea (*Cajanus cajan* L.). DNA Research doi:10.1093/dnares/dsr007
- Dudley JW (1993) Molecular markers in plant improvement: manipulation of genes affecting quantitative traits. Crop Science 33:660-668 Lee M (1995) DNA markers and plant breeding programs. Advances in Agronomy 55:265–344
- Duke JA (1981) Handbook of legumes of world economic importance. Plenum Press NewYork
- Dunwell JM (2000) Transgenic approaches to crop improvement. Journal of Experimental Botany 51:487–496
- Durstewitz G, Polley A, Plieske J, Luerssen H, Graner EM, Wieseke R and Ganal MWV (2010) SNP discovery by amplicon sequencing and multiplex SNP genotyping in the allopolyploid species *Brassica napus*. Genome 53: 948–956

- Endo M, Kokubun T, Takahata Y, Higashitani A, Tabata S and Watanabe M (2000) Analysis of expressed sequence tags of flower buds in *Lotus japonicas*. DNA Research 7: 213–216
- Eujayl I, Sorrells M, Baum M, Wolters P and Powell W (2001) Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. Euphytica 119:39–43
- Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, Hansen M, Steemers F, Butler SL, Deloukas P, Galver L, Hunt S, Mcbride C, Bibikova M, Rubano T, Chen J, Wickham E, Doucet D, Chang W, Campbell D, Zhang B, Kruglyak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J and Chee MS (2003) Highly parallel SNP genotyping. Cold Spring Harbor Symposium on Quantitative Biology 68:69–78
- Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, Hansen M, Steemers F, Butler SL, Deloukas P, Galver L, Hunt S, Mcbride C, Bibikova M, Rubano T, Chen J, Wickham E, Doucet D, Chang W, Campbell D, Zhang B, Kruglyak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J and Chee MS (2003) Highly parallel SNP genotyping. Cold Spring Harbor Symposia on Quantitative Biology 68:69–78
- FAO (2009) http://www.fao.org. Production database
- Fatokun CA, Menancio-Hautea DI, Danesh D and Young ND (1992) Evidence for orthologous seed weight genes in cowpea and mung bean based on RFLP mapping. Genetics 132:841–846
- Feltus FA, Singh HP, Lohithaswa HC, Schulze SR, Silva TD and Paterson AH (2006) A comparative genomics strategy for targeted discovery of single-nucleotide polymorphisms and conserved-noncoding sequences in orphan crops. Plant Physiology 140:1183–1191
- Flandez-Galvez H, Ford R, Pang ECK and Taylor PWJ (2003) An intraspecific linkage map of the chickpea (*Cicer arietinum* L.) genome based on sequence-tagged microsatellite site and resistance gene analogue markers. Theoretical and Applied Genetics106:1447–1456
- Forment J, Gadea J, Huerta L, Abizanda L, Agusti J, Alamer S, Alos E, Andres F,Arribas R, Beltran JP, Beltran JP, Berbel A, Blazquez MA, Brumos J, Canas LA, Cercos M,

Colmenero-Flores JM, Conesa A, Estables B, Gandia M, Garcia-Martinez JL, Gimeno J, Gisbert A, Gomez G, Gonsalez-Candelas L, Granell A, Guerri J, Lafuente MT, Madueno F, Maros JF, Marques MC, Martinez F, Martinez-Godoy MA, Miralles S, Moreno P, Navarro L, Pallas V, Amador-Perez MA, Perez-Valle J, Pons C, Rodrigo I, Rodriguez PL, Royo C, Serrano R, Soler G, Tadeo F, Talon M, Terol J, Trenor M, Vaello L, Vincente O, Vidal Ch, Zacarias L and Conejero V (2005) Development of a citrus genome-wide EST collection and cDNA microarray as resources for genomic studies. Plant Molecular Biology 57:375–391

- Foster MW and Sharp RR (2004) Beyond race: toward a whole genome perspective on human populations and genetic variation. Nature Reviews Genetics 5: 790–796
- Frahm MA, Rosas JC, Mayek-Pérez N, López-Salinas E, Acosta-Gallegos JA and Kelly JD (2004) Breeding beans for resistance to terminal drought in the lowland tropics. Euphytica 136:223–232
- Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB and Tanksley SD (2000) fw2.2: A quantitative trait locus key to the evolution of tomato fruit size. Science 289:85–88
- Fredslund J, Madsen LH, Hougaard BK, Nielsen AM, Bertioli D, Sandal N, Stougaard J and Schauser L (2006b) A general pipeline for the development of anchor markers for comparative genomics in plants. BMC Genomics 7:207
- Fredslund J, Madsen LH, Hougaard BK, Sandal N, Stougaard J, Bertioli D and Schauser L (2006a) GeM prospector – online design of cross-species genetic marker candidates in legumes and grasses. Nucleic Acids Research 34: 670–675
- Fridman E, Pleban T and Zamir D (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proceedings of the National Academy of Sciences, USA 97:4718–4723
- Gaitán-Solís E, Choi IY, Quigley C, Cregan P and Tohme J (2008) Single nucleotide polymorphisms in common bean: their discovery and genotyping using a multiplex detection system. The Plant Genome 1:125–134

- Galeano CH, Fernández AC, Gómez1 M and Blair MW (2009) Single strand conformation polymorphism based SNP and Indel markers for genetic mapping and synteny analysis of common bean (*Phaseolus vulgaris* L). BMC Genomics 10:629
- Garg R, Patel RK, Jhanwar S, Priya P, Bhattacharjee A, Yadav G, Bhatia S, Chattopadhyay D, Tyagi AK and Jain M (2011) Gene discovery and tissue-specific transcriptome analysis in chickpea with massively parallel pyrosequencing and web resource development. Plant Physiology Preview doi:10.1104/pp.111.178616
- Gaur PM and Slinkard AE (1990a) Inheritance and linkage of isozyme coding genes in chickpea. Journal of Heredity 81:455–461
- Gaur PM and Slinkard AE (1990b) Genetic control and linkage relations of additional isozyme markers in chickpea. Theoretical and Applied Genetics 80: 648–656
- Gaur PM, Krishnamurthy L and Kashiwagi J (2008) Improving drought-avoidance root traits in chickpea (*Cicer arietinum* L) – current status of research at ICRISAT Plant Production. Science 11:3–11
- Gaur R, Sethy NK, Choudhary S, Shokeen B, Gupta V, et al (2011) Advancing the STMS genomic resources for defining new locations on the intra-specific genetic linkage map of chickpea (*Cicer arietinum* L) BMC Genomics 12: 117
- Geervani P (1991) Utilization of chickpea in India and scope for novel and alternative uses In: Uses of Tropical Grain Legumes: Proceedings of Consultants Meeting 27-30 March, 1989 Andhra Pradesh,India, Patancheru: ICRISAT Center, pp 47–54
- Gong L, Stift G, Kofler R, Pachner M and Lelley T (2008) Microsatellites for the genus Cucurbita and an SSR-based genetic linkage map of *Cucurbita pepo* L. Theoretical and Applied Genetics 117:37–48
- Gowda SJ, Radhika MP, Kadoo NY, Mhase LB and Gupta VS (2009) Molecular mapping of wilt resistance genes in chickpea. Molecular Breeding 24: 177–183
- Gujaria N, Kumar A, Dauthal P , Dubey A, Hiremath P, Bhanu Prakash A, Farmer A, Bhide M, Shah T, Gaur PM , Upadhyaya HD , Bhatia S, Cook DR, May GD, Varshney RK (2011) Development and use of genic molecular markers (GMMs) for construction

of a transcript map of chickpea (*Cicer arietinum* L.) Theoretical and Applied Genetics122:1577–1589

- Gupta PK (2009) Single-molecule DNA sequencing technologies for future genomics research. Trends in Biotechnology 26: 602–611
- Gupta PK and Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Functional and Integrative Genomics 4:139–162
- Gupta PK and Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185
- Gupta PK, Mir RR and Kumar J (2008) Wheat genomics: present status and future prospects. International Journal of Plant Genomics 896451
- Gupta PK, Varshney RK and Prasad M (2002) Molecular markers: principles and methodology. In: (eds) Jain SM, Ahloowalia BS and Brar DS, Molecular Techniques in Crop Improvement, Kluwer Academic Publishers, The Netherlands, pp 9–54
- Hall AE (1993) Is dehydration tolerance relevant to genotypic difference in leaf senescence and crop adaption to dry environments? In: (eds) Close TJ, Bray EA, Plant response to cellular dehydration during environmental stress. The American society of plant physiologists
- Hanriot L, Keime C, Gay N, Faure C, Dossat C, Wincker P, Scote-Blachon C, Peyron C and Gandrillon O (2008) A combination of LongSAGE with Solexa sequencing is well suited to explore the depth and the complexity of transcriptome. BMC Genomics 9: 418
- Harushima Y, Yano M, Shomura A, Sato M, Shimano T, Kuboki Y, et al (1998) A high-density rice genetic linkage map with 2275 markers using a single F2 population. Genetics 148:479–494
- Helbaek H (1959) Domestication of food plants in the old world Science 130: 365–372
- Hiremath PJ, Farmer A, Cannon SB, Woodward J, Kudapa H, Tuteja R, Kumar A, BhanuPrakash A, Mulaosmanovic B, Gujaria N, Krishnamurthy LK, Gaur PM, KaviKishor PB, Shah T, Srinivasan R, Lohse M, Xiao Y, Town CD, Cook DR, May GD and Varshney RK (2011) Large-scale transcriptome analysis in chickpea (*Cicer*

arietinum L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. Plant Biotechnology Journal 9:922–931

- Ho CL, Kwan YY, Choi MC, Tee SS, Ng WH, Lim KA, Lee YP, Ooi SE, Lee WW, Tee JM, Tan SH, Kulaveerasingam H, Alwee SSRS and Abdullah MO (2007) Analysis and functional annotation of expressed sequence tags (ESTs) from multiple tissues of oil palm (*Elaeis guineensis* Jacq). BMC Genomics 8:381
- Hou X, Li L, Peng Z, Wei B, Tang S, Ding M, Liu J, Zhang F, Zhao Y, Gu H and Qu LJ (2010) A platform of high-density INDEL/CAPS markers for map-based cloning in Arabidopsis Plant Journal 63: 880–888
- Huang X and Madan A (1999) CAP3: a DNA sequence assembly program. Genome Research 9:868–877
- Huang X,Wang J,Aluru S, Yang S and Hillier L (2003) PCAP: A whole-genome assembly program. Genome Research 13: 2164–2170
- Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular Ecology Resources 8:3–17
- Hüttel B, Winter P, Weising K, Choumane W, Weigand F and Kahl G (1999) Sequence-tagged microsatellite markers for chickpea (*Cicer arietinum* L). Genome 42:210–217
- Hyten DL, Song O, Fickus EW, Quigley CV, Lim JS, Choi IY, Hwang EY, Pastor Corrales M and Cregan PB (2010) High-throughput SNP discovery and assay development in common bean. BMC Genomics 11: 475
- Hyten DL, Song Q, Choi I-Y, Yoon M-S, Specht JE, Matukumalli LK, Nelson RL, Shoemaker RC, Young ND and Cregan PB (2008) High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. Theoretical and Applied Genetics 116:945–952
- Ignacimuthu S and Prakash S (2006) Agrobacterium-mediated transformation of chickpea with α-amylase inhibitor gene for insect resistance. Journal of Bioscience 31: 339–345
- Iruela M, Rubio J, Barro F, Cubero JI, Millán T and Gil J (2006) Detection of two quantitative trait loci for resistance to *Ascochyta* blight in an intra-specific cross of

chickpea (*Cicer arietinum* L.): development of SCAR markers associated with resistance. Theoretical and Applied Genetics112:278–287

- Jacomini E, Bertani A and Mapelli S (1988) Accumulation of polyethylene glycol 6000 and its effects on water content and carbohydrate level in water-stressed tomato plant. Canadian Journal of Botany 66:970–973
- Jayashree B, Bhanuprakash A, Jami A, Reddy SP, Nayak S and Varshney RK (2009) Perl module and PISE wrappers for the integrated analysis of sequence data and SNP features. BMC Research Notes 2:92
- Jayshree B, Buhariwalla HK, Shinde S and Crouch JH (2005) A legume genomics resource: the chickpea root expressed sequence tag database. Electronic Journal of Biotechnology 8:129–133
- Johnson GC, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, Ueda H, Cordell HJ, Eaves IA, Dudbridge F, Twells RC, Payne F, Hughes W, Nutland S, Stevens H, Carr P, Tuomilehto-Wolf E, Tuomilehto J, Gough SC, Clayton DG and Todd JA (2001) Haplotype tagging for the identification of common disease genes. Nature Genetics 29: 233–237
- Journet E, Tuinen DV, Gouzy J, Crespeau H, Carreau V, Farmer MJ, Niebel A, Schiex T, Jaillon O, Chatagnier O, Godiard L, Micheli F, Kahn D, Gianinazzi-Pearson V and Gamas P (2002) Exploring root symbiotic programs in the model legume *Medicago truncatula* using EST analysis. Nucleic Acids Research 30:5579–5592
- Kahl G, Mast A, Tooke N, Shen R and van den Boom D (2005) Single nucleotide polymorphisms In: (eds) Meksem K, Kahl G, The handbook of plant genome mapping: genetic and physical mapping, Wiley-VCH, pp 75–104
- Kar S, Basu D, Das S, Ramakrishnan NA, Mukherjee P, Nayak P and Sen SK (1997) Expression of cryIA (c) gene of Bacillus thuringiensis in transgenic chickpea plants inhibit development of pod-borer (*Heliothis armigera*) larvae. Transgenic Research 6: 177–185
- Kashiwagi J, Krishnamurthy L, Crouch JH and Serraj R (2005) Variability of Root length density and its contributions to seed yield in chickpea (*Cicer arietinum* L) under terminal drought stress. Field Crops Research 95:171–181

- Kashiwagi J, Krishnamurthy L, Crouch JH and Serraj R (2006) Variability of root length density and its contributions to seed yield in chickpea (*Cicer arietinum* L.) under terminal drought stress. Field Crops Research 95:171–181
- Kashiwagi J, Krishnamurthy L, Serraj R, Upadhyaya HD, Krishna SH, Chandra S and Vadez V (2005) Genetic variability of drought-avoidance root traits in the mini-core germplasm collection of chickpea (*Cicer arietinum* L.). Euphytica 146: 213–222
- Kazan K, Muehlbauer FJ, Weeden NF and Ladizinsky G (1993) Inheritance and linkage relationships of morphological and isozyme loci in chickpea (*Cicer arietinum* L) Theoretical and Applied Genetics 86: 417–426
- Keim P, Diers BW, Olson TC and Shoemaker RC (1990) RFLP Mapping in soybean: association between marker loci and variation in quantitative traits. Genetics 126:735–742
- Kholova J, Hash CT, Lava Kumar P, Yadav RS, Kocova M and Vadez V (2010) Terminal drought-tolerant pearl millet [*Pennisetum glaucum* (L.) R. Br.] have high leaf ABA and limit transpiration at high vapour pressure deficit. Journal of Experimental Botany 61:1431–1440
- Koebner RMD (2004*)* Marker assisted selection in the cereals. The dream and the reality. In: Gupta PK, Varshney RK (ed.) Cereal Genomics, pp 317–329
- Konieczny A and Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant Journal 4:403–410
- Kota R, Varshney RK, Prasad M, Zhang H, Stein N and Graner A (2008) EST-derived single nucleotide polymorphism markers for assembling genetic and physical maps of the barley genome. Functional and Integrative Genomics 8:223–233
- Kota R, Varshney RK, Thiel T, Dehmer KJ and Graner A (2001) Generation and comparison of EST-Derived SSRs and SNPs in Barley (*Hordeum Vulgare* L). Hereditas 135:145–151
- Kottapalli P, Gaur P M, Katiyar SK, Crouch JH, Buhariwalla HK, Pande S and Gali KK (2009) Mapping and validation of QTLs for resistance to an Indian isolate of *Ascochyta* blight pathogen in chickpea. Euphytica 165:79–88

- Krishnamurthy L, Vadez V, Jyotsna Devi M, Serraj R, Nigam SN, Sheshshayee MS, Chandra S and Aruna R (2007) Variation in transpiration efficiency and its related traits in a groundnut (*Arachis hypogaea* L.) mapping population. Field Crops Research 103:189–197
- Kumar J and Abbo S (2001) Genetics of flowering time in chickpea and its bearing on productivity in semiarid environments. Advances in Agronomy 72:107–138
- Kumar J and Rao BV (1996) Super early chickpea developed at ICRISAT Asia Center. International Chickpea and Pigeon pea Newsletter 3:17–18
- Kumar J and Rheenen HA (2000) A Major Gene for Time of Flowering in Chickpea. Journal of Heredity 91: 67–68
- Kumar J, Pannu RK and Rao BV (2001) Development of a short-duration chickpea for the subtropics. International Chickpea and Pigeonpea Newsletter 8:7–8
- Kurata N, Umehara Y, Tanoue H and Sasaki T (1997) Physical mapping of the rice genome with YAC clones. Plant Molecular Biology 35:101–113
- Labdi M, Robertson LD, Singh KB and Charrier A (1996) Genetic diversity and phylogenetic relationships among the annual *Cicer* species as revealed by isozyme polymorphisms. Euphytica 88:181–188
- Larher F, Leport L, Petrivalsky M and Chappart M (1993) Effectors for the osmo induced praline response in higher plants. Plant Physiology and Biochemistry 31:911–922
- Lawrence CJ, Dong Q, Polacco ML, Seigfried TE and Brendel V (2004) Maize GDB, the community database for maize genetics and genomics. Nucleic Acids Research 32:393–397
- Leport L, Turner NC, Davies SL and Siddique KHM (2006) Variation in pod production and abortion among chickpea cultivars under terminal drought. European Journal of Agronomy 24: 236– 246
- Li W, Sun D, Du Y, Chen Q, Zhang Z, Qiu L and Sun G (2007) Quantitative trait loci underlying the development of seed composition in soybean (*Glycine max* L. Merr.). Genome 50:1067–1077

- Li YD, Wang YJ, Tong YP, Gao JG, Zhang JS and Chen SY (2005) QTL mapping of phosphorus deficiency tolerance in soybean (*Glycine max* L. Merr.). Euphytica 142:137–142
- Lichtenzveig J, Bonfil DJ, Zhang HB, Shtienberg D and Abbo S (2006) Mapping quantitative trait loci in chickpea associated with time to flowering and resistance to *Didymella rabiei* the causal agent of *Ascochyta blight*. Theoretical and Applied Genetics113:1357–1369
- Lichtenzveig J, Scheuring C, Dodge J, Abbo S and Zhang HB (2005) Construction of BAC and BIBAC libraries and their applications for generation of SSR markers for genome analysis of chickpea, *Cicer arietinum* L. Theoretical and Applied Genetics 110:492–510
- Lister R, Gregory BD and Ecke JR (2009) Next is now: new technologies for sequencing of genomes, transcriptomes and beyond. Current Opinion in Plant Biology 12: 107–118
- Lokko Y, Anderson JV and Rudd S et al. 2007. Characterization of an 18,166 EST dataset for cassava (*Manihot esculenta* Crantz) enriched for drought-responsive genes. Plant Cell Reports 26:1605–1618
- Luo M, Dang P, Guo BZ, He G, Holbrook CC, Bausher MG and Lee RC (2005) Generation of Expressed Sequence Tags (ESTs) for gene discovery and marker development in cultivated peanut. Crop Science 45:346–353
- Lynch M and Walsh B (1998) Genetics and analysis of quantitative traits. Sunderland, MA, Sinauer
- Mackill DJ, Nguyen HT and Zhang J (1999) Use of molecular markers in plant improvement programs for rainfed lowland rice. Field Crop Research 64:177–185
- Mantri NL, Ford R, Coram TE and Pang CKE (2007) Transcriptional profiling of chickpea genes differentially regulated in response to high-salinity, cold and drought. BMC Genomics 8:303
- Mardis ER (2008) Next-generation DNA sequencing methods. Annual Review of Genomics and Human Genetics 9: 387–402
- Matsui T and Sing BB (2003) Root characteristics in cowpea related to drought tolerance at the seedling stage. Journal of Experimental Agriculture 39:29–38

- May GD, Lekha PT, Kashiwagi J, Huntley JJ, Farmer AD, Cook DR and Varshney RK (2008) Whole transcriptome shotgun sequencing for variant detection and transcript profiling in chickpea (*Cicer arietinum* L). In Plant Animal Genomes XVI Conf San Diego USA P385 (http://wwwintl-pagorg/16/abstracts/PAG16_P05f_385html)
- Mayer MS, Tullu A, Simon CJ, Kumar J, Kaiser WJ, Kraft JM and Muehlbauer FJ (1997) Development of a DNA marker for fusarium wilt resistance in chickpea. Crop Science 37:1625–1629
- McIntosh GH and Topping DL (2000) Food legumes in human nutrition In: Knight R (ed), Linking Research and Marketing Opportunities for Pulses in the 21st Century. Kluwer Academic Publishers, Dordrecht, The Netherlands pp 655–666
- Mefford HC, Cooper GM and Zerr T (2009) A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. Genome Research 19:1579–1585
- Melchinger AE (1990) Use of molecular markers in breeding for oligogenic disease resistances. Plant Breeding 104:1–19
- Menancio-Hautea D, Fatokun C, Kumar L, Danesh D and Young N (1993) Comparative genome analysis of mungbean (*Vigna radiata* [L] Wilczek) and cowpea (*V unguiculata* [L] Walpers) using RFLP mapping data. Theoretical and Applied Genetics 86: 797–810
- Meyer E, Davies S, Wang S, Willis BL, Abrego D, Juenger TE and Matz MV (2009) Genetic variation in responses to a settlement cue and elevated temperature in the reef-building coral *Acropora millepora*. Marine Ecology Progress Series 392: 81–92
- Millan T, Clarke HJ, Siddique KHM, Buhariwalla HK, Gaur PM, Kumar J, Gil J, Kahl G and Winter P (2006) Chickpea molecular breeding: new tools and concepts. Euphytica 147:81–103
- Millan T, Rubio J, Iruela M, Daly K, Cubero JI and Gil J (2003) Markers associated with *Ascochyta* blight resistance in chickpea and their potential in marker-assisted selection. Field Crops Research 84:373–384
- Millan T, Winter P, Jungling R, Gil J, Rubio J, Cho S, Cobos MJ, Iruela M, Rajesh PN, Tekeoglu M, Kahl G and Muehlbauer FJ (2010) A consensus genetic map of chickpea (*Cicer arietinum* L) based on 10 mapping populations. Euphytica 175:175–189

- Mitra J (2001) Genetics and genetic improvement of drought resistance in crop plants. Current Science 80:758–763
- Möhring S, Horstmann V and Esch E (2005) Development of a molecular CAPS marker for the self-incompatibility locus in *Brassica napus* and identification of different S alleles. Plant Breeding 124:105–110
- Moinuddin and Chopra RK (2004) Osmotic Adjustment in Chickpea in relation to seed yield and yield parameters. Crop Science 44: 449–455
- Morgante M and Salamini F (2003) From genomics to breeding practice. Current Opinion in Biotechnology 14: 214–219
- Muchero W, Diop NN, Bhat PR, Fenton RD, Wanamaker S, Pottorff M, Hearne S, Cisse N, Fatokun C, Ehlers JD, Roberts PA and Close TJ (2009) A consensus genetic map of cowpea [*Vigna unguiculata* (L) Walp] and synteny based on EST-derived SNPs. Proceedings of the National Academy of Sciences of the United States of America 106:18159–64
- Muehlbauer FJ and Kaiser WJ (1994) Using host plant resistance to manage biotic stresses in cool season food legumes. Euphytica 73: 1–10
- Muehlbauer FJ and KB Singh (1987) Genetics of chickpea In: (eds) Saxena MC, Singh KB, The Chickpea CAB International, Wallingford, Oxon, OX10 8DE, UK, pp 99–125
- Mullikin JC and Ning Z (2003) The Phusion assembler. Genome Res 13: 81–90
- Mutlu N, Miklas P, Reiser J and Coyne D (2005) Backcross breeding for improved resistance to common bacterial blight in pinto bean (*Phaseolus vulgaris* L.). Plant Breeding 124:282–287
- Nadimpalli RG, Jarret RL, Phatak SC and Kochert G (1994) Phylogenetic relationships of the pigeon pea (*Cajanus cajan* L.) based on nuclear restriction fragment length polymorphism. Genome 36:216–223
- Nasu S, Suzuki J, Ohta R, Hasegawa K, Yui R, Kitazawa N, Monna L and Minobe Y (2002) Search for and analysis of single nucleotide polymorphisms (SNPs) in rice (*Oryza sativa, Oryza rufipogon*) and establishment of SNP markers. DNA Research 9:163–171

- Nayak SN (2011) Identification of QTLs and Genes for drought tolerance using linkage mapping and association mapping approaches in chickpea (). Ph.D thesis, Osmania University, Hyderabad, India (http://ec2-50-19-248-237.compute 1.amazonaws.com/118/1/merged_document-6157.pdf)
- Nayak SN, Balaji J, Upadhyaya HD, Hash CT, Kavi Kishor PB, Chattopadhyay D, Rodriquez LM, Blair MW, Baume M, McNally K, This D, Hoisington D and Varshney RK (2009) Isolation and sequence analysis of DREB2A homologues in three cereal and two legume species. Plant Science 177: 460–467
- Nayak SN, Zhu H, Varghese N, Datta S, Choi H-K, Horres R, Jüngling R, Singh J, Kavi Kishor PB, Sivaramakrihnan S, Hoisington DA, Kahl G, Winter P, Cook DR and Varshney RK (2010) Integration of novel SSR and gene-based SNP marker loci in the chickpea genetic map and establishment of new anchor points with *Medicago truncatula* genome. Theoretical and Applied Genetics 120:1415–1441
- Nayyar H, Bains T and Kumar S (2005) Low temperature induced floral abortion in chickpea: relationship to abscisic acid and cryoprotectant in reproductive organs. Environmental and Experimental Botany 53:39–4
- Neil AM, Andrew F, Stephen FK, Raymond JL, Faye DS, et al (2008) Management of high-throughput DNA sequencing projects: Alpheus Journal of Computer Science & Systems. Biology 1: 132–148
- Nelson M, Phan H, Ellwood S, Moolhuijzen P, Hane J, Williams A, O'Lone C, Fosu-Nyarko J, Scobie M, Cakir M, Jones M, Bellgard M, Ksiazkiewicz M, Wolko B, Barker S, Oliver R and Cowling W (2006) The first gene-based map of *Lupinus angustifolius* L—location of domestication genes and conserved synteny with *Medicago truncatula*. Theoretical and Applied Genetics113:225–238
- Nguyen HT (2004) Transcriptional profiling and proteomic analysis of maize roots under drought stress. Proceedings of Gordon Research Conference on salt and water stress in plants, Hong Kong, June 13–18
- Novaes E, Drost DR, Farmerie WG, Pappas GJ, Grattapaglia D, Sederoff RR and Kirst M (2008) High-throughput gene and SNP discovery *in Eucalyptus grandis*, an uncharacterized genome. BMC Genomics 9:312

- NP (2003) Management of drought in chickpea-a holistic approach. In: Saxena NP (ed) Management of agricultural drought-agronomic and genetic options. Oxford & IBH Publishing, New Delhi, pp 103–122
- Or E, Hovav R and Abbo S (1999) A major gene for flowering time in chickpea. Crop Science 39:315–322
- Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N and Weigel D (2008) Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Research 18:2024–2033
- Paran I and Zamir D (2003) Quantitative traits in plants: beyond the QTL. Trends in Genetics 19: 303–306
- Parsons BL and Heflich RH (1997) Genotypic selection methods for the direct analysis of point mutations. Mutation Research 387:97–121
- Pfaff T and Kahl G (2003) Mapping of gene-specific markers on the genetic map of chickpea (*Cicer arietinum* L). Molecular Genetics and Genomics 269:243–251
- Pilon-Smits EAH, Ebskamp MJM, Paul MJ, Jeuken MJW, Weisbeek PJ and Smeekens SCY (1995) Improved performance of transgenic fructan-accumulating tobacco under drought stress. Plant Physiology 107:125–130
- Pratijit PR (2010) KASPar: SNP markers linked to *ascochyta* blight resistance genes in chickpea (http://wwwvirtualsciencefairorg/2009/prat9p3/indexhtml)
- Price AH and Tomos AD (1997) Genetic dissection of root growth in rice (*Oryza sativa* L) II: mapping quantitative trait loci using molecular markers. Theoretical and Applied Genetics 95:143–152
- Price AH, Cairns JE, Horton P, Jones HG and Griffiths H (2002) Linking Drought resistance mechanisms to drought avoidance in upland rice using a QTL approach: progress and new opportunities to intergrate stomatal and mesophyll responses. Journal of Experimental Botany 53:989–1004
- Prioul JL, Quarrie S, Causse M and de Vienne D (1997) Dissecting complex physiological functions through the use of molecular quantitative genetics. Jounal of Experimental Botany 48:1151–1163

- Qi LL, Echalier B, Chao S, Lazo GR, Butler GE, Anderson OD, Akhunov ED, Dvorak J, Linkiewicz AM, Ratnasiri A, Dubcovsky J, Bermudez-Kandianis CE, Greene RA, Kantety R, La Rota CM, Munkvold JD, Sorrells SF, Sorrells ME, Dilbirligi M, Sidhu D, Erayman M, Randhawa HS, Sandhu D, Bondareva SN, Gill KS, Mahmoud AA, Ma XF, Miftahudin, Gustafson JP, Conley EJ, Nduati V, Gonzalez Hernandez JL, Anderson JA, Peng JH, Lapitan NLV, Hossain KG, Kalavacharla V, Kianian SF, Pathan MS, Zhang DS, Nguyen HT, Choi DW, Fenton RD, Close TJ, McGuire PE, Qualset CO and Gill BS (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics 168:701–712
- Qing-shan C, Zhong-chen Z, Chun-yan L, Da-wei X, Hong-mei Q, Da-peng S, Cai-yun S, and Guo-hua H (2007) QTL Analysis of major agronomic traits in soybean. Agricultural Sciences in China 6:399–405
- Quarrie SA (1996) New molecular tools to improve the efficiency of breeding for increased drought resistance. Journal of Plant Growth Regulation 20:167–178
- Quyen L, Gutierrez-Marcos JF, Costa LM, Meyer S, Dickinson HG, Lorz H, Kranz E and Scholten S (2005) Construction and screening of subtracted cDNA libraries from limited populations of plant cells: a comparative analysis of gene expression between maize egg cells and central cells. The Plant Journal 44:167–178
- Radhika P, Gowda SJM, Kadoo NY, Mhase LB, Jamadagni BM, Sainani MN, Chandra S and GuptaVS (2007) Development of an integrated intraspecific map of chickpea (*Cicer arietinum* L) using two recombinant inbred line populations. Theoretical and Applied Genetics115:209–216
- Rafalski JA (2002) Novel genetic mapping tools in plants SNPs and LD-based approaches. Plant Science 162:329–333
- Rahangdale SL, Dhopte AM and Wanjar KB (1994) Evaluation of chickpea genotypes for yield stability under moisture deficit. Annals of Plant Physiology 8:179–184

- Rajesh PN and Muehlbauer FJ (2008) Discovery and detection of single nucleotide polymorphism (SNP) in coding and genomic sequences in chickpea (*Cicer arietinum* L). Euphytica 162:291–300
- Rajesh PN, Tullu A, Gil J, Gupta VS, Ranjekar PK and Muehlbauer FJ (2002) Identification of an STMS marker for the doublepodding gene in chickpea. Theoretical and Applied Genetics 105:604–607
- Raju NL, Gnanesh BN, Lekha P, Jayashree B, Pande S, Hiremath PJ, Byregowda M, Singh NK and Varshney RK (2010) The first set of EST resource for gene discovery and marker development in pigeonpea (*Cajanus* cajan L). BMC Plant Biology 10:45
- Rakshit S, Winter P, Tekeoglu M, Juarez Muñoz J, Pfaff T, Benko-Iseppon AM,Muehlbauer FJ and Kahl G (2003) DAF marker tightly linked to a major locus for *Ascochyta* blight resistance in chickpea (*Cicer arietinum L*.). Euphytica 132:23–30
- Ratnaparkhe MB, Santra DK, Tullu A and Muehlbauer FJ (1998) Inheritance of inter-simple sequence-repeat polymorphisms and linkage with a *fusarium* wilt resistance gene in chickpea. Theoretical and Applied Genetics 96:348–353
- Ravi K, Vadez V, Isobe S, Mir RR, Guo Y, Nigam SN, Gowda MVC, Radhakrishnan T, Bertioli DJ, Knapp SJ and Varshney RK (2011) Identification of several small main-effect QTLs and a large number of epistatic QTLs for drought tolerance related traits in groundnut (*Arachis hypogaea* L.).Theoretical and Applied Genetics122:1119–1132
- Rehman AU (2009) Characterization and Molecular Mapping of Drought Tolerance in Kabuli Chickpea (*Cicer arietinum* L.) PhD thesis, University of Saskatchewan, Saskatoon, Canada
- Reiter RS, Coor JG, Sussman MR and Gabelman WH (1991) Genetic analysis of tolerance to low-phosphorus stress in maize using restriction fragment length polymorphisms. Theoretical and Applied Genetics 82:561–568
- Ribaut JM, Banziger M, Betran J, Jiang C, Edmeades GO, Dreher K and Hoisington D (2002) Use of molecular markers in plant breeding: drought tolerance improvement in tropical maize. In: (ed) Kang MS Quantitative Genetics, Genomics, and Plant Breeding, CAB International Publishing, Wallingford, USA, pp 85–99

- Rodi CP, Darnhofer PB, Stanssens P, Zabeau M and van den Boom D (2002) A strategy for the rapid discovery of disease markers using the MassRRAY system. BioTechniques 32:62–69
- Rodi CP, Darnhofer PB, Stanssens P, Zabeau M and van den Boom D (2002) A strategy for the rapid discovery of disease markers using the MassARRAY system. BioTechniques 32:62–69
- Romo S, Labrador E and Dopico B (2004) Water stress-regulated gene expression in *Cicer arietinum* seedlings and plants. Plant Physiology and Biochemistry 39:1017–1026
- Rostocks N, Ramsay L, MacKenzie K, Cardle L, Bhat PR, Roose ML, Svensson JT, Stein N, Varshney RK, Marshall DF, Graner A, Close TJ and Waugh R (2006) Recent history of artificial outcrossing facilitates whole genome association mapping in elite inbred crop varieties. Proceedings of the National Academy of Sciences of United States of America 103:18656–18661
- Rozen S, Skaletsky HJ and Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers In: (eds) Krawetz S, Misener S, Bioinformatics Methods and Protocols: Methods in Molecular Biology Humana Press, Totowa, NJ, pp 365–386
- Rubio J, Moussa E, Kharrat M, Moreno MT, Millán T and Gil J (2003) Two genes and linked RAPD markers involved in resistance to *Fusarium oxysporum* f. sp. *ciceris* race 0 in chickpea. Plant Breeding 122:188–191
- Ryan JG (1997) A global perspective on pigeonpea and chickpea sustainable production systems: Present status and future potential In: (eds) Asthana AN, Masood Ali, Recent Advances in Pulses Research, Indian Society of Pulses Research and Development, Indian Institute of Pulses Research (IIPR), Kanpur, India, pp 1–31
- Sabaghpour SH, Mahmodi AA, Saeed A, Kamel M and Malhotra RS (2006) Study on chickpea drought tolerance lines under dryland condition of Iran. The Indian Journal of Crop Science 1:70 –73
- Salvi S and Tuberosa R (2005) Genomics based approaches to improve drought tolerance of crops. Trends in Plant Science 11:406–412
- Salvi S and Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends in Plant Science 10:297–304

- Sambrook J, Fritschm E F and Maniatis T (1989) Molecular cloning Volume I-III Cold Spring Harbor Laboratory Press, Plainview
- Santra DK, Tekeoglu M, Ratnaparkhe M, Kaiser WJ and Muehlbauer FJ (2000) Identification and mapping of QTLs conferring resistance to ascochyta blight in chickpea. Crop Science 40:1606–1612
- Sato K, Nankaku N and Takeda K (2009) A high-density transcript linkage map of barley derived from a single population. Heredity 103:110–117
- Sauter A, Davies WJ and Hartung W (2001) The long-distance Abscisic Acid signal in the droughted plant: the fate of the hormone on its way from root to shoot. Journal of Experimental Botany 52: 1991–1997
- Saxena NP, Krishnamurthy L and Johansen C (1993) Registration of a drought-resistant chickpea germplasm. Crop Science 33:1424–1426
- Schmitt ME, Brown TA and Trumpower BL (1990) A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. Nucleic Acids Research 18:3091–3092
- Schneider KA, Rosales-Serna R, Ibarra-Perez F, Cazares-Enríquez B, Acosta-Gallegos JA, Ramirez-Vallejo P, Wassimi N and Kelly JD (1997) Improving common bean performance under drought stress. Crop Science 37:43–50
- Semagn K, Bjornstad A and Ndjiondjop MN (2006) An overview of molecular markers for plants. African Journal of Biotechnology 5:2540–2568
- Serraj R, Bidinger FR, Chauhan YS, Seetharama N, Nigam SN and Saxena NP (2003) Management of drought in ICRISAT cereal and legume mandate crops. In: Kijne JW, Barker R, Molden D (eds) Water productivity in agriculture: Limits and opportunities for improvement. CAB International, Wallingford, UK pp 127–144
- Serraj R, Hash TC, Buhariwalla HK, Bidinger FR, Folkertsma RT, Chandra S, Gaur PN, Kashiwagi J, Nigam SN, Rupakula A and Crouch JH (2005) Marker-assisted breeding for crop drought tolerance at ICRISAT: Achievements and prospects In: Proceedings of the Green-Gene Conference, Bologna, Italy, pp 217–238

- Serraj R, Hash TC, Buhariwalla HK, Bidinger FR, Folkertsma RT, Chandra S, Gaur PN, Kashiwagi J, Nigam SN, Rupakula A and Crouch JH (2005) Marker-assisted breeding for crop drought tolerance at ICRISAT: Achievements and prospects In: Proceedings of the Green-Gene Conference, Bologna, Italy, May 2003, pp 217–238
- Serraj R, Krishnamurthy L, Kashiwagi J, Kumar S, Chandra S and Crouch J (2004) Variation in root traits of chickpea (*Cicer arietinum* L) grown under terminal drought. Field Crops Research 88:115–127
- Sethy N, Shokeen B, Edwards K and Bhatia S (2006b) Development of microsatellite markers and analysis of intraspecific genetic variability in chickpea (*Cicer arietinum* L). Theoretical and Applied Genetics 112:1416–1428
- Sethy NK, Choudhary S, Shokeen B and Bhatia S (2006a) Identification of microsatellite markers from *Cicer reticulatum*: molecular variation and phylogenetic analysis. Theoretical and Applied Genetics 112:347– 357
- Sethy, NK, Shokeen B and Bhatia S (2003) Isolation and characterization of sequence-tagged microsatellite sites markers in chickpea (*Cicer arietinum* L). Molecular Ecology Notes 3:428–430
- Sharma HC and Ortiz R (2000) Transgenics, pest management, and the environment. Current Science 79 :421–437
- Sharma KD, Chenand W and Muehbauer FJ (2004a) A consensus set of differential lines for identifying races of *Fusarium oxysporum* f.sp. ciceris. International Chickpea and Pigeonpea Newsletter 11: 34–36
- Sharma KD, Winter P, Kahl G and Muehlbauer FJ (2004b) Molecular mapping of *Fusarium oxysporum* f. sp. *ciceris* race 3 resistance gene in chickpea. Theoretical and Applied Genetics108:1243–1248
- Shendure J and Ji H (2008) Next-generation DNA sequencing. Nature Biotechnology 26: 1135–1145
- Shoemaker R, Keim P, Vodkin L, Retzel E, Clifton SW, Waterston R, Smoller D, Coryell V, Khanna A, Erpelding J, Gai X, Brendel V, Raph-Schmidt C, Shoop EG, Vielweber CJ, Schmatz M, Pape D, Bowers Y, Theising B, Martin J, Dante M, Wylie T

and Granger C (2002) A compilation of soybean ESTs: generation and analysis. Genome 45: 329–338

- Silva FG, Iandolino A, Al-Kayal F, Bohlmann MC, Cushman MA, Lim H, Ergul A, Figueroa R, Kabuloglu EK, Osborne C, Rowe J, Tattersall E, Leslie A, Xu J, Baek JM, Cramer GR, Cushman JC and Cook DR (2005) Characterizing the grape transcriptome Analysis of expressed sequence tags from multiple *Vitis* species and development of a compendium of gene expression during berry development. Plant Physiology 139:574–597
- Simon CJ and Muehlbauer FJ (1997) Construction of a chickpea linkage map and its comparison with maps of pea and lentil. Journal of Heredity 88:115–119
- Singh A, Singh IK and Verma PK (2008) Differential transcript accumulation in *Cicer arietinum* L in response to a chewing insect *Helicoverpa armigera* and defence regulators correlate with reduced insect performance. Journal of Experimental Botany 59: 2379–2392
- Singh KB (1997) Chickpea (*Cicer arietinum* L). Field Crops Research 53:161–170
- Singh KB and Reddy MV (1991) Advances in disease resistance breeding in chickpea. Advances in Agronomy 45: 191–22
- Singh KB, Pundir RPS, Robertson LD, van Rheene HA, Singh U, Kelley TJ, Rao PP, Johansen C and Saxena NP (1997) Chickpea In: (eds) Fuccillo D, Sears L, Stapleton P, Biodiversity in Trust Cambridge University Press, pp 100–113
- Singh RJ and Jauhar PP (2006) Genetic Resources, Chromosome Engineering, and crop improvement. In: Genetic resources, chromosome engineering and crop improvement – grain legumes [eds Singh R, Jauhar P] CRC Press: Boca Raton, FL pp 185–214
- Singh, A, Singh IK and Verma PK (2008) Differential transcript accumulation in *Cicer arietinum* L in response to a chewing insect *Helicoverpa armigera* and defence regulators correlate with reduced insect performance. Journal of Experimental Botany 59: 2379–2392
- Smithson JB, Thompson JA and Summerfield RJ (1985) In: (eds) Summerfield RJ and Roberts EH. Chickpea (*Cicer arietinum* L). Grain Legume Crops Collins, London, UK, pp 312–390

- Soltis DE, Soltis PS, Chase MW, Mort ME, Albach DC, Zanis M, Savolainen V, Hahn WH, Hoot SB and Fay MF (2000) Angiosperm phylogeny inferred from 18S rDNA, rbcL, and atpB sequences. Botanical Journal of the Linnean Society 133: 381–461
- Song QX, Liu YF, Hu XY, Zhang WK, Ma B, Chen SY and Zhang JS (2011) Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing.BMC Plant Biology 11:5
- Sorrels ME and WA Wilson (1997) Direct classification and selection of superior alleles for crop improvement. Crop Science 37:691–697
- Sreenivasulu N, Kavikishore PB, Varshney RK and Altschmeid L (2002) Mining functional information from cereal genomes-the utility of expressed sequence tags. Current Science 83:965–973
- Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M, Kota R, Varshney RK, Perovic D, Grosse I and Graner A (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theoretical and Applied Genetics 114:823–839
- Steppuhn H and Raney JP (2005) Emergence, height, and yield of canola and barley grown in saline root zones. Canadian Journal of Plant Science 85: 815–827
- Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, van Montagu M, Sandberg G, Olsson O, Teeri TT, Boerjan W, Gustafsson P, Uhlen M, Sundberg B and Lundeberg J (1998) Gene discovery in the wood-forming tissues of poplar: Analysis of 5,692 expressed sequence tags. Proceedings of National Academy of Sciences United States of America 22: 13330–13335
- Stoddard FL, Balko C, Erskine W, Khan HR, Link W and Sarker A (2006) Screening techniques and sources of resistance to abiotic stresses in cool-season food legumes. Euphytica 147: 167–186
- Szittya G, Moxon S, Santos DM, Jing R, Fevereiro MP, Moulton V and Dalmay T (2008) High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families. BMC Genomics 9:593

- Takahashi Y, Shomura A, Sasaki T and Yano M (2001) Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the α subunit of protein kinase CK2. Genetics 98:7922–7927
- Tanksley SD (1993) Mapping polygenes. Annual Review in Genetics 27:205–233
- Tanksley SD, Young ND, Paterson AH and Bonierbale MW (1989a) New tools for an old science. Biotechnology 7:257–264
- Tanksley SD, Young ND, Paterson AH and Bonierbale MW (1989b) RFLP mapping in plant breeding: new tools for an old science. Nature Biotechnology 7:257 – 264
- Tar'ran B, Warkentin TD, Tullu A and Vandenberg A (2007) Genetic mapping of *Ascochyta* blight resistance in chickpea (*Cicer arietinum* L.) using a simple sequence repeat linkage map. Genome 50:26–34
- Tegelstrom H (1992) Detection of mitochondrial DNA fragments In: Hoelzel AR (ed) Molecular genetic analysis of populations: a practical approach IRL Press, Oxford, pp 89–114
- Tekeoglu M, Rajesh PN and Muehlbauer FJ (2002) Integration of sequence tagged microsatellite sites to the chickpea genetic map. Theoretical and Applied Genetics105:847–854
- Thiel T, Kota R, Grosse I, Stein N and Graner A (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acids Research 32: e5
- Thiel T, Michalek W, Varshney RK and Graner A (2003) Exploiting EST databases for the development and characterization f gene-derived SSR-markers in barley (*Hordeum vulgare* (L). Theoretical and Applied Genetics 106:411–422
- 'tHoen PA, Ariyurek Y, Thygesen HH, Vreugdenhil E, Vossen RH, de Menezes RX, Boer JM, van Ommen GJ and den Dunnen JT (2008) Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Research 36:e141
- Thompson JD, Higgins DG and Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,

position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673–4680

- Thomson MJ, Zhao K, Wright M, Reynolds A, Rahman MA, Ismail AM, KL McNally, Bustamante CD and McCouch SR (2010) Application of Illumina BeadXpress 384-plex SNPs sets for diversity analysis and genetic mapping in rice. Plant & Animal Genomes XVIII Conference, January 9-13, San Diego, CA (http://www.intl-pag.org/18/abstracts/P05b_PAGXVIII_249.html)

- Thudi M , Bohra A, Nayak SN, Varghese N, Shah TM, Penmetsa RV, Thirunavukkarasu T, Gudipati S,Gaur PM, Kulwal PL, Upadhyaya HD, KaviKishor PB, Winter P, Kahl G, Town CD, Kilian A, Cook DR and Varshney RK (2011) Novel SSR Markers from BAC-End Sequences, DArT Arrays and a Comprehensive Genetic Map with 1,291 Marker Loci for Chickpea (*Cicer arietinum* L.). Public Library of Science 6: e27275

- Thudi M, Senthilvel S, Bottley A, Hash CT, Reddy AR, Feltus AF, Paterson AH, Hoisington DA and Varshney RK (2010) A comparative assessment of the utility of PCR-based marker systems in pearl millet. Euphytica 174:253–260

- Toker C and Cagirgan M I (1998) Assessment of response to drought stress of chickpea (*Cicer arietinum* L) lines under rainfed conditions. Turkish Journal of Agriculture and Forestry 22:615–621

- Toker C and Canci H (2003) Selection of chickpea (*Cicer arietinum* L) genotypes for resistance to Ascochyta Blight [*Ascochyta rabiei* (Pass) Labr], yield and yield criteria. Turkish Journal of Agriculture and Forestry 27:277–283

- Trick M, Kwon SJ, Choi SR, Fraser F, Soumpourou E, Drou N, Wang Z, Lee SY, Yang TJ, Mun JH, Paterson AH, Town CD, Pires JC, Lim YP, Park BS and Bancroft I (2009) Complexity of genome evolution by segmental rearrangement in *Brassica rapa* revealed by sequence-level analysis. BMC Genomics 10: 539

- Tuberosa R, Frascaroli E, Salvi S, Sanguineti MC, Conti S and Landi P (2005) QTLs for tolerance to abiotic stresses in maize: Present status and prospects. Maydica 50:559–570

- Tuberosa R, Salvi S, Sanguineti MC, Landi P, Maccaferri M and Conti SS (2002) Mapping QTLs regulating morpho-physiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. Annals of Botany 89:941–963
- Tullu A, Muehlbauer FJ, Kaiser WJ, Simon CJ, Mayer MS, Kumar J and Kraft JM (1998) Inheritance and linkage of a gene for resistance to race 4 of *Fusarium* wilt and RAPD markers in chickpea. Euphytica 102:227–232
- Turner NC (1986) Crop Water deficit: a decade of progress. Advances in Agronomy 39: 1–51
- Turner NC, Wright GC and Siddique KHM (2001) Adaptation of grain legumes (pulses) to water limited environments. Advances in Agronomy 71:193–231
- Udupa SM and Baum M (2003) Genetic dissection of pathotypespecific resistance to ascochyta blight disease in chickpea (*Cicer arietinum* L.) using microsatellite markers. Theoretical and Applied Genetics 106:1196–1202
- Udupa SM, Sharma A, Sharma RP and Pai RA (1993) Narrow genetic variability in *Cicer arietinum* as revealed by RFLP analysis. Journal of Plant Biochemistry and Biotechnology 2:83–86
- Upadhyaya HD, Thudi M, Dronavalli N, Gujaria N, Singh S, Sharma S and Varshney RK (2010) Genomic tools and germplasm diversity for chickpea improvement. Plant Genetic Resources 1-14. doi:10.1017/S1479262110000468
- Vadez, V, Rao S, Kholova J, Krishnamurthy L, Kashiwagi J, Ratnakumar P, Sharma KK, Bhatnagar-Mathur P and Basu PS (2008) Roots research for legume tolerance to drought: *Quo vadis? J Food Legumes* 21:77–85
- Van der Maesen LJG (1972) *Cicer* L. a monograph of the genus, with special reference to the chickpea (*Cicer arietinum* L.), its ecology and cultivation. Mededlingen landbouwhogeschool (Communication Agricultural University) Wageningen 72-10 342 p.
- van der Maesen LJG (1987) Origin, history, and taxonomy of chickpea In: (eds) Saxena MC and Singh KB. The Chickpea. CAB International Publications, UK

- Van Ooijen (2006) JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations Kyazma BV, Wageningen, Netherlands
- Varshney RK (2009b) Gene-based marker systems in plants: high throughput approaches for discovery and genotyping. In: (eds) Jain SM, Brar DS, Molecular Techniques in Crop Improvement, Springer,The Netherlands, pp119–142
- Varshney RK (2010) Gene-based marker systems in plants: high throughput approaches for discovery and genotyping. In: (eds) Jain SM, Brar DS. Molecular Techniques in Crop Improvement, Springer, The Netherlands, pp 119–142
- Varshney RK , Hiremath P, Lekha P, Kashiwagi J, Balaji J, Deokar AA, Vadez V, XiaoY, Srinivasan R, Gaur PM, Siddique KHM, Town CD and Hoisington DA (2009a) A comprehensive resource of drought- and salinity- responsive ESTs for gene discovery and marker development in chickpea (*Cicer arietinum* L). BMC Genomics 10:523
- Varshney RK and Dubey A (2009c) Novel genomic tools and modern genetic and breeding approaches for crop improvement. Journal of Plant Biochemistry and Biotechnology 18: 127–138
- Varshney RK, Hoisington DA, Upadhyaya HD, Gaur PM, Nigam SN, Saxena K, Vadez V, Sethy NK, Bhatia S, Aruna R, Gowda MVC and Singh NK (2007a) Genomic Assisted Crop Improvement Vol II: Genomics Applications in Crops. In: (eds) Varshney RK and Tuberosa R, Molecular genetics and breeding of grain legume crops for the semi-arid tropics. Springer, The Netherlands, pp 207–242
- Varshney RK, Bertioli DJ, Moretzsohn MC, Vadez V, Krishnamurthy L, Aruna R, Nigam SN, Moss BJ, Seetha K, Ravi K, He G, Knapp SJ and Hoisington DA (2009d) The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.). Theoretical and Applied Genetics doi 10.1007/s 00122-008-0933-x.s
- Varshney RK, Chabane K, Hendre PS, Aggarwal RK and Graner A (2007b) 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 173:638–649

- Varshney RK, Close TJ, Singh NK, Hoisington DA, and Cook DR (2009e) Orphan legume crops enter genomics era!. Current Opinion in Plant Biology 12:1–9
- Varshney RK, Graner A and Sorrells ME (2005a) Genomics-assisted breeding for crop improvement. Trends in Plant Science 10: 621–630
- Varshney RK, Graner A and Sorrells ME (2005b) Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23:48–55
- Varshney RK, Hoisington DA and Tyagi AK (2006) Advances in cereal genomics and applications in crop breeding. Trends in Biotechnology 24: 490–499
- Varshney RK, Hoisington DA, Nayak SN and Graner A (2009f) Molecular plant breeding: methodology and achievements In: (eds) Somers D, Langridge P, Gustafson PJ, Plant genomics: methods and protocols The Humana Press, Totowa, pp 283–304
- Varshney RK, Nayak SN, Jayashree B, Eshwar K, Upadhyaya HD, Hoisington DA (2007c) Development of cost-effective SNP assays for chickpea genome analysis and breeding. SAT eJournal 3 (1)
- Varshney RK, Nayak SN, May GD and Jackson SA (2009g) Next generation sequencing technologies and their implications for crop genetics and breeding. Trends in Biotechnology 27:522–530
- Varshney RK, Penmetsa RV, Dutta S, Kulwal PL, Saxena RK, Datta S, Sharma TR, Rosen B, Carrasquilla-Garcia N, Farmer AD, Dubey A, Saxena KB, Gao J, Fakrudin B, Singh M N, Singh BP, Wanjari KB, Yuan M, Srivastava RK, Kilian A, Upadhyaya HD, Mallikarjuna N, Town CD, Bruening GE, He G, May GD, McCombie R, Jackson SA, Singh NK and Cook DR (2009) Pigeonpea genomics initiative (PGI): an international effort to improve crop productivity of pigeonpea (*Cajanus cajan* L.). Molecular Breeding 26: 393-408
- Varshney RK, Thiel T, Sretenovic-Rajicic, Baum M, Valkoun J, Guo P, Grando S, Ceccarelli S and Graner A (2008) Identification and validation of a core set of informative genic SSR and SNP markers for assaying functional diversity in barley. Molecular Breeding 22:1–13

- Varshney RK, Thiel T, Stein N, Langridge P and Graner A (2002) *In Silico* analysis of frequency and distribution of microsatellites in ESTs of some cereal species. Cellular and Molecular Biology Letters 7:537–546
- Vavilov NI (1926) Studies on the origin of cultivated plants. Lenningrad 129–133
- Vavilov NI (1951) The origin, variation immunity and breeding of cultivated plants. Chronica Botanica 13-1/6:26–38, 75–78, 151 (1949-50) New York
- Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, Marden JH (2008) Rapid transcriptome characterization for a non model organism using 454 pyrosequencing. Molecular Ecology 17: 1636–1647
- Vettore AL, da Silva F, Kemper EL ,Souza GM, da Silva AM, Ferro MI, Henrique-Silva F, Gigloiti EA, Lemos MVF, Coutinho LL, Nobrega MP,Carrer H, Franca SC, Bacci M, Goldman MS, Gomes SL, Nunes LR, Camargo LEA, Siquera WJ, Sluys MAV, Theimann OH, Kuramae EE, Santelli RV, Marino CL, Targon MLPN, Ferro JA, Silveira HCS, Marini DC, Lemos EGM, Monteiro-Vitorello CB, Tambor JHM, Carraro DM, Roberto PG, Martins VG, Goldmann GH, de Oliveria RC, Truffi D, de Rosa VEC, Rossi M, Colombo CA, de Araujo PG, Sculaccio SA, Angella A, Lima MMA, Siviero F, Coscrato VE, Machado MA, Givret L, Mauro SMD, Nobrega FG, Menck CFM, Braga MDV, Telles GP, Cara FAA, Pedrosa G, Meidanis J and Arruda P (2003) Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Research 13:2725–2735
- Voorips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93:77–78
- Walker AR and Walker BF (1984) Glycaemic index of South African foods determined in rural blacks--a population at low risk of diabetes. Human Nutrition-Clinical Nutrition, 38: 215–222
- Wang J, Wong G K, Ni P, Han Y, Huang X, Zhang J, Ye C, Zhang Y, Hu J, Zhang K, Xu X, Cong L, Lu H, Ren X, He J, Tao L, Passey DA, Yang H, Yu J and Li S (2002) RePS: a sequence assembler that masks exact repeats identified from the shotgun data. Genome Research 12: 824–831

- Wang S, Basten CJ and Zeng ZB (2007) Windows QTL cartographer 2.5 http://statgenncsuedu/qtlcart/WQTLCarthtm
- Weber KA, Achenbach LA and Coates JD (2006) Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction. Nature Reviews Microbiology 4: 752–764
- Weeden NF, Timmerman GM, Hemmat M, Kneen BE and Lodhi MA (1992) Inheritance and reliability of RAPD markers In: Proceedings of Joint Plant Breeding Symposium Series Applications of RAPD Technology to Plant Breeding. Crop Science Society of America, American Society for Horticultural Science and American Genetic Association 12–17
- Weiland JJ and Yu MH (2003) A cleaved amplified polymorphic sequence (CAPS) marker associated with root-knot nematode resistance in sugarbeet. Crop Science 43:1814–1818
- White JA, Todd J, Newmann T, Focks N, Girke T, Ilarduya OM, Jaworski LG,Ohlrogge and Benning C (2000) A new set of arabidopsis expressed sequence tags from developing seeds- The metabolic pathway from carbohydrates to seed oil. Plant Physiology 124: 1582–1594
- Wicker T, Narechania A, Sabot F, Stein J, Vu GTH, Graner A, Ware D and Stein N (2008) Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. BMC Genomics 9: 518
- Wilkinson S and Davies WJ (2002) ABA based chemical signaling: the coordination of responses to Stress in plants. Plant,Cell and Environment 25: 195–210
- Williams PC and Singh U (1987) The chickpea-nutritional quality and evaluation of quality in breeding programs In: (eds) Saxena MC and Singh KB. The Chickpea, CAB International Walling-ford, pp 329–356
- Winter P, Benko-Iseppon AM, Hüttel B, Ratnaparkhe M, Tullu A, Sonnante G, PfaV T, Tekeoglu M, Santra D, Sant VJ, Rajesh PN, Kahl G and Muehlbauer FJ (2000) A linkage map of the chickpea (*Cicer arietinum* L.) genome based on the recombinant inbred lines from a *C. arietinum* X *C. reticulatum* cross: localization of resistance genes for Fusarium races 4 and 5. Theoretical and Applied Genetics 101:1155–1163

- Winter P, Pfaff T, Udupa SM, Hüttel B, Sharma PC, Sahi S, Arrequin-Espinoza R, Weigand F, Muehlbauer FJ and Kahl G (1999) Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (*C arietinum* L). Molecular and General Genetics 262:90–101
- Wu J, Maehara T, Shimokawa T, Yamamoto S, Harada C, Takazaki Y, Ono N, Mukai Y, Koike K, Yazaki J, Fujii F, Shomura A, Ando T, Kono I, Waki K, Yamamoto K, Yano M, Matsumoto T and Sasaki T (2002) A comprehensive rice transcript map containing 6591 expressed sequence tag sites. Plant Cell 14:525–535
- Xiaolei Wu, Chengwei Ren, Trupti Joshi, Tri Vuong, Dong Xu and Henry T Nguyen (2010) SNP discovery by high-throughput sequencing in soybean. BMC Genomics 11:469
- Xiong L and Zhu JK (2003) Regulation of abscisic acid biosynthesis. Plant Physiology 133: 29–36
- Xu K and Mackill DJ (1996) A major locus for submergence tolerance mapped on rice chromosome 9. Molecular Breeding 2: 219–224
- Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC and Mackill DJ (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442:705–708
- Yadav OP, Mitchell SE, Fulton TM and Kresovich S (2008) Transferring molecular markers from sorghum, rice and other cereals to pearl millet and identifying polymorphic markers. Journal of SAT Agricultural Research 6
- Yadav SS, Turner NC, Berger J, Kumar J, Hegde VS and Kumar S (2004) Development of widely adapted kabuli cultivars In: Proceedings of International Chickpea Conference, 20-22 January 2003, Raipur, Chhatisgarh, India, pp 20–27
- Yamamoto K and Sasaki T (1997) Large-scale EST sequencing in rice. Plant Molecular Biology 35:135–144
- Yamamoto YY, Kondo Y, Kato A, Tsuji H and Obokata J (1997) Light-responsive elements of the tobacco PSI-D gene are located both upstream and within the transcribed region. Plant Journal 12:255–265

- Yan J, Shah T, Warburton ML, Buckler ES, McMullen MD, et al (2009) Genetic Characterization and Linkage Disequilibrium Estimation of a Global Maize Collection Using SNP Markers. Public Library of Science 4(12): e8451 doi:101371/journalpone0008451
- Yang H, Boersma JG, You M, Buirchell BJ and Sweetingham MW (2004) Development and implementation of a sequence-specific PCR marker linked to a gene conferring resistance to Anthracnose disease in narrow-leafed lupin (*Lupinus angustifolius* L.) Molecular Breeding 14: 145–151
- Yang H, Shankar M, Buirchell BJ, Sweetingham MW, Caminero C and Smith PMC (2002) Development of molecular markers using MFLP linked to a gene conferring resistance to Diaporthe toxica in narrow-leafed lupin (*Lupinus angustifolius* L.). Theoretical and Applied Genetics 105: 265–270
- Young ND (1996) QTL mapping and quantitative disease resistance in plants. Annual Review of Phytopathology 34:479–501
- Young ND, Cannon SB, Sato S, Kim D, Cook DR, Town CD, Roe BA and Tabata S (2005) Sequencing the genespaces of *Medicago truncatula* and *Lotus japonicus*. Plant Physiology 137: 1174–1181
- Yu JK, Sun Q, Rota ML, Edwards H, Tefera H and Sorrells ME (2003) Expressed sequence tag analysis in tef (*Eragrostis tef* (Zucc) Trotter). Genome 49:365–372
- Yuan C, Cunningham Daniel Rios F, McLaren WM, Smith J, Pritchard B, Spudich GM, Brent B, Kulesha E, Marin-Garcia P, Smedley D, Birney E and Flicek P (2010) Ensembl variation resources. BMC Genomics 11:293
- Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468
- Zhang H, Sreenivasulu N, Weschke W, Stein N, Rudd S, Radchuk V, Potokina E, Scholz U, Schweizer P, Zierol U, Langridge P, Varshney RK, Wobus U and Graner A (2004) Large-scale analysis of the barley transcriptome based on expressed sequence tags. Plant Journal 40:276–290
- Zhu C, Naqvi S, Breitenbach J, Sandmann G, Christou P and Capell T (2008) Combinatorial genetic transformation generates a library of metabolic phenotypes for the

carotenoid pathway in maize. Proceedings of the National Academy of Sciences of the United States of America 105:18232–7

- Zhu JM, Kaeppler SM and Lynch JP (2005b) Mapping of QTLs for lateral root branching and length in maize (*Zea mays* L.) under differential phosphorus supply. Theoretical and Applied Genetics 111:688–695
- Zhu JM, Kaeppler SM, and Lynch JP (2005a) Mapping of QTL controlling root hair length in maize (*Zea mays* L.) under phosphorus deficiency. Plant and Soil 270:299–310
- Zhu YL, Song QJ, Hyten DL, Tassell CPV, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND and Cregan PB (2003) Single-nucleotide polymorphisms in soybean. Genetics 163:1123–1134

Annexure 1

Details on primer pairs designed for homologous (chickpea) as well as heterologous (related legume) species

S. no.	Primer ID	Accession number/ ID[a]	Forward primer sequence (5'- 3')	Reverse primer sequence (5'- 3')	Product size (bp)	Source of genes/ markers*
Homologous genes						
1	Ca2S071618_2464_3037	Ca2S071618_2464_3037*	GCACACCAGCTTCCCTAGTA	CCTAGGCTTTAGACATGGCA	112	A
2	Ca2S219622_1285_1092	Ca2S219622_1285_1092*	GTTCAAAGGTACCCTCAGCC	TCCAACTTTCGTTTCTGCTC	103	A
3	Ca2S224282_3544_2314	Ca2S224282_3544_2314*	CGATGTGGAGACGGAGTAGT	CGGTTTGCAGTACTATTGGG	122	A
4	Ca2S375228_1350_1378	Ca2S375228_1350_1378*	GCCTTCAGAATCCAGATTGA	TGCGCTTAAATCAGTCACCT	128	A
5	Ca2SFE671114	Ca2SFE671114*	TGCAGTGCTAAATTTTGGTG	TGTTCATATGGATCCTGGCT	125	A
6	Ca2C19231	Ca2C19231*	AATGGAATGGAATGGAGTCA	ATTCCACACGGGTTCATTC	128	A
7	Ca2C19867	Ca2C19867*	GCTCTCACCGTAACTCAACC	CATCTTCAGAAGGACCGAAA	149	A
8	Ca2C22370	Ca2C22370*	ATTGCTTGAGCAACTCCTTG	GTCCATATTACATGGCCCCT	145	A
9	Ca2C28890	Ca2C28890*	ATAGCCCCACCGTATAGACC	ATCAACGCAGAGTTCCTTCA	101	A
10	Ca2C29585	Ca2C29585*	ACGACGATTCCCTGTCAGTA	TCTTTCCGTTACCAGCATTT	163	A
11	Ca2C30120	Ca2C30120*	TAAGCCAACATCCCAACTGT	ACCAAAGGAAACCAAACTCC	169	A
12	Ca2C476	Ca2C476*	AAAAAGGAAATGTGCAAGGA	CAGAACCAAAATGTTCAGCA	120	A
13	Ca2C33686	Ca2C33686*	TTGCGTGTCTCAGGAAAATAG	CTTCATCCCTCACTCCTCCT	249	A
14	Ca2C36189	Ca2C36189*	CAGATTTCCACTTGGGATTG	AAATGCAGAATCACAGGCTC	143	A
15	Ca2C39635	Ca2C39635*	CACCGCATTACTTGGGATAC	AACCCCAGGAAAATAAAACG	215	A
16	Ca2C42216	Ca2C42216*	CTGCATATGCGATACGAGAA	ATTACTTCCTCCACCCCTTG	199	A
17	Ca2C2507	Ca2C2507*	TACCCCGAGGAAAAGAAATC	CTGCTTTCCAGGAGATTCAA	145	A
18	Ca2C4135	Ca2C4135*	ACCAGTCTTGTGCATCCAAT	ATAAGCCTAGGGAGGACGAA	178	A
19	Ca2S095423_2383_0398	Ca2S095423_2383_0398*	GGACCTTCCCGTGTTAAGAT	TCGTTTGAAATTGGGTTTGT	160	A
20	Ca2S155734_1421_0905	Ca2S155734_1421_0905*	GATCTTCCAGCTCCAACAAA	CACCTGATTCCAATTCCAAG	124	A
21	Ca2S209838_1352_0399	Ca2S209838_1352_0399*	CTGAAATTGAACGATCTGGG	ACCTTTCCCTCACGGTACTC	188	A
22	Ca2C21305	Ca2C21305*	AACCTACTTTGTGCCCTTCC	TGCAGTTACAATTGATGATGG	112	A
23	Ca2C8356	Ca2C8356*	GGTGAGTAACACGTGGGAAC	ACCCACCAACAAGCTAATCA	134	A
24	Ca2C8776	Ca2C8776*	TAAGGCTTTGCCAAGAGAAC	AAGGAACAAGCCATGTCAAA	149	A
25	Ca2C10890	Ca2C10890*	ATTGTTTCAAAAAGGCATCG	AGGGCAACACATGTCCTAAA	192	A
26	Ca2C17525	Ca2C17525*	ACACGTTTCCATGCGTTATT	CATGAGAGTTTGATCCTGGC	144	A

27	Ca2C20566	Ca2C20566*	ACGGAAGCCACATGATTCTA	TCACTTCTTCAACCAGCTCC	130	A
28	Ca2C44722	Ca2C44722*	TTTTGCGGCAATAAATTGAG	TGTTCATTTGATTTTTGGGG	191	A
29	Ca2C25224	Ca2C25224*	CCTTTGTTTTGATTCCATTCAT	CAATGTCACTAGTCCATTTCCA	186	A
30	Ca2C26959	Ca2C26959*	AGTCCACGCCGTAAACTATG	TAATCTTGCGACCGTACTCC	104	A
31	Ca2C30564	Ca2C30564*	CCTCGTGGTTTTGTCATTGT	GCCTTGTTGGTTGAAGAAGA	158	A
32	Ca2C32076	Ca2C32076*	CAAGGAGAAACAAGCGATGT	GGCGAGTAATTTCTGTGGAA	109	A
33	Ca2C33756	Ca2C33756*	TCTGTATGCAAGCGATTCTG	TTTTTGACCTCTCACTATCAAATG	117	A
34	Ca2C35272	Ca2C35272*	TTTTCAGTTCAGTTCAATTCATGT	GCAGCCATGTATTTGATTTTC	146	A
35	Ca2C35789	Ca2C35789*	CTCACCCCAATTTCTTGTTG	TATATTGGAGTGGGGTCCCT	192	A
36	Ca2C36777	Ca2C36777*	TTTTCCGGCTTCTTTACCTT	GTTGTAACCATTCCCCTGTG	100	A
37	Ca2C41925	Ca2C41925*	GGATTGGGATTTCCTTTGAC	GGAAGGAGATGAATCGGAGT	186	A
38	Ca2SGR396802	Ca2SGR396802*	CAACACACCTGAACGTACCA	CTACATGAGCGGGAAAGAAA	195	A
39	Ca2C5540	Ca2C5540*	TAACACATGCAAGTCGAACG	GATAAATCTTTCCCCCGAAG	140	A
40	Ca2SFE670434	Ca2SFE670434*	AAAATAGCCATCCTTGTAGTTCA	CAATACCCTGCATCCTTGTT	201	A
41	Ca2C42902	Ca2C42902*	TAAGCAGTGGTATCAACGCA	TAAGCAGTGGTATCAACGCA	210	A
42	Ca2S137764_1568_3224	Ca2S137764_1568_3224*	TAAGCAGTGGTATCAACGCA	TCTTGTTGGGGGTGAAACTA	298	A
43	Ca2C37366	Ca2C37366*	TAAGCAGTGGTATCAACGCA	TAAGCAGTGGTATCAACGCA	308	A
44	Ca2C30103	Ca2C30103*	AGGTTGTGGAACAGCATCAT	TTCCTTCAGTGAGCCAAGAC	421	A
45	Ca2S276045_1443_3617	Ca2S276045_1443_3617*	ATCCAATGTTCTCGAATGGA	TTTCCGTTTGGATGTTGATT	114	A
46	Ca2S290507_2844_3854	Ca2S290507_2844_3854*	GCAAAGCGTGAATCAATTTT	CCTGAACCGTGAATCAAATC	397	A
47	Ca2S114521_2553_3482	Ca2S114521_2553_3482*	TCACGGATCTCACCTTTAGC	GTGGAGCCTTTTCCTTCTTC	161	A
48	Ca2C32482	Ca2C32482*	TCACTCTCGTTTGGAAGAGG	TTTGCCATCATGTAAAGGGT	571	A
49	Ca2S203245_0585_2299	Ca2S203245_0585_2299*	TATCCTTCCCATCCCTTTTC	CAGACAGGCCTTTGAAAGAA	156	A
50	Ca2C392	Ca2C392*	TCAATTTGGTTTCATGTCCA	CCCCGAATCTTTGATCACTA	218	A
51	Ca2S382168_3591_0769	Ca2S382168_3591_0769*	AATCATGACCATGTCCCATC	CATAGACAGTGGGGGTCTTG	171	A
52	Ca2S389296_2488_3865	Ca2S389296_2488_3865*	GTGGGGAAAATGGCTTTAAT	GTAGCGAAGCATACCTCCAA	470	A
53	Ca2C31040	Ca2C31040*	TCACCTTTTCATTATGGGGA	TTTTAAGTTTGGGGTGGTGA	176	A
54	Ca2S393415_0677_2561	Ca2S393415_0677_2561*	TTGAGGCAATTGTTCCAGAT	TTATGAGGCGCTTTTGAAAC	162	A
55	Ca2C35129	Ca2C35129*	ACTAGCCATCGTCTTGTTCG	TTGAAATCTGAAGCGAAACC	323	A
56	Ca2S125676_2925_1604	Ca2S125676_2925_1604*	TGTCGCTAATCCTCTTGGAC	GGAAGGAGTGATTGTGATGG	367	A
57	Ca2S322579_3448_3365	Ca2S322579_3448_3365*	GTGGAGGGTGGTTATTGTTG	CTCTTGCACCGAAGGTAAAA	179	A
58	Ca2C20752	Ca2C20752*	ATAAAACAAAGGCAGGAGGG	ACGAGGGCAGTCATATGGTA	101	A
59	Ca2C25175	Ca2C25175*	AGGGGTTATAGGACCACGAC	GCGTACGGTTTCAGGTTCTA	268	A
60	Ca2C27851	Ca2C27851*	ACACCCCAGTACAACCACAC	GAATAAGGAGAAGGCCAAGG	138	A
61	Ca2C35356	Ca2C35356*	CTTCACGGAAGAACCTGAAA	ATACTGGTGGAGGAGTGCAA	218	A

62	Ca2C40247	Ca2C40247*	TGTTTGTGGAGCTCAGTTCA	TAAGCAGTGGTATCAACGCA	251	A
63	Ca2S322940_1303_1036	Ca2S322940_1303_1036*	CAAAAGCCACCAGAAAATCA	GTTTGAAAGGGTATGCTGGA	166	A
64	Ca2C9915	Ca2C9915*	TGAATGTGCTTTGGGTAGTG	TGAGCATTAGCAACCAAACA	118	A
65	Ca2C24932	Ca2C24932*	TTGGAGCACAAAGAAAAAGG	GGAGGGTGCAGTATTTGATG	328	A
66	Ca2S053379_3256_2980	Ca2S053379_3256_2980*	CCCACACCAATCTTCTCAAC	CATAATTTCCGGCGACATAC	359	A
67	Ca2S062448_4035_1045	Ca2S062448_4035_1045*	AGTGCTCTAATAGCTGCCGA	AGCGTGAAACCTACCACAAG	165	A
68	Ca2C40663	Ca2C40663*	TGGATAACGGGTAGTTTGGA	AGAGGAAGAAGGCGAGCTAC	256	A
69	Ca2C23801	Ca2C23801*	TTAAGGAAACAGCTGCCAAC	ATCATGTGCCGTCATCTTCT	101	A
70	Ca2S384100_2509_3706	Ca2S384100_2509_3706*	GAAAATGCTGAAATAGGCCA	CCTCCATTTTGGTACATTGC	244	A
71	Ca2C33173	Ca2C33173*	TGGCAACTACCGTAACCATT	ATTGACTTCACTGTCAGGGC	160	A
72	Ca2S089707_2585_1357	Ca2S089707_2585_1357*	CATGAGTGGTAGGAGAGCGT	GACTGGCGCAGGAATATTTA	233	A
73	Ca2C42645	Ca2C42645*	TTTCAGGCTTTTGTTAAGCG	TAACAAGTTCGGGACGGTAA	144	A
74	Ca2C19301	Ca2C19301*	AGGACAGAATGAACCAAAACC	TCATCATCCCAAAGGTCATT	139	A
75	Ca2C21276	Ca2C21276*	AGCTCCTAGAGCCTCATGGT	TCCTGTTGTGGATTTGTCAG	220	A
76	Ca2C40872	Ca2C40872*	TTTTCGTTTTGTCAAGCCTC	GAAACTGGTGGATGTTCAGG	521	A
77	Ca2C29315	Ca2C29315*	CTGCTTTTGCTTTGGAATGT	GAATCGGACGAACTTCATTG	177	A
78	Ca2C38786	Ca2C38786*	ATCAAGGGTCCAATAGAGGC	ATGTTGGCTATTCCACCTCA	336	A
79	Ca2S082278_2819_1307	Ca2S082278_2819_1307*	ATATCATTCCCCGGTTGTTT	CAAAGCATCTCACACAGCAC	131	A
80	Ca2S290488_1071_3554	Ca2S290488_1071_3554*	ATTTGCCTGCAAAAATAACG	ACGAGATGCCAAAATCAGAG	168	A
81	Ca2C18785	Ca2C18785*	AATTATCAGGGCCACACAAA	TTTGCAAGTCACAGGCACTA	173	A
82	Ca2S407092_0729_3445	Ca2S407092_0729_3445*	CCGAACCAAACACAATCTTC	CAATGATCTTCCCCAAGGTT	302	A
83	Ca2S021522_3215_0169	Ca2S021522_3215_0169*	AGCAAAACGACACCGTAAAG	TTCAAATCCCTAAGAACCCC	379	A
84	Ca2S237458_1029_2001	Ca2S237458_1029_2001*	CTCCAGCTTGACCACCTCTA	GCCGATGAAGATTGAGATTG	128	A
85	Ca2S289800_2555_0786	Ca2S289800_2555_0786*	AGAGGGAGGGGATAGGAGTT	GAAACCAGTTCCTCACTGGA	232	A
86	Ca2C18618	Ca2C18618*	ACCGACAAAAAGATGAATACAA	TTCCTCTACATCCATGTGAAAAT	160	A
87	Ca2S013999_1790_3142	Ca2S013999_1790_3142*	ACTGACAACGAAGCAAGGAG	AACACCTGGATCTGTGGAAA	196	A
88	Ca2S341337_1018_2975	Ca2S341337_1018_2975*	GTGGCAAATCAACCTGTTTC	TCCCATCATGTGATACACCA	145	A
89	Ca2C43432	Ca2C43432*	TAGCTGGTCTGAGAGGATGG	AAGAAGTTTACGACCCGGAG	153	A
90	Ca2S408548_3100_3291	Ca2S408548_3100_3291*	GAACCGGTTGAAGATGAATG	ATCCTCCTGAACGCTTTTCT	126	A
91	Ca2C3599	Ca2C3599*	TCATAACCTTAGCGCATTCC	ATGACTGGACGAAACCAAGA	272	A
92	Ca2S383454_2176_1945	Ca2S383454_2176_1945*	TCTGCGTTAAGGAAAACTGG	TTCTTAATCTGGCGGTTTTG	306	A
93	Ca2C44586	Ca2C44586*	TTATTTCAAATGGAACGGGA	TTTTCGTTGAAAGCTTGGAG	233	A
94	Ca2C26568	Ca2C26568*	ATCCATAAACAGCGAAGCAG	GCATCCCAACTGTCTGAAAC	266	A
95	Ca2C39678	Ca2C39678*	CAGGAGCAATTCCTTCTTCA	GCATTCTAAGCGTGGAAAAA	313	A
96	Ca2SGR409670	Ca2SGR409670*	TTCCTATTGAAGCATGCAAA	ACACAGGGCTTTCTGTTCAC	130	A

97	Ca2SGR391017	Ca2SGR391017*	GCAACAAATGAACCGGATAC	CCAAGGTGAGCGACACTACT	447	A
98	Ca2S348934_3400_0654	Ca2S348934_3400_0654*	GGGATGGAACAGTAGGTCAA	CTGTTGTTTTACGCATTAGCAG	100	A
99	Ca2S114987_0832_2041	Ca2S114987_0832_2041*	CCAGAGCTATTTATCGACGC	ACTTTGGCTAAAAAGCCTCC	139	A
100	Ca2S220861_2577_3780	Ca2S220861_2577_3780*	TCAGGAAAGAAAAGGCGTTA	AAGCAGCAGGTCGAAAACTA	201	A
101	Ca2C44452	Ca2C44452*	TCAGTTACAGAATTTGAACACACA	TGGAGCTTTCACAGATCCTC	126	A
102	Ca2S090699_3800_2828	Ca2S090699_3800_2828*	CAAGAGCTGGGACCTTATCA	ACTTCGAACTTTACCGGCTT	213	A
103	Ca2C12526	Ca2C12526*	TTGAGGAGGAGAAACACAGC	GACAGGCTTCCAGAAACTGA	243	A
104	Ca2S387716_1005_2344	Ca2S387716_1005_2344*	TAAGCAGTGGTATCAACGCA	GCACTTGCCATAAGCTCACT	461	A
105	Ca2SFE671074	Ca2SFE671074*	TAAGCAGTGGTATCAACGCA	TTGCATGAAGGTGATGATTG	625	A
106	Ca2SGR395698	Ca2SGR395698*	TAAGCAGTGGTATCAACGCA	AGCAACTTGATAGTCGTGCC	421	A
107	Ca2C9498	Ca2C9498*	TGAGCTGGAATTTGAGTATTGA	AAATGACTGCAACACGATCC	102	A
108	Ca2SGR397002	Ca2SGR397002*	ACGTGATAGCCTGCGATAAG	CACCTAGTACCACCACCGAC	304	A
109	Ca2S124718_0933_2144	Ca2S124718_0933_2144*	GTCGTGGCCGTTAAATTATG	TACCAATGATTTGCGATTCC	279	A
110	Ca2C12244	Ca2C12244*	GATGTGGAAGCACACACACT	CCACATGTACTACACACAAGCC	100	A
111	Ca2C8312	Ca2C8312*	GTTGGATTGATCAAATTCGC	GCATCATTTTCTTCCCAAAC	167	A
112	Ca2S076469_0294_0429	Ca2S076469_0294_0429*	TTCCACCACCTATTCTTCCA	CACAAGGGGTTTACCTTCCT	277	A
113	Ca2C30	Ca2C30*	CACATCGTCATCCTCAAACA	TTGACGAAAGCATGCACTAA	130	A
114	Ca2C43322	Ca2C43322*	CATGCCATCAGTCAAAAACA	TGGTATGATTGGGAGCATCT	136	A
115	Ca2C104	Ca2C104*	ACAAAAGGGGTTTAATTCGC	ACTCATTTCTGTCCCCATTG	125	A
116	Ca2C228	Ca2C228*	CTCAACAAAACCCAATCCAC	TCCCATTAAGTGCTTTTCCA	274	A
117	Ca2C309	Ca2C309*	CAGAGTGCATCGGTCATGTA	CAACCGAGACTCCCGTAGTA	219	A
118	Ca2C495	Ca2C495*	GAGTCCCAGAACTTGAAGCA	CTCTAGCAGTAGCAGCCGTC	278	A
119	Ca2C629	Ca2C629*	TCAAACCACATCATGAAAGC	TAGCATTTGGGCACATTACA	129	A
120	Ca2C718	Ca2C718*	AATACGGAAGCCACTTCCTC	GTGTAGGCCATCCAACTCTG	129	A
121	Ca2C895	Ca2C895*	TTCAAGCACTTCCAAAAAGG	CCGTATCCATGTCTTCGTCT	134	A
122	Ca2C937	Ca2C937*	GTTCGCGACAAAATTCAAGT	TATGTGAAAGGGCCAAAGTG	107	A
123	Ca2C1177	Ca2C1177*	TCTGCTGGTGTTGGATTTTT	GCCAAACCAAGTTATCCTGA	122	A
124	Ca2C1600	Ca2C1600*	ACTACAGTGAGCCCCAAGTG	AAGTAACACGGTCATGAGGC	111	A
125	Ca2C1891	Ca2C1891*	GATACAACACCAGGTTTGGC	GAATCAAAGGGGGTCTGAGT	103	A
126	Ca2C1947	Ca2C1947*	GAACTGGAGAGGGGAAAGAG	TCTTTGAGGACCACTGAAGC	549	A
127	Ca2C3096	Ca2C3096*	TGCCAACTCTGCTAGATTCC	GGTATGTTTGTGGCAGAAGG	102	A
128	Ca2C3281	Ca2C3281*	ATCCGGTAAGAGAAGAAGCC	CATTCTTGGATCTGACACTGC	105	A
129	Ca2C4224	Ca2C4224*	GCACACCACGCATGTATAAG	TTGAATGTTTGGTGCCTTCT	165	A
130	Ca2C4763	Ca2C4763*	GGAAAGCTGATTTGGTCAGA	GGAATAAAGCCACCAATCAA	115	A
131	Ca2C4876	Ca2C4876*	AACTAAAATATGGCCAGCCC	CCAAAAATTGGTTTGCTCAC	121	A

132	Ca2C2481	Ca2C2481*	GACACGAATTTCCATTTCCTT	TTGAGTCTTCTTCTCCACTTTTG	180	A
133	Ca2C4958	Ca2C4958*	TAAAAGTGAACCCCACCTCA	GGTGTTCTTCCCTTCCTCTC	137	A
134	Ca2C5792	Ca2C5792*	GTCAATGTTTCGATGAACCC	TTGTTAAGTGCGTGTAGCCA	240	A
135	Ca2C6058	Ca2C6058*	TGGAGTTTTGGGTTCCACTA	CCCAAATTTACAAGCACACC	130	A
136	Ca2C6533	Ca2C6533*	TAGCACCAAATTCCCCATTA	TTGTGTTGACAATTGGTTCG	202	A
137	Ca2C7508	Ca2C7508*	TACTGTTGGTGATGCCAATG	TGCTTGAATCACAACCTTTG	161	A
138	Ca2C7909	Ca2C7909*	TTCCTGTGGCCAAGTAAAAA	CGAGGAATTAGCCTTGATGA	152	A
139	Ca2C8226	Ca2C8226*	CCAAGAAGTGGGAATAGGGT	TTTTGTCCAGCTTTGCCTAC	299	A
140	Ca2C8240	Ca2C8240*	TAGGAGGAGCTGTTCCCTTT	TAGCCTGTTTCGTCTGAACC	293	A
141	Ca2C10202	Ca2C10202*	CAGCTAGCCTTGCTCAACTC	TGATCAGTGGGTAGGGAAGA	268	A
142	Ca2C8636	Ca2C8636*	TATAACGGGAGCGAGTGAAG	AAATAATCAAGGAATCGGGC	100	A
143	Ca2C8663	Ca2C8663*	TCAAACCCTCCCATCACTAA	TCACCTGGGTTGTTGACTTT	222	A
144	Ca2C9452	Ca2C9452*	AAGTGTGGATTCATGGCTGT	TAACGACGTTGGCTTGGTAT	198	A
145	Ca2C10102	Ca2C10102*	AAAGGTTAACTGCTTGGCCT	AGCCAACTCACCAACATCAT	133	A
146	Ca2C10519	Ca2C10519*	TGTGCTAAGGCAATGAAACA	TTGGGAAAACATCTTGCAGT	167	A
147	Ca2C10999	Ca2C10999*	TCCACCACCATATGAACACA	GGTGGAGGTGGAGACTTGTA	115	A
148	Ca2C11150	Ca2C11150*	TTGCCGGAAAATTTCATTTA	CAAGTGGGGTTTATCTCGTG	125	A
149	Ca2C11438	Ca2C11438*	TTTTTCTCCCCCTGGTAAAC	GCTCCACCTTAACGACAGAA	115	A
150	Ca2C11620	Ca2C11620*	ATGGATGTTCAGGAGTGGAA	TCTGCTTGGAGTGCTTCTTT	110	A
151	Ca2C11901	Ca2C11901*	TCACAGATTAACCCGATTTCA	AAAGCGTATTCGACAAGCAT	143	A
152	Ca2C12236	Ca2C12236*	AGTTGAAACTCCTTGCACCA	AATTCAAGGGAGGTCAAACC	202	A
153	Ca2C13663	Ca2C13663*	TCCTTGGCTCTTTCAAACAC	ATGGAATCCAAGTCAAACGA	212	A
154	Ca2C13971	Ca2C13971*	GTTGCAAGCTTTGGAGGATA	AGCCAATCAATTCTACAGCG	203	A
155	Ca2C14371	Ca2C14371*	AGCAATGGGAAGTAGCCTCT	ACCGCCAGTAGCCTTAAAGT	121	A
156	Ca2C15002	Ca2C15002*	CAGTGGACTCTTCAATGCCT	TTAGTTGTTTTGCCGACTCC	111	A
157	Ca2C15489	Ca2C15489*	AAGCAGAGGAAAAGGGTTGT	TTACGGACGAACGAAAGAAG	285	A
158	Ca2C15544	Ca2C15544*	TGGTTTGAAGTCCATCACCT	TTTATAATCCTGCGCGACTC	166	A
159	Ca2C15563	Ca2C15563*	CAAACAGCAGAATGCAAATG	GGAGATAAAGGTGCTGGACA	192	A
160	Ca2C15856	Ca2C15856*	TTCAACGCCTTCTCCACTAC	AAAATTGCCTAGCGTGTTTG	142	A
161	Ca2C15915	Ca2C15915*	GGTGTGATCGGTTTCTTGAC	GTGATAAACTGTGATCCGGC	103	A
162	Ca2C16301	Ca2C16301*	TCCCGAATCAGATTTACATGA	TACATGGGATTTGTTGTCCC	120	A
163	Ca2C16735	Ca2C16735*	TTGCCCTTCAAGATATTCAGA	TTGGATGCAAGAATCAAGGT	109	A
164	Ca2C17050	Ca2C17050*	CAACCACCTCAATCCATTTC	ATGTACACCACCTCCTTCCA	221	A
165	Ca2C17309	Ca2C17309*	TCAGAGTTTGAACAATGGCA	TGCAAACATTGGTTTCAACA	109	A
166	Ca2C17503	Ca2C17503*	TCCAAATTGCATTGAGGTTT	GAATCATGCAGGAGCAACTT	251	A

167	Ca2C17694	Ca2C17694*	CCCTCGTGGGGTTCTACTAT	CAAACTCAGATCCCACAACC	113	A
168	Ca2C18093	Ca2C18093*	GCAAGCTAGCACACTAAGGC	ATTCCTACCAACTTCCCGAC	253	A
169	Ca2C43617	Ca2C43617*	GATGGTTATCTTCAGCGTGG	AAACCTCCACAATCCTGACA	292	A
170	Ca2C18525	Ca2C18525*	TTCGTGTCTTTGGACCAGAT	TTGATTCAGCTTCATCAGCA	165	A
171	Ca2C18951	Ca2C18951*	GCAGCTGCATAGATATGGCT	TGTGCTTACGCTATGCTCAG	101	A
172	Ca2C19263	Ca2C19263*	ATTTCAAGTGGTGGTTCCAA	GCTTCTTCCATTGCTGGTAA	166	A
173	Ca2C19340	Ca2C19340*	TTTGTTCACGTGCTTTCTGA	CGATGCCACTGACAACATTA	170	A
174	Ca2C20500	Ca2C20500*	AGAGATGGACATCCTTGCTG	CAGTCTTTTCTCCTTGTCCG	124	A
175	Ca2C25383	Ca2C25383*	ACATTGGAACTCCAGCACAT	GATCTTGTCATTGGTGGCTC	184	A
176	Ca2C20993	Ca2C20993*	GTGTTATCACAGCTCCCACC	TTATGTGATCTTGTGCCCCT	368	A
177	Ca2C21034	Ca2C21034*	GTGGAAGAGGGTTCGAAGAT	CCCCTACGTTCTTGTCCTCT	162	A
178	Ca2C28092	Ca2C28092*	AAAGGTACCTCCACCCAAAG	ACATAATTTTCCCCCTGCTC	245	A
179	Ca2C25135	Ca2C25135*	GATCAGCTACCAGGAAGCAA	AGAAAAAGCAATCCAACATCA	167	A
180	Ca2C25529	Ca2C25529*	TTCCAACATCTTGGCTTCTC	GAAGTGCAGGAGGCAATAAA	198	A
181	Ca2C25573	Ca2C25573*	AGACTGTGTCGTGTGGGATT	TGCAGTCTCACAAGTTGAGC	105	A
182	Ca2C25632	Ca2C25632*	GGGGGAATAGAAGAGGATGA	AGCAAACAATTGGCAAGAAG	140	A
183	Ca2C26543	Ca2C26543*	TGCAAAAGCAGAAAAGGAAC	CGATCCATCTGTGGATGACT	108	A
184	Ca2C26702	Ca2C26702*	TCTCTCAGGATTGCAGCTTT	CATACCAGACCCTTCAGGTG	117	A
185	Ca2C27208	Ca2C27208*	AAGAATCGCATCGACTTCAA	ACAAAGGGTTTTGGTGTTTG	108	A
186	Ca2C29710	Ca2C29710*	TTATACGAGGCTGACTTGGC	TCGGACAAAAGATCTTGGAG	119	A
187	Ca2C27301	Ca2C27301*	CAGAAACTAAGTTGTTGACCCA	GAGATTGTGATCAAGCCAAGA	141	A
188	Ca2C28151	Ca2C28151*	GGAGTTCAGTCCCAATGACA	GAGCTGGAGGATTACACCAA	151	A
189	Ca2C28957	Ca2C28957*	AACGTGTAGAGGGAGTCACG	TTTTTCGATTCAAACCGTCT	313	A
190	Ca2C29208	Ca2C29208*	GCCGCATCTCTTCACACTAC	CTGGAATTGAAAGTGGATGG	125	A
191	Ca2C29302	Ca2C29302*	GCTTTGGAGCACACGAGTAT	GAATGTCCGAGATGATCAGG	349	A
192	Ca2C30432	Ca2C30432*	AGAAAACCGCGTTTGTTAAG	TCGAAACTTAACCGGAGAGA	148	A
193	Ca2C18912	Ca2C18912*	GTCCAACAATCCTCTCCCTT	CCTCCTAGAGCGTGGTTACA	378	A
194	Ca2C20537	Ca2C20537*	GATAGAGTCATCCTCCTATCCA	CTTTCCCACTTGGAACTTTG	725	A
195	Ca2C24067	Ca2C24067*	TATACCCAGTGAGGTTGTGAG	GGTAATAGGGTTTAGGGTGAAC	487	A
196	Ca2C9868	Ca2C9868*	CTGTATCCACTTTGTGCATC	CCTGACCCTGATGCTCATTT	468	A
197	Ca2C31575	Ca2C31575*	AACTAATGCTCGGACTGCTG	AGGCTTGTGTCCTTTCCTCT	219	A
198	Ca2C31704	Ca2C31704*	AGCAAGTTTGAGGAACACCA	TTGACATAGGAGGGGATCAA	266	A
199	Ca2C31732	Ca2C31732*	GGAGGTAACAGCAACGTGAT	ACAATTACTCTCAGCCACGC	143	A
200	Ca2C31975	Ca2C31975*	TGTCCTCAACACTGTCGCTA	TCTTTTCAAGCCAATGCTTC	215	A
201	Ca2C32155	Ca2C32155*	CCTTTGAGATTCACTGCGTT	GAAAAAGCAAACGGTTGAGA	129	A

202	Ca2C25298	Ca2C25298*	TGTGCACTGGAAGACAAAGA	TTCAGACTTTGGTGCAGACA	100	A
203	Ca2C32781	Ca2C32781*	GACCAACAGCATTAACGAGG	GTAAAGGTTCATTCCCCGTC	130	A
204	Ca2C33004	Ca2C33004*	TTGCAACAACAACCTGAAGA	CCCAAGTTGCATCTAAGGAA	257	A
205	Ca2C33287	Ca2C33287*	CTAGTCGGGTGAACCATCAG	CTTGTCCCAATTGTCTTCCA	127	A
206	Ca2C33518	Ca2C33518*	TCCACATTTTCAGAAGAGGC	CTGCAACTCCAACACCTTCT	132	A
207	Ca2C34104	Ca2C34104*	GGTGACAACTCCCAGCATAG	CACAAACAACAACCACAGGA	276	A
208	Ca2C34140	Ca2C34140*	AAAATTTTAAGCTACGCACCA	TAGCAGCTTTGCTCTGTGAA	127	A
209	Ca2C34487	Ca2C34487*	GAGAAGCCTCTCCGGTAGTC	TTGGATGCCCTTATGTTGTT	111	A
210	Ca2C34663	Ca2C34663*	TTTTGTTAGTTGACCGGGAC	TTCTCCCTGAATTTGACACG	123	A
211	Ca2C34997	Ca2C34997*	TACCTTTTGGCTTTGGAAAA	GCAAAAGGGAATTGTGTCAG	129	A
212	Ca2C35634	Ca2C35634*	AAACCAAACCGGAAAATCTC	TCGTTGACAGTGTTTCCTGA	141	A
213	Ca2C35738	Ca2C35738*	TGCTATGAGTCAACTTGGCA	GGTCGATGTTATCCTGCAAC	510	A
214	Ca2C36438	Ca2C36438*	CTTTTACGGCTCCGTCAGTA	ATGCAGATCGTGGGTTATGT	102	A
215	Ca2C36468	Ca2C36468*	AGCTTCAGGTTGATTTTCCC	TGAGCTTGAAAGCATTGACA	192	A
216	Ca2C36478	Ca2C36478*	ACATTTGACACAACTGCCCT	CGGTATTGACGCTGCTACTT	255	A
217	Ca2C36741	Ca2C36741*	CTAACCTTGACGCTCGTGAT	CTGGACCACGTTAGGATGTC	362	A
218	Ca2C37084	Ca2C37084*	TGGAGAAAGAAGCCTCTGAA	CCAATCCTGCATTTCCATAG	106	A
219	Ca2C37086	Ca2C37086*	ACCACTCTATCCTTGTGCCA	TGATCAAATCACGGTTGATG	101	A
220	Ca2C37158	Ca2C37158*	CACGAACAGAAAGCCAACTT	TCAGGAGCCATTTGCTTTAG	103	A
221	Ca2C37361	Ca2C37361*	CAAACCACACAAAAACGACA	TCTCTCCATCTTTGATTTCGAT	108	A
222	Ca2C37892	Ca2C37892*	TTGTCTCCGAAGGATCAAAG	AATAACCTCCCAACCTGCAT	180	A
223	Ca2C38116	Ca2C38116*	GAGAGCTGGAGCAGAGTTTG	TTCCACAATCAATCAAAGCA	166	A
224	Ca2C38259	Ca2C38259*	GCTGAAGCGTATCTTGTTGG	TAAGCCCTCTCGCCTCTAAT	125	A
225	Ca2C38337	Ca2C38337*	AACAATTCCATGGCCACTAA	TGTAGCCATTGAAGACACCA	256	A
226	Ca2C42668	Ca2C42668*	AGGCACTTGTCATCCCTGTA	CATCTGAGACATCTTCCGCT	144	A
227	Ca2C38636	Ca2C38636*	TTCTCTGTGCTGTCCAACAA	TGGCGTAGCAATAAGTGACA	229	A
228	Ca2C38995	Ca2C38995*	GCTTATCCCCAAACTTGGTT	CTGACATCCTGACCAGACCT	192	A
229	Ca2C40609	Ca2C40609*	CGCGAGAGACAAAGAAAGAG	ACCCACTGGGAGATAAGACC	106	A
230	Ca2C40649	Ca2C40649*	GTTTCGACATGGCAAAATTC	GATACGTCCTTTCCCATCCT	290	A
231	Ca2C41045	Ca2C41045*	CCAATCCTATTTCCGAACCT	GATTACGACGACAGCGAAGT	231	A
232	Ca2C41861	Ca2C41861*	GTTGTTAATGGTCGTGAGGC	AGCGTGCTTACTGGTCAAAC	359	A
233	Ca2C41871	Ca2C41871*	CACATGCATTCAACCAAAAA	TGGATAAGGGACCAAGAACA	195	A
234	Ca2C42103	Ca2C42103*	AAAATGGGTTGGAAAGAAGG	GATTGGGTCTGTGACGGTAG	193	A
235	Ca2S002345_0787_1458	Ca2S002345_0787_1458*	AACGTATTCGGCCCTTTATT	ACTGAACATTTTAACTCGGTTTG	100	A
236	Ca2S040884_0954_0411	Ca2S040884_0954_0411*	TGGTGTTTTAGCTGTCTCCAG	TGAAACCTCATCTTCATCCC	133	A

237	Ca2C22191	Ca2C22191*	GGATCCAACTAGGCTTTGCT	CACGAAAACATAACACGCAA	114	A
238	Ca2S047648_1460_3780	Ca2S047648_1460_3780*	CACAATGTAAGGGGATGAGC	GGGAGTCAAGCTTCATAGCA	154	A
239	Ca2S073454_1132_0586	Ca2S073454_1132_0586*	GCTACTCAATGCCAAATGCT	CAAAACTCCAAAGCCTCAGA	220	A
240	Ca2S117228_1922_1005	Ca2S117228_1922_1005*	TTTGGTAGCAGAAACCTTGC	GCAGTAAGATGTTGGAAACACA	186	A
241	Ca2S126415_1648_0587	Ca2S126415_1648_0587*	TGCATCCATAGATACAACAAACA	CAAGAAGCATTTTGTCCTTACA	106	A
242	Ca2S111188_1421_0319	Ca2S111188_1421_0319 *	GGATGATGCAATCCAAATTC	ACTTCGAACTTTACCGGCTT	133	A
243	Ca2S170394_1286_0297	Ca2S170394_1286_0297*	GAATTCCCTCGTTTTCCTTC	AAATGAATTGTGGAAAGGCA	114	A
244	Ca2S195010_0658_0713	Ca2S195010_0658_0713*	CATCTTGAGAAAGGCCAGAA	CATGTCAACAGCGCAAAAC	103	A
245	Ca2S211594_0676_0883	Ca2S211594_0676_0883*	ATGCAGACAGTCCTGGTCAT	GCATCTTTCACCAACCATTT	116	A
246	Ca2S216025_0094_0932	Ca2S216025_0094_0932*	CAAGGTGACGACATTTCTGA	GATCCGGTGGCTATGTTAGA	145	A
247	Ca2S138620_0847_0184	Ca2S138620_0847_0184*	ATGTTTGAAGAAGGATGCCA	ATCATCATCCCACTCGTCAT	111	A
248	Ca2S295752_0698_0195	Ca2S295752_0698_0195*	AAGTTGCCAATTCCAGACAG	TTTATCGGGTTTGTGATGCT	163	A
249	Ca2S386630_0787_3874	Ca2S386630_0787_3874*	TTTTGTTGTCCATCAATTCG	TTCGAAGATTAATCCAAAGCA	101	A
250	Ca2S390799_0796_3257	Ca2S390799_0796_3257*	TTTTGCATTTACCACGTGAA	TTCCATGCAAGAGTCATTAGG	121	A
251	Ca2C16174	Ca2C16174*	AGAAAATGTCCAAGCACAGG	ATTGGTTTTGGTGGTTGTTG	132	A
252	Ca2S092962_4005_1590	Ca2S092962_4005_1590*	TTTAAGGAAGGGGGAGTCAA	CATGAGTACATGACCCACCA	103	A
253	Ca2S094966_2754_2656	Ca2S094966_2754_2656*	TGAAAACACGAGTCAACAAAA	GTTTAGAAACCTTCGGGCAT	119	A
254	Ca2S110494_3324_2020	Ca2S110494_3324_2020*	TCTGACATCAATGAACAAGCA	TTTCCGTTTGTAGTGGTGGT	177	A
255	Ca2S125919_3413_2502	Ca2S125919_3413_2502*	CACAATGTAAGGGGATGAGC	TTTTGGATAAATTGACACCTCTG	129	A
256	Ca2S141149_3853_3766	Ca2S141149_3853_3766*	GAAATACCACGACACGCTTC	CGCCTCTAAGTCAGAATCCA	131	A
257	Ca2S143601_3004_0194	Ca2S143601_3004_0194*	GCCAAGTGGAAAAGAAAGAA	ATCCCATTGAAGAACAGCAG	166	A
258	Ca2S164036_3651_2669	Ca2S164036_3651_2669*	AGAGTAATAGTGGGGGTGGC	GTATGAAAGGCCAGCAGCTA	120	A
259	Ca2S180175_2544_1644	Ca2S180175_2544_1644*	GGAAGCAGCCATCTAAGGAT	CCAAACCTCAAATAGGGGTC	148	A
260	Ca2S191429_3805_3465	Ca2S191429_3805_3465*	GCTCAAGCCAGAGAAGTTTG	AACAGAGAATCATCCCATGC	122	A
261	Ca2S198222_2727_4051	Ca2S198222_2727_4051*	GCCACGTTAATACAACAGGG	CGGGGGAATGTCTTTTAACT	102	A
262	Ca2S217931_2289_0990	Ca2S217931_2289_0990*	GAATGTAGGTGCTGGAGGAG	CAGGAAGGTCTGAAGATCCA	119	A
263	Ca2S224825_3107_2806	Ca2S224825_3107_2806*	GTGCTGGTTTCCATTTAACG	CGCCATTAGAGAAAGATGGA	112	A
264	Ca2S231566_3879_0955	Ca2S231566_3879_0955*	GTAGGGATACCCGCTGAACT	CCTATGTGACGTCCTGTTCC	197	A
265	Ca2S232393_3220_2305	Ca2S232393_3220_2305*	AGTGACTTGGTCAAATGGGA	AATATCAGCATGGTTTCCGA	136	A
266	Ca2S242431_3223_1815	Ca2S242431_3223_1815*	AGAGCACATGACCCACTCAT	GAACAAAACGTGACAAGGCT	100	A
267	Ca2S299215_3871_3476	Ca2S299215_3871_3476*	TGCTACTTATCCATGGCCTC	AAACTCCCCATTTTCATATTCA	111	A
268	Ca2S314748_3499_3314	Ca2S314748_3499_3314*	GCCAAACGGAGATTAGATGA	ATTCCCCTATTTGGGTTCAC	109	A
269	Ca2S139542_0526_0682	Ca2S139542_0526_0682 *	TGAAATTGAATGAGGGGAAA	TGGCTCCTCCACATAGTCAT	110	A
270	Ca2C43237	Ca2C43237*	TTTGGCCTGAAGTGAAAAAG	GTGTTTGGTTCTACGTTGGG	160	A
271	Ca2C39196	Ca2C39196*	CTTGCAATCCTCTGGTTCAT	GTGTGGCACTGAAGGGATAG	143	A

272	Ca2S417306_2592_2701	Ca2S417306_2592_2701*	CCCTCTTGTACTGCTCTCCA	TATCCGGCGAGTATTGTGAT	129	A
273	Ca2C10187	Ca2C10187*	GCCCAATCCGATTCACTTTA	CCCCACGCTAATTCCAGTAA	189	A
274	Ca2C13532	Ca2C13532*	TGCTGTTCTTGGTGTTCCTG	CCTACCATGCGGAAACACTT	527	A
275	Ca2C13931	Ca2C13931*	TTGGCAACAGAGAAGCATTG	TCACAACAACACCTTTCCCA	844	A
276	Ca2C13931	Ca2C13931*	TTGGCAACAGAGAAGCATTG	TCACAACAACACCTTTCCCA	844	A
277	Ca2C15380	Ca2C15380*	TGGAGGGGTCTTCCTTTTCT	TCCCCAGCAGTATTCTCCAC	272	A
278	Ca2C15992	Ca2C15992*	GCATGCAACAACCAAGTCAT	AAGCAGGTGGATGGATTGTC	182	A
279	Ca2C16244	Ca2C16244*	TGCTCTCCAGCAGTGTACCA	TTTGGGTTTGATCCTCTTGG	176	A
280	Ca2C16326	Ca2C16326*	AAGAACAAGGGCATTGATGG	TGCGACCTTAACGAAATCCT	177	A
281	Ca2C18042	Ca2C18042*	TTGCATGGTTTGGTTCTTGA	TGACTTTGCTATTCCCACCC	361	A
282	Ca2C19645	Ca2C19645*	GTGAGATCGTGAAACTGCGA	GCAACCCATAGTGGGCTAAA	362	A
283	Ca2C19645	Ca2C19645*	GTGAGATCGTGAAACTGCGA	GCAACCCATAGTGGGCTAAA	362	A
284	Ca2C20444	Ca2C20444*	CTCGCCCCTTGAAGTAATTG	TTTGCAGGTGCAGTGATGAT	304	A
285	Ca2C20877	Ca2C20877*	ATTTTACATTGGGGGCGTTT	TCGAGGGGTTAATGGATCAG	182	A
286	Ca2C21617	Ca2C21617*	GCCGGAGAATTTCTTGTTTG	CCTCGAAGAATAGCCTGTGC	267	A
287	Ca2C22325	Ca2C22325*	TGAGATGATTCGTGAGGCTG	TTGTTGCATTGAAGCAGGTC	258	A
288	Ca2C22909	Ca2C22909*	TTTTCAGCACGAGTCCATTG	ATGAAAGAGCAGGCCATCAC	230	A
289	Ca2C23993	Ca2C23993*	TTGGCATACACATCAGTGGAG	TTTGTTCACCAGGAGGCTCT	150	A
290	Ca2C25794	Ca2C25794*	AGAGGCGGACAAGAGATTGA	TGCAAGACGCTCATAACCAG	288	A
291	Ca2C28364	Ca2C28364*	GGTCGTCGTAGCAGGAAAAG	CAGTGACGGTTGAGTGGAAA	157	A
292	Ca2C28387	Ca2C28387*	AGAGCTTTTGACCTTGCTCG	AGAGCAGTGCAACCTGGATT	166	A
293	Ca2C29331	Ca2C29331*	TTCACACCTCTCGGTCACAA	GTGGAAGGTTGTTTGCCACT	256	A
294	Ca2C30125	Ca2C30125*	AGCAAGAGAAGGACCATTCG	TCCTTTGGTCCCAAAAACAG	183	A
295	Ca2C31913	Ca2C31913*	GCAAGCACACCAGCTTTGTA	CAACTTTGCCAGCACAAGAA	313	A
296	Ca2C32365	Ca2C32365*	CCGGCAGGAATACTGTTGAT	GGGTCCCGTACTTTGTTCCT	352	A
297	Ca2C32365	Ca2C32365*	CCGGCAGGAATACTGTTGAT	GGGTCCCGTACTTTGTTCCT	352	A
298	Ca2C32639	Ca2C32639*	TGCCACTATGGCACTCTCAG	GAGCTCCGCAGTGTGGTAAT	263	A
299	Ca2C32801	Ca2C32801*	ATGGCACTTGCTTCTCGTCT	TTCGCATAATGAGCAACAGC	254	A
300	Ca2C33338	Ca2C33338*	TTTCTCCGGAAACACTACCG	CATGTCAGTGGAATTGCTGC	372	A
301	Ca2C34020	Ca2C34020*	ACCAAGGCACCAGATACCAG	GCATGAGGCAACAGAGTTGA	177	A
302	Ca2C3452	Ca2C3452*	AAGGCACCAAAACTGATTGG	GGCTTGTTGCTGTTGTCAGA	229	A
303	Ca2C3559	Ca2C3559*	ATGTTTGGTGGACTTGGAGC	TACCCAACATGGCACTTGAA	285	A
304	Ca2C3559	Ca2C3559*	ATGTTTGGTGGACTTGGAGC	TACCCAACATGGCACTTGAA	285	A
305	Ca2C35818	Ca2C35818*	CAAAGGCTCCTCCTTCACTG	TGCCTTGTTCAGAGACGTTG	206	A
306	Ca2C38039	Ca2C38039*	CGGTAGGAACAGGGAAAGTG	GCCCTGAAGCTGTTGATCTC	153	A

307	Ca2C38039	Ca2C38039*	CGGTAGGAACAGGGAAAGTG	GCCCTGAAGCTGTTGATCTC	153	A
308	Ca2C38830	Ca2C38830*	CTCCATCCATGACCTGGTTT	ATCATGAAAATGCCCGTGTT	270	A
309	Ca2C40036	Ca2C40036*	GCAGAGCTGACCCGAGTTAC	CGCTTGAAACCAATGAGGAT	320	A
310	Ca2C4012	Ca2C4012*	TACGGCGTCGTCTAGGTTTT	TCTTCTGTTCCGGGTAATCG	297	A
311	Ca2C40740	Ca2C40740*	TCTCGCGGATTTTCATAACC	GGAATACCCGGAAACACCTT	823	A
312	Ca2C40740	Ca2C40740*	TCTCGCGGATTTTCATAACC	GGAATACCCGGAAACACCTT	823	A
313	Ca2C40740	Ca2C40740*	TCTCGCGGATTTTCATAACC	GGAATACCCGGAAACACCTT	823	A
314	Ca2C41295	Ca2C41295*	CACTTGCTCCCGATCTTCTC	CATACGAGGCTCTGGCAAAT	162	A
315	Ca2C41312	Ca2C41312*	TGCTCCACCAACAAACTGAA	GTCGTCCTGGATATCGGCTA	211	A
316	Ca2C42415	Ca2C42415*	TGCTTCAGTTCCCACTCCTT	GCCAATGTGACTCCCAAACT	151	A
317	Ca2C42428	Ca2C42428*	CACCGTAGGCTATTCGAAGC	CGGCGGAAGTAGAACAAGAG	152	A
318	Ca2C43714	Ca2C43714*	CCATCGCTATGTTTGGAGGT	TGACTTCTCCCCTATGTGCC	330	A
319	Ca2C43714	Ca2C43714*	CCATCGCTATGTTTGGAGGT	TGACTTCTCCCCTATGTGCC	330	A
320	Ca2C4756	Ca2C4756*	TTTTTGGTTCGGTGAAGACC	CAATTCGTGCCAACATGAGA	192	A
321	Ca2C5423	Ca2C5423*	AGGTTTGGTTGCATCAGAGG	TCTGGCCTTGACAGTAGCCT	152	A
322	Ca2C5918	Ca2C5918*	GAGGTCCAACAGCTGCTTTC	TTCATCGTATCCATTTCGCA	492	A
323	Ca2C6285	Ca2C6285*	GGTCCCACATTTCATCCAAC	CGTTGTTCCGGTTTCTGTTT	253	A
324	Ca2C6628	Ca2C6628*	GTACTTGGGTTCCAAGCAGG	TTTACGAATCGATGGCACAA	174	A
325	Ca2C6628	Ca2C6628*	GTACTTGGGTTCCAAGCAGG	TTTACGAATCGATGGCACAA	174	A
326	Ca2C6628	Ca2C6628*	GTACTTGGGTTCCAAGCAGG	TTTACGAATCGATGGCACAA	174	A
327	Ca2C7769	Ca2C7769*	GTGATTTTGAAGGGGCTGAA	CACAGCCTCTCATGCACATT	450	A
328	Ca2C7850	Ca2C7850*	AAGAGATGGTGAAACGGGTG	GAGACCAGCCAAATTCCAAA	368	A
329	Ca2C8619	Ca2C8619*	AGGGGCTATCTAGGGTTGGA	AAAACAGCTCTTCCTTGCCA	263	A
330	Ca2C9884	Ca2C9884*	GCTCTTGCACGGTCAAGTTT	GGAAACACGGCTAACGAGAG	164	A
331	Ca2S024238_0616_1943	Ca2S024238_0616_1943*	TGAATCAATTGTTGTGCCGT	CCTATGCGGAGGACATCATT	163	A
332	Ca2S029795_3667_1892	Ca2S029795_3667_1892*	GGGGGAAAACAAAACATTGA	GCTTCAACAAGATGAGCACC	234	A
333	Ca2S032873_1646_3001	Ca2S032873_1646_3001*	GAAACTCAAGCTCAGCCGAC	TGCAAAGTCATCTGCTTCTGA	158	A
334	Ca2S036352_0372_2258	Ca2S036352_0372_2258*	TTGCGTGTCTCAGTGCTTTT	TTTGAAGGTGTTGGTGATGG	263	A
335	Ca2S037055_0416_1164	Ca2S037055_0416_1164*	AGTAGCTCCCCAATCCGTTT	AACGTAAGCAACATCCTGGC	154	A
336	Ca2S039865_2799_2981	Ca2S039865_2799_2981*	AAGCCATTCGATGGTTTGAC	AACAACTAACGCCGTAACCG	208	A
337	Ca2S046433_0281_0387	Ca2S046433_0281_0387*	CCTGAAGGGAAATTACCAACAA	TCCTGGTGGAACAAAAATGG	150	A
338	Ca2S051745_1193_3063	Ca2S051745_1193_3063*	GAACAGCCTTCCAATTTCCA	GGGGGAACACATTCCTAACA	241	A
339	Ca2S065160_2442_0289	Ca2S065160_2442_0289*	ACTCACGCTCAACACAAACG	CACGTCCCATTGCTTTAAGA	186	A
340	Ca2S082575_1212_1008	Ca2S082575_1212_1008*	ACCATGTTTCCCAGAACTCG	CGCACTATGGAAATGGGAAG	150	A
341	Ca2S107358_1454_0701	Ca2S107358_1454_0701*	TGTCTCATACGAGCTTCCCA	TGCGGATATCAATTTCAACG	177	A

342	Ca2S108294_3225_2835	Ca2S108294_3225_2835*	ACTGCCCTTCCGGAAAATAC	CGCTTTTGAATCTTCCAACC	152	A
343	Ca2S115524_2391_2662	Ca2S115524_2391_2662*	TCCCTTTTCTCAGGATCAATG	GAAAGGGAATCCTGCAATCA	172	A
344	Ca2S120778_2784_3503	Ca2S120778_2784_3503*	CCCGGTGACACTACCTCAGT	GTGCCAATCCTTTCCTTTGA	156	A
345	Ca2S128786_0400_2573	Ca2S128786_0400_2573*	AGATGAGAGGAAAGGGTGGG	ATCACAAACATGCGTTTCCA	206	A
346	Ca2S130035_1355_1630	Ca2S130035_1355_1630*	GAGTTGAGCCCTAAGCACAGA	GGAATCGGCAAGAAGATCAA	175	A
347	Ca2S136450_3946_0702	Ca2S136450_3946_0702*	GCAGTTAAGGAACAAAGCCG	GCGGTAAGAGGAACGAGGAT	156	A
348	Ca2S155008_0540_1008	Ca2S155008_0540_1008*	CAAAGCAGAGGCAATCACAA	ATTCGCGGAGATTGTCTGAT	177	A
349	Ca2S162925_1065_2547	Ca2S162925_1065_2547*	TGGGAAGAGCTCCTTGTTGT	TCCCACACCTTAATAATCCTGC	150	A
350	Ca2S183433_2797_2969	Ca2S183433_2797_2969*	GCAGAGGAGCACAGTGATGA	TCAAACCACTTCCACACAGG	197	A
351	Ca2S191027_2342_1347	Ca2S191027_2342_1347*	CCTCGGGAAAACGAATATAGG	GCTTGAAACCAGCCCTTGTA	151	A
352	Ca2S206370_3058_3496	Ca2S206370_3058_3496*	GATGGAGAGTGTCGTGAGCA	TTCTTCTATTGGAGGAGGAGGTC	153	A
353	Ca2S207240_3200_3124	Ca2S207240_3200_3124*	AATTCATTGCCCTTGCACTC	AAACAACTTGCAGAAAGGCG	205	A
354	Ca2S214308_3212_4056	Ca2S214308_3212_4056*	TCTGCACGTTTTTCTGTTGC	TCTTAAGGGAGGATTTTCGGA	154	A
355	Ca2S217264_1973_0317	Ca2S217264_1973_0317*	GAGATCCACACAAAGCAGCA	CCGGATAACAGAAACGGTTG	193	A
356	Ca2S219318_3121_0370	Ca2S219318_3121_0370*	TAAGGGACCTTTGTGGGAGA	GTTTTCCTCCTCCCTGCTTC	227	A
357	Ca2S223369_3770_0255	Ca2S223369_3770_0255*	GAGCCAAGGTTTGCATTCTT	TCCACCCTTTTCTTCACAATG	158	A
358	Ca2S229641_1261_3358	Ca2S229641_1261_3358*	TAGGGGGAAAGCCACTAGGT	CCCACTCACAACATCAGGTG	242	A
359	Ca2S230684_0187_1035	Ca2S230684_0187_1035*	TTCACACGTTTTGCAGGAAG	AGCTGGCAATGGTACTGAGG	217	A
360	Ca2S230835_3257_2715	Ca2S230835_3257_2715*	GCTGGTTTGTTTGAGAAGGC	ATGCTAACCGAAGAGAACGC	157	A
361	Ca2S253115_3155_0954	Ca2S253115_3155_0954*	AGGAAGGGTAAAATCGGGAA	AAATGGTTCTGGACACTGGC	170	A
362	Ca2S254168_0423_2095	Ca2S254168_0423_2095*	TCTCGGAAATAGATGGTCCC	TTGACTTGATGTTGTTTTTGTTCA	159	A
363	Ca2S265948_3183_0035	Ca2S265948_3183_0035*	TGCTTTCCCTACAGCCATTT	GCATTGGCTTGCAACTTAAA	158	A
364	Ca2S273250_3787_2637	Ca2S273250_3787_2637*	GGTGCATTTGATGCTCTCAC	GGACAAGGAGAATCTTGGCA	164	A
365	Ca2S276275_0929_3483	Ca2S276275_0929_3483*	CCAATGGCCAATCTTCAGTT	TATGGCTTTCGATTCCTTGC	176	A
366	Ca2S290239_3184_3746	Ca2S290239_3184_3746*	CATTGGAAAACATCAACCCA	CCTTTGGGCCATAATGTTGA	218	A
367	Ca2S309712_2419_0589	Ca2S309712_2419_0589*	AGATGGGTGTGCTAAATGCC	GCGAAGGTTTTTCACCTCAT	193	A
368	Ca2S325480_0441_3276	Ca2S325480_0441_3276*	AAGCAGTGGAACAACTAGGCA	TTTCCGATCTTCTTGCTTCG	155	A
369	Ca2S325752_3151_3658	Ca2S325752_3151_3658*	TGGAGTGCAAGAGAACAACAA	AGGCATGAGGGATCTGAGAA	154	A
370	Ca2S330532_1714_1726	Ca2S330532_1714_1726*	TGTCTTCAAGGGGGTCTTTG	AACAAAACAATGCCCAGGAA	185	A
371	Ca2S334387_0551_2663	Ca2S334387_0551_2663*	GAAATCGATCATTTGAGACGC	AAAAGCCATCATTTGTTGCTG	150	A
372	Ca2S345228_3347_0971	Ca2S345228_3347_0971*	TCTCCTGTTTTCGTCTGTTAATG	AAACAAGCTTTAGCGAGGCA	153	A
373	Ca2S356591_1831_0898	Ca2S356591_1831_0898*	CAACTGATCAGACCAATCAAGC	ACATGTATGGGTGAGGAGCA	150	A
374	Ca2S370669_2947_2194	Ca2S370669_2947_2194*	TGAGCAAGAAGCATCCACAG	TCCACCCCTTAAACCCTTCT	185	A
375	Ca2S385516_0911_3815	Ca2S385516_0911_3815*	AGAAAAAGGACCGGGAGTTG	GAATGGAGCGCCATAGATGT	168	A
376	Ca2S386953_3402_3333	Ca2S386953_3402_3333*	GGTGACAAACATTCGAGTCTCA	CGATTCACGAGAGAACCACA	183	A

377	Ca2S414575_3358_3526	Ca2S414575_3358_3526*	CCTGCTGCTCCAACTTCTCT	CTCCCAAAACCAGACAGAGC	196	A
378	Ca2S424015_2237_3887	Ca2S424015_2237_3887*	AGCACACACATACGGACGAG	GCACCATGAGTTTTGGGTCT	240	A
379	Ca2S431102_0894_2012	Ca2S431102_0894_2012*	TTTGTACGAACTCTTTCCGTGA	GAAGTGGTGCAACAATGTCAG	151	A
380	Ca2S432511_1751_3047	Ca2S432511_1751_3047*	CTGCAGCCACTAATGAACGA	CCACCAAACCTCAAAAGCTC	204	A
381	Ca2C10257	Ca2C10257*	GTGTTGGCGATTCCGTTACT	CCCAACAGCAACCTCATTTT	247	A
382	Ca2C10322	Ca2C10322*	CCATTTGGGAGAGAGGACAA	ACTGCAACCCTGAGTGGAAT	181	A
383	Ca2C10324	Ca2C10324*	GGAGTGTCATTGGAAGGGAA	CGCAGCTGACAGAATGAAAG	230	A
384	Ca2C1043	Ca2C1043*	GAGGCAAATGTCCCCATAGA	AAAATGGTCGTCTCTCACCG	230	A
385	Ca2C10766	Ca2C10766*	CCCGTCAAAGGAGAGCATAG	GCACATGAAGAGGGTGGAAT	153	A
386	Ca2C10894	Ca2C10894*	TTTTTGCAACACAAACCCAA	CAACCGTTGATGTCATGAGG	290	A
387	Ca2C11129	Ca2C11129*	GCAATCCTGAATGGAAGGAA	ATGTCCGTCAGGGAGATCAG	182	A
388	Ca2C12266	Ca2C12266*	CGACCCTTTCACAAGAAGGA	AAGCAAACGGAAGTGGATTG	175	A
389	Ca2C12268	Ca2C12268*	GTGATGGCTTGTTGGGAGAT	GAGACCTTTTGCGTTTGCTC	514	A
390	Ca2C12786	Ca2C12786*	CCCGACAACATCCTCAACTT	TCGATAATTACTCCACCGCC	516	A
391	Ca2C12857	Ca2C12857*	GCCATCTTTTGGTGTGGAAG	CACACACATATAAACAACCCCAA	285	A
392	Ca2C13027	Ca2C13027*	GAAGCGTATGGGTGGTTGTT	AAGCCCTTCTCTCAGGCTTC	162	A
393	Ca2C13088	Ca2C13088*	TTCATGGAGCAAAGGAGCTT	AATCCTTGCACCCTTCACAC	258	A
394	Ca2C13152	Ca2C13152*	GCCACACAAAAAGGAAGGAA	ACAATGTGGCCTCCTTGAAC	286	A
395	Ca2C1329	Ca2C1329*	ATGGCGTTATCACCTCCAAC	TCAGAATGACAAACCCCACA	158	A
396	Ca2C13554	Ca2C13554*	CGGACCCTCACTTACAGGAA	GATGACAAATGCTTGGCTGA	159	A
397	Ca2C13997	Ca2C13997*	GTGGCGGAAACAAAACAAGT	CTGCGTATTGGAGTTGCTGA	332	A
398	Ca2C14433	Ca2C14433*	GAGACAGAGACCATGGGGAA	GCTACTCAAACACCCCCAAA	421	A
399	Ca2C1448	Ca2C1448*	TGCATGAAATTGGCTTGAAA	GGAAGCACGGTGAGAAGAAG	317	A
400	Ca2C14686	Ca2C14686*	GCCTTCCCATAGCAATACCA	CGGTGAATGGAGCAAAGAAT	190	A
401	Ca2C14713	Ca2C14713*	CCATTTCCTGAACCACCATC	ATTCTACACCCAATTCCCCC	607	A
402	Ca2C1520	Ca2C1520*	TGGGATGAATGGGTTGTTTT	TCATTCTTTCGTTTCCTGCC	164	A
403	Ca2C15242	Ca2C15242*	GCAGAAATAGCGTGGCTGAT	ACAACAATTCCGCAACAACA	239	A
404	Ca2C15575	Ca2C15575*	TGTGCATTTTGATGACCTCC	TGAAGGGCTTGTGAAAGAGTC	195	A
405	Ca2C15786	Ca2C15786*	TGTGTGTGACTGGTGCTTCA	GAGGGTCGTCGAAAACAAAA	264	A
406	Ca2C16069	Ca2C16069*	CCCGAAATAGAGGAACAAGC	ATCCCTCACTTTTCTTCGCA	228	A
407	Ca2C161	Ca2C161*	GGACAACCGAAGAGGATCAA	GTTGTCCGTCCTTCCTGGTA	257	A
408	Ca2C16129	Ca2C16129*	GCACGATGGCCTGACTATTT	TTGTCGCTGACGTTTTCAAG	191	A
409	Ca2C1614	Ca2C1614*	CCGGTAAGAAGCAGTCGAAG	TTTCCATTCTTTACCGCCAG	167	A
410	Ca2C1635	Ca2C1635*	GAAGCCAAACGTGGTGAAAT	TGGACAGTTTTCTTGCACACA	199	A
411	Ca2C16445	Ca2C16445*	CCAAGAAACCAAAGGCAAAA	GCAATGCCTGTGTGTGAGTT	264	A

412	Ca2C16757	Ca2C16757*	CCTGGGAGGCTGATAATTCA	ACAAGGGACGGAATCTTGTG	224	A
413	Ca2C17163	Ca2C17163*	CTTGTTGGTGATGATCCGTG	TACAAGCCCAATATTCCCCC	371	A
414	Ca2C17213	Ca2C17213*	TTTGTGTGAGGCACTTTCCA	ACAGCCCCATCATCAAAAAG	456	A
415	Ca2C17777	Ca2C17777*	ATGCGAACATGAGAAGACCC	AATACTTTGGGCCCAGCTTT	205	A
416	Ca2C17798	Ca2C17798*	GAAGGACAACCCGAGTAGCA	ATACGCTCGCTCGAAGACAT	184	A
417	Ca2C1781	Ca2C1781*	ATTTGAAGCCATTTGCTTGG	TGCAATGAATGGTGAGCAAT	314	A
418	Ca2C19051	Ca2C19051*	GCCACACTTGCTTCCTCTTC	CTATCACAGGCCTCACAGCA	517	A
419	Ca2C19555	Ca2C19555*	CGTATTTCTGCAGCCCATTT	GTCTCACGGCCTCTGAACTC	229	A
420	Ca2C19584	Ca2C19584*	ACAAGTTGTTGGTTGGCCTC	ATCTCCCAAGGTACATCCCC	372	A
421	Ca2C19589	Ca2C19589*	TCCCATTTGGACTCACACAA	GCTCGTCTTCCACTCTCCAC	290	A
422	Ca2C1977	Ca2C1977*	GAGCCAGTGCGCATGTAATA	TATTCAAGGTCTGGGGGTTG	155	A
423	Ca2C20014	Ca2C20014*	GTTCCTTCTCAGCAAGTGCC	GACCATGTGTTGACACTGCC	782	A
424	Ca2C20094	Ca2C20094*	GAAGTCAGGGTCTGCTTTGC	TCCAAGAACCCCTAAAGCCT	899	A
425	Ca2C20158	Ca2C20158*	CCTTCTCATCACAAGCACGA	ACCACCATCATCATCCACCT	204	A
426	Ca2C20319	Ca2C20319*	TTCCTGGCTCTGAGTTTCGT	TTTAGGCCTTTAGGGGGAGA	721	A
427	Ca2C20659	Ca2C20659*	GCCGAGCTCTCTTATTGGTG	GAAGAGGACGATGGATTGGA	211	A
428	Ca2C20818	Ca2C20818*	TGCAAGGTCTGTGAAACGAG	TTTCAAGATTGTTTTCCGGC	242	A
429	Ca2C21072	Ca2C21072*	TGGGCCATTGCTTTCTTAAC	TCCTTGGGTCCACTCTCAAC	182	A
430	Ca2C2110	Ca2C2110*	AGCAATCATCGGTAAGGGTG	GTTCCGGCAATATGAGGAGA	553	A
431	Ca2C21371	Ca2C21371*	TCCCATTTTCATTCGAAACC	ATGAAGCTGCTCTTTGCCAT	333	A
432	Ca2C21936	Ca2C21936*	CTGCTAGAACGTCAATGCCA	ACTGGGTATGGAGCAACTGG	669	A
433	Ca2C22121	Ca2C22121*	AAAGGACCAACCAAGCCTCT	ACGCAACGAAAACTTGAAAA	187	A
434	Ca2C22239	Ca2C22239*	ATGCTTATTGGCGATGTTCC	TGAAGAAACAACATTTTCCCC	256	A
435	Ca2C22608	Ca2C22608*	CCGGTGGCTATAATGCTGTT	CGATTGATGGAGCAAAAGGT	234	A
436	Ca2C22716	Ca2C22716*	ACGCAGAGTGTGTGGAGTTG	ATGATGGAGGTTGTAAGGCG	219	A
437	Ca2C22906	Ca2C22906*	GGTCACGTGTTGATGATTGG	GCTGTGACCCCACTACCATC	175	A
438	Ca2C23022	Ca2C23022*	CACCACATTGAGGACATTGC	AACGACCTCCTCCTCCTCAT	188	A
439	Ca2C23048	Ca2C23048*	TTTCTAGTGTGGGACCCAGG	AATATCAACCTCACCACCGC	202	A
440	Ca2C23299	Ca2C23299*	CGACTTATCCCTGAGCTTCG	CCCTTTTGCCTTGGTCTACA	320	A
441	Ca2C23306	Ca2C23306*	GCATACGCTTTCCTTGAAGC	GAGGTTGGAGTTGTACCGGA	188	A
442	Ca2C23349	Ca2C23349*	AGAGCACTTGCTGCTGTGAA	AACATTTCTCTGCACCTGGG	364	A
443	Ca2C23417	Ca2C23417*	GCTGCAACAAGGAGGATAGC	CAGCAAATGGAAGGCAACTT	158	A
444	Ca2C23559	Ca2C23559*	TTGCTGGCCTTATCATTTCC	TCATCATGGGTGGCACTAGA	634	A
445	Ca2C23568	Ca2C23568*	GGCCAAAGAATGTTACCAGC	ACAAGCTAGATAGGCCGTGG	288	A
446	Ca2C23666	Ca2C23666*	ACACAAGCACAGGCAGATTG	TTTGCAATGTTATGGTGACTACG	165	A

447	Ca2C23858	Ca2C23858*	GTCAAGGGAAAGCTGGTCTG	AGCAAGGAAGTTTCATCCCA	336	A
448	Ca2C24052	Ca2C24052*	GGAACAGATGTGCCAAGTGA	CCCAATCTGGTAGTTCATGCT	151	A
449	Ca2C24440	Ca2C24440*	CCAATGGTCCGAAAACACTT	TGTTCCCGTTTGCTTCTTCT	164	A
450	Ca2C25292	Ca2C25292*	AGACCCTGCTCCTTGAGTGA	TGCTGCAGTCATGTCAACCT	356	A
451	Ca2C25471	Ca2C25471*	CTCTCAAGTTCCTTGGCAGG	GGGGATAAGAAGGATGGGAA	279	A
452	Ca2C258	Ca2C258*	ATTGGGTGAGCCACAAGAAC	TCATGTGTACCCTCGTTCCA	596	A
453	Ca2C26825	Ca2C26825*	TGGAGATGGGGAGAGATTTG	ATCACAACCAGCACTCCACA	230	A
454	Ca2C26976	Ca2C26976*	AAGTCCCACTGGCAATGAAG	GGCAGGAAAATGAGAAGCTG	320	A
455	Ca2C27366	Ca2C27366*	CACACGATTCCGAAACACAC	CCCAAACAAAAGTTGGCACT	478	A
456	Ca2C2748	Ca2C2748*	TCATACATGGCTCGTGGAAA	CCCTCTTGCATCCAAGTCAT	521	A
457	Ca2C2794	Ca2C2794*	TTTTGGCCTCAACATTGTCA	TGACCATAGATGCATCGGAA	216	A
458	Ca2C28516	Ca2C28516*	AATCATCGAAAGGCGAATTG	ATTCTTTCAACGCCACCATC	231	A
459	Ca2C29392	Ca2C29392*	TGGCAAGATGGACAAACAAA	AGGATACACCTGACCGTTGG	316	A
460	Ca2C29749	Ca2C29749*	AAAGCAAGCACAGAGCCAAT	ATAACACGCGATCCTGAACC	899	A
461	Ca2C29875	Ca2C29875*	GGCATTTCTGAGTCTCTGCC	AAACTTCACAACGCCAACCT	179	A
462	Ca2C30045	Ca2C30045*	CCCCAAACGTGCTTATCATT	ACGGGAATATCGTCCATCAA	261	A
463	Ca2C30118	Ca2C30118*	TGCTCATTTCAAACCCACAA	TGGAGGACCATCATCAAACA	448	A
464	Ca2C30361	Ca2C30361*	CGATCGATCCTTCAGATGGT	GAGGTTTCTGAACACGTGGA	166	A
465	Ca2C30401	Ca2C30401*	CAGCAGCAACAAAACTCCAA	AACGTCTCCAACTTGTGCCT	222	A
466	Ca2C30537	Ca2C30537*	CAAGTTCTCCTTGAGGCTGG	ACATCTATCGGGGTTTTCCC	169	A
467	Ca2C30864	Ca2C30864*	CTTGCAAGTGTGGCATAAAA	CGGAAGCAATGTGACAAATG	173	A
468	Ca2C31185	Ca2C31185*	AGAGTCGCCGAAGAAACAAA	ACCTGTATGCTGTATCGCCC	152	A
469	Ca2C31186	Ca2C31186*	GGTGACTGTGTTTTGATGCG	AAAGCTCCTTTGCTCCATGA	174	A
470	Ca2C31438	Ca2C31438*	AAACCAGTCTCCCACCACAG	AGGATAATTCCACGAACCCC	270	A
471	Ca2C31996	Ca2C31996*	ATCTTCTTCGTTGCGCTGTT	TTGGTCTTTCATCCCCTCAC	511	A
472	Ca2C32031	Ca2C32031*	ACACCGTTGGAAGACGAAAC	CCGGACTTTCACGATTCACT	293	A
473	Ca2C32254	Ca2C32254*	GAAGTTGAGACCGAAACCCA	TGATCAGGCAGAATCAGCAC	241	A
474	Ca2C3233	Ca2C3233*	GGGATCCAATTGCTTTCTCA	TCAACACTTGTTTGTCGCGT	300	A
475	Ca2C32463	Ca2C32463*	GGGGAGGGCTCTATTTCTTG	CATTGTCCGTTCTTCCTGGT	418	A
476	Ca2C32732	Ca2C32732*	TTTGGACACATGCCAAAGAT	GGTAATGGGTGTTGATGCAAG	150	A
477	Ca2C33121	Ca2C33121*	GAATGTCACGTGGTGCTGAC	AAATCACGAATTTGGCTTGG	160	A
478	Ca2C33863	Ca2C33863*	GGACCAAATGATTGGCTTGT	AGAAAATCTGGACATGGGGA	177	A
479	Ca2C34020	Ca2C34020*	ACCAAGGCACCAGATACCAG	GCATGAGGCAACAGAGTTGA	177	A
480	Ca2C34047	Ca2C34047*	GCAGCGTCTCGTTATTAGCC	TCACGGCGTAACTCTGACTG	971	A
481	Ca2C34206	Ca2C34206*	TGGCTGAGATGATTCCTGGT	TCAGTCCAAGGAACACCTTTC	229	A

482	Ca2C34413	Ca2C34413*	CAATCCTCGCCTAAAAACCA	CAAGCTAGTGCCATGGTTCA	245	A
483	Ca2C3452	Ca2C3452*	AAGGCACCAAAACTGATTGG	GGCTTGTTGCTGTTGTCAGA	229	A
484	Ca2C34583	Ca2C34583*	ATGCAGCAGAAGCAAGTGAA	CCGACGATACTCGACGTCTT	281	A
485	Ca2C34655	Ca2C34655*	GGCTGTTTTTCACCGTTCAT	AAAGCATTGCAGCATTTGTG	495	A
486	Ca2C34824	Ca2C34824*	ATACTTTGCCAAATGGCTGG	TGTCTTTTGGGATACCCGAC	200	A
487	Ca2C35101	Ca2C35101*	GCTGCTGGGTCAAAAATCAT	TTGCTAGTGATTGTGCGTCC	211	A
488	Ca2C36235	Ca2C36235*	TCCATTCCAGGAGAGACCAC	CTGACCTGTGCCCGTTATTT	353	A
489	Ca2C36237	Ca2C36237*	GCCACAATTGGTTGTTGTTG	AATCGTGAATCGGCAAGAAG	329	A
490	Ca2C36298	Ca2C36298*	TCCCATGGAGTTGATGTTCA	ACTCCATCTTTTTCGTCCCA	161	A
491	Ca2C36391	Ca2C36391*	TGACACGTACCAAGAGCAGG	TCGAGGAAATGCCCAGTATC	320	A
492	Ca2C36571	Ca2C36571*	GGCGTTGAACCTGTGAAAAT	GGACGCCCATCTTCATAAAA	191	A
493	Ca2C36636	Ca2C36636*	CATTGAGAGGGTGCTTGGAT	GAATGAGAGGTCGACCGAAA	243	A
494	Ca2C37037	Ca2C37037*	GTCAGGCTGCTCCAAAGTTC	CTTGAGTAATCAGTGCGCCA	258	A
495	Ca2C37368	Ca2C37368*	AAGACTCCGCAATGGAGAGA	CCTAGAAGGGGCCCTAACAC	215	A
496	Ca2C37682	Ca2C37682*	ACCCTTTGATCGATGTTTCG	AGCAGAAAGATTCGAAGGCA	185	A
497	Ca2C3775	Ca2C3775*	TCAAACTGAAACCTAGCGCC	CCGCAATGTTAGGTGGAACT	186	A
498	Ca2C37998	Ca2C37998*	GCGAGGAAACAAGTGGAGAG	TTCCCCCATTGGTAAAAACA	276	A
499	Ca2C38118	Ca2C38118*	ACATCCATGGCAACATCAGA	GGGAAATGGGTATCGGAGAT	242	A
500	Ca2C38128	Ca2C38128*	TAAAAACGCGTACCGAATCC	GTCGGCTTTTCAAGCAACTC	588	A
501	Ca2C3892	Ca2C3892*	TCATGCCAAATGAAAGTGGA	TGGATCAGTGATGGGAACAA	239	A
502	Ca2C39133	Ca2C39133*	CTTGACCACCTATTTCCCCA	GGTGGCGGGAGACAGTATAA	170	A
503	Ca2C39203	Ca2C39203*	CATCATCGAGCTGGTAGCAA	TTGTCTGAACCAGGGTTTCC	176	A
504	Ca2C39439	Ca2C39439*	AAGAGAGAGGCGGTCACAAA	AGGGGGTATGTTGTCAGCAG	235	A
505	Ca2C39641	Ca2C39641*	ATCGTGCCAGGAAGCTTAAA	CCTTTCGCATTGTCATCTCA	156	A
506	Ca2C39968	Ca2C39968*	TGTTGAGAACTCTGATGGCG	GAGCAGCGTTCATCTTTTCC	244	A
507	Ca2C40131	Ca2C40131*	ACAGGTGGTGAAAGAGGTGG	GAACAACTGCACGAAGAGCA	173	A
508	Ca2C40244	Ca2C40244*	AACAGCGCAACGAAGAAGAT	GCTTTTCAATTTGCTCCCTG	175	A
509	Ca2C40350	Ca2C40350*	CCCATTTTGAGGAGCTTTTG	GCAAAGAAGGCAAGAGATGG	211	A
510	Ca2C40359	Ca2C40359*	CAAGGGTCAGAAGAGTTCCG	TCAGCAGTCAAACCCTCCTT	249	A
511	Ca2C40459	Ca2C40459*	TAATTCGACAACGCTTGCTG	AACCTGAGGTCCAGATCACG	357	A
512	Ca2C40580	Ca2C40580*	CTCACCCTTCCACCTTACCA	AGCACCAAAAGCCAGGAATA	633	A
513	Ca2C40759	Ca2C40759*	TCACCTTGCCCGATGATATT	TTGAAATGAGGACGTTGCAG	156	A
514	Ca2C40858	Ca2C40858*	TTGCAGCACTACATCCTTCG	GAACGGCGTTTCAGGATTTA	187	A
515	Ca2C41121	Ca2C41121*	CACGCTTCCAACAACAGAGA	CCTCCTCTTCCCAATTTTCC	348	A
516	Ca2C41406	Ca2C41406*	TGCGATAAATGGGGTTGATT	GTTATAACCCTCGGCTGCAA	248	A

517	Ca2C41582	Ca2C41582*	CATCTTCGCACTCTTCTCCC	GATCAGCAGCAACCACAAGA	341	A
518	Ca2C41827	Ca2C41827*	TGTCTCAGCACGTGGAAATC	TGATCAGACCAACCCCATTT	298	A
519	Ca2C42088	Ca2C42088*	CATGGATTGTCGACGTGTTC	CAGTTCCACACCAAAACGTG	306	A
520	Ca2C42261	Ca2C42261*	TACGCTGAAGGTACGCCTCT	CTGCGTCTGTCAGATTGCTC	203	A
521	Ca2C42337	Ca2C42337*	CCCTCGTTTGCATCAGAAAT	GTGCAACTAAAGCGACACCA	214	A
522	Ca2C42894	Ca2C42894*	TAAGCCAAGGATGATTTGCC	AATTTTAGGCAGCCAGAGCA	161	A
523	Ca2C4302	Ca2C4302*	TGTCTCCACTGGATCTGCTG	GAAGAGCTGGAACATTTGGC	212	A
524	Ca2C43176	Ca2C43176*	GAATCCTGTTATGGCGGAGA	ACATCCACCTGAGGAACTGG	415	A
525	Ca2C43246	Ca2C43246*	CTTTCTGCTGCTGCTCCTCT	GCGAAGGCTGTTCAGAAAAC	302	A
526	Ca2C44034	Ca2C44034*	TCTTTTCGCTTCAACTCCGT	TGTTGCAAGCCTTCTCATTG	590	A
527	Ca2C44050	Ca2C44050*	AATGAACGAAAAATCGCTGG	CGCCTCCATCAAAATCTCAT	284	A
528	Ca2C44267	Ca2C44267*	ATGGCACGTAAGTCCCACTC	GGGATTGCAGGGTGTAATTG	319	A
529	Ca2C44719	Ca2C44719*	TCGAACCCAAAACTCCAAAC	CGATTTTGGGCTGAACAGAT	332	A
530	Ca2C4778	Ca2C4778*	ACTCAACCGCATGAAAAAGC	TTGGATGACCAGGATCCATT	165	A
531	Ca2C5197	Ca2C5197*	GTGGTTGTGATGATCGTTGC	TGTTTTCCAATGCACTGTCC	189	A
532	Ca2C5918	Ca2C5918*	GAGGTCCAACAGCTGCTTTC	TTCATCGTATCCATTTCGCA	492	A
533	Ca2C6096	Ca2C6096*	AGCCCTAGGAGATGGGTTGT	TTGCAAAAGCATTTGAGTCG	315	A
534	Ca2C6216	Ca2C6216*	CGCTTCCGGTGATCTAATTC	GGCCTTCCTTAGGCTTCAAC	164	A
535	Ca2C6324	Ca2C6324*	AACGGATTCGATATGGACGA	TCTCTTGCGACCCTTTCTGT	364	A
536	Ca2C7029	Ca2C7029*	GTGCATTGCGAGGGAGTTAT	TTTTGCATCTACGTTGCCAG	155	A
537	Ca2C7044	Ca2C7044*	GGAGACAGTGGGGAACAAGA	TGGCACAACAGACCTAAGCA	274	A
538	Ca2C7054	Ca2C7054*	TAACAACAGCGGCAATGGTA	TGGTCAGCATTGGTTGAAAG	166	A
539	Ca2C7102	Ca2C7102*	TCCAAAAACCCTCACACCTC	CCTCTGAAGGAGTTTTCCCC	400	A
540	Ca2C7141	Ca2C7141*	GGACCATGGACTCCTGAAGA	TGAAGCCTTCTTTCCCATTG	413	A
541	Ca2C7194	Ca2C7194*	CGAAAGGTTCAGGATGGAAA	ATTCCTTGCAATGCTCCACT	184	A
542	Ca2C7257	Ca2C7257*	ACGGCTTTTCACAAAGTGCT	TTCAGCTCACGGTCATCTTG	337	A
543	Ca2C7446	Ca2C7446*	AGAAGACGGCTGAGAAACCA	CACGCTTATGCGTCAACAAT	173	A
544	Ca2C7756	Ca2C7756*	TCCGGGACTGACTTTGTACC	TCTTCAAGCTTTCGCTGGTT	182	A
545	Ca2C7909	Ca2C7909*	ATGGCATTGTCAGTCCTTCC	CCTGGTCGAATAGGGAAACA	176	A
546	Ca2C8398	Ca2C8398*	AGCAGGAGTGCTCAACCAAT	AGAGAGCCTGGCATTTTCAA	288	A
547	Ca2C8766	Ca2C8766*	TGGAGCTCAAATGGCATACA	TGTCGTGTCTCGTATCTCGG	152	A
548	Ca2C881	Ca2C881*	CAATACAGCAGCTATGCGGA	ATCATTGCAGGCCTCAAAAC	595	A
549	Ca2C8964	Ca2C8964*	CAGCAGATCCAGTGGAGACA	TTGCAGACTGGGGATATGTG	175	A
550	Ca2C9004	Ca2C9004*	CAAACCCTTCCTTTCTGTGC	TGCCTGTTGCTTGAGTTTTC	235	A
551	Ca2C9440	Ca2C9440*	GCCCGAGAATGAAATTCAGA	GGCTGGCTATGAAGAATGGA	271	A

552	Ca2C9649	Ca2C9649*	AGGAGACAGCCCAATTCGTA	TCTGCTTCAAGCGGGTTAAT	155	A
553	Ca2C9701	Ca2C9701*	GACCTCCTGGTAATGCATCC	GGGGAAAAACTCTGCATTGA	166	A
554	Ca2C9943	Ca2C9943*	CGCTTGTTCAAGGGAGAAAG	TTGGACTATGGTTGCTTCCC	163	A
555	Ca15712CD24E12	AB025007.1	CGCAGGAAAAGAAAACGAAG	TCTATGGCCCAACTTTCCAG	149	A
556	Ca147613127884	AJ005948.1	GTAAGAATTGCGGGGTGAGA	GCCACAATCACACACTGGAC	113	A
557	Ca93773928149	AJ131049.1	TTGTCATGGTGGTTGCATCT	GACATCGCCTTCTCAAGCTC	211	A
558	Ca114752961299	AJ225027.1	AAAAGGACGCTGCTCAAGAA	CATTTTAGGACCTTTGGGCA	287	A
559	Ca17912CD35D01	AJ275309.1	GCAGTGGTATCAACGCAGAG	TGAAGATTTCATATTCCTATCCG	117	A
560	Ca147547635493	AJ400861.1	CAGCATCAATTCCTTTGGGT	AACCGCAATGCGTTAAAAAC	190	A
561	Ca1064020975617	AJ487038.1	CACCCCAATTTGAAGTTGCT	ACAGAGCTCCAACGAAGGAA	145	A
562	Ca1448960219076	AJ884609.1	GAGGTTCCCGATCCTTCTTC	AGCCCTCATCAAACCATCAG	282	A
563	Ca1683047832555	AY312269.2	TCGAGTTCACAGTCTCCGTG	TCACCAAATGCACCATCAGT	186	A
564	Ca1141050872749	CD051357.1	TGGACTCGTTCCCAGAAATC	CACCGAAGGGACAGTATGGT	228	A
565	Ca1101947832352	CK148664.1	TCTGAGAAGCGTGTCCTCCT	GGAGATTCCAAAATGGAGCA	236	A
566	Ca16689169746141	CK148867.1	ATTTATCCACCGACGAAAGC	TTCACTGTCTCAAGTTGGCCT	102	A
567	Ca1695759709748	CK148894.1	GCTCCTGAGGTATTGAAGCG	GCTCAGTTTCTGCCCAAAAG	112	A
568	Ca17637169745184	CK149017.1	CCCGCCTGTTTCCTATTGTA	CAGAGTGGAGAAGGGTTCCA	237	A
569	Ca15332105635868	DY475098.1	GGTCGGGTGACCAAGTTTTA	AGGTGTTCAAGCATAGAGGCA	212	A
570	Ca11057105636058	DY475152.1	TGGTGCAAACTAAGGATGGA	GGTCTGTGTTCCTCAAAACCA	216	A
571	Ca16509169748554	DY475291.1	AAACCCCACATGTGAAGGAA	CCATGCACGTCGTAAAGCTA	136	A
572	Ca14751105637020	DY475497.1	ATCCCGGTCTTTTTAATGGG	TTTGCAAAACACAAAGGGAG	126	A
573	Ca10119105637170	DY475526.1	ATTTTTCCCTTGCGTGATTG	AACCAAAGATGTGGATTTTGC	113	A
574	Ca11147150173597	EH058895.1	TAAAAACCCAAGGCAAATCG	TGATCATCGTGCTGGAAAAG	182	A
575	Ca12257150173782	EH058956.1	TCACCTCCGCAACATGATTA	TTGCATGGAATGAGCAAGAG	243	A
576	Ca12164150173955	EH059041.1	AATATTGCTCGCTAGCCCCT	GCCAGATCAGTGTGAGGACA	117	A
577	Ca15356150174078	EH059084.1	TTGTTGGTCCCAAGAAAAGG	CTTGTGTTGTCTGGAAGCGA	141	A
578	Ca14402150174189	EH059181.1	CTTGTGTTCCTTTGGTGGCT	GGAGCTTGATGCTGAGGTTC	232	A
579	Ca10904150174283	EH059254.1	ACTAATTGCAACCGGCATTC	GGACTGAATCCTTGGGGTTT	213	A
580	Ca15861CD06F10	EH059300.1	TGGTATCAACAAGGAGTGATGG	TTGAATTCGAAACATGATGTCC	105	A
581	Ca10402169743789	FE668632.1	TGGCATATAATCGGGTCCAT	AATGGCTGCTCCAAAGGTTA	257	A
582	Ca157594586605	FE669222.1	TTGGCTCTTTCCATTCCAAC	CCAAAGCCTGGGAAAATGTA	173	A
583	Ca11371169745895	FE670067.1	CAGGATTCGACTTCCCACAT	AAAGGCTTCCGTTTACGCTT	257	A
584	Ca14849169744854	FE670216.1	ACCATCTCTTGCATAACCGC	ATTAATGCTTCGCCGATGAG	119	A
585	Ca17630CD21C09	FE670270.1	GCGATGCTCTTCCTGACATT	TTGGTCACCTTGTTGGTGAA	286	A
586	Ca9298169745219	FE670293.1	CATGCGTTCGCTTTAACAGA	TAGCAGCTTGGGGTAGCTTG	299	A

587	Ca10556169746056	FE670468.1	AACGTCAACTGGAAAAACGG	GGTATCGGAGGGTTTGGAAT	198	A
588	Ca15006169745568	FE670784.1	ACCCTTTTGTTCCACACGAG	GTCGGGCGAAAGTGAAATAC	232	A
589	Ca12505169746120	FE670880.1	AAGAATCAGGCCAAACACTGA	GAGCTTGCAACTGGGAAAAG	136	A
590	Ca16577105636435	FE670889.1	GCGGGTATGAATCTGATTTG	AAAGTATATCCCCCGATGCC	228	A
591	Ca14576169747085	FE671079.1	TCGAATTCATTCCTTCGTCC	AAGGCGTGGATTGTTTATCG	231	A
592	Ca15194169747796	FE671178.1	GCATCTCAATGGGAAATGCT	ACTGGAAACAATTTGCCTGG	229	A
593	Ca1728247832582	FE671262.1	TTCAAAAGCAAGCTCAGCAA	CTGCAACCTTGTCCAAGTGA	184	A
594	Ca11688169748274	FE671272.1	CGGCAGAAAAATTCGAAAAG	AACATCGTAAAGCCAGGGTG	155	A
595	Ca15873150174352	FE671299.1	CGGTGGAGAGTGAGGAAGAG	ACCCCCTTCAAGTACAACCC	167	A
596	Ca10699169746343	FE671367.1	TCCTTTTCACCGCTCTCATT	GACCGGTCCAACAGAACACT	282	A
597	Ca13382169746951	FE671389.1	GGAGCGAGCAGAATAAGGTG	TGCATAGCATATGGGGATGA	227	A
598	Ca9563169748070	FE671524.1	AAATGAACCTCTTCCTGCCA	CTTGCAGGCTGTGTCGAATA	204	A
599	Ca12784169748099	FE671553.1	TCCAACTCTTTGCCTTTCGT	CCTTTTCACCCTCAACCTCA	181	A
600	Ca12387169748596	FE671642.1	TTTTGGGCAGCAAAGATAGG	GGAAGTTTTGGTCGCAAGAA	254	A
601	Ca15265169746577	FE671697.1	TTCATAACCCCCTTCCCTTT	ATGGCATCGAGTCCAAGTTC	184	A
602	Ca10769169746624	FE671744.1	GTGTTTGGATATTTTCGCCG	GACAAGAACCACCACGGTTT	236	A
603	Ca15256169747801	FE671939.1	ACATCTGAAGGCCATTCCAC	GCTGAGTGGAGAATGAAGCC	282	A
604	Ca10058169748238	FE672016.1	ATCTGCAATGGGGAAAACAG	TACCCTTTGTTTTTACGCCG	118	A
605	Ca15994169746275	FE672059.1	ATGGAATGAGCAAGGGTGAC	TCACCTCCGCAACATGATTA	239	A
606	Ca10490169748746	FE672104.1	TCTGCTTCTGTTGCTGCTGT	ATTTGATTTGGCAAAGGTGG	254	A
607	Ca14937169746819	FE672191.1	AGTCAACATTTCCAGGCACC	AGGCCAAAAAGCAAGAACAA	246	A
608	Ca12182169746897	FE672233.1	CTTTCCCTGCTTGAACTTGC	TTCAAATGCAAATGGATGGA	297	A
609	Ca13529169748367	FE672397.1	CCTGTGATCCAAGATGAGCA	TGCAGCTTGTGTGTGAACAA	135	A
610	Ca12586169746506	FE672514.1	TCGATTCGAATTCCGAGTATTT	TTCAAATGCAAATGGATGGA	121	A
611	Ca11817169747544	FE672726.1	TTTTGTGGACTTTTCACCCC	AGGAGCAACCTGTGGAGAGA	193	A
612	Ca18041169747341	FE672757.1	GCACCACACCTTCACAACAC	GTTACGCCAATTGTCTCCGT	223	A
613	Ca16267169746520	FE672986.1	TCATAGCATGTCATGTGTAACTTCA	CACCACCGCCTTACAAGAAC	153	A
614	Ca14710169746503	FE672991.1	AAAACACATTTCACCGGAGC	AGCAAGAGCTTTAGCAACGG	121	A
615	Ca16237169748701	FE672996.1	CAACCATCTTCAGCCTTGGT	GCAGTGGAGATTCCATGGTC	254	A
616	Ca9905169746859	FE673071.1	CGTTTTGCAACCTTCTCCTC	GTTGACCAACACAATGCTGC	161	A
617	Ca12411169747340	FE673134.1	ACTGCCAATCCTACCCTCCT	ACAGCTGGAGCATATTGCCT	281	A
618	Ca180066470362	FE673135.1	GGCTTCACTCTTCATTGATGG	TGGTTCCATTGTACGCTTTG	256	A
619	Ca12863169747399	FE673193.1	CATGCAAAACAATTCCATCG	CTTCTGAAACATAAGCCGCC	283	A
620	Ca11491169747952	FE673230.1	GCACTAAGCCATCCTTGGAA	TCTTGCCCCTTTTACCCTTT	281	A
621	Ca9972169748026	FE673304.1	GAACATGGAGGAATATGGCG	TCATGAATCGTTCGGTTGAA	129	A

622	Ca10880169748548	FE673346.1	TTTTTGCCCAAGACAAGACA	TTCTCTTTTTGGTGGCTGCT	190	A
623	Ca16411CD18G08	FE673352.1	ACGTTTTCTGCCTCAGCAGT	CACCAAACAAACAAGCTGGA	144	A
624	Ca9299169748624	FE673374.1	TTTGCCACCTCAGTTTTTCC	CAACAAATCTTTGTCCTCAAGGT	288	A
625	Ca15352169747422	FE673504.1	TTTGGGGGTACACATTCGAT	AAACTATGCCGCCTGAGATG	157	A
626	Ca9351169747446	FE673528.1	ATCGCCAAAATGACCAAAAC	CCGAAAGCAATCTTCACCAT	268	A
627	Ca12874169747449	FE673531.1	GACTTCCGAAGCAGCAAAAC	AACAGGTACCAAGCGCATCT	271	A
628	Ca114201212811	X95875.1	AACCAAGCTGTCAAGGCTGT	TCATGATGAGGAATCGTGGA	196	A
629	Ca10101CD17D05	HO214215	GACAGGAACGACAAATCCGT	CTCGTCAATGCTCAGTCCAA	144	A
630	Ca10273CD20F12	HO214247	TGAGCAGAAATTGCAACAGG	GCGCTTTCCAAGAACTATGC	280	A
631	Ca10314CD21F06	HO214216	AATTTTCTTCACGTCGCCAT	ATGGGCCGTCAGTTAGTGTC	168	A
632	Ca11346CD35C09	HO214248	TCGGTCAAACTATCATCACCA	GCCAGCAAGAGTTGGAAGAG	140	A
633	Ca11651CD15D01	HO214217	CATTGATATTGGAGCGTGAAAA	TTTCAGGATGCAGAGTGCAG	156	A
634	Ca12038CD10B02	HO214249	GCAGTGGTATCAACGCAGAG	TCTTGGAAGCTCCAAAAACAA	130	A
635	Ca12284CD20E07	HO214218	TCCGGAGTCAAAATTTCAGC	TTCGGGAATGAGGAGTTAGG	136	A
636	Ca13487CD21H10	HO214250	CCTCCCTGCGAACAAAATAA	AACGCAGAGTGAAAACAGCA	189	A
637	Ca14038CD25A07	HO214219	TACTCTCTTGCGTCGGGACT	CATCCGAGGGTTCTCTTGAA	208	A
638	Ca14088CD18C06	HO214251	TGAGCAGGCACAATACCAGT	TTTCTGTTTTGGGTCGTTCC	124	A
639	Ca14366CD06G08	HO214220	GCAAATCTATGCCTGCTTCA	AAGGCGTGTTGAGGTAATGG	148	A
640	Ca15419CD38G10	HO214252	GCAGTGGTATCAACGCAGAG	CTGCAAGTGACAACGGAAAA	102	A
641	Ca15452CD38E03	HO214221	GTGTGCTGTCTCAGCATCGT	TGCTAACTCAATTGGCCTTG	297	A
642	Ca15775169744160	HO214253	ACCTCAACCACCCTCCTTCT	CTCTTCCCAGTCGCTGAATC	192	A
643	Ca16272169746498	HO214222	ATTCGGGGAAAAAGAGTTGG	AATCGCCGATGTGATAGGTC	158	A
644	Ca17387169748252	HO214254	TTCACAAGGCTGAGGAGGAT	CGGCAGTTGGGTTATCTGTT	113	A
645	Ca1773847832705	HO214223	ATGGGTGCAGGTTTCATAGC	TGAAGCAGAGCACTTTGTGG	150	A
646	Ca17898CD15C05	HO214255	ATCAGGCATGCGTTAAAAGG	TCCATGGGATCAAAATTTCC	271	A
647	Ca18263169747863	HO214224	TGAGAATGCCAGAGACGTTG	TAGCTTTTCCTTCCCTGCAA	121	A
648	Ca18271CD15D02	HO214256	CCTTTTCTCTGGTGGGTCAA	CAGGCATTGATATTGGAGGC	115	A
649	Ca18414CD19C11	HO214225	GCAGTGGTATCAACGCAGAG	TTTTGCTCCATTTCCTGTGA	147	A
650	Ca18446CD24F10	HO214257	GGACAATTTGGTGTCGGTTT	TCTTAGTTCCCCTGCCAAGA	166	A
651	Ca18532	HO214226	AGTTGGATCTGGTGGGTTTG	GCAACACGAAGCTGAAAACA	247	A
652	Ca18959	HO214258	AATTGCAGCAGCACTTGTGT	GATGCTGCCTGGAAATTTGT	114	A
653	Ca19001	HO214227	GCCAGTTGCATTAAACGGAT	TTTCCAAGGTCCTTTTGTCG	183	A
654	Ca19157	HO214259	GGGTTCCTTCAGGTTGTTGA	GGTGAACAATCGGCATCTTT	293	A
655	Ca19378	HO214228	CCTTTCACCCAAAACTCCAA	CAAGGCTTTCGGCATGTAAT	272	A
656	Ca19487	HO214260	CAAATGGAGCGTTTTTGGTT	TTTTGTCCAGGACCACATGA	157	A

657	Ca20010	HO214229	AACGCCAACTCCTCTCTTCA	GGGTTTGGAGTTTTTGGGTT	255	A
658	Ca20156	HO214260	ACCCAATCGTGATTGTCCAT	CCCACGGTGACTACTGTCCT	213	A
659	Ca20356	HO214230	GATTCACGTGACGCAGCTTA	CGCTTACGTCAAACCCCTAA	135	A
660	Ca20534	HO214262	GAGGATCCAGTGTCGTGGTT	CAGTCCAAGAGCGCATGTAA	135	A
661	Ca15409CD29B10	HO214231	ATTCGGGAGTGATTTGCTTG	AGAAAATTAGCAGCCCCACA	118	A
662	Ca20788	HO214263	AATAAGGGGTCGTGGGTTGT	TTGGGCCTACCAATGTCTTC	104	A
663	Ca21040	HO214232	GCAGTGGTATCAACGCAGAG	CCAACAGATGATAGCCTCGC	161	A
664	Ca21047	HO214264	GTGGGAACAGCCATCAAAGT	TGCTTTTGGCACATCTTCTG	285	A
665	Ca21249	HO214233	CAACCGTTGAAACCTTCGAT	GTCTCCTCCGTTTTCCCTTC	298	A
666	Ca21417	HO214265	TGCAAAACAATTTCCAACCA	AGAAGCTGGCGATTCACATT	278	A
667	Ca21567	HO214234	GGATTCCACCGACAAGAGAA	CTGATTCACGCTCAGGAACA	269	A
668	Ca21878	HO214266	CGTCATCAAAAACCGTCCTT	CATCGGTAGGAAGCTTGAGG	263	A
669	Ca21977	HO214235	GGCAACATCTGACCACCTTT	TGCATATCAGGTGGAAGCAA	254	A
670	Ca22344	HO214267	CCCTTCGGTCATCAACGTAT	TTTTCAAGAGGATTTTGCCG	154	A
671	Ca22434	HO214236	GCCAGAGCATGCTTCACATA	CTTGGTGGCTTGCTTTCTTC	203	A
672	Ca22684	HO214268	GGGTTTAGACCGTCGTGAGA	GCGTTCCAGCCATAATCCTA	188	A
673	Ca9445CD16A11	HO214237	GTACCAGACAGCAGGCGTTT	GATCGTGGTTCAAGCCAACT	171	A
674	Ca9914CD18G11	HO214269	GCAGTGGTATCAACGCAGAG	TTTTGCAATAATAAGCCAACGA	277	A
675	CaESTCg11	HO214238	TCCATATTGTAATCGCCAGAAA	CGGGAATTCGATTAAGCAGT		A
676	CaESTCg13	HO214270	GGCGGCACATCTGTTAAA	GCGTTCAGCCATAATCCTACAGA		A
677	CaESTCg2	HO214239	TCTCGCGACCCCTACTCTAA	AGCAGTGGTATCAACGCAGA		A
678	CaESTCg3	HO214271	GTGGACTGGCTCACTGACAA	AGGCAACTCCTCTACGGTCA		A
679	CaESTCg5	HO214240	TGGCGCTTCAAAAGGTAGTT	TGTGGAATTGTGAGCGGATA		A
680	CaESTSn12	HO214272	GCACCATTTGAGGCATGTTT	ATCCTCCAATTGATGTCCAC		A
681	CaESTSn16	HO214241	GGCCTAAACTTGGAACATGG	CACCAAACAAACAAGCTGGA		A
682	CaESTSn18	HO214273	AATCTTGCTTTTGGTTCAGCA	GGAAACAGCTATGACCATGATT		A
683	CaESTSn2	HO214242	TGGTGGAGATATTGATGAGCA	CAAAGGATCAAAGCCAATGC		A
684	CaESTSn22	HO214274	CGACTCACTATAGGGCGAATTG	TCAAGCCAGACGGGTCTTAT		A
685	CaESTSn28	HO214243	GAAAACACATAGGGGTATTAGATTATG	AGCAGTGGTATCAACGCAGA		A
686	CaESTSn3	HO214275	AGCAGTGGTATCAACGCAGA	GGAGCATTCAACCCGACATA		A
687	CaESTSn30	HO214244	TGTGAATACATGACGTGAACCT	AGTGCTGGGATTGAGATGGT		A
688	CaESTSn32	HO214276	CCGCAATTAAGAACTATTTGACA	TTTATGCTTCCGGCTCGTAT		A
689	CaESTSn33	HO214245	TCCTACCAATAACCGGGTTTC	AATTGGCAATCGGTAAAACG		A
690	CaESTSn5	HO224282	GCAGTGGTATCAACGCAGAG	CTGTTCTGGTGTGAGGACGA		A
691	CaESTSn9	HO214246	CGACTCACTATAGGGCGAATTG	CGGGCAACGTCTCAATACTT		A

692	AGLC1	HO214284	AACATCATCAAGGTCTCCTGGGTA	GGTGATGAAGTTACTGATGGTGGA		B
693	AGLC10	CK149125	TCCCATCAATGTTCTCACCA	GTTCAAATGATGATGGCAAAGA	600	B
694	AGLC100	CK148706	TCCAGGTGGAGGAGTCAGAT	TCAAACGTCTTCCACCTTCA	351	B
695	AGLC101	CK148993	TCCAGCAAGTGATAACCCTAAAAC	GACGCATCAAATCCACCTTT	317	B
696	AGLC102	CK149141	TTCCTTCCACCATTGGAAAC	TAATTCACTTTCGGGTCACG	662	B
697	AGLC103	CK149114	CAAAGGATTGCCATTGGACT	CCCATTTGCATTGATGGCTTA	425	B
698	AGLC104	CK149103	GCAAGACATGGTGGTTGTTG	CACTTCCCTAGTTGCCCTCA	282	B
699	AGLC105	CK148783	TCCAATTGGTGGCTCTTCTT	CCACCAAATCCAAACGGATA	590	B
700	AGLC106	CK148751	GGACGCTACAAGGCAGAATC	ACATTGCGCTCTTCTGGAGT	660	B
701	AGLC107	CK149104	GAAGATGTCTGATCAAGTTTTGC	TGCAAAACAATTTCCAACCA	554	B
702	AGLC108	CK149109	CCCAAGAAGGGAATTGTTTG	AAGTTGTAAAGCTGCATGTAAACC	180	B
703	AGLC109	HO214278	CAGCCCTTTTATGCCTGATCC	AAGCCTTGGTTTCTTGGACAG	370	B
704	AGLC11	HO214312	ACTCACCACCATGTCCCTTC	AAAATTCAGCCATATGTCAAACC	160	B
705	AGLC110	HO214285	AACATCATCAAGGTCTCCTGGGTA	GGTGATGAAGTTACTGATGGTGGA		B
706	AGLC111	HO214313	ACTAGTCCTGCAGGTTTAAACGA	CCTCTTCCCTCAATTTTCCTCACA		B
707	AGLC112	HO214286	CGACTCCCTCATCACCTCCA	CTTTGGGTCTCTGTTGTTGCTGA	222	B
708	AGLC113	HO214314	TGTCCAAAATTGGGATCAGAGA	AGAACGACTTCAGCAGCAGCA		B
709	AGLC114	CK148985	GGTAGGTCGCGTTGTTGCA	GAGATTGTTGGTGAGAGAAGCA		B
710	AGLC115	HO214287	TTATCATGTTTGCAACATACTCCA	GGGTCTCTGCTTCTGTCACCA		B
711	AGLC116	HO214306	CTTCACCTCTACTGCTGCTACTACTC	GAGAAACTCAGACCCATGTTAATG		B
712	AGLC117	HO214315	GCAAAGCATCCTTCACCTCT	CCTCCAGTGTGTGTGAGATTG		B
713	AGLC118	HO214288	CCGCTGTGTGTTGCAAAG	GAGCACTACTAGCATTACACTCAGTAA		B
714	AGLC119	HO214279	CTGTTGCAAAGCATCCTTCA	TGTTGGTGAGAGAAGCAGGA		B
715	AGLC12	CK148942	GCAAAGCATCCTTCACCTCT	TCCCTCCCACTTATATGTATGC		B
716	AGLC120	HO214307	TTGTGTTGTGGAGCTCTCCTT	CCTATCCAACCTCCCAAGTG		B
717	AGLC121	CK148973	ACCCTTTCGGTTGCAGCTGA	TGTTCGGATGATTGAGGCAGGA		B
718	AGLC123	CK148704	TCAAAAACCGTTCCTTCCTC	TGGTGTGTTGCTTGAGTTCC		B
719	AGLC124	CK148970	ATATGGGATTTCCTGGCACA	ATCTGATCGTTTGGGCTGAC	327	B
720	AGLC125	CK149149	CAAGGTCGTTGACACCAATG	TCTGACCCAACAGGGAAATC	155	B
721	AGLC126	CK148905	TTCACAACAATGGCTGAACC	ATGCCTTCTTTGCTGCAGAT		B
722	AGLC127	CK148903	AACGAATTCGCCCTTTAAGC	GGCAAAGCAGGGGTTAACTT		B
723	AGLC128	CK148681	GTCAATGGCAGTAACATGCAG	TTGGGTATGGAGTTCCACAAG		B
724	AGLC129	HO214280	TGCAGTTTTTGACCTTGGTG	CAAAGCAAGCCGATCCTTAG		B
725	AGLC13	HO214316	TTTCATAGCCTGCCACACAG	GCTTTTGCCTATGCATGTTG		B
726	AGLC130	HO214308	ATACGCTTACGCTGCTTTGC	TATGAAGGTGGTGACGGTGA		B

727	AGLC131	CK149036	ATGGCTTCAGCTAGGGAAGAA	ACCCATACCACCCAAATTGA		B
728	AGLC132	CK149029	GGCTCCCTCCTGCAAATCCA	GAAGTAATTCAGGTAAGTGGCGAA		B
729	AGLC133	CK149037	AAGGTGCAGAGAGTGCGAGT	CAGATGAGCGACGTGAGAGA		B
730	AGLC134	HO214289	TTCGATTTCGGGTCATTTTC	CAGGAATCCTCCCATTCTCA		B
731	AGLC135	CK149090	TTCTCGAGCCGAAAGATTTG	CTTCCTCTGCAGGGTCAGTC		B
732	AGLC136	CK148997	TCGATCGCAGTTTGAATCAC	AGGACAGAGCCACGAAGAAG		B
733	AGLC137	HO214317	GATGCAGGGTCGTTTCAAAT	AGCGATCAACACCGAGAGAT		B
734	AGLC138	CK149099	ATCCAGAATCGTGGCTCAAT	AGGCCTCTTCTCATGGGTCT	592	B
735	AGLC139	CK148972	TTGGAGGGTACAGACGGAAG	CATTGTTGTTTTGCGTGGAC		B
736	AGLC14	CK149086	CTCGCTGGACCTCTCATCTT	AGTGCAAGCCACGAGAAAGT		B
737	AGLC140	CK149002	AAGAGCGTAATGCTGCTGGT	TGCATGTGCAGACTCTGATG	532	B
738	AGLC141	CK148981	CAACTCTAAGGTGTTTAGGTGGTA	ATCCAAAACAGCTCATTGCTCA		B
739	AGLC142	HO214290	GTGTGGACCCAGAAAGGAAA	GACCAAGCCCCGAAATAAAT		B
740	AGLC143	HO214281	TGGAATTGGTGTTGGAAATG	CAGCAACATGAAATGCATCA	137	B
741	AGLC144	CK148893	GTTGCCCCATCTGGAAACT	TTGGGGGACCAACAACTCTA		B
742	AGLC145	HO214318	TGGTGAGATCGCAACAAAAG	CAGCACCAATTTCCTCAGGT		B
743	AGLC146	HO214291	TCAACTGCCATGGAATGAAG	CACAGGAAGAAGGCCTCGTA		B
744	AGLC147	HO214319	AGACCGAGGTCCTGTGTTGT	CCTTCACCAGAGGCCATATC		B
745	AGLC148	HO214292	GCCATGTCCTTTATGAGACCA	CACCTCGCCTCAAATACCTG	438	B
746	AGLC149	CK149098	AGAAGGTTTTCGTTAGGGTTGA	ACCTTGGCACAAGAGGTTTC	386	B
747	AGLC15	CK149086	CCCTCTTCCCTCCGTTCTAA	CTTGACCACCGGACAAAACT		B
748	AGLC150	CK148891	CACTTTCCATGACCCACAAA	AGGTGGAAGGGGAGAGAGAG		B
749	AGLC151	HO214309	GCAGCAACTATTTACACTGGTA	CTCTCTGGGAGAAAGCTCGGAA	453	B
750	AGLC152	HO214320	GAGATCGCTTTTGCTGGACT	GGCTCAAAGGCACAAGTGAT		B
751	AGLC153	CK148886	GGTGCCATTCTTTTTGAGGA	AAATCGAGCACCTTGCTTGT		B
752	AGLC154	CK149145	TAGTCCGGCGAAGAGAGAGA	TTGAATGTCCACCGGAAGTT		B
753	AGLC155	CK148715	TTGTATCAAGCAAGCGGAAG	CGCCAAGCTCAGAATTAACC		B
754	AGLC156	CK148828	TGTACAGGTCCCAGCTTTCC	CACCTCCACTTCCTCTTCCA		B
755	AGLC157	CK148714	CGCATGACTCGTTCCTCTCT	CGCATGCAGCAAATCATAAG		B
756	AGLC158	CK148829	GGCAGCTCAAACAATGTCAA	TTCTGCTTGTTTGTTGGGTTT		B
757	AGLC159	CK148831	GAGAGGGATGTGGTTGTTTCA	CCCTTTGTTCCTTGAAGCTG		B
758	AGLC16	CK149087	AACCAATCCGCAAAATCAAA	TGCATGGGTTAGGAATGAAAG		B
759	AGLC160	CK148713	TGAAGAAAGCCAAGCCAGTT	GGACAACCTTGGTCTTTGGA		B
760	AGLC161	CK148653	ACTGATCAAGGTCTCTTCTAGACA	CCCAACAAACTGGACAAAGCAGA	465	B
761	AGLC162	CK148838	AGCTTGCAGTGGGCTTACAT	CGCCAAGCTCAGAATTAACC		B

762	AGLC163	CK148881	AGTCCTGCAGGTTTAAACGAAT	ATAAGGGGCAAACTCCATGA		B
763	AGLC164	CK148846	GGTCCAAGCGTTGTAAAGGA	GGAGAGCATAGGCACATATCG		B
764	AGLC165	CK148970	CCTTTGCTTCGGAAAGTTGA	CGGAAGACTTGAGGAAATCG		B
765	AGLC166	CK148844	GGGAGGATGAAACAAATCCA	TTGCCACAATCAGCTACGAC		B
766	AGLC167	CK148884	TTCGGATTCAGGAAAAGACG	TCCATGAATCGTGTCTGGAG		B
767	AGLC168	CK148848	TGAGAGGCTTAGGGATGAGC	CCTCCCCTTCATTATCACCA		B
768	AGLC169	CK148685	CAAAATGTCCATAAGCAGCAGT	TGAAGGAGAATTTTGATTTTAGGC		B
769	AGLC17	HO214293	AGTGAGTTGGTTCGGAAACG	AACATGCGCTCAAGTTCAGA		B
770	AGLC170	CK148832	GGTGGTCAAAAGGGAAACAA	ATCGCGATTATCGGTAGCAG		B
771	AGLC171	CK148887	GAGTACTTGCCAACTAGCTTAGGA	TTGGATATAACAGATGACGGGGAA		B
772	AGLC172	CK148823	CACCGAAACCAACACTAGCA	GGGTCCCTGAGCAAAGAATA		B
773	AGLC173	CK148650	TCGGCATTAACCTCAGTTCC	TGAGCCATGATTTTCGTTGA		B
774	AGLC174	CK148745	CCAACATAGCTTCCCACGAT	GGAGGGTTGCTTTGTGTGAT		B
775	AGLC175	CK148708	ACCAGGCCAACCAATACAAA	CCGGGGCAGGTACACTACT	81	B
776	AGLC176	HO214321	GCAAGCTTGGATGATGAAACT	CAAGGCTCGATGACAACTCA		B
777	AGLC177	CK148736	TCGATCGCAGTTTGAATCAC	AGGACAGAGCCACGAAGAAG		B
778	AGLC178	CK148827	GCCACGGTTGGTCTGTTCGA	CCGTTAAGGTTGCCGGACGA		B
779	AGLC179	CK148729	TCAAAATCCTGGTGGAGGTC	TGCCACTGCTGGTAAAGAGA		B
780	AGLC18	HO214294	GACCGGTTAGTGACCCTTCA	TTTGCTGCTGCATGTCTCTT	523	B
781	AGLC180	HO214282	TGCACCATGTTGGAGAAAGA	CGTCCTGGATTCATCGAAGT		B
782	AGLC181	CK148723	CACAGCATTATGGCCAACAGCA	TGTCAGGGGTTTTGACAAATCTCA		B
783	AGLC182	CK148719	CAGCAAAGCAAAGGTTCACA	TAAACCCTGAGCCCATTCAG		B
784	AGLC183	CK148717	TTAAGGGCTGCCTGGAAGTA	GAAATCCGGCACTTAATGGA	969	B
785	AGLC184	CK148737	TTGTGCTTGAATCTGCGAAC	GAAATCACCGGAGAAATGGA		B
786	AGLC185	CK148860	CTTCGCCGGCACTTAGTTAG	AGGTGATGGACTCGATTTGG		B
787	AGLC186	CK148690	ATCGCGAGGCATCTACCTTA	TGCGACCATTATGAGTGAGC		B
788	AGLC187	CK148688	TGCTTTGAGGAGGGAAAGAA	TGGGGTCCAAATTTTGAGATA		B
789	AGLC188	CK148687	ACCAACAAGGGTTTGATTGC	TGACAAGTTGGAAGCCAATG		B
790	AGLC189	CK149018	GTGCGCGCGTATTTACTCTGA	TGGACGAATTCGCCCTTTCGA		B
791	AGLC19	CK148924	CATCCCACACTACTTTTTACCTCA	CCTCCTCCGACAAATTCATCACA	279	B
792	AGLC190	CK148679	TCGTTGAATGAAGGTTTGGA	TCACATGGATGAACAAATGGA	149	B
793	AGLC191	CK148678	CGTTTGGGCTGACAGTTTGGA	GCCATGACATCGGATATGATAGCA		B
794	AGLC192	CK148672	TCGTCACTCGAAACCAATCA	CGCCCAACGAAGAATTAGAC		B
795	AGLC193	CK148880	GGAACATATGTATTGCGTGCAT	TGAATGTGTGTCTGAAAATTGATG		B
796	AGLC194	CK148740	AACAGGACGAACAATCCCTCA	AGTCCTGCAGGTTTACAAGGA		B

797	AGLC195	CK148840	ATGCTGGGCAACACCATTAG	ATGATTCCGACGATGATTCC		B
798	AGLC196	HO214322	TGATAAATTGCAGGTGGAGAGAGA	TAAGTTAGACTCCCAGGCAAGGTA		B
799	AGLC197	CK148702	AATGGCAGATCCGTTGTCAT	AAGATCATCGATCCGACCTG		B
800	AGLC198	CK148699	GTGAGGGCGTAGCTTCACAT	CGAGCAGTTGAGGACCAAAT		B
801	AGLC199	CK148767	TGTCCATGTCGCTTGAGAAA	AGCTCAATGGTTACGGGAGA		B
802	AGLC2	CK148953	GCAGCAGCCAGCCTAAGTAT	GGTCTGTGTTGGGCTTGTTT		B
803	AGLC20	CK148978	GTCACTGGCCGTCGTTTTAC	TGATGGAAGGACAGTTGCAG		B
804	AGLC200	CK148962	GCATCCTTCCCACTTCTTTGCA	GAATGGACTCGGATGTCTTAAGCA		B
805	AGLC201	CK148977	CTTGCATGAGGATCATAACAACGA	AGTTCCTAGTGAGTTTTGGGGAGA		B
806	AGLC202	CK148766	TCAGAATCCCCATTAGTGCAG	GCTTTGGGATAGGATTTCCAG		B
807	AGLC203	CK148811	CATGCTTCCTTTCGCTTCTC	CCTGCCGATCATATTTCCAG		B
808	AGLC204	CK148685	CGAATGACACGATACCCAGGA	AAAGGGCTCTTGTTATAGGAGGA	689	B
809	AGLC205	CK148797	CGCAATTTTACGCTGAACCT	TCCCCTGCCATCATAGAGTC		B
810	AGLC206	HO214310	CACAGCTCCTCCAGTGTTGA	ATGGCATGGGGAGACTATTG		B
811	AGLC207	CK148808	GCCTTCGCGATACTCTGTCT	GCAAGAAAGGTCGTGGTCTC	95	B
812	AGLC208	CK148872	TCTCCTTCGTCCGATCTGTT	ATCGCCTGAATCCTATCACG		B
813	AGLC209	CK148845	GAATTGGACCTGGACTGACG	ATCGGTTGGGAACCTATGTG	458	B
814	AGLC21	CK148885	TTCTGATTCTGCGTCGTTTG	ATGTTCCAGCCCAGAAAACA		B
815	AGLC22	CK148709	TGTCAGACTGAGCTGTGTATGAGA	TTGCCCGTATGGTTATGTTAGGAA	415	B
816	AGLC23	CK148947	AATGGTGATTCGTCAGTCGCCTA	CTGTCTGAAGAAAGTGAACGAA		B
817	AGLC24	HO214283	ATGCATTGGATGCGGTATTT	CCAATTCCATTTGTCCCAAC		B
818	AGLC25	HO214311	GCCTATGGACTGCTCATGCA	GCAGCAGTGTCCAATACTGCAA		B
819	AGLC26	CK148951	AGCCAGGGATTTCAGAACCT	GCAACACTACCACCACCACA	605	B
820	AGLC27	CK148933	GTCGCTTCTTGCATCAAGGT	GATGGCAAGAACATCAGGGTA		B
821	AGLC28	CK148934	TGCACCATGTTGGAGAAAGA	CGTCCTGGATTCATCGAAGT		B
822	AGLC29	CK148945	GCTGATCGCCTTTTGGAACTCA	TCGAAGGCAAGGTAAAAGTTAGCA		B
823	AGLC3	CK148968	GTCCAGTTCGCCCAATTCTA	ATGGAAGGCCGTGTCAATAA		B
824	AGLC30	HO214295	TCGTGCAGCTCATCCATATT	TTGTCCACGCATCATCTCAT		B
825	AGLC31	CK149117	GGAAACTACCGCATTCCTGA	TTCGTGAATGTTGCAAGAGG		B
826	AGLC32	HO214323	ATGCTCAGAGACAGGCAACA	GCCTTCACCTCAAAAACACC		B
827	AGLC33	HO214296	CTCCTGTAGTGGCATATCTTCGAA	TGGTCCATTTATGCCGCTGGTA		B
828	AGLC34	CK149140	TGCAGCTTGTCCGGATGCA	TAGGTCCGAGAGGCATCAGAGA	815	B
829	AGLC35	HO214324	CCAAGGGATCAACATAACGATCCA	GCAAAGAAGCATTTCAAGCCAA	813	B
830	AGLC36	HO214297	ACTAGTCCTGCAGGTTTAAACGA	GTGACAGTATTTTGGAGGAGTCA		B
831	AGLC37	HO214325	TAGTCCTGCAGGTTTAAACGA	GGTTGCAGCATTGCTCGA		B

832	AGLC38	HO214298	CAAGTGCCACAACTCTAAATCCAA	CATCTTCCAATGTGAATGACCCAA		B
833	AGLC39	CK149136	CAAATTTCTGTTCTTCCACCCCAA	GGCGATCTTCGAGTCCATCGA		B
834	AGLC4	HO214326	GCTAAACCTTAGAGCAATGACTCA	CCTTGCTTGTGCCTTATCTTCCA	260	B
835	AGLC40	HO214299	TCTTCAACACCTCCATCTAACCTA	GACATGAAACCAAAGCATCACA		B
836	AGLC41	HO214327	TGCTCTGCCCCATCTGAGGA	ATCACATGGTGGTGCTGGTCA	850	B
837	AGLC42	CK149134	TCTCTGAAACACTCTAGCAAGTGA	CGGCTTTGGGGAACGAAGGA	266	B
838	AGLC43	HO214300	TAGTCCTGCAGGTTTAAACGA	TAACATGGGTCTCTGCTTCTCTCA		B
839	AGLC44	CK148880	CTTTACCAAAACCACCTTCACCAA	GCAGGTCGCGTTGTTGCAA		B
840	AGLC45	CK149070	ACTAGTCCTGCAGGTTTAAACGAA	TAACATGGGTCTCTGCTTCTCTCA		B
841	AGLC46	CK148932	CTTTACCAAAACCACCTTCACCAA	TCTCTCTCTCTCTCTTCTGTTCCA		B
842	AGLC47	CK148819	ACTTTACCAAAACCACCTTCACCA	TCTCTCTCTCTCTCTTCTGTTCCA		B
843	AGLC48	CK148960	CTTTACCAAAACCACCTTCACCA	CTCTCTCTCTCTCTTCTGTTCCA		B
844	AGLC49	CK148683	GCAGGTCGCGTTGTTGCAA	ATCGTTGAACCTGAAGTGTGA		B
845	AGLC5	HO214328	AACTAGTCCTGCAGGTTTAAACGA	TAACATGGGTCTCTGCTTCTCTCA		B
846	AGLC50	CK148682	GCCGAGGTACACTTTACCAA	TCCTCACACTTCAGGTTCAACGA		B
847	AGLC51	CK148677	TTCTCAGACTTCAATCCTAGCA	TTGGTCCAACTTATGACTTCCA		B
848	AGLC52	CK148718	CTTTACCAAAACCACCTTCACCA	CAGGTCGCGTTGTTGCA		B
849	AGLC53	CK148806	CTCTTTCTTTCCCTCTAGTTTCCA	GTGCTTTTTTCTGAGGTTCAGGTA		B
850	AGLC54	CK148742	GACTAGTCCTGCAGGTTTAAACGA	TCTCTCTCTCTCTCTTCTGTTCCA		B
851	AGLC55	CK149133	ACTAGTCCTGCAGGTTTAAACGA	ATTCCAAAATGGAGCAGGTGCA		B
852	AGLC56	CK148721	CGAATGACACGATACCCAGGA	AAAGGGCTCTTGTTATAGGAGGA	689	B
853	AGLC57	CK148900	CTCCTCTTCTCCGTCGTAGCA	CTGGTCCTTGACGGGAGTGA		B
854	AGLC58	CK148795	TCTACCCGCCGCTCAATAGGA	CCTTCGATGTCGGCTCTTCCTA		B
855	AGLC59	CK148800	GTTTACATCATGACCCGCCCTA	TCACCAAGACCAGAACGTTCCA		B
856	AGLC6	CK148967	TGCCCAACGGTTTCTTTTACCA	TCAGAGATACTCGCCCACCAA		B
857	AGLC60	CK149124	TAGTCCTGCAGGTTTAAACGA	ATGGGGACAGAATTGGGGTGA		B
858	AGLC61	CK148802	CGGCCGAGTACAATTTCTTCCA	ATTTGCTGATGATTGCGTTCCA		B
859	AGLC62	CK148805	CCTCTCTCTCTCTCTCTTCTTGCA	AATCCTTTAGGAGGAGAATGTGGA		B
860	AGLC63	CK148809	TCTTTGAGCAGCATTCATTCCACA	GAGTGCTACCTTCAAAGACTGCA		B
861	AGLC64	CK149128	CGATCAAGAACCCAGTTTTGCAA	AAGATCGACAGGCGATCTGGTA	337	B
862	AGLC65	HO224279	CACTCTCCGTTCCGGTTCCA	CTGTCCATGCCCTTGTCCA		B
863	AGLC66	CK149070	ACCAACAATCTCCCTCTTCCCTA	GCGAGGTACACTTTTCCCCAA		B
864	AGLC67	CK149102	CAGGTCGCGTTGTTGCA	GGCCGAGGTACACTTTTCCA		B
865	AGLC68	CK128906	GGTCGCGTTGTTGCAAAGCA	GTTGTGTGAGAGAACGCACAGA		B
866	AGLC69	HO214301	TTCATCTGGCACTAGCATATCTGA	CGACAATTCTTGCTTCAACAACCA	642	B

867	AGLC7	CK148966	TAATCATCGGTCATGAGTCTGTCA	CAAAATCGAAGATCTGCATCTGCA		B
868	AGLC70	CK149074	GCCGAGGTCAGTAGGAGAGA	CTTGCTTACGGATCTGGTCCAA		B
869	AGLC71	CK149095	GTCGTGAAAAGCCTTGGACGA	ATCAACCTTTCAATATCGCGCAGA		B
870	AGLC72	HO214302	CATGTTTTCTACCCTCACAATGCA	TACTCACTTGTTGTTCCAGACA		B
871	AGLC73	CK148849	TTCGATCCTCCGACCCCGAA	TTCGCTAGATCTGGATACTTCTCA		B
872	AGLC74	CK149006	CAGGTCCGCGTTGTTGCAA	GGAAGAGTGAGATTGTTGCGTGA		B
873	AGLC75	CK148853	CATGATTGGAACTTGAGTCGTA	TCAGTTGCTTCCCTTTTTCTGGTA		B
874	AGLC76	CK148862	TCTTCTTCTTCTTCTTCAGCCACA	GTGGATTGGGAAATGTGAATGTCA		B
875	AGLC77	CK149101	GCAGGTCGCGTTGTTAGCA	ATTACTATGCTTCCTTCTCCTCCA		B
876	AGLC78	CK149132	CCACAAAGGACGACAACAACGA	CCCAACACGAACCACACGA		B
877	AGLC79	CK149080	ATCCATCACAACCCTCAACTCA	CTCCGTCAACCTTTCCGCAA		B
878	AGLC8	HO224280	TGTTGTCTCGCCAATTCAAAGCA	CGTTTGGTGGCATTCCTGCA		B
879	AGLC80	CK149038	GGTCGCGTTGTTGCAAAGCA	TGCTTCCTTCTCCTCCATTACCAA		B
880	AGLC81	CK149043	CAAACTCCTCAATAGCAGGCACA	GCTGTATCGGAGAGTGGTCAGA	262	B
881	AGLC82	CK148871	CCGAGGTCTTGCCATTGGTA	CAGATTCGTTATTGCCTTCCCGTA		B
882	AGLC83	CK148894	CGCCATCGTTACTTTCTCTTACCA	AGTGCAGGGCACCAATCACA		B
883	AGLC84	CK149089	TTTAATTACGCGGTTTCCACGA	GAAGACTTGAGACATGGGCACA		B
884	AGLC85	CK148761	GATTTGCTTGGTGATGATGCTGA	CCTCGTGGTCCACCATAGCTA		B
885	AGLC86	HO214303	CGTGGGATTGAAAAAGTTGCTA	CACTACCAGCCAAAGCACTCA	319	B
886	AGLC87	CK149100	CAACAACAACCTATCCGAACCTCA	ACTATCCCTAACCTTCCATCACCA		B
887	AGLC88	CK148768	CATGAGTGGTAGTGGGAGTGGA	GTTCGTTTGAGTCGTTTACTGGAA	364	B
888	AGLC89	HO214330	CTAGACAGGAATGTTGTCTAGAGA	GAGATTGGGGGATGACAAACACA	578	B
889	AGLC9	CK148885	TCAACAACGCTACCCGATCCAA	TTCTCAAGAGCACCACAAAAGAGA	291	B
890	AGLC90	CK149113	CGGCGGCTATATTGGTTTTGCA	TCCTAAACCCCACTTATCTCCCTA	253	B
891	AGLC91	HO214304	GACCCCCAAAAATGAAAAAGCA	TTGCCCATACATTCTTCACCCAA	469	B
892	AGLC92	CK148772	TCCATCTTTGAGTTGGCATTACCA	CGCGGTCGAAAGAACGCAA		B
893	AGLC93	CK148904	CTTCAAGTTCTTCGTTTGACGCAA	CCTTTCTCCCACAACCTCTCCA		B
894	AGLC94	CK149093	TTTGTGATGGTCCTGCTCTCTCA	ACCGCTTCAGGATCAACTCGA	216	B
895	AGLC95	HO224281	TCTTCCGATCCTAAGAAAGAGCAA	ACCAATATGGAGAGCACCAGTCA		B
896	AGLC96	HO214305	CCACCTTCCATCTCCAATTCCAA	GACTGAATCGGAGAAGGTTTCTCA	361	B
897	AGLC97	CK149125	CCAGCTTCTAATGTAGGTCTGCA	CAGCAGCAGCAGAGAGAGCA		B
898	AGLC98	CK149016	TAATCCCCAAACAGGTTACACTGA	AGGGCAAGCCAAGGAAATCCA		B
899	AGLC99	CK149119	TTGGTGCGATGGCAGCA	ACAATCATCGGCGGGCAGA		B
900	AGLC122	HO214331	GCCACGTTAATACAACAGGG	CGGGGGAATGTCTTTTAACT	1106	B
901	AGLC210	TC100902	ATGCAGACAGTCCTGGTCAT	GCATCTTTCACCAACCATTT		B

902	AGLC211	TC230969	CAAGGTGACGACATTTCTGA	GATCCGGTGGCTATGTTAGA		B
903	AGLC212	TC95432	ATGTTTGAAGAAGGATGCCA	ATCATCATCCCACTCGTCAT		B
904	AGLC213	TC108358	AAGTTGCCAATTCCAGACAG	TTTATCGGGTTTGTGATGCT		B
905	AGLC214	HO214277	AGAAAATGTCCAAGCACAGG	ATTGGTTTTGGTGGTTGTTG	862	B
906	AGLC215	TC106321	TTTAAGGAAGGGGGAGTCAA	CATGAGTACATGACCCACCA	370	B
907	AGLC216	TC96248	AGAGTAATAGTGGGGGTGGC	GTATGAAAGGCCAGCAGCTA		B
908	AGLC217	TC96248	GGAAGCAGCCATCTAAGGAT	CCAAACCTCAAATAGGGGTC	1676	B
909	Aldol	AF416480	CCCAAGAAGGGAATTGTTTG	AAGTTGTAAAGCTGCATGTAAACC	180	C
910	est	AJ746343	GAATGTAGGTGCTGGAGGAG	CAGGAAGGTCTGAAGATCCA	1639	C
911	ge13bg	AJ131047	GTGCTGGTTTCCATTTAACG	CGCCATTAGAGAAAGATGGA	425	C
912	kpi1	AY635930	GTAGGGATACCCGCTGAACT	CCTATGTGACGTCCTGTTCC	310	C
913	lrr	AJ609275	AGTGACTTGGTCAAATGGGA	AATATCAGCATGGTTTCCGA	662	C
914	ntp	AJ489612	AGAGCACATGACCCACTCAT	GAACAAAACGTGACAAGGCT	425	C
915	rgr2	AY356156	TGCTACTTATCCATGGCCTC	AAACTCCCCATTTTCATATTCA	370	C
916	rgr4	AY356154	GCCAAACGGAGATTAGATGA	ATTCCCCTATTTGGGTTCAC	160	C
917	SHMT	AF416481	TGAAATTGAATGAGGGGAAA	TGGCTCCTCCACATAGTCAT	554	C
918	tk	AB025004	TTTGGCCTGAAGTGAAAAAG	GTGTTTGGTTCTACGTTGGG	660	C
919	tpi2	AJ276262	CTTGCAATCCTCTGGTTCAT	GTGTGGCACTGAAGGGATAG	317	C
920	chs	AJ012690	GAATTCCCTCGTTTTCCTTC	AAATGAATTGTGGAAAGGCA	282	C
921	CaHa134	EL585364	TTACAAGCTCCTTTGCAGGT	GACCAAATTCAGGTGCTGAG	145	D
922	CaHa140	EL585394	GCTGGTGAATGTGGAAAATC	ACCAGACATTTTGGCAGTGT	186	D
923	CaHa141	EL585365	AACAACAAGGCAAGTTCAGC	GTCTTGGTGGTCATCCTCTG	105	D
924	CaHa149	EL585366	CGATGAAGTGGATGGAAAAC	GACCCTAGGGGTAAAAACCA	178	D
925	CaHa156	EL585367	CCACACAAACTGGATCACAA	AAGGAGGAGTGGAAGAGGAA	187	D
926	CaHa159	EL585368	GTATCGGTTGGCTCAATGTC	CCACGAGAAGAGGAACAAGA	312	D
927	CaHa22	EL593260	GCTTGTCGACGAATTCAGAT	AGGAAGTTTCTTCAAAGCTCG	353	D
928	CaHa23	EL593261	GCAGGTAACCAAAGAAAGCA	TCCCAAGAGTAACAGCCAAG	105	D
929	CaHa245	EL585369	ATGCGACTAACGCTGCTACT	GACACATCAAAGCCAAATCC	151	D
930	CaHa265	EL585370	AGCATTAAGATCGGTTCGTG	TCTGAAGCTCCAATTCCTTG	242	D
931	CaHa269	EL585390	CGGAAAATTTTGTTGTTTGG	GAGGTCATTGCTAAGGCAGA	142	D
932	CaHa270	EL585371	CTGATGGCTTGGCTGTTACT	TCTTGCTTGCAAACTGAACA	126	D
933	CaHa271	EL585372	GTGCTGTTGGCTTCTGTTTT	TGCTTACGCATTTCAACAAA	268	D
934	CaHa274	EL585373	TTTGAAGTGATTGTCTGCGA	CAACGGAGGAACCTCCTAAT	117	D
935	CaHa287	EL585391	ATGATGAATGGCATGAAACC	AGTAGCAGCGTTAGTCGCAT	144	D
936	CaHa297	EL585386	GACCAAATTCAGGTGCTGAG	TTACAAGCTCCTTTGCAGGT	145	D

937	CaHa299	EL585357	GATCCAAACCAAAACACTGC	CAAGCCTTCAATAAAAGCCA	390	D
938	CaHa302	EL585374	AGTGTGGCAAACAACCATTT	TCACAAGATCACCGACACAC	217	D
939	CaHa32	EL585350	GCACAATCTTAAGCTCCCAA	TGGTCTCAAATTCCCAGGTA	202	D
940	CaHa34	EL585351	AGACTGTTGCTGAGGAGTGG	ATTGTCGATTGCAGCTTCTC	115	D
941	CaHa35	EL585352	AACCCTTTGTGAGTAACCCC	TCTCAGTTTGTGATGAGCCA	152	D
942	CaHa36	EL585353	TCTCAGTTTGTGATGAGCCA	CAATGTCAGCCAAATCAACA	171	D
943	CaHa41	EL585354	TAAAAACCAGTGCCGAAAAG	AGTTACTGGCATGCTTCTGG	251	D
944	CaHa43	EL585355	AACCTTTCACCAACCCTTTC	AGCTCAACAAATATGGCTGC	129	D
945	CaHa458	EL585393	CGGTCTTGTGGAGACATCA	CGTGAGCGGATAACAATTTC	244	D
946	CaHa470	EL585388	CTGATGGCTTGGCTGTTACT	TATTGTTGGTCGAGGTTGGT	264	D
947	CaHa478	EL585375	GGTGAAAAGATGGTGGTGAG	GCCAATACTTCACACGTTCC	196	D
948	CaHa480	EL585387	TGAACTTGGTAGGGCTTCTG	CTCAGCATAGCTCTTGGCTC	189	D
949	CaHa488	EL585376	TTGTGGTCCAGAAGATGGAT	GCATGAGAAACTCCGAAAAA	263	D
950	CaHa49	EL585356	CAAAGAGAAAGCAAATCCCA	TTTTCACATTGCTAGGAGCC	206	D
951	CaHa490	EL593262	TGCTCATGAATCACAACACA	CGAGGTACGAGACTTTTAAGGTT	131	D
952	CaHa495	EL585389	AACAACAAGGCAAGTTCAGC	TAACAACTGAGGGAAGCTGG	316	D
953	CaHa500	EL585377	TTATGCTGACTGTTCTGCCA	CACACACATCTCCAAACCAA	110	D
954	CaHa506	EL585378	GATCCAAACCAAAACACTGC	CAAGCCTTCAATAAAAGCCA	390	D
955	CaHa508	EL585379	GGCTCGCTATGCTGCTATTA	AAGAAAACCTCAGCCCAAAC	132	D
956	CaHa511	EL585380	ACATGTTGCTCTTAGCAGGG	GACCCTAGCTGCTGATGAAA	110	D
957	CaHa515	EL585381	TAATTTCGGCTGCAGTTACC	CAGCTAAACCTGTCCCAGAA	190	D
958	CaHa520	EL585382	TGGTCCTTTACTTCCCATCA	AGGACATTCCAAGTATGGCA	408	D
959	CaHa522	EL585383	AATACGGTGCTATTCGGTCA	ATGATCGACGGCAGTTTTAG	108	D
960	CaHa55	EL593263	ACGATGAGATTGACATGGCT	AAATCATGAGAAAGGGCTCC	152	D
961	CaHa584	EL585384	CGCTTTCTTTGTCTTCCTTG	ATATGGAGCTTTTTCGGCTT	138	D
962	CaHa599	EL585385	GAGGTCATTGCTAAGGCAGA	CGGAAAATTTTGTTGTTTGG	142	D
963	CaHa609	EL585392	TTTACACCATGCACAACCTAGA	TTCAAGTCAAACAGAAAAATGC	104	D
964	CaHa61	EL585358	CACAACAACAACCCACAAGA	GCCTTAACTTGCTTGGCATA	246	D
965	CaHa66	EL585359	TGCTCCAGAAGCAGAGAAGT	ATGTCATGCTTGCTTATTGC	201	D
966	CaHa67	EL585395	AAGATGCAAGGATACCCACA	ATCGTCCTGGTAGTGATGGA	108	D
967	CaHa81	EL585361	TTCAAGGCACTTGAGGGTTA	GGCAACCAAATTAAGCACAC	120	D
968	CaHa82	EL585362	AGAGGCTTCAGTTGGCCTAT	TTGCGGTTCACACACATTAC	237	D
969	CaHa91	EL585363	CGGAAAATTTTGTTGTTTGG	GAGGTCATTGCTAAGGCAGA	142	D
970	CaHa75	EL585360	GAGGAAGCCGATAAATCTCC	GGTTTCTTGATGGTTACTTCAAT	101	D

Heterologos genes

1	LUP1	CA410658.1	GCAACAGAGAGAGATGTATATGAG	GCCAACATTATCTTGAACCT	425	E
2	LUP10	BG153990.1	GATCCTATTGTGTTGGCCT	TTCTCATAAGGTGCCAGC	455	E
3	LUP120	CA411405.1	GTTTGACTCCATCACAGATTG	CTGCTCTCAACAAGAATCAAC	733	E
4	LUP132	CA410507.1	ATTTCGAGAGGCTCAACACTG	TTGCACACAGAACGCGTTAT	460	E
5	LUP144	CA410561.1	CCAAAGAATATCCTCATTACTG	GGTGATGAGAAGGTAGTTGACA	461	E
6	LUP169	CA409561.1	AGTGCCGATATGCTGTTTAT	GTCCATCTCAGTAGGATCAGTT	497	E
7	LUP171	CA409550.1	ATGCTGAAGTAGTTGATGAAGC	AACTCTTAAGTGCAACATGACC	416	E
8	LUP19	CA409626.1	TTCTCCAAGAAGGATTGGTA	TTCCTGAATGAATGATCCTC	487	E
9	LUP204	BG154069.1	GAGGACATTCAACCTCTTGTT	ATTACACCAGTATGACGAGGTC	476	E
10	LUP211	CA409681.1	TCCGAGTACAAGAAGCCATA	ACAGATCCTTCTCCTGGACT	461	E
11	LUP221	CA410151.1	GGTATCAAGGAGTCTGATACTGG	GCAATCTGCTTCACATTGAT	476	E
12	LUP227	CA410275.1	ACCATAACTTGACTGAGGAAGA	ACACTTCAGCTTGTCCATAAC	412	E
13	LUP235	CA410523.1	AGAGAACATGAAGCCATGAA	CCATCTTCAGGAATAGCAAG	447	E
14	LUP240	CA410832.1	GTCTGTTCATTCCTTCAGAATC	CACCATACTTCACTCCTTCAC	509	E
15	LUP241	CA410838.1	CTCCTTCCTCCTCGTATCAT	CTCATTCCTATCATGCAGTCTA	450	E
16	LUP242	CA410847.1	TTGTCTCTGGTAATTCTCACAC	AGACATGTATGACACCTGATTG	466	E
17	LUP244	CA410975.1	CAGATCTCTGAGTTCAAGGAAG	CATTAGGTTGAGGAATTCTGG	395	E
18	LUP246	CA411057.1	ACTAATGCAGATGAGAGGTCTT	ATCCTTGTCAGAGAGAGTTCC	431	E
19	LUP255	CA411216.1	GTGTCTACACCGATGTTAATGT	TTCTTGACAGGCTTGAGAAT	730	E
20	LUP258	CA411319.1	AGGCTAACTACTCAAGAAGCAG	GTGTCAAGTACTCCAACAATCA	704	E
21	LUP263	CA411389.1	CAATCATCTATGGTGATGTGAG	CTTCTACTTCGTCCTTGATAGC	713	E
22	LUP274	CA410212.1	AGTTTCCCATGTTGTTGCT	CTGATTCCAGGAGGTTTATATG	479	E
23	LUP276	CA410228.1	GGAAACTTAACACCAAAGAGAG	TTGCTCCATACACTTCTTGTAG	471	E
24	LUP280	CA410338.1	GGATGTGACAAACAAGGATT	GAACAGACTTTCTCCTACGTTCT	400	E
25	LUP302	CA410731.1	TTAGTTGGGAATACAGCACC	GGGACAAACAAATGTGAAGT	529	E
26	LUP307	CA410781.1	CACTGATGAACAGATCTCTGAGT	CAACCTCATTAATCATGTCCTG	474	E
27	LUP318	CA410959.1	CCTAGTGATTCTGGGATTACAT	GTTGCTTCAATGTCACCTGTAT	396	E
28	LUP325	CA411064.1	GTGTTCTAGCAGAAGATTGTGAC	GCCAGTTACCTTCCTAGCTT	464	E
29	LUP326	CA411066.1	CTCACAGGAAATTCGAACAC	CTCTCTAACAATGTGAGTCATACC	420	E
30	LUP327	CA411083.1	GCTAGCACTGGAATTGTTCTT	AAAGCTAGCTGTGACATCTCTT	454	E
31	LUP36	CA4101341	CTGATATCCATTCACCTCAAG	AACACCTGGAAGTCCAACTA	474	E
32	LUP51	BG104139.1	TAGGTCTAGATACTGGTGTCCC	GGTTTGTAGTTCTTGTTTGGAG	468	E
33	LUP60	BG154069.1	GAGGACATTCAACCTCTTGTT	ATTACACCAGTATGACGAGGTC	476	E
34	LUP8	BG149164.1	CCAAGAAGAAGGACCTCAT	TACACACTCCTTAGCGGATT	432	E
35	LUP80	CA410141.1	AGTGCCGATATGCTGTTTAT	GTCCATCTCAGTAGGATCAGTT	505	E

36	LUP91	CA4107721	TCTTTAAAGGTGTAGCTCTGGT	AGGTTGGAGTCAGTCTTAGAAA	505	E
37	LUP92	CA410807.1	GCGTGCTGATTACGATTAC	GTGAAAGAACCATCAGAGAATC	526	E
38	LUP94	CA410832.1	ACTACCATTGAGGAAGTCAGAG	GATGTCGTTGAAGTTTGAGTAG	509	E
39	LG1	TC221394	CTGCAAAAGTTTACCTTCAGGGC	TTCCTTGAGCAAGGATGGTTCC	590	F
40	LG10	YPR_010_F09_076Ab1	TTTCTCTTGAACTTGGGGTGTTCC	GGACGGAGGCAGAGACGTGG	640	F
41	LG100	YPR_001_A03_017Ab1	CCACGGGGAGTTTTCACATAACC	GGGTTACAAGGCTGGTATGACTCAC	720	F
42	LG101	YPR_001_A03_017Ab1	GTGAGTCATACCAGCCTTGTAACCC	GAAGCGTGCTGCTCGTCACAG	560	F
43	LG102	CO979090	GCCACACCTTGGTTCTAGAGGGC	GGCGTTTGAAGGAGCGTAAAGC	590	F
44	LG103	TC211005	GTGCATTCCATTGCAATAGCATCC	TGGACACTGTTGACCTGCACCC	860	F
45	LG104	TC212299	CCGTTATGGACCATCCTTTGCG	CCAAGGCAACATTTCACAACCTG	570	F
46	LG105	TC229445	TCCACGGTCGCATGAGGCC	CGAAATACAGCCTCTTGCACAGG	330	F
47	LG106	TC229445	GGGTGGTTCCATCATTACTGAAATG	TCATGCACTGCCTCAATTTCTCC	720	F
48	LG107	TC229445	CAGGCCTTCCTCCTCACATATCTG	GGTACTTGTGATAATTGGTTGGCTCC	860	F
49	LG108	TC229445	ATATGGGGAAGATGGTGAATGCC	GCATGAGTTTGCTCTGGGGTTCC	510	F
50	LG109	TC208738	CGTCCACCCCAACTTGCACC	GGGAGGAAATTTCAGCTGTGTGG	580	F
51	LG11	TC205974	CGTACATAACCATCTTCACCTCCAC	AGAAGAAATTGGGGGTGTGAAAGG	990	F
52	LG110	TC213906	TTGTTCAATGGTCAGGAGACAATCC	GGCTGATGGCAAGATCTCCTGG	300	F
53	LG111	TC209212	GAGCATGAAATTCTACCTCCTCTTCC	TAATTCTCTTTGGGGGAGAAGGC	380	F
54	LG112	YPR_008_D03_027Ab1	TTCATGGTAAGCCTTCTCAGCAGAG	AAGCTATCTATGACATTTGCAGGCG	380	F
55	LG12	TC205974	CCTTTCACACCCCCAATTTCTTC	GGAGGTGGTCAGGATGCCTCAG	200	F
56	LG13	TC205974	CTGAGGCATCCTGACCACCTCC	CAGTCAATGCAGTTACCATGTCACC		F
57	LG14	TC205974	TGGTAACTGCATTGACTGGGCG	GAGTCCGACAAGGAATCTGGCC	310	F
58	LG15	TC205974	GGCCAGATTCCTTGTCGGACTC	ATCAGTGCCGGCGAAGATGC	870	F
59	LG16	CO979052	TGACCAGACCGCAAGTCTGTTCC	GGGACATGGCTCCTTACCAAGC	480	F
60	LG17	CO979052	GCTTGGTAAGGAGCCATGTCCC	TTCAAAGGGAATCAAACTTGGACC	330	F
61	LG18	BM308011	CCTTGTCCCACGAAACATCAAGC	TCAGAAGAGCATCCATGCTGGC	600	F
62	LG19	BE660669	GCTGGTGCCACAGCTGATGTCC	CCAGCAGGACCAATGCCACC	360	F
63	LG2	TC208087	CCTCTCCTCAGCTTGTATCTCCTTG	TCGTGTATTTGCCTTGAGGGTTC	300	F
64	LG20	BE660669	GGTGGCATTGGTCCTGCTGG	TCCAGGGCCCATACATCAGCC	410	F
65	LG21	TC228354	GAGTTCGGCTGCATCACAGAGC	CCATCCACGAGGCTTTAAGAGGTG	630	F
66	LG22	TC228354	GTGCACCATCAACTGCTGGTGAC	CAAGCTGTGCTGTCCCCGATC	570	F
67	LG23	TC228354	TCGGGGACAGCACAGCTTGG	CGATGCCATGTCCAAATTCCG	330	F
68	LG24	TC228354	CGGAATTTGGACATGGCATCG	AGCATTTGGTCGCAATCCTTCC	310	F
69	LG25	TC228354	TTGCGACCAAATGCTAGAGGTCC	AGCAAAACGCTCAAACCCATCC	870	F
70	LG26	TC228354	TGGATGGGTTTGAGCGTTTTGC	CGACCTCTAGTTGGCGTTGGTG	510	F

71	LG27	TC220227	GACCTGGAACAGCCACTACCCC	TCTTCTTATTTGCCGAGCTTCCTG	240	F
72	LG28	TC220227	TCAGGAAGCTCGGCAAATAAGAAG	CGGCACCTTTTCAGGAAGCTGC	550	F
73	LG29	TC220227	CCTGAAAAGGTGCCGGAGCAG	TTCCTGGAATGCTAGAGCGGAC	270	F
74	LG3	TC208087	CTCAAGGCAAATACACGAGACGG	CCACCTGCAACGTTTCCTCG	870	F
75	LG30	TC220227	TCCAGGAAATGAGGCCAGGG	CGCCTCAATGGACTCTGGATCC	930	F
76	LG31	BQ273547	CCTTCATACGTTTCATAGGAAGGGTC	CCGTTTCTGCCATCAATTCCG	550	F
77	LG32	TC228615	CCCTGACTTCTTACTGCATGCAGTG	CACTTTCGCGCCTCATATCAGC	840	F
78	LG33	TC228615	GCTGATATGAGGCGCGAAAGTG	GCCCCGCAGAGGGCAGAGG	480	F
79	LG34	TC228615	TGCCCTCTGCGGGGCAGG	TTCTCCTCAGCCTCCCGTCG	810	F
80	LG35	TC228227	GCACACGAGGAGGGCATTCG	CATGAGGATATGTCAACGTGGATGC	390	F
81	LG36	TC228227	GCATCCACGTTGACATATCCTCATG	GATTGGTGGCACAGCCTCGG	760	F
82	LG37	AW830293	GGAGCAACAGCCCTTAATGACCG	CTCTCACTTCCTGCCAAGTCAACC	240	F
83	LG38	TC229193	GTTGAAACTAGTGGAGCCCTTAACAC	CCTCACACTGGGTCGGGTTAACC	410	F
84	LG39	TC229193	TTAACCCGACCCAGTGTGAGGC	ACTGTGATTGCAACATGGTTTCCC	640	F
85	LG4	TC208087	CGAGGAAACGTTGCAGGTGG	CGTGAAGGGTTTGGCCAAGGC	500	F
86	LG40	TC229193	ATGTGTTCAAGCCAATGATTGCC	GAAGCAGACGAATCTCCAAGCAG	690	F
87	LG41	TC229193	TGCTTGGAGATTCGTCTGCTTCC	CAACTGTTGCTGCCTTGTCATAACC	330	F
88	LG42	TC229193	TGACAAGGCAGCAACAGTTGCC	AATCAGATGGTCCAAGCATTTTCTC	420	F
89	LG43	YPR_009_H05_044Ab1	CACGCATTCTCCTCATCCAAAGG	AAGAGGCCAAGAGAAAAGGACGTC	540	F
90	LG44	TC215655	TGGGAAGGTGCTCACATACTCTGC	GAGGATTTTCAGTGGGACATGGC	480	F
91	LG45	TC215655	GCCATGTCCCACTGAAAATCCTC	CCACGGAAGTTCCACACAGCC	360	F
92	LG46	TC215655	GGCTGTGTGGAACTTCCGTGG	GACACTTGCCAGTTAAAATGTTGCTG	600	F
93	LG47	TC212235	ATTGGCATAATCGTTTCATGCCTC	GCCATCTTCATCAACCACTTGCTC	660	F
94	LG48	TC212235	GAATAACATTGATGGCAGGTGGTG	TCTCAATATCATCAAGGTCCAGCTG	580	F
95	LG49	BQ299641	CAACAAGCCAGTCGCTAGTGCTAC	CATACATGCATTTTCAAAGGCCTC	1040	F
96	LG5	TC215834	AGTAATCCAGATATCAGCCACAGGG	AAGTGAATGTGAGGGAAATGTGCC	840	F
97	LG50	TC210291	CAAGTTCATCGGTTCATCATACTCCAC	GATACCATCTGTGATGGATTGCACC	225	F
98	LG51	TC210291	CGTTGTCCCATGCTCCAGTTTGC	GTCAGTTGCAAGAGGCTGTCCAAG	370	F
99	LG52	TC222088	AAGTGAAAGGCCATAAGAGAGAGCC	GCAAGTCTTGTTTGTCTCCATACCG	620	F
100	LG53	TC218263	CACTGGCTGCTGAGGAATCTGC	CCAGGAGGAAGCTTCTCAGCCAG	210	F
101	LG54	TC206681	TCTGCTTATGACCGCTACCTTCAG	AATTGACATTCCGTCCATTTCGTG	300	F
102	LG55	TC206681	TGTTGAAGGTCTCCCATCTGACAG	GCTGCACAGGCCGGATTTGC	340	F
103	LG56	TC206681	CGGCCTGTGCAGCAACTGC	CCGCAAGTGACTAGACTCAGGATTG	720	F
104	LG57	TC231967	CCTACCTAGCGTGATGACTTGTGC	CCTTGACCCTCTGTTATGGCATAC	200	F
105	LG58	TC219831	GTGAATTAGCAGGATTTCTTGGCC	GCCCTCGTTTGGAAGCTTATGG	300	F

106	LG59	TC221234	TGTTGGCCATGTTCCAAAGACC	GGCATTGTACACTTTGCTGGCC	760	F
107	LG6	TC230161	GGCAGCTGAAGACTTTATCTTCTCC	TATGGAAGCATGGGAGGTGGTC	500	F
108	LG60	TC217408	CCTCCCTGCTCTGCCAGAACG	TGCCCGGGGCATGGATTTC	340	F
109	LG61	TC217408	AATCCATGCCCCGGGCAAC	TTTCCTCCAAAGCAAGGAGCG	540	F
110	LG62	TC217408	GCTCCTTGCTTTGGAGGAAAACC	GGAAAACTTCTAGCAATTCGTCAAAG	580	F
111	LG63	TC217408	TGACGAATTGCTAGAAGTTTTCCTTC	GATGAATCGGATAAGCTATTTGAGCC	360	F
112	LG64	TC226066	GCAGCAGCGATCTCGAAGTTTG	GCACACAAGGAGCCACACTCCC	870	F
113	LG65	TC208741	AGGATACATACCTTCTTGGAAGGAGTC	TGAGCCTTGGGAAATGCCATC	690	F
114	LG66	TC208741	GCCCACACCAAAGAATCTGCC	TACACCTCCAGCAATAGCAGCTTC	600	F
115	LG67	TC226707	ATTGCTTGAAGGGCATCCTTCAC	GGCTGTCCTTGGAGAAGGTGCTAC	210	F
116	LG68	TC226707	CGCACAGGATTTGGTGCAGTTCC	TCCTGGATTCATTTGGAAGACATTG	560	F
117	LG69	TC226707	ACAATGTCTTCCAAATGAATCCAGG	TTGTTCTTGGTAAAGATGGTGGAGC	530	F
118	LG7	TC230161	TTGACGTTGATCTTTACAGCCAGG	TGCAAGAAGAACCTCTTGGACCAG	640	F
119	LG70	TC226707	GGATTCCCATTCCCTGCAAACC	GATGAACAAAGTCCATCAGGAAAGG	520	F
120	LG71	TC226707	CCTTTCCTGATGGACTTTGTTCATC	GCAAAGCAGGATTCGAGTGACCTC	480	F
121	LG72	TC226707	GAGGTCACTCGAATCCTGCTTTGC	CAGGGCTCAACAGGGGTTGTG	260	F
122	LG73	TC226707	CACAACCCCTGTTGAGCCCTG	AATCACAGTGTTCATGTCTTGACACG	390	F
123	LG74	YPR_008_C06_039Ab1	CCACCCCACTCTTCGCTGTTCC	TCCTTGGAGCCGCAGCTGC	290	F
124	LG75	TC206767	GCGTGTGGCGCATGGAAGG	AACGTGGCCTTGATCAGAATTCC	540	F
125	LG76	TC206767	TCAAGGCCACGTTCTTGAGGC	AGGACCCCATCATCACCCAAAG	955	F
126	LG77	TC206767	GGGCTTGGTTCGCAGCATAGG	GAATCACGCTGGAAGTATGGATGAG	750	F
127	LG78	TC206767	CTCCACTAGGAGGTGCTGCAGC	TCCTTTGAATTCCCCAGCGAAG	250	F
128	LG79	TC219033	ACCAGCCATGTGTTTCCACAGC	TGGTTTGAAATTGTCTCCTCTCCAC	350	F
129	LG8	TC234849	GGTTCCACCGATCTCGCTCC	ACGCAGAAACCTCGGAATTGC	260	F
130	LG80	TC219033	TGTGGAGAGGAGACAATTTCAAACC	GGTGGAAGCATTGGATTGGTGC	550	F
131	LG81	TC219033	CCACACAGCGATAGTGGTGCTCC	CATGCCGTTGGTTATCTCACTGG	860	F
132	LG82	TC216103	CTGCGCGATCCAAAGAGCG	GTTGTTCCATGGATTGAAGGCG	500	F
133	LG83	BG653534	GGAGAACTGGCTCGGTATGCTGC	TGGGAATGTTGTGATGCTTCAACC	420	F
134	LG84	TC214052	GTGGCCGTATCTTGGCTGGAAC	GGGATCACTTTCAACACATTGAACC	420	F
135	LG85	YPR_010_D02_015Ab1	GTGCCAGTTCCACCATCATAGCC	CAGCTCCAATCCAGCATCTTGC	680	F
136	LG86	TC232891	CACAAAATGAGACAACAGGATCTGC	TTAGGGACTGGTGAGCTGTACTTGG	230	F
137	LG87	TC217560	CCAATGGCTGAGGAATCTAGCAC	AACAACTGATTCAGGTGCAGGGAG	580	F
138	LG88	TC217560	GCTCATGGAGGCAATGATGCTG	CATTGCCGCGACACAGAGGAC	490	F
139	LG89	YPR_009_B12_093Ab1	CCCTTCTCTACAGCTTCCTTCCACC	GGGTTTTCACTGACCAAGCATTACC	540	F
140	LG9	YPR_010_F09_076Ab1	AGTGAAGTGAATGTAAAGGTTCCCC	TCACTGAGGCTTTATGTGGCTTCC	290	F

141	LG90	TC216765	TTGGAATGCCTTCTCTCTCACCTAC	CCTTCCTCCATACATTCCATTTCAAC	320	F
142	LG91	TC216765	GTTGAAATGGAATGTATGGAGGAAGG	GACAATGAGCCCTTCAGGTCCTG	655	F
143	LG92	TC216765	TGTCAAGGCCGACATTGATGATTAC	TGCCGAAACTTCTTTAGCACCAG	1230	F
144	LG93	TC216765	GCATTGGATTATACCGACATTCCTG	GCTTCTTGCAATGAATTGAGTTGGG	570	F
145	LG94	TC216765	CCCAACTCAATTCATTGCAAGAAGC	CTTTACGGAGAGCCAAAGCAGCAG	710	F
146	LG95	TC216765	CATGGGCAAATGATTATATTCGCC	CCTCTTCCCCTATTGTAGAGAGGCC	300	F
147	LG96	TC216765	GGCCTCTCTACAATAGGGGAAGAGG	GCACCTGAACCCGGAACGAC	420	F
148	LG97	TC216765	GTGACCCTCAGCTGTGATCATCG	AATGCTTTTAACCATTCGGCACC	480	F
149	LG98	BM094229	CCCGCTCTCTCGGATCCTTATG	TGGCTGTGGAATTGGATGGTTC	220	F
150	LG99	YPR_001_A03_017Ab1	CAGCATCAATAGGAACATGGTTCTC	CAACTGTTGTCCGTGTTTTGGC	870	F
151	Lj10112	TC10022	GGACATGTTGAGCCCGTAGT	TCAACAGGAGTCTGAGCTGC	215	F
152	Lj10813	TC10689	GGAAACAAGTTCCGGATGTC	CCTTCACGGAAATGATGTAGAG	100	F
153	Lj11437	TC9986	GCTGCTGTATCTGGCCTTTC	CATGCTATTGTGTGGAAGCAA	206	F
154	Lj15033	TC11978	CGAGTGCAAGCAAGAAGATTT	TGTGGTTCCAAGCTTCATCA	152	F
155	Lj16346	TC12311	CTGGAGTGGTTGATGGAAGG	AGCCAATGCATTTACTTCAACA	226	F
156	Lj1799917	TA4408_34305	TCGAATTCAGAAAGTTCCGC	TTTCCACTCGTTGCAAAATC	133	F
157	Lj1800895	TA4629_34305	TATGGAACAACATGGGCTGA	TCAACTTCTTGACGCCAGTG	139	F
158	Lj1802004	TA1113_34305	TTGGAGGACTTCCTCAATGG	AGAGCAGCTGAAGGCAGAGT	247	F
159	Lj1802179	TA1613_34305	GTTTCAAGGTGAAGCCATGC	TTCCTTGGGTCTCTCCTCCT	100	F
160	Lj1805738	TA3627_34305	ACCCTGATAAGAACCCCACC	AAATTGCTCTCTTTTGTGGATCA	102	F
161	Lj1806583	TA9435_34305	TGATGAATCAAATTCAATGGATG	TCTCAGGCAAATCCTCTTCC	188	F
162	Lj1808289	TA10917_34305	TTCAAACTCCTTTTCTGGCTG	GCTCAGCATCTGTACCACCA	239	F
163	Lj1811422	TA11373_34305	GGTTCCTTCAGCCCTTTTTC	ATAGCCAAAGGGAATACGCC	269	F
164	Lj1812368	TA60_34305	CATCGACAACGTGAAGGCTA	AGGTGGAGGGTGGACTCTTT	141	F
165	Lj1812530	TA208_34305	CTGATACTGGTGGCCAGGTT	AAGGGCACTCGAAGAATGTG	211	F
166	Lj1812738	TA339_34305	GTCCTTGAGCTTTCTCGCAG	CATCCAGCATGTCCACAAAG	295	F
167	Lj1813057	TA550_34305	CGCTGTAAGGACAGCAACAG	TGAAACCCCATTTCCTGCTA	106	F
168	Lj1813457	TA931_34305	TGAATCAATCATGTTTGGGC	GTTCTTCATCGCTGCCTTTC	130	F
169	Lj1813598	TA87_34305	TGTGAAGACCCTTACTGGCA	CTCTGTTGGTCTGGTGGGAT	114	F
170	Lj1813711	TA375_34305	CCAATGCCTCGAAAACCTAA	ATGGCATGGGTTACTTCAGC	285	F
171	Lj1815087	TA13334_34305	AATGAGGCCACCTTCACAAC	AATGAACGGAGAAGTGGTGG	246	F
172	Lj1815296	TA13063_34305	CAAAACGAGGAGGCCATAGA	AGGGTTCGTAGAACGAAGGG	139	F
173	Lj1816457	TA2125_34305	TGACATTCTTCCTCCAGCCT	CATGGTGTGGAGCTGAGAGA	116	F
174	Lj1816906	TA2779_34305	CAAGGCCAGATGAGAGATGG	CAATAAGCAGGCATCTTGGC	133	F
175	Lj1817267	TA2865_34305	TCACTAGAGCTTGGTGGGAAA	TGACCATTTGTCCAGAAGCA	104	F

176	Lj1817518	TA2018_34305	GCATCCTCTCCCACACATTT	TCTCAGGAGGTCTGTCAGCA	162	F
177	Lj1817597	TA2552_34305	GATTTGCTTGGGTTTTGTGG	CTCCAACTGATGAGCTCCAA	115	F
178	Lj1823017	TA6981_34305	TTGCTTGCAATTTAGAAATGGTT	CACATTGCACAGAAGGATCA	141	F
179	Lj1823572	TA6515_34305	TCATGGCTCCAAAGAAGGAT	AACGCCATGTTGTTGACCTT	128	F
180	Lj1823627	TA6008_34305	TCCGCCTACAACCCTTAGAA	TTCCAGGATGCCTTCAATTC	214	F
181	Lj1826423	TA7106_34305	TTGAAAGGTGTAACTCCCCG	GCTTCGCCTCTGCTTTCTTA	163	F
182	Lj1828530	TA8794_34305	CAATCGTGATCGTTCTGGAC	TTGGTCCATCCAAACAACAA	165	F
183	TC92979	TC103219	ATTTGCTGAGGAATGCAACC	AGTAGGGGATGAATTTGGGG	226	G
184	TC91137	TC98483	CCATGATTCCAGATGCTCCT	AATCCTTCCACCAACTGCAC	152	G
185	TC87946	TC108386	GTGCGTCTTTGGAAACCAAT	ACCTTTGTTGGTCACCTTGC	323	G
186	TC86390	TC107316	AGACGAAAACAGGAGGCTCA	TTTGGAGCCATAGTGAAGGC	183	G
187	TC89690	TC96823	AAATGGGCTTCAAGTGTTCG	TGGTGAAGGTGATGGTGAAA	297	G
188	TC77131	TC94416	GCTGCTTTTATGCTTGCTCC	CTGGAGGGCAATGTTTGTTT	202	G
189	TC78634	TC108137	AGCTGAAGCTACCGCCAATA	CTTCCCTAAAGGCACCATCA	331	G
190	TC89442	TC102108	CGGTGGAGTGGTTGAAAAGT	ACCAGTACCGATCCAAGCAC	167	G
191	TC76798	TC100486	GGTGACCACCGTGCTAAGAT	GTTGGGTCGAACCAGAGGTA	314	G
192	TC88183	TC95978	CTGTATGCTCCAGGCTCACA	TCCTTACACGATGCAGCAAG	160	G
193	TC83271	TC105147	TGCTCAAGTGTGCTACTGGG	CGAGTTGATGGTCTGTCCCT	196	G
194	TC79417	TC108677	ACATCAGGGCCATGTTCTTC	TCCAACCCAATTTCCATTGT	231	G
195	TC84246	TC111882	GAAGCATTGGCAAAAGAAGC	CCATGTCCATCAAAAATCCC	200	G
196	TC87949	TC95999	ATGGAGTTATGCGAAGGTGG	CATGTCGCGATGTACCAAAC	150	G
197	TC87314	TC101814	TTCATGGCTGCTACACTTGC	CATATGAAGGCGTTCGGTTT	193	G
198	TC81224	TC103907	CCGGGGAAGTTGTAGCATTA	GCCAAGCCAAAATCACAAAT	373	G
199	TC77943	TC101250	TTAGCCAACCGTAACGAACC	CGCTGATTTCTTAGCGGAAC	348	G
200	TC88902	TC109408	TGAGTTGATCACTGGCCAAA	TTTCAATTGCATTCGGTGAA	200	G
201	TC86143	TC107102	TCGTGGAATCTCTGAAACCC	TGAACCTGACCCAAACACAA	157	G
202	TC90918	TC98305	TTCTCTTCCCCAATGGACAG	TCTCTGATACCCATTTGCCC	277	G
203	TC81463	TC103804	CAGCAGCAGTGTCACCATCT	GCAACTTTGAGTGGTCCCAT	370	G
204	TC89082	TC96726	GCACCACCAATATTCATCCC	CTGTCCTTGGTGGATCGTTT	358	G
205	TC90495	TC110561	CGGTCACGAGGATAGTGGTT	ACCACCTGAGGTCTTCCTCC	157	G
206	TC80869	TC103320	AATGAACTTCGCTTGGGATG	CCACCATTTGGGTTGTTAGG	246	G
207	TC91644	TC103483	AGAGATGGATCATCATCGGC	TTGCAGCATTGAGTTTCGAC	205	G
208	TC89212	TC109615	AGGCCTGAGACCTACAAGCA	TCCCTGCAAAACCCACTTAC	162	G
209	TC92972	TC112335	TTCCGGTGGAGAAAGTATGG	TAGTTAAGCAGGCACACCCC	389	G
210	TC79935	TC102662	TTGGCTGGGTACATTTGACA	AATCCCACGTCGAACAAGTC	364	G

211	TC86255	TC100830	TTTCATTACTCCGGTACGGC	AACCGTCATACGAATCAGGC	256	G
212	TC109615	TC109615	AGGCCTGAGACCTACAAGCA	TCCCTGCAAAACCCACTTAC	162	G
213	TC86292	TC94874	GCCTTGAAGGACCAAATGAA	TGTGATTGTTGTGCATGGTG	298	G
214	TC78754	TC101694	ATGGCCGGGGATATACTTTC	CTGCTTTGGCTAGGGCTATG	364	G
215	TC81471	TC103194	AGCAACAACCTGCAACTCCT	GAAGCAGTGGTGCAACGATA	301	G
216	TC82573	TC98428	ATCCACCACCAACAACCATT	TACAGCAGAAGCACCACCAC	323	G
217	TC76339	TC106364	AACAACAACAATGCAACCGA	AGGGAAGTTGGTTTTGGCTT	198	G
218	TC83285	TC111411	TTACGAAGAGGGCGTGAACT	GGAAAATCGGTGTGCAAGAT	371	G
219	TC77877	TC107549	ATGAATCACTGGGGAACAGG	TGATGGCATGGTGGAACTAA	362	G
220	TC88758	TC109362	ACGAGGCAACACATCACAAA	GCCACGAGAGAAAATGGAAG	225	G
221	TC79053	TC96103	TCGCGACCAATTCATTATCA	TGGTACTTCCGTTACGAGGG	199	G
222	TC79268	TC108512	TCATCAACCTTTTCCAAGCC	TCCCTCTCTGGCTCAGATGT	309	G
223	TC77213	TC94630	GACTCTCCAGCACCGTCTTC	CTCAACGGCATCATCAGAGA	262	G
224	TC80232	TC102808	CTTTTTCAGGCAGGCTTTTG	CCACGGTTCCTCAAATCCTA	197	G
225	TC78257	TC107542	GAGGCTGAGGTGAAACCAAG	TCAAGACCAAAAGTCCCACC	243	G
226	TC93616	TC102711	CAACAAATGGGATTTTTCCG	TCCTCACAACCGCATTAACA	221	G
227	TC95045	TC95045	GTGCACTCCATCACCCTTCT	TTGTTGTTGCTTGTTGCTCC	313	G
228	TC76474	TC106458	CAACCACTCCGATTGTTCCT	TTGAGTTGATTTTGCACCCA	242	G
229	TC78648	TC108324	GACGGCATCGAAATCTGAAT	TGTCCATCCAAATCCCTCTC	285	G
230	TC77169	TC100735	GAGAGTTTGGAGGAACTGCG	TTGCAAGAACACCAAACAGC	250	G
231	TC90930	TC104541	AGGTGGATTATTGGGAAGCC	GATCGTTGTAGTGGCCAGGT	266	G
232	TC77955	TC94970	CAATTGTTTGGGGACTTGCT	TGGTCCTGCAAACATCACAT	386	G
233	TC87800	TC95616	GGCATTGGCACCTATCATCT	TGGCTGCTGTAGAAGAAGCA	167	G
234	TC77346	TC107122	AATCAGAGACCATTTTGGCG	AGGGATTGGCCCAGTTACTT	235	G
235	TC96859	TC96859	GTCATGCCCTAGACGGTTGT	CTTTTCCTTCTGTTCCTGCG	398	G
236	TC96130	TC96130	TGAACTCCCTCCTGGCTTTA	TGCCTTGTCTGTGCCAGTAG	274	G
237	BE324413	NF016F06PL1F1048	AAAACTGCCGACGTTACTGG	GAGGCTGGAACTCTCACTGC	194	G
238	BG582680	TC108268	ATCATGCAAGAAACCAAGGC	TCCCTTGTCACCTTTTGTCC	236	G
239	BI262275	Pc_009_02367_Mar06	TCACATGCTCCTTGTTCTGC	GTGTCCCTGATGCCAAGATT	374	G
240	AW686309	NF036D10NR1F1000	AAAATTCCATTGCTACCCCC	TGTGTGCATTTTGGTGTCCT	198	G
241	AW980681	EST391834	TTCCAAGGTGAGTTTTTGGG	TAGTGAGGGAGGTTGTTGGG	276	G
242	TC106447	TC106447	ACCGAGCCAGTCGCTTAGTA	GAGCTCCTGTGCTTTATCGG	227	G
243	TC94373	TC94373	TGGGTTGATTGGGATGATTT	GGTCTGTGCTGCAAGATTCA	150	G
244	TC101888	TC101888	GCTGAAATGGTTCTGGCATT	CCGGAATCTGTTGAAGGAAA	293	G
245	TC108267	TC108267	GGACAAAAGGTGACAAGGGA	GGTGCAGAAGAAGAAGGTGC	227	G

246	TC94589	TC94589	TGCTAATGCGAAGCAATCAC	CTGCTACTTGAGCCACTCCC	399	G
247	TC100851	TC100851	TGCTCATTTCAATCCTGCTG	GTGGCAACTCCAGAAATGGT	249	G
248	TC94653	TC94653	GTTGCAGGTTACGTTGGGTT	AAGCATGTGAATCCACACCA	306	G
249	TC100817	TC100817	GTGTTCGAAACATCATTGCG	TTTTCACTGTCTCCCCAACC	218	G
250	TC89224	CK149070	CGCAGAGCACCACTTTACAA	TAATTGCTCCTTCCAACGCT	380	G
251	TC91637	CD051336	AACTCCTGCAAAGCGTCACT	CCTCTTCGCTATTACGCCAG	298	G
252	TC77624	DY475357	CACCACAAGAAATGAAAAGGAA	GGCTCAAGGAAAACGTTGAT	154	G
253	TC81572	CK148896	CTCTCTCTGCTGCTTCCACA	CGGCGTTCTTCTTCTAATGC	150	G
254	TC77707	AGT-5-B08-D9	TTCTGTTCTTCCACCCCAAC	CGAGCAAGAATCGAACACAA	368	G
255	TC81109	EST0936	TCCAAACTCTGCCAATTTCA	CGAGGTACTTGAGGGTGGTC	219	G
256	TC77515	CA0877	GACAACTGTGCAGGGATTGA	CAGCTGCTCAGAAACCATGA	218	G
257	TC86205	Ca_002_00130_Nov06	AGGAAAAATTGGGAGGGAGA	ACCATGATTACGCCAAGCTC	157	G
258	BF650547	TC111370	TCAGTGGGTTCAGAAGGTCC	TGCTGCTTTGTGTGGATCTC	156	G
259	TC85206	TC100248	AATGGCTTCGATTCAACGTC	CAATTCCCTAGTCTGCGCTC	286	G
260	TC81664	TC102925	AATCAACAACAGCAGCAGCA	CCCACTTTCCAAAATTTCCA	159	G
261	TC83451	TC105481	GAACTCAGTGCCCGGAAATA	AAGCTCAGACAGCAGCAACA	170	G
262	TC77133	TC94417	ACATTTTGCTGTTGATGCCA	ATATTGATATGCCTTGCCGC	213	G
263	TC77134	TC94419	ATCAATCGCTGCTCGAAGAT	TCACAAAACCACCCAAACAA	239	G
264	TC80526	TC97208	TTGCTAACAACCCCAACACA	TGCAGTATCTTTGCAGCACC	340	G
265	TC77457	TC94826	CAACCAGCAGAAGAAGGAGG	ATGCACCACCTCCTTGAAAC	331	G
266	TC91210	TC97820	CCTCCTTCTTCCCCAAACTC	CTGCACAGGCTTTTTCATCA	246	G
267	TC80715	TC103537	TGGGAGGTTTGTGAGTGTGA	AGACGTGAATCTCCCAATCG	210	G
268	TC76964	TC94372	TTTGGGCTTGGATCTGAAAC	TCATCCCAATCAACCCATTT	347	G
269	TC84328	TC112152	TTCTCGTCATTCCAGCCTCT	TTCGGGGACAAAAATACAGC	167	G
270	TC87769	TC101930	CACTCAGCGTTGATTGAGGA	ATGGCCAAGACACCAAGAAC	365	G
271	TC79044	TC101857	CTGATTTTGGGTTTGGCTGT	TCTGAAATTTTAGGCCACGG	257	G
272	TC86177	TC94746	CTGAAGGCAAGAAACCAAGC	GAACCGTTCACAGTTTCGGT	238	G
273	TC87719	TC95946	GGTTGCAGTTAAACGCCATT	CATCGTTCGGAAACAACCTT	386	G
274	TC88481	TC108586	GGATGCGGAGATAGATTGGA	CCAATACTTTGCTGTCCGGT	283	G
275	TC85561	TC94311	GTCGCAAATGCCTAATTGGT	TTGATCTGCCATTGTTGCAT	188	G
276	TC77244	TC100193	AGGGCACTCTTCTGTGAGGA	AAGTCATCAACACCCCAAGG	233	G
277	TC79908	TC96614	GCTTGGAGCAGTGGAGTTTC	ATTCTCCATCATGTTCCCCA	362	G
278	TC81755	TC104316	GGTCAAAAGAATCCTGGCAA	AGATGGTGACACTGCTGCTG	331	G
279	TC88643	TC96868	GTGGGAGATAGGGGTTCCAT	AAACATGCTGCGAAACACAG	274	G
280	TC85891	TC94572	TGGATCCAGAAATCTACGGC	TTTTCCACGGAATCTCAAGG	206	G

281	TC91874	TC98794	GCAATTCCAAGAGCAAGAGG	TCGAAACTACCCATGAAGCC	268	G
282	TC80369	TC103195	CGGTTTGTTTTCAACGGTCT	GGATCCAGGGAGTTGTTTCA	170	G
283	TC78889	TC95992	GGCCTTTGCATTGTGATCTT	AAGGAAATAGCTTCGTCGCA	295	G
284	TC88396	TC102752	CTGAAATCAAATCCCTCCGA	ATGAAGAACCGTGTATCCGC	293	G
285	TC90657	TC109632	TGATCCTGCACAAACTGCTC	AGCCACCTTCTCCAGTGCTA	223	G
286	TC89380	TC96964	CCCTTCCTTCCTAGCAGCTT	TTGGATGATTTGGAAAAGCC	158	G
287	TC87284	TC101655	TGGCAATGGCAAATACTGAA	TTGTCACTTCTGGTGGTGGA	336	G
288	TC77687	TC100985	GCTTCACCGGAATCATTTGT	CAACCCATCCAATATCAGGG	340	G
289	TC89819	TC109710	GTTCCATTATATGGCCGTGG	GTTGAATTGCCGAGGATGTT	162	G
290	TC91696	TC105034	GTGCGGACTTGCTTTTTGAT	GCAAATCTTGGTTCTCTCGG	160	G
291	TC79937	TC95744	GAACTTCGTCGTGGAGAAGC	GTGTGAATCCAGCCATTGTG	397	G
292	TC81674	TC97423	TCAACACCATGTTCCCTTCA	TCGCTAACGTCATCACCAAG	182	G
293	TC77204	TC100770	TCATGGGCTTTTTCTCCATC	TCCATTAAACAAACGGCACA	208	G
294	TC88476	TC108669	CATGTTTCTCCTTCCCAACAA	GGTCAGTAGGGGCAAAGTCA	303	G
295	TC77278	TC94707	AATCGTGCAAGAAACCAAGG	TCTTTTTGACAGGGCAGCTT	204	G
296	TC86176	TC94757	CGAAACCACAAACACCTCCT	CTGATCCAGGGTTTTTGGAA	178	G
297	TC84899	TC111883	CAATGGAAGAGTCCTGGGAA	TCCCAACAAACTCAACACCA	186	G
298	TC82040	TC110666	CAACAAAAACTGGTGAAACCC	ATCTGGATGTTGCATGCGTA	192	G
299	TC86533	TC95045	GTGCACTCCATCACCCTTCT	TTGTTGTTGCTTGTTGCTCC	313	G
300	TC88498	TC108500	GGGCACGGTTGACCTACTTA	TCACCCATCGAAGTTCATCA	273	G
301	TC78267	TC101636	GGTCTTTCAAGGTGTGGGAA	AAGGGCCATTGTTGTTTGAG	355	G
302	TC76467	TC106447	ACCGAGCCAGTCGCTTAGTA	GAGCTCCTGTGCTTTATCGG	227	G
303	TC78268	TC101634	CCAGATACGGTTGCTGGAAT	GGCCATTGGTGTCTGAGATT	369	G
304	TC79699	TC102905	ACAATGGCTGTGCCTCTTCT	CTGGCTTCTTAACCTGGCTG	393	G
305	TC88206	TC102516	CTGCTGATCTCTTCCTTGGG	GCGGAACGATAAATCCTGAA	378	G
306	TC92821	TC99519	TTATTGCTGGGCCTTTCAAC	GAGCAAACATCCAACAGCAA	161	G
307	TC79250	TC96191	TGCTGATTCGTGTCCATTGT	AGGCTCGAACTTCTCATCCA	164	G
308	TC79250	TC96191	TGCTGATTCGTGTCCATTGT	AGGCTCGAACTTCTCATCCA	164	G
309	TC86162	TC107178	CAAGTGAGATGCCAAGCAAA	AAGAGTGAAGCCCAAGCAAA	221	G
310	TC95605	TC95605	CATGCAAGTTCGTATGGTGG	GAGGCATTGCAGAATTCCAT	191	G
311	TC103469	TC103469	AGCGAACTGGCAAGAGTTGT	CATTATTAGTGGCGGTGGCT	343	G
312	BF644787	NF016A03EC1F1020	CATGGAGGAAGAGGTTTGGA	TGCAATTGTTGACCACCTGT	262	G
313	BG451955	TC101529	ACGAAAAGGACCATGGACAG	TCTTCTCCTGCCTCTTTGGA	386	G
314	BG586643	TC98397	ACAGGCTAACGCCACTGTCT	TTTCCAGTCCATAACAGCCC	177	G
315	AW586039	EST317662	TTGATGGAATTGTTGGGGTT	CATGTCGAAAGAGCCTCACA	158	G

316	AW686806	TC94555	AGCCTCTGAAAATCCAAGCA	GGTCGAAATTGATCGGAAGA	295	G
317	BE205257	EST397933	CAGCAATTCTGGAAGCAACA	AACCAAACGTTGAACCCTCA	332	G
318	TC106458	TC106458	CAACCACTCCGATTGTTCCT	TTGAGTTGATTTTGCACCCA	242	G
319	TC94377	TC94377	CAGAGCATATCGATTCGGGT	TTCGGGGACAAAAATACAGC	352	G
320	TC101889	TC101889	TTTCCTTCCAACAGATTCCG	TAGTCCCATTTAACGGCGTC	203	G
321	TC108268	TC108268	GGACAAAAGGTGACAAGGGA	GGTGCAGAAGAAGAAGGTGC	227	G
322	TC94590	TC94590	TGCTAATGCGAAGCAATCAC	TGCTACTTGAGCCACTCCCT	399	G
323	TC94651	TC94651	AAAAATTGCGGTGTCTGGTC	AAGCATGTGAATCCACACCA	336	G
324	TC100816	TC100816	AGAATACCACGGTTTCGTGC	AAGCCTGTCCAACGCTCTAA	151	G
325	TC101057	TC101057	CTTGGTGGACACAAACGATG	TTGATTCCGGCGAGATTAAC	236	G
326	TC78447	CD051312	GGTGAAGAGATGCGGGAATA	TTTATGACCCTTCCACCAGC	288	G
327	TC85165	CK148902	AAGGACCTGTTTGGCTACGA	TAACGTGTTTGGCGAGTTTG	274	G
328	TC85165	Ca_002_00086_Nov06	AAGGACCTGTTTGGCTACGA	TAACGTGTTTGGCGAGTTTG	274	G
329	TC103928	CK148909	GCGGACGATGGTAAAAGTGT	CATCAGCTTCCTCTTCCTCG	268	G
330	TC86258	AGT-7-C10-G12	GCACAGGACAATTGCCTACA	GCATAGGAAGGGGAATCACA	161	G
331	TC86332	EST902	CTGAAGAAGCTGCTTTGGCT	GCTTCCCATTGTGGTTCAGT	386	G
332	TC85414	23_A1	GTCAGGGTGTGCTTTTGGTT	ATCACTTGGGTCAAGATCGG	188	G
333	TC76606	Ca_002_00144_Nov06	AAATGTCGGACGAGGAACAC	ACCATGCTTCCCAGTTTTTG	164	G
334	TC95048	EST0307	ACCTTGGCAGCTCCAAACTA	GATTTCTGGGAACGAGTCCA	167	G
335	BQ137035	Ca_004_00187_Nov06	GCTGCCCTCAGTGAAGAAAC	TGCAGGTTTAAGAATTCGCC	169	G
336	TC83550	TC111370	TCAAGGTTGTAAAGCACGCA	CCATGGCATGCAACTATCTC	315	G
337	ABE92614(P)	DUF862	CAAGTCACTTCCGTGAAGCA	ACATCCTCAGCCCTAGCAGA		G
338	ABE83065(P)	N211178e	ACGGAAGCCGACATGTTTAC	AGGAATGTCGCAGCTGTCTT	255	G
339	Mt101026	TC99893	ACCACCTGGTGTTCCAGAAA	TTCATGTTTGCCTCAGGTGT	216	G
340	Mt106087	TC100816	CAACTGTCTTTGAGCTTGCG	CAGAATTCATCGGTGCCATA	250	G
341	Mt106141	TC100603	GAGACTGCTGTTAGGGACCG	CCCTTGAGCCAGCAAGATAC	256	G
342	Mt106193	TC100565	GATTCGATGCATGTGTTTGG	AACAGGCCATTTCAGGACAC	199	G
343	Mt106212	TC100459	TGGGATCCCTTCAAGGATTT	AGGCAATCTGAACCTCCTCA	291	G
344	Mt106421	TC100715	CCAGCTGAGCTTGAGAAGAGA	AGTTACCGCACACCACAACA	145	G
345	Mt106570	TC100662	GCATCAAGAATCATTGGGGT	CAACCATCGTGGACACATTC	212	G
346	Mt106613	TC100630	GAGAAAAGCTCGTGTCTCCG	GCTGGTGGTAGTGGATGGTT	237	G
347	Mt106628	TC100810	TGTCCTGGTTCTTCACTTGCT	TTGCACTTCAATGGTTTAGCC	158	G
348	Mt106886	TC100499	CCAGACATTGCACCATTTGA	CAATCCCATCAGAAACTGACA	133	G
349	Mt107012	TC100669	TGCGTAACAGCTCACTGGTC	ATCCTTGGTCTTTCGGTGTG	216	G
350	Mt107464	TC100209	ACTTCACTTGGTGCTCCGTC	CTTGAATCTTGGCCTTGACA	125	G

351	Mt107563	TC100666	ATGTGAATTACTCCGACCCG	TTTTCAACTTGCCATTGCTC	246	G
352	Mt107694	TC100127	AAATAATTCCAGAGGATTGGTGA	AGCAAATAAATAGAATAGCTCGACA	100	G
353	Mt107714	TC100813	GCGTGTGCACTTTGTAAGGA	GTTGGACATTTCCTCACGCT	172	G
354	Mt107786	TC100213	TTCTCATGGGTGGTGTTTCA	CTCATCGGCCATTGTTCTTT	164	G
355	Mt112050	TC101243	TCAAAGCTAAGCCACCCAAC	GGGTCGGCAGAATTATCAAC	285	G
356	Mt112737	TC101540	GCTGCTGTGAAGATGGTGAA	TTGATTTGAATTTGGCACGA	289	G
357	Mt112969	TC100999	ACGAAAAGTCGTGAACCAGC	CCCCATTCACCACTCCTAGA	108	G
358	Mt113149	TC100967	CTTCAGATCATCAATGCCGA	AAGCAGGAGTAGTGCCCGAG	100	G
359	Mt113180	TC101198	AAGCAACCGATCCAACTGAG	ATGGTGCCAAAATCCAACAT	234	G
360	Mt113537	TC101049	AATCTTCGCTTGCTCGTGAT	GGCACGGGACCACTACTATG	127	G
361	Mt113634	TC100926	CGAGGGTTTGGAGTCTCTTG	TGAACATATCTCCGCCAACA	167	G
362	Mt113657	TC101388	CGAGTTTCTGGTGTTTGTCG	AGGTCAAGAATCTGCAGGTCA	224	G
363	Mt113698	TC101432	TGTGGCATGCTTTATGTGGT	TTGGCCAAATCAACTCTTCC	298	G
364	Mt113732	TC101413	CTGAAGGGAAAAGCATTGGA	CTATGTACCGCCTCTCTGCC	235	G
365	Mt114205	TC102166	TTTTTGCTCCGAAGTTGCTT	GCCAAAGATCCCTCAAAGGT	293	G
366	Mt115084	TC101970	AACCTTGGAGCAAGGAAAGC	GGACCAACCTTTCGGGTAAT	225	G
367	Mt115640	TC101920	GAGAAGACTGGTGGTCCTGC	CATCTAGCTCGCGCCTAAAC	217	G
368	Mt115771	TC102147	TCGCGCTTTATTCACACTTG	ATGCGCTCTCCTTGTTTTGT	176	G
369	Mt116561	TC103225	AAGCCTGGTGATTTTGTTGC	CAGCCCCCTTCTTACCTTTC	203	G
370	Mt117159	TC103761	GCGAGAGAGAAATTTACCTGG	GGATGTAGATCAACAAACATCTCC	104	G
371	Mt119073	TC104017	TGATGGATCTGTTTGGTGCT	AAGGTTGAAGAATCTGGGGG	162	G
372	Mt121172	TC105748	GCTTGACAGCAGTCCCATTT	GTGCTCTTGGAGTGTGACGA	207	G
373	Mt121549	TC105410	AGAAGAGGCTGTCCTTGCAG	CCCGATCAAGAGTTTTTCTTTC	101	G
374	Mt122051	TC106339	TATTTGGGAAGAGGGCTTGA	ATGTCTCTCCATCCCACCAG	130	G
375	Mt122060	TC106830	TGGTGCAGAAGCAAAAGTGA	CCAACAACCACAACTGGATTC	240	G
376	Mt122067	TC106556	TTTGTTCAAGGTTGTGACGC	TTCAGCAGCAAGAGCAAGAA	186	G
377	Mt122134	TC106835	TTAAGGTGAAGGGGTTGTGG	ATCTGACTGACGCGCTCTTT	139	G
378	Mt122467	TC106365	AACCACCGGTCTACAAACCA	GCTTCTCAACTGGTGGCTTA	183	G
379	Mt122523	TC106729	ATGAGATATGGGCTGCTCGT	TTTGTGTGAGCACCAGCTCC	118	G
380	Mt122547	TC106426	GATCAGCGGATGTTGCTTTT	AAAGGCACAGATCCCAACAG	193	G
381	Mt122619	TC106445	CACCACCACCAAAGCCTTAC	TATTTATATGGCGGCTTCGG	211	G
382	Mt122761	TC106245	TGGGCTACTTACAACGAGCA	TTCTGGTCCTTGTTCTCCCTT	107	G
383	Mt122930	TC106630	TGGAAGCTTTTAATTGCGCT	CATTTGCAAACCTTGACCCT	250	G
384	Mt123109	TC106324	TGTGAAAGTCAACCAATGGA	TCCAGAAATGAGAAGTTCCTGA	132	G
385	Mt123162	TC106871	TCTGGCAGAAGCTTTCACAA	AGCCAACAAAACAAGAGGGA	265	G

386	Mt123213	TC106567	GAGGTTTAGATGTTGTGAATCGAA	CAGAGGAAATTTCAGCAGCA	105	G
387	Mt123231	TC106515	AACTTGGGATCAAAGCATCC	TGACATGTGTGTACCTTGCC	181	G
388	Mt123274	TC106540	CCTCAATTTTGGCAGATGGT	CCGAATGTATGAGCACCTGA	193	G
389	Mt123479	TC106350	AGAGGGGATACGGCAAAGTT	GTTTCCAGCATCACCACCTT	239	G
390	Mt123620	TC106726	CCCGAGGTATACCCTCTGTTT	CTGCTTGTTCACCCTGACTTC	108	G
391	Mt123632	TC106558	GTGAAAAAGGGTTTGCTCCA	TCTGTGGAAGCCACATTGAC	242	G
392	Mt123772	TC106564	CGAATCAAAACAGCAGTGGA	GGTTGAAAGCTGGACTTGGA	190	G
393	Mt123991	TC107253	ATGAAGTCGCTGGAAAATGG	CCAAGCTGAAGCAACAATCA	191	G
394	Mt124079	TC106965	CCTAATGAAGAGTCTCGCCG	CAGCCTTAAGCTCAGCACCT	127	G
395	Mt124227	TC106943	ATGATGAGAAAGAGCGCGAG	TTCAACAATTGCAACCCTGT	181	G
396	Mt124229	TC107101	TGTGGGAGAAGAAGATTGGG	CATAAGCCTCAACCCCAAAA	183	G
397	Mt124331	TC107674	GGGTTCTCTTATGGCAGGGT	GGGCATGGTCATTGACTTCT	211	G
398	Mt124817	TC107439	GAGGCCATGTTTGCTGATTT	CTTGCCCAATATATGCCGTC	103	G
399	Mt124917	TC107524	AAGGTTCCCAAAACAAAAACC	CCTAATGTCGAGCAAAAGAACA	236	G
400	Mt124935	TC107534	CTTTAACCTCGGGACCTCAA	TAAAAAGCCCACTTCCACCA	216	G
401	Mt124972	TC107055	TTGTCAGCATTCCATGTGGT	ATTTGGCAATGGTTTCTGGA	300	G
402	Mt125035	TC107007	GGGAGTTTGCATTGAAGCAT	CTTTTTGGCCTTTCAAGTGG	271	G
403	Mt125058	TC107685	CATTCCCATGAATGAGTTTCC	CAACCTAACATCAGGTCGCA	163	G
404	Mt125196	TC107836	GGGATGATGAAGCTATCCGA	TAAATGGCTGGTTGTTGCTG	133	G
405	Mt125375	TC107117	TATTGCTGCTGCACTGAAGG	TGTTCTCCCTTGAATGGTCC	158	G
406	Mt125500	TC106935	GGAGGTGCAAAGAGGCATAG	CCATAGAGGGTTCTTCCCTG	260	G
407	Mt126176	TC108243	TCCTAGAGGCGATTAAGGCA	TTCCTTTGCTTTGGTCTTGG	110	G
408	Mt127625	TC108849	GTTGGGTATCCACACAGCGT	GAGAGTCTCAGCACGAGGTTG	250	G
409	Mt127626	TC108242	AAGGGATGAGGATGTTCGTG	TCTGCGCTTTCTTTCTCCTC	154	G
410	Mt127721	TC108147	TCGTAACTTCGGTGTTGGTG	TCAGGAACAATCTTCCCAGC	216	G
411	Mt127793	TC108308	TTTTGAAGACCTGCCTGAGC	TGGAAATCCTGGTGGTACAAT	121	G
412	Mt127898	TC108101	CGGCGATGAGCTTCTTTTAC	ATTGAGAGCATGTCCCTTGG	112	G
413	Mt128366	TC109724	GAAACAGATTAGCGCAAGGC	GTTCCCGTTGCTTTTGTTTC	272	G
414	Mt129135	TC109739	TCAAACCGGTGGAAAACACT	ATTTCAACGGAACACCTTCG	284	G
415	Mt129339	TC109043	TCCGTGAGATTGGAAGACAA	CCAGCCAAGACCAAGGATAA	293	G
416	Mt129970	TC110297	GTTTTGCGAGATGGAAAGGA	GGAGTGTAACCTTCAGGCGA	170	G
417	Mt131413	TC110526	CAAGTCTGGTGACTCTGCCA	GTCGCTTATTTCTTGCCAGC	210	G
418	Mt132016	TC110919	ATGGGAAACACTCCCAAACA	AAGCTTGGAATCACGTCCAC	109	G
419	Mt133126	TC111196	GCCGAGATTGGTGAGTTTGT	TGGAGGAACACTTGACAACA	240	G
420	Mt6799803	TA31726_3880	AAGGCAAGCCTAAGCTGTTG	TGATCCATTTGGCATTAGCA	290	G

421	Mt6800312	TA31948_3880	ATATGGCGGCTCAAGACAAT	AACCTTTCCGGTAGGCATCT	112	G
422	Mt6800685	TA31536_3880	TCGACTCCTCAAAATGGGAC	GATGATATCTTCTGGCGGGA	263	G
423	Mt6802114	TA25011_3880	TCGCGCTTTATTCACACTTG	CATGCTCCCCAACTCTTGTT	101	G
424	Mt6802223	TA25022_3880	GGTGGAAGGGAAACAAATGA	GCTTTCCTTGCTCCAAGGTT	166	G
425	Mt6803180	TA24645_3880	GGTCATGAAAATTGATGGTGAA	TTTGGAAGCTCCACAAGAAAA	192	G
426	Mt6806704	TA29438_3880	GCTGATGTTGCACTTTGTCC	CCCACAACACTCACACCAAG	281	G
427	Mt6808376	TA29285_3880	TCTGGGAACTGAAAGATCCG	TTGGCTGGAAGCCTTCTTTA	213	G
428	Mt6808874	TA20602_3880	CCTTATGCTCTTGACCAGGC	GTTATGGGGAGTATCTCCTAGATG	271	G
429	Mt6808925	TA20775_3880	TCCGCTAACATGGTTCTTCA	TATTTGCCGAATGTTCTCCC	167	G
430	Mt6811198	TA26736_3880	CCAAAGAACCACCTCGTCAT	TCTTCAACGATGGGAATAGGA	151	G
431	Mt6812211	TA27437_3880	GGAAAGAGGAACAAGCTCGTAA	ACTTCTTTCTCCGAGCAGCC	183	G
432	Mt6814147	TA35233_3880	CAGCGGTGGGGAGATATACA	GGTTTCCAGTTCTCTGCTGC	134	G
433	Mt6815261	TA31068_3880	TCGTGCGAAAATGGACTATG	CAGAGCGCGTTTCTCTTCTT	142	G
434	Mt6815341	TA30776_3880	CGATGGACGAGGAAATTGTT	TTTCCCTTCTTTCTGTCCCC	217	G
435	Mt6816580	TA22862_3880	TGACTCTCCCCATCTTCCTG	AACCATAGCTCCTGAAACGG	245	G
436	Mt6817377	TA22826_3880	TGACGTGAGGTGTGGTGTTT	TGTATGACCTCTCCCTTGGC	235	G
437	Mt6817477	TA23280_3880	CCCGATGAAGAGGTTGACAT	TTCCTCGATCTCTTTCAGCC	129	G
438	Mt6817719	TA22622_3880	CTGCAAATAGGCCAGCAGAT	CCGATGTATCCAGCCCATAC	195	G
439	Mt6817754	TA22828_3880	GACGTGAGGTGTGGTGTTTG	TGGCCACCAATATGAGAACA	116	G
440	Mt6818065	TA22833_3880	TGGGTGGCTTAGCTTTGAAC	AAAATTGCCCCCTCCATTAC	222	G
441	Mt6818511	TA25682_3880	AACCGTCGCTAGAGTTGCAG	TCCTTGGTTTCAATTACACGG	188	G
442	Mt6819871	TA25927_3880	AAAGTTCGTATGAATGGGCG	TGAGCTGCTGCTCTTGTGAT	273	G
443	Mt6819924	TA25787_3880	AGGCACTGGAGCCTACACAT	TCCTTACACGATGCAGCAAG	297	G
444	Mt6821544	TA19088_3880	CTTTAGCGCCATCCATTTTC	CCTGTTCCACGAAGTGCATA	297	G
445	Mt6822333	TA18634_3880	GAGAAACCATCCAAAGCATCA	GCCAATGCAGTAGTGCAACA	273	G
446	Mt6823024	TA32864_3880	AACGCGAACTGGATGAAATC	TTAGCGTTGGTTTACGTCCC	297	G
447	Mt6823880	TA33468_3880	CCTACAAAACGTTTGCCGAT	ACAGTTTTCACTTCCGGGTG	103	G
448	Mt6826338	TA36562_3880	GGCGACGATTGTCTGTTTCT	CGCTAACGTGTTCACAAACG	280	G
449	Mt6826498	TA20246_3880	CTGACAAGGGTGGTGATCCT	GAAATCCACCTCCACCAAAG	189	G
450	Mt6829188	TA28510_3880	GAAAGGGCCAGTTTTTGTCA	TTGCTTTTGCCCCATAGAAC	159	G
451	Mt6831056	TA22311_3880	TTCATCTTTGGCACAGCATC	GGTTGTTGTAATAGTCGATTCCTTC	292	G
452	Mt6831737	TA22493_3880	TGCTGGAACAACCCGAGTAT	GCAACATTCATTGGCATCTC	252	G
453	Mt6831918	TA22310_3880	GGCGGCTTCTTAAGTCCTCT	CCGGTGCAATTTGGATTAAA	191	G
454	Mt6832813	TA35895_3880	TTAAAAACAATGAAGCCCGC	CCTTAGGACCCGACTAACCC	134	G
455	Mt6836309	TA33702_3880	TGGTCACCAAAGGGTGATTT	TTATCCGGCCTACAAAATCG	144	G

456	Mt6836854	TA23880_3880	GAGAATCTGGGTCTCCTCCC	ACAGTGAACTCGACCCCATC	238	G
457	Mt6837948	TA23825_3880	GGGACCATTGGCGTAATCTA	CCATAATACCGTTTTCCCGA	232	G
458	Mt6838070	TA24460_3880	CACGTAACGGGGAGCTGTAT	TGTTCGATGATCTCAGCAGG	160	G
459	Mt6839703	TA37603_3880	CGGTGCTGAGGTTGGAATAA	ACTTGTCCGTTCCCTCTCAA	105	G
460	Mt6839746	TA38438_3880	GAACAAGGGGTGGTGAGAGA	GTTCAACCTCGCTGCCTTTA	282	G
461	Mt69028	TC94103	ACTACAAATCTCCCCCACCC	TAGGAGATGGAGGAGGAGGG	222	G
462	Mt96345	TC98558	GAACAAGGGGTGGTGAGAGA	TTTCCTTCCATCCTCTTCCA	120	G
463	Mt97521	TC97977	CTCTGTCTTCGCATTCCCTC	CTTCAATCCACCCATCCAAT	169	G
464	Ms1739606	Ms_002_00442_Mar06	GTGTTTCATCTGTTGGGGCT	CAAATGATGGATGCAACAGC	108	H
465	Ms1739717	Ms_004_00775_Mar06	AGTTTTTCGTGGGTGACGAG	ACAGAACAAATTGATCCGAAA	195	H
466	Ms1739770	Ms_002_00405_Mar06	GGTGCAGGGCTTGTCTATGT	TTTGATAGGATCCGTTTCGG	160	H
467	Ms1740381	Ms_002_00313_Mar06	TCAATCTTAATGGTTACAAGTCAGG	AGCCACCAATGGTTTCATTC	207	H
468	Ms1740646	Ms_002_00197_Mar06	CACAATTCGGAGACGAATGA	TGGTGGCACTATATCCCTGA	153	H
469	Ms1741518	Ms_002_00481_Mar06	CAACGTTGATGGAAAGCAGA	CATCCCTCAATTTCATGTCG	275	H
470	Ms1741519	Ms_003_00735_Mar06	GCAGATCAATAAATTGAGGGCT	TGAGGACAAAACAACTTCCAA	226	H
471	Ms1742284	Ms_003_00740_Mar06	GCAAAGTCTTATCTTCGCCG	ATGGTGTCGGAGCTTTCAAC	177	H
472	Ms1742627	Ms_003_00640_Mar06	GCAAAGTCTTATCTTCGCCG	GAAGACGGAGGACAAGGTGA	101	H
473	Ms1742727	Ms_002_00221_Mar06	GCTGCTGTATTCCACAATGC	CCACTTGTCCATTCCCTCTC	163	H
474	Ms6943512	TA2250_3879	AGATGGACAAATCACCGAGC	TAAACACCGCTCTCATCCAA	123	H
475	Ms6944057	TA1845_3879	GGCTTTATGTAAGGGGCACA	ACTGTTACCATGAGGCCTGG	218	H
476	Ms6944246	TA1792_3879	TGCTGTAAAAGTGAGCTGCAA	ACAGAAAGCAACCTTGGCTC	107	H
477	Ms6944323	TA1598_3879	GCAAAGTCTTATCTTCGCCG	TTTCACGTTGTCGATGGTGT	190	H
478	Ms6944384	TA2161_3879	TTTGCAACGCAAATGGAATA	CCAGGGATCAAACGATTCTG	292	H
479	Mt100596	TC99129	GCAAGAACTTTCCAAACCCTT	TGCCTTGCTTCTTTTGGAAG	300	G
480	Gm103068	TC231846	GAAACCTGATCCAATGGTGG	AAGCCAGTGGGAATTTCCTT	234	I
481	Gm108977	TC233844	ATGGTATCGTCCGCTAAACG	CCGTGGCTCGACTAAAAGAG	148	I
482	Gm109991	TC234688	CTTCTCATTACCCGCAGGAG	CCAAGCGAACCCACTTAAAA	160	I
483	Gm2069450	TA69424_3847	CAAGTCTCCACCTCCACCAT	TCCACAACAGCACCTTTGAG	273	I
484	Gm2070282	TA70023_3847	CAGCAATTCCCCAATCCTAA	GGGGGCATAATCTCCTCCTA	277	I
485	Gm2072360	TA75158_3847	CACAATACCCCCAGAATTGG	AGTGAACTCCACCCCATCAG	262	I
486	Gm2074968	TA41222_3847	GCCCAACACTTTTGGAAACA	TCAACCTCCTCCATCTTGCT	144	I
487	Gm2075028	TA41043_3847	TTATAGCAATCAGTGGTTTTGGA	TCTGAACATGACAGTACCTGAATTT	101	I
488	Gm2076047	TA41057_3847	ACAGCTCAGGTTGCATCTCA	CCTTCTCCCCTGAACCTGAT	107	I
489	Gm2076094	TA41211_3847	TGCTGTTTGAGACTGCTCTTCT	TCGTCAATGCTCAGTCCAAG	106	I
490	Gm2076345	TA41221_3847	AGGTTCTTGGCAGGGGTACT	TCCTCCACCTTACCCTCCTC	213	I

491	Gm2077934	TA64642_3847	CCCATCAATGAGGCTCAAAT	GCTCCCACTCTGCTTCTTTG	265	I
492	Gm2081341	TA72123_3847	TCTTTGCAGCAGCTGTCTCT	ACTGCTCAAGAGGAGGACGA	187	I
493	Gm2083007	TA66118_3847	CATCAAAGAGGGTCCTCAGC	AAAAGCCAAAGCAAGCTCCT	132	I
494	Gm2083723	TA65368_3847	CGTGGCGTTCTTCATGTCTA	AGAGCCGAAGCTGAAGATCA	142	I
495	Gm2084313	TA65750_3847	TGGTTCATAAATCAAAGAAAAAGG	GGTTACTTCTTTCACCTGCACC	103	I
496	Gm2084815	TA72543_3847	CAGCGAAGGTGGAGGTTAGT	TGAACATCTGTGGAGGAGGA	277	I
497	Gm2086576	TA61843_3847	AATTCCACAAGACTGCTGCAT	GGCTCGACTTCATCATCTTCA	100	I
498	Gm2086731	TA61517_3847	TGGAATATGCATGGAGATGAA	GAATGACCAAAGGCAACTTATC	251	I
499	Gm2087485	TA61613_3847	CACCACCTTCACCATCTCCT	GGTGGCGCAGTCTTGTAATA	295	I
500	Gm2088328	TA61897_3847	TCAGGATGACACATTGACCG	TTTTGCAAACCTTTTAAGAGCA	191	I
501	Gm2088903	TA70498_3847	GTCATTCAAATACGCCTGGG	CCTGGGAAGTTCCAGTAATCA	170	I
502	Gm2089279	TA70567_3847	CAATCCTGACACAGGGAGGT	CGAGCGTTTTAACTGCAACA	197	I
503	Gm2091985	TA53930_3847	AAGCAAAGGATGTTGTTAAAGG	CCACAGTTCTTTATGATTGTCTCC	106	I
504	Gm2093361	TA60542_3847	CTCATCTCCCACTGCACAGA	TAGCACGAGGTCTCTTGGGT	206	I
505	Gm2095976	TA51590_3847	CCCTACTGCATTGTCAAGGG	TTGGCTTGAGATTTGGAACC	209	I
506	Gm2096212	TA51695_3847	GGAAGAAGAATATCTTTAAGTTGGC	CATTATGCCTTCTGAAGCAGC	191	I
507	Gm2096909	TA51398_3847	TGTTCAAGTCCATCCCCTCT	TTTCCATTTTTCCCTTTCCC	130	I
508	Gm2097970	TA75594_3847	TTGATTATGATGGAGTCTATGACCA	GATCAGATTTTGGATATATTTTGGC	101	I
509	Gm2098781	TA75760_3847	GGGAAGAAACTTCCTCGACC	GTTCATTCGGCAGGTGAGTT	275	I
510	Gm2099161	TA63078_3847	TGTTGGGTGGAAATGTGTTT	CGGTAAACCTCTGATAACTGCC	115	I
511	Gm2099239	TA62850_3847	TCAACAATGCTATTGACGGC	CAATTATTGGGTTTGTTTGGG	130	I
512	Gm2100275	TA62851_3847	GGCTCATGTTGGTGTTATGG	CATCATGATCAGCCGTCAAT	212	I
513	Gm2101935	TA56273_3847	TGACTGATAAAGCAGTGAGCAAA	TGTGGTGGCTCTCCATACAA	112	I
514	Gm2103135	TA56235_3847	AGTTGGAAAGACTGCACATGG	CATCTGGCCTAGTTCCCGTA	255	I
515	Gm2105159	TA41334_3847	TCCACCACCACCTAAAAAGC	TCTGGTGGAGGAGGAGAATG	168	I
516	Gm2105178	TA41299_3847	ACCACCACCTAAAAAGCCCT	TGGAGGTGGAGGAGAAGAGT	298	I
517	Gm2105266	TA41346_3847	TACGGTTTGATCATTGCGGT	TAGGATGATGCCAACGATGA	264	I
518	Gm2105278	TA41291_3847	CACCACCCTACAAGTACCCTTC	TTGTAAACTGGTGGTGGTGG	112	I
519	Gm2105282	TA41586_3847	GGGTTTAGACCGTCGTGAGA	CTGATCTAACGGTTCCTCTCG	103	I
520	Gm2106080	TA41655_3847	GGGGCACTTCATTACATTGAC	GCCACGGAACTCTACATCGT	295	I
521	Gm2106350	TA41761_3847	TGGTGGTATGCAGATTTTCG	TCCTGAATCTTGGCCTTCAC	102	I
522	Gm2106486	TA41665_3847	AGATGAGGCTGATGAAATGC	ACCAGGATTCTCACTGGCTT	165	I
523	Gm2107210	TA44456_3847	AAGCTTGGCGATAAGCAGAG	ACAGCAGATGGGATACGACC	224	I
524	Gm2107472	TA44892_3847	TAGAGTGGCTATTGGGCAGG	CTAAACTTGGTGAAGCCCCA	150	I
525	Gm2107723	TA44459_3847	AAGATTGGGTTGTTTGGTGG	TTCACTCTGCTTATCGCCAA	193	I

526	Gm2107741	TA45116_3847	AGAGTGTGGGAAGTGCTGCT	CACCAGACACCAAAGATCCA	246	I
527	Gm2108001	TA44366_3847	CATTCCTACCAAGATTAACAAGGG	ACAGAAAGAACCACAAGCCC	150	I
528	Gm2113547	TA48058_3847	TTGTGAAACGAGGATGGGAT	CATTCCAGCATAGCGGATTT	246	I
529	Gm2115926	TA50507_3847	TTCAATGCTGGCATGTTTGT	TGTTCTCCTCCTTCCCAGTG	274	I
530	Gm2116554	TA50841_3847	ACTTAGGCTTGGAGGGTGGT	CTCCGTTAGTAGTGGCGAGC	188	I
531	Gm2117177	TA49898_3847	AAGCAAAGCGGATTGAAGAA	CTGGAGGCAATCCAAGTCTC	153	I
532	Gm2117188	TA50114_3847	GTGGTATGAGCACCTTGGCT	ACTTTTTCGCCAAGAACACG	246	I
533	Gm2117810	TA49631_3847	GAAAATCTGATGGTGGCAGG	CAGCTAACAACTTTCCATCCG	283	I
534	Gm2118217	TA50261_3847	TTTTAGTGCATGGCCTTGTG	TCTCTTGCCTGCTGGAATCT	184	I
535	Gm2118543	TA50258_3847	TGGGCAGTATGGGGAGTTAG	GAAAGCCGAACATTTTGTCC	144	I
536	Gm2118770	TA49578_3847	CTGGTCCATCACCCCAGTAT	AACCAGAGGAAGGTGGTTGA	240	I
537	Gm2118923	TA49436_3847	CCCTTTTTATGGAAGCTATATGGA	CTCCAACTGCAACACTCTGTTT	260	I
538	Gm2119552	TA43857_3847	ACCGTTTTTGATGCTAAGCG	CTCCCCCTTGTAGTTCACCA	138	I
539	Gm2119675	TA43922_3847	AGCCTCCTGAGGTATGGGAT	TATCCAAGGGCTTCAAGGTG	154	I
540	Gm2119760	TA43865_3847	TTTTCGAGGTGAAAGCCACT	TGGCATTTCCACTAATATCCTTC	126	I
541	Gm2120555	TA43891_3847	CCTGAGAAGTTTAAAGAGCTTGC	CACCACCCATTCCTTCCTTA	112	I
542	Gm2120918	TA43974_3847	CTCAATCTGGTGGTGTTAGGC	TACTGCATCATCAAGCTCCA	200	I
543	Gm2121126	TA44183_3847	TCGATGAGAAATGATGGTGG	TCCATGTTAGAAGCTGGCCT	251	I
544	Gm2121560	TA46726_3847	TGAGGCACTTCCAGACAATG	CATCGATGTGAACCACACCT	101	I
545	Gm2121580	TA47134_3847	CCCTCTTTCACTTCTCCTTCTG	CAAATGTTAACCCCACAACG	107	I
546	Gm2122482	TA46569_3847	CGGATATTCCGGAAGTGATCTTA	TAGAGCAGCAGCACGTTCCT	103	I
547	Gm2123242	TA54661_3847	GAAAGAGAAGCTTCAAGAAAATGC	TTCCAATCAGCACTCCCAAT	218	I
548	Gm2124498	TA55180_3847	TGATGAACCTGATGCCAAAA	TCCTTGGATTTGGATTTCCC	188	I
549	Gm2125123	TA54469_3847	AAATGCGATCTGCCTGAGTT	CAAAGGTTTTCTCGGTGCTA	235	I
550	Gm2125343	TA59662_3847	AAGTCACCTCCTCCACCATC	GGTGGAGGGGATTTGTAAAC	110	I
551	Gm2125837	TA59962_3847	GGAAGTCTTGGTGGCATGTT	GTGCATCATCTTGCTTCCCT	238	I
552	Gm2126602	TA59631_3847	GGGATTTGGACAAGCATTTG	GGAAGAGGGTGTTGTGGAAA	123	I
553	Gm2127090	TA59381_3847	AGATTGGTGCAGGAAACAGG	CTGCACTACTCTTGCGCTTG	133	I
554	Gm2127540	TA63696_3847	AGCTGCGCTTTCTTGCAAAT	AAAGTCTTTGAAACCACAACTACAA	100	I
555	Gm2129282	TA42643_3847	GTACCCGTTATGGTGCCAGT	CTCAACCTCCTGATGGTGCT	217	I
556	Gm2129530	TA42954_3847	CATCCCCTCCTCCACCTTAC	TGGTGGGGGTGACTTGTAAT	281	I
557	Gm2131909	TA45810_3847	GGGAGAACAAGATTGGCAAG	GGTCACCCTTCACGAACACT	135	I
558	Gm2132124	TA46177_3847	GGGTTGGATTGGAGAGTCTG	GGCAACAATGTCCAACTGTG	241	I
559	Gm2132596	TA45459_3847	GTGGCGGTGAAAAGAAAGAG	GAACAAGCATTGGGGTTCTC	295	I
560	Gm2134837	TA48681_3847	GAGCAATTCCTCTGCTGGAA	GCAATTCTTTTTGGAAAACTCC	104	I

561	Gm2139030	TA57993_3847	AAAGACGAAGTGCTGGCATC	CCATCACAGGCTTACCCATT	142	I
562	Gm43427	TC208796	AAAAAGCAATGTCAGCAGCC	TGCATCGTTTAATGCATCTTG	296	I
563	Gm46291	TC210766	AGTGACTCTCGACGGGAGAA	TGCACTGTCTTTTCGCACTT	189	I
564	Gm54398	TC214405	AAGAAGCCCCACGTGAACAT	GTTCTTGACGTAGTCGGCGT	249	I
565	Gm66847	TC219948	CTGTTGGGAAGTCTTGGTGG	GTGCATCATCTTGCTTCCCT	244	I
566	Gm67517	TC219546	CAAGCGTTTAATCGGTCGTA	TGAAGTACGCAGGGACAGTG	245	I
567	Gm74712	TC221774	TGATGCTGCAATTCCTTTTG	ACACAACAGCTTTCACCTGC	101	I
568	Gm82725	TC224406	AAAGGTGCTGCTGAGGAAAA	TGCTTCTCTGCTTCTTCCTTCT	104	I
569	Gm83141	TC225240	AGATTCCAGCAGGCAAGAGA	GCATAGCCAGGATAACTGCC	280	I
570	Gm83582	TC225023	TCGATGGTAGGTTATTCGCC	TCGTGTTGGGATTGGGTAAT	138	I
571	Gm84542	TC226190	TCTGTTCCATTGTTGTTGCAG	AACAAATTCGTGCTCCCTAA	115	I
572	Gm85836	TC225502	AGGAAACAAGCCGAGCAGTA	TACTGCATCATCAAGCTCCA	294	I
573	Gm95423	TC230072	GTTGGAAAGACTGCACATGG	CATCTGGCCTAGTTCCCGTA	254	I
574	Gs1686171	Gs_002_01118_Mar06	AAAAGAATCGGTAGGGGAGC	CGTCCTTGAACCGATAACCA	297	I
575	Gs6846017	TA2828_3848	CATGACCCGTAGGCCTTCTA	AGCCACACTGGGACTGAGAC	122	I
576	Gs6846772	TA2810_3848	TCGATGAGAAATGATGGTGG	GCTCGTTGCCTTCTCCATAA	103	I
577	Tp6849300	TA1927_57577	GCATTGACATGGGGGTACTT	ACCGGTCCAGTTTCAGAGTC	124	J
578	Tp6849610	TA1520_57577	GAATCCCAGTCGAGACCTTG	CTGAATACCCTTTGGGGGTT	247	J
579	Tp6849647	TA1869_57577	CTTGCGATCCCTGCTTCTAC	GACAGGGACAGGGACAGAGA	180	J
580	Tp6849710	TA1694_57577	GGAGCTCATCCAAGACCATC	GATAGTTGGCTGCCAGATCC	227	J
581	Tp6849720	TA1214_57577	CTGGGGATGGTGCTTACCTA	AACGGTTGATCAAATGGAGC	106	J
582	Tp6849831	TA1934_57577	CTCCCTTTTGGGCAGAAACT	GCATTATCAGAAACCTTTGGC	100	J
583	Tp6849839	TA1488_57577	GTCTAGGGAGGCTGTGAGCA	CCTCCAAAGCTAATTTTGATGTTT	129	J
584	Tp6850270	TA1914_57577	TGGCATCATCCTATCTTCCC	CAAAGCCAACTCCACAGTGA	115	J
585	Tp6850680	TA1896_57577	CACATGATTCCTCCTCCACC	GCTTCTTTTCCATCCTTCCC	104	J
586	Tp6850763	TA1722_57577	CATCACGAGCTTGAAACTGC	CCATTCCAACAATATGGGCT	244	J
587	Tp6851192	TA650_57577	CTTCATTTGGTGCTCCGTCT	ATCAACCTTTGCTGGTCTGG	161	J
588	Tp6851314	TA635_57577	AATACATTGAGGCTGGTGCC	CGGATGTATCCTTCAATGGG	117	J
589	Tp6851412	TA106_57577	TATAAGTCACCTCCTCCGCC	TGGTATTTGTAAACGGGTGG	104	J
590	Tp6851668	TA849_57577	ATTGAAGGGATAGGTTGGGG	GCAGTGTCCTTTGTGCATGT	214	J
591	Tp6851762	TA320_57577	CACCTCCTCCACCATATGTCT	AGGAGGTGGTGGTGACTTGT	161	J
592	Tp6851794	TA265_57577	CCACCTCCTCCACCATATGTC	GTGGAGGTGGCGACTTGTAG	100	J
593	Tp6851919	TA555_57577	ATTGGGTTCGCTGTGTTCTT	ACTGTGTGGTAAAGGGCTGC	182	J
594	Tp6851957	TA197_57577	ATGGAGGATCTGTAAATGAAGC	CAGGTTCTTTAGAAAACTGAGAGG	295	J
595	Tp6852047	TA720_57577	TACTCCCACCATGGGGTCTA	CTTTCCAGAGAAGGGTGCAG	119	J

596	Tp6852097	TA215_57577	TGCTGGTGATGATGCTCCTA	TTTTCTCTGTTGGCCTTTGG	291	J
597	Tp6852284	TA278_57577	TTCCATCGTTTACGGCTAGG	CTGGACAGGTGGTGGAAACT	200	J
598	Tp6852466	TA192_57577	GCGGATTTCGACTTCATCAC	GGAGGTGGCGACTTGTAGAC	246	J
599	Tp6852529	TA88_57577	ACCACCGTCCTGCTGTCTTA	GGAGGTGGCGACTTGTAGAC	225	J
600	Tp6852940	TA363_57577	GCAAGATGATGACTTGCACC	AAAACAGGTCTGTGATGCCC	193	J
601	Tp6853858	TA2158_57577	TGAGTGCTGCAGAGTTGGAT	CCACGGTGGTTTCTGATTTT	190	J
602	Tp6854083	TA2923_57577	GCAAGGAAAGTTGCAAAAGC	TTTTCTCCTCTGGCTGTGGT	238	J
603	Tp6854346	TA2990_57577	ATCAGCAAAGACAAGGGTGC	CTTACAATCCTCCTGGTGGC	287	J
604	Tp6854715	TA2257_57577	ATTCCCTGGTGAAAATGCAG	GGAGGCATCTTTTTGTGGAA	165	J
605	Tp6854787	TA2719_57577	TACACCACCTTCGACACCAA	GGTAACACCGGAAGGTTGAC	183	J
606	Tp6855079	TA2865_57577	TGTAATTCTGGTATGGAAACATGC	AGATTTGGCAACCCCTTCTT	135	J
607	Tp6856169	TA3815_57577	GAATTGCTCGCAAATGGTTT	CTGGTGCCATTTCTCCATTT	232	J
608	Tp6856320	TA3391_57577	ATGTTCGGCCTTGATATTGC	TCTGACCAGTTGGCTGAATG	217	J
609	Tp6856412	TA3197_57577	GTTGGAGTTGCTGTTGGGAT	GGGATCTGTTGCGTGAAAAC	169	J
610	Tp6856829	TA3863_57577	ATTGGCTGACTTGGATTTGG	CATCCTTTTTCGCCGATTTA	206	J
611	Tp6857034	TA3954_57577	TTGTGGCACTGAACAAGAGG	TCATGATGCATTGCTCCTTC	246	J
612	Tp6857139	TA3002_57577	TGAGAGGGTGTTCCTTTTTCA	TCTTTGGCCAGCAAGAGAGT	257	J
613	Tp6857294	TA3098_57577	GCTGATATGGCACATGCACT	AGTTCGAATACCAACGCCAC	255	J
614	Tp6857470	TA3531_57577	TGCCACTTTGTTGGAATTGA	TACCCTCAGCAAAACCATCC	232	J
615	Pc1709272	Pc_009_02407_Mar06	AGAAGCTCGAGATCGACGAC	TATCCACGGAAGCAAGGAGT	220	K
616	Pc1710424	Pc_027_02671_Mar06	AAAGACCCTTGGAGAATGGG	CTTCCCCAAAATCCTTGACA	104	K
617	Pc1711976	Pc_002_00557_Mar06	CGACGATGCCATAACTGTGT	ATCGTTCCAATAAGCGATGC	206	K
618	Pc1714501	Pc_007_02308_Mar06	AAGAGGTGCTTGGTGATGCT	CACCCACGAACAGACTTCCT	205	K
619	Pc1717650	Pc_003_01625_Mar06	GGAAGAGGAAGAGGGAGAGG	AGAGGGAGGAGGAGGAGGAG	130	K
620	Pc1722835	Pc_002_00773_Mar06	CCCTACTGCATTGTCAAGGG	GGGATTTGGAACCCATGATT	202	K
621	Ps1768102	Ps_012_00624_Mar06	TGTTGAAGGAGGTGGTAGGG	CACTGGTGAAGGTGGTGATG	298	L
622	Ps1768379	Ps_003_00390_Mar06	TCACCCTGTCTACCACTCCC	GGTGGTAAGGTGGAGGAGGT	248	L
623	Ps1768382	Ps_002_00296_Mar06	CTCCGTGGAGGTATGCAGAT	GCAACACAAGGTGAAGGGTT	226	L
624	Ps1768451	Ps_011_00616_Mar06	TAAAGGAGAGAAAATGTACATCCA	TACTCGATCCAATCGATGCC	177	L
625	Ps1768764	Ps_005_00538_Mar06	CAGAAGGAATCCACCCTTCA	TTCCATCCTCCAATTGCTTC	206	L
626	Ps1768922	Ps_002_00304_Mar06	TCTCTCTCGCGTTTTCCCTA	CCTTTGTCCTGAATGCAACC	247	L
627	Ps1769865	Ps_003_00404_Mar06	CAGAAGTTCGAGACCAGCCT	CTGGAGTGCAGTGGTGTGAT	183	L
628	Ps1770278	Ps_002_00092_Mar06	TCCTTCGTTTGGATCTTGCT	GCAGCTTCCATTCCAATCAT	292	L
629	Ps1770807	Ps_005_00534_Mar06	TGGAAGGATCTTCCTGATTCC	GGAAGCTGTCTATGGCCTTG	130	L
630	Ps1770858	Ps_003_00425_Mar06	CAATCCCCATAATTGGCAAC	CGTAGGTTACGCCAATGGTC	272	L

631	Ps1770909	Ps_003_00405_Mar06	CACCACCACCACACAAGAAG	GGTGGTAAGGTGGAGGAGGT	168	L
632	Ps1771689	Ps_003_00376_Mar06	GGTTTTGAGAAGCCATCAGC	AGATAATCACCAAGTGCCCG	230	L
633	Ps6869023	TA539_3888	TCACCCTGTCTACCACTCCC	TATGGCTTCTTGTGTGGTGG	105	L
634	Pv1608	TC135	CCAGGATGAAGGAAGGACAG	TCCTTAGTCGCAGACACCAA	196	M
635	Pv1806	TC784	TCAAGGTTGTTAGGCTTTCG	CGTAGGTCAAGGGGCAAATA	112	M
636	Pv2301	TC1959	CAGAAAGTAGTGAAAGGAAATCCAA	TATCATGTGCGGCTCTCTCA	100	M
637	Pv327	TC265	ATTGTTAGCAGGAAGTGGGG	GCTTGGCATTAACACCATCA	100	M
638	Pv3790	TC1161	CACCTGTCCGCACTACAAAA	GGCACTAAATGGGAGAAATCC	104	M
639	Pv4540	TC2144	TTGGAGGACTTCCTCAATGG	ATGAACAAAGCAGCAGAGGG	253	M
640	Pv4675	TC2159	TTTCAATGGCATCCTTTCTCT	TCCTAGACCCACCAGAAGGA	273	M
641	Pv6865724	TA3297_3885	AAGACAAGGCAAGGTTGGTG	CCATTCACATGAAGCTCTGG	177	M
642	Pv6867715	TA4138_3885	AAGGAAAAGAAGGAGAAGAAGGAG	TCATTGGGTTTGTCCTTTCC	248	M
643	Pv885	TC525	GAAAAGAAAAAGAAAGGAAAAGAAA	ATCTCTGGTTCTGGTCGTGC	104	M
644	Rp1788051	Rp_005_00252_Apr06	TGGTGAACGTTCCAAAGACA	CCCTTCTTGTCACCACCAAT	295	O
645	Rp1788317	Rp_003_00207_Apr06	ACCAGACCAGCAGAGGCTTA	TCCTTGTCTTGGATCTTGGC	219	O
646	Rp1788514	Rp_002_00131_Apr06	TCCAAATCAAACCAGTACCCA	TACCGTGAGGCCTTGTAACC	175	O
647	Ah1756637	Ah_002_00205_Mar06	TTTCCCTCAGGATAGCTGGA	TTGTGGGGTCTAGGTTAGCG	252	P
648	Ah6927938	TA1302_3818	GACTCAGGCAAGTCGACCAC	GTCTCGAACTTCCACAAGGC	197	P
649	Ah6928256	TA1780_3818	CGCAAGAAGCAGTCAGATGT	GTAACACCCTGGTTGGTTGG	236	P
650	Ah6928397	TA1660_3818	TGCTCAGCATGTTGAAGAGG	CCAAGAGACGAGGAATCTGC	164	P
651	Ah6928505	TA1435_3818	CACAGAAGATCAGGGTCGGT	TCTGACATGTGTGCGAGTCA	108	P
652	Ah6929189	TA1242_3818	GCAGATTTTCGTGAAGACCC	AGGGTGGACTCCTTCTGGAT	198	P
653	Ct6874464	TA612_3832	TGATGCATACATAAATGGACATGA	GGTTTTGGTGATTTGCACAG	187	Q
654	Ct6874739	TA235_3832	CAAGGGGAAAACAAGCAAAA	TCTTCTTTGCACGTTTGTGG	170	Q
655	Ct6875390	TA128_3832	TCCATTCCATGTTCTCAGGA	GGAACAATCCTGTTCTCCGA	300	Q
656	Ct6875951	TA957_3832	GTGGTGCAGTGGTGTTATGG	CGAAGGCATGAGTTGAGGAT	103	Q
657	Ct6876038	TA151_3832	CCTCGAGGTCGAATCTTCTG	CTTCTGGATGTTGTAATCGGC	148	Q

[a]Sequneces for Accession IDs with '*' are available in the Legume Information System (LIS, www.comparative-legumes.org) at http://cicar.comparativelegumes.org/data/2011/58da8857f0f21afded122214cd604b9f/transcript_contigs.fa.gz

* **Source of genes/ markers** A: Illumina sequencing/ 454 transcript sequences; B: Chickpea EST; C: Rajesh and Muehlbauer, 2008; D : Singh et al. 2008; E: *Lupinus* spp.; F:*Lotus japonicas*; G: *Medicago truncatula* ; H: *Medicago sativa* ; I: *Glycine max* ; J: *Trifolium pratense* ; K: *Phaseolus coccineus* ; L: *Pisum sativum* ; M: *Phaseolus vulgaris* ; O: *Robinia pseudoacacia* ; P: *Arachis hypogaea* ; Q: *Crotolaria tenuifolia*

Annexure 2

List of chickpea EST-SNPs markers derived from UG-IV set of unigenes showing primer sequence, and product size(bp) details for each marker.

S.no.	Primer ID	Forward primer	Reverse primer	Expected Product Size(bp)
1	Ca2C42878	AAACAAGGGCAGGGAACTTT	ACATTCCTCGTCCCAAGATG	246
2	Ca2C42611	GGAGGTGCTATCATTTCCGA	CCAGTCAGCAGCATCAGGTA	672
3	Ca2C43782	GGGAAAGCAAAAGGAGAAGG	AGGTGTTTGTAGCCCTGGTG	236
4	Ca2C19195	CAATGCATGGTGACACACAA	TCCCAGAATTCCTCAACCTG	498
5	Ca2SGR399431	GGGTTTTTGGGTTTTAGGGA	GATAAAAAGGGGCTGGAAGC	679
6	Ca2C718	CGCTGGAGACAAGCTAAACC	CTGCTGTACCAAGCCACTGA	379
7	Ca2C585	CCATCAACAACCATGACTCG	TCCACCTGCATTATTCCCTC	201
8	Ca2C11361	CTCATTGCTTGAAGATGGCA	CAATGGCATCATCTGGTACG	207
9	Ca2C13051	AATCCTTCACCATTTCAGCG	AGATTTCCGTTGAATCACGC	461
10	Ca2C42782	TGGAGACATGGACCAAAACA	AGGCAACTCCTCTACGGTCA	289
11	Ca2SGR399333	GCGATGTAAAACGAGGGTGT	CCACCAAAAATTCCATACCG	404
12	Ca2C43336	GCTGCTGTTGGGATTTCATT	TTATTGCCACAGTTGGTGGA	214
13	Ca2C42692	TCATCTCGATATGCTCGCTG	ACATGGACTGACATGGACGA	168
14	Ca2C43936	CCGGAGGAATCTACGAATGA	TTTTTCTCATCTTCGGTCGG	272
15	Ca2C21271	ATGACGTGCAGACAAAGCAG	ATTTGGGCAATTAGCCCTTT	275
16	Ca2SGR409670	TGAAAATCAATGAAAGGGCA	CGGGACATAACTGAACACCC	189
17	Ca2C39414	CATGCCAAGTCCAATCACTG	GCTGCTCAAACAACACCAAA	356
18	Ca2SGR399716	GGGGGTTTTCCCTTTTGTTA	CATATGCTCAGCTTGCCAGA	150
19	Ca2SGR399815	CATGGTCGTGTTACAGTGGC	CAACAAGCTTCCAGCAATCA	328
20	Ca2SGR409100	CTGAGCATGCTGTCTTGGAA	GCGCTCTGAACAAACATGAA	261
21	Ca2SGR399970	CTCCTCAAACGAAATCGACA	TCAGAACTGTGTCTGGCTCG	215
22	Ca2C42582	GTTATTGGCTCTCTGCAGGC	CTTCGATTCTTCCCCGTGTA	220
23	Ca2C43484	GTAGCCTGCGAAAAGCTACG	TCGCCACTACTTTGGGAATC	188
24	Ca2C43499	GGCCCAAGGTTCAACTACAA	TTACAGCGAAACTGCGAATG	568
25	Ca2C43467	AGACATCAACCGAGATTCCG	GTTCTTTTCGCCTTTCCCTC	268
26	Ca2C43375	GGGTCAAGATATAAAGGGCG	ATGGATACATGGGGTTTCCC	154
27	Ca2C8893	CCTTTGTGACACGATGTTGG	CCGTTGGTGAAAGTGGAGAT	405
28	Ca2C42817	TCCGTGGTTGGTTTTACCTC	AGGTTCAGGGCATTCTCCTT	163

29	Ca2C27483	TTTGATTTCCAGTGCTGCTG	TAAAGCCCATGTCATCGTCA	155
30	Ca2C43919	GTTCTGGAGACGCCACTAGC	TGCGACATTCATCACTCTCC	153
31	Ca2C43916	AGCCGACGTTGAAATACCAC	CAGAACCACCGGATCACTTT	248
32	Ca2C42666	TCGCTTCTCCCACTTATGCT	CGAGATTCCCACTGTCCCTA	154
33	Ca2C42621	CCCCATACGCAAACAACTCT	TCGGTGTTCAACCATCAAAA	799
34	Ca2SGR408900	AACATCGGGAATGTTTCGAG	CGATGAACACCCGAAGATTT	530
35	Ca2SGR408994	TCCGTTTTCAACAACCATGA	GGTAGGACAAACGCGAAAAA	239
36	Ca2C42867	TAGGTCCATTTCACAAGCCC	TGGAGCTGAATTGCATGAAG	234
37	Ca2C35363	GCTCCATTAGCATTGCCTTC	AAAGGGTCTTGTTGGTGTGC	201
38	Ca2C43662	AAAAGATGGACGCCACATTC	CTAGGGTTTTGAGGCTGCTG	490
39	Ca2C16048	CAACCGGAACTACGGTGTCT	TCGTCCCCATCCTCATACAT	412
40	Ca2C43545	AAGGAAAGGAGGAGGACCAA	GCAAGCAATTTTCCAAGCTC	282
41	Ca2C30733	TGGTTTCATTCTCCAATCACA	CGATCGATGCATAAATCACAA	484
42	Ca2C20880	CGTTCAGTGGGAGCTTTCTC	CGTTGCAACTCCATCATACG	420
43	Ca2C25517	ACTGGCCAACAATCACACAA	GGAGGCAGCACAAGAGAAAC	217
44	Ca2C33267	ACAACAGAATTGCCCTGACC	AGGCCCGAGAAATTTCCTTA	696
45	Ca2C43828	TCTCTCTTACGATGGGTGGG	GAACATGAAATTGCCTCCTCA	456
46	Ca2C12943	AAACGCGTACCCTTGACATC	GAAGAAGACCCTGAACTGCG	237
47	Ca2C34423	AGGAAGCAGAGGGTGGAAAT	GCCTTCTCGTTCTCATACGC	232
48	Ca2C24803	ATCGGAGCCTTACTGGGTTT	TCGTTTGCAGCCTTAGCTTT	194
49	Ca2C31127	TACCAGAACAAGACCTGCCA	TGTGATTTGGCATCCTTTGA	289
50	Ca2C43744	CGGTGAAGAGGAAAAAGCTG	ATGTACCACCCTAAAGCCCC	249
51	Ca2C43650	GGGGGTTGATTATTCATTTGC	TAATTGGCATGTTGGAACGA	337
52	Ca2C43539	CCCACAGGTATGGGATAACG	ACCCAACATGAGATGCAACA	163
53	Ca2C43058	TGGTGAGCAAAAGTGGAGTG	GCAATCCTATGAGCGGTGTT	246
54	Ca2C33657	CACATACCATGAATGTGTGCC	GCCAATTTTGATCCAACCAC	421
55	Ca2C23395	CAAGTGGAGAACGGTGACAA	GGGCATGCAAAGAAAGTGAA	370
56	Ca2C42642	TTACTTTCCGTCACCCGTTC	TTATCGGCCTTAGCAGCAGT	595
57	Ca2C22854	AGCGTGCTCAGATAAAGGGA	GAACATCATGGACAAACCCC	461
58	Ca2C43505	GCCTTCTTCAAACCAAATTCA	AATTGGCTCTGTAACGGCTG	236
59	Ca2C43222	GTTTCAATGAGAAGCGGAGG	TTTCCTGTTGTGGTTTGCTG	155
60	Ca2C34533	CAGGCTTCCAGAAGAAGGTG	CTTTGCAGCAGGTTTCTTCC	206
61	Ca2C42710	TAGGCTTCACCCAGCAACTT	AGGCAAGTCCCAATGAGCTA	245
62	Ca2C5199	AACCGACCTCCACTCTCCTT	AATGTTAAGCATTCACCGCC	194
63	Ca2C8942	ACCAGTTGATGACCGAAAGG	CTTTCCTTTTTGCAGCCTTG	290

64	Ca2C42587	GCACAAACAATTGCTCCAGA	AATAACACGCCCCATTCAAG	386
65	Ca2C20777	TTAACGGTGCCGTACACAAA	TAATGACCTCACCCTTTGCC	249
66	Ca2C5572	AGTTGGATCTGACTGGGGTG	TCCCAGCTTAAATCCATTGC	426
67	Ca2C23573	CAATTTACGCTTTTGCCCAT	GGGTACGGTACGGAGAGTCA	282
68	Ca2C44220	GTTGTGGGACAAAAATGGCT	CAAGCTTGGTTGGGAAGGTA	307
69	Ca2C17306	TCTGCATCAAGGCACTGAAC	GGGGTTCCCTACGGAATAAA	489
70	Ca2C44110	CTTCAACAGCAAGAGGCACA	AGAAGAGCAAGGCTGATCCA	237
71	Ca2C3917	TCGAGGGAGCTGAAATGAGT	CAAAGCAACCATTGCTCTCA	295
72	Ca2C9564	GCCCTGCTGGCTACTATCTG	ATTGAAGCAAGGAAGCCTGA	503
73	Ca2C19564	TCAATCACATGGCTTTGGAA	TTGGCGGTGGAGATAGTTTC	244
74	Ca2C28874	CGCCTCCATATTATCGTCGT	TCCAAAACACGTGGAACAAA	283
75	Ca2C14402	AATGACACGCCGTTACATGA	GAAGAGGTGAGCTGGTTTGC	220
76	Ca2C22017	ATACATACGGGCTACGGCTG	TGGCTTCTTGCAAATCAGTG	408
77	Ca2C44166	CGGGGAGTTGAAAATAAGCA	CTTCATGATGGATCAGCCCT	658
78	Ca2C27784	GCTTGCCAGCAAAAATAAGC	GAAGGCGAAAATTCAGGACA	296
79	Ca2C2842	GACCCAAGCAAAGACCACAT	AAGGGAGAGAAGAAGCCAGC	473
80	Ca2C2529	TGGCATATAATCGGGTCCAT	CCTCCCTCGAAGAATCATCA	437
81	Ca2C44338	CAATGGCTTCCTACACCGTT	GAGATTCATCAAGGGACGGA	509
82	Ca2C33309	TTGCTCAAACAGATGGTGCT	ACTGTCACTACGGGTCCCAC	238
83	Ca2C18912	TTGTTATAACGAGGGCCCAG	ACGACTCTTTGCTCCTTCCA	261
84	Ca2C44635	AAGGAGAGTCCAGTGCCAGA	ATAATGCTCTGCCAAATGCC	298
85	Ca2C44705	GCTGAGGAAGTGAATGAGGC	CAGCCAAATCAACAACCCTT	223
86	Ca2SCK148696	TGCTTCTCGAATTCCTTCGT	TTCCAGTGATGCATCTTCCA	185
87	Ca2C16596	TGGGAGGTTGGAATCTTGAG	CCAAGGGTTATCAATGGGTG	351

Annexure 3

Details on sequence diversity estimates of 264 genes in chickpea germplasm

			General features for all genes					All genotype			
S.No.	**Primer ID**	**Gene source***	**Total no.of genotypes**	**Number of cultivated genotypes**	**Number of wild genotypes**	**Sequence length (bp)**	**No.of indels**	**Number of SNP**	**Number of haplotypes**	**Haplotype diversity**	**Nucleotide diversity (π)**
1	Ca10273CD20F12	A	3	2	1	448	0	4	2	0.666	0.006
2	Ca10490169748746	A	3	2	1	186	1	1	2	0.666	0.0036
3	Ca14038CD25A07	A	3	2	1	607	0	1	2	0.666	0.0011
4	Ca157594586605	A	3	2	1	319	0	2	2	0.666	0.0042
5	Ca18532	A	3	2	1	140	0	2	2	0.666	0.0095
6	Ca20156	A	3	2	1	221	0	4	2	0.666	0.0121
7	Ca21249	A	3	2	1	149	0	2	2	0.666	0.0089
8	Ca14402150174189	A	3	2	1	208	0	1	2	0.666	0.0032
9	Ca21567	A	3	2	1	72	0	1	2	0.666	0.0093
10	Ca22434	A	3	2	1	175	0	2	2	0.666	0.0076
11	Ca12863169747399	A	3	2	1	736	0	3	2	0.666	0.0027
12	Ca1773847832705	A	3	2	1	830	0	5	3	1.001	0.004
13	Ca9351169747446	A	3	2	1	129	0	1	2	0.666	0.0052
14	Ca93773928149	A	3	2	1	602	0	3	2	0.666	0.0033
15	Ca1683047832555	A	3	2	1	226	0	3	2	0.666	0.0088
16	Ca1448960219076	A	3	2	1	628	1	5	2	0.666	0.0053
17	Ca10556169746056	A	3	2	1	119	0	2	2	0.666	0.0112
18	Ca10880169748548	A	3	2	1	101	0	1	2	0.666	0.0066
19	Ca13382169746951	A	3	2	1	219	0	1	2	0.666	0.003
20	Ca15775169744160	A	3	2	1	107	0	2	2	1.167	0.0125
21	Ca15994169746275	A	3	2	1	195	0	1	2	0.666	0.0034
22	Ca12411169747340	A	3	2	1	427	0	2	2	0.666	0.0031
23	CaESTCg3	A	3	2	1	169	0	1	2	0.666	0.0039
24	CaESTSn16	A	3	2	1	539	0	5	2	0.666	0.0062
25	CaESTSn30	A	3	2	1	166	0	2	2	0.666	0.008
26	Ca2C228	A	6	4	2	227	0	1	2	0.334	0.0019

27	Ca2S126415_1648_0587	A	6	4	2	61	0	1	2	0.6	0.0072
28	Ca2C1600	A	6	4	2	162	0	2	3	0.8	0.0054
29	Ca2C8663	A	6	4	2	222	0	10	3	0.6	0.0197
30	Ca2C10102	A	6	4	2	101	0	1	2	0.533	0.0043
31	Ca2C12236	A	6	4	2	196	0	2	3	0.733	0.0045
32	Ca2C43617	A	6	4	2	282	0	2	2	0.634	0.0031
33	Ca2C42668	A	6	4	2	97	0	1	2	0.334	0.0045
34	Ca2S143601_3004_0194	A	6	4	2	108	0	1	2	0.533	0.0041
35	Ca2S295752_0698_0195	A	5	3	2	118	0	3	3	0.7	0.0122
36	Ca2S211594_0676_0883	A	5	3	2	66	0	1	2	0.4	0.0073
37	Ca2S117228_1922_1005	A	6	4	2	189	0	3	3	0.9	0.007
38	Ca2S242431_3223_1815	A	6	4	2	65	0	1	2	0.533	0.0067
39	Ca2S110494_3324_2020	A	6	4	2	166	0	1	2	0.533	0.0026
40	Ca2S232393_3220_2305	A	5	3	2	210	0	1	2	0.6	0.0023
41	Ca2S124718_0933_2144	A	6	5	1	213	0	2	2	0.334	0.0041
42	Ca2C17503	A	7	5	2	202	0	5	2	0.476	0.0101
43	Ca2C33173	A	4	4	0	108	0	1	2	0.5	0.0051
44	Ca2C3599	A	6	5	1	215	0	1	2	0.6	0.002
45	Ca2C20537	A	6	5	1	725	0	2	2	0.334	0.0012
46	Ca2C18093	A	6	4	2	78	0	1	2	0.334	0.0056
47	Ca2C36478	A	6	4	2	257	0	2	3	0.833	0.0034
48	Ca2C38337	A	6	4	2	204	0	1	2	0.334	0.0021
49	Ca2C41871	A	6	4	2	194	0	2	2	0.533	0.0045
50	Ca2C42103	A	6	4	2	166	0	1	2	0.334	0.0026
51	Ca2C4224	A	6	4	2	139	3	4	2	0.334	0.0126
52	Ca2C6533	A	6	4	2	196	0	2	2	0.634	0.0045
53	Ca2C10202	A	6	4	2	225	0	1	2	0.334	0.0019
54	Ca2C28092	A	6	4	2	159	0	2	3	0.6	0.0055
55	Ca2S290507_2844_3854	A	6	5	1	328	0	3	2	0.334	0.004
56	Ca2S125676_2925_1604	A	6	5	1	281	0	1	2	0.334	0.0016
57	Ca2C21276	A	2	1	1	210	0	1	2	1	0.0048
58	Ca2S289800_2555_0786	A	5	4	1	193	0	1	2	0.4	0.0025
59	Ca2SFE670434	A	6	5	1	146	0	2	3	0.6	0.006
60	Ca2C9868	A	6	5	1	477	5	1	2	0.334	0.0009
61	Ca2C24067	A	6	5	1	356	0	4	4	0.8	0.0049

62	Ca2C12526	A	6	5	1	153	0	3	3	0.733	0.0086
63	Ca2C3559	A	5	4	1	907	0	5	4	0.9	0.0026
64	Ca2C3892	A	7	6	1	543	1	10	3	0.833	0.0075
65	Ca2C6285	A	5	4	1	253	0	2	2	0.4	0.0038
66	Ca2C6628	A	6	5	1	227	0	1	2	0.334	0.0019
67	Ca2C25794	A	5	4	1	729	1	5	2	0.4	0.0033
68	Ca2C38039	A	5	4	1	378	0	4	4	0.9	0.0051
69	Ca2C30045	A	5	4	1	217	2	1	2	1	0.0022
70	Ca2C30401	A	5	4	1	128	0	1	2	0.4	0.0038
71	Ca2C32639	A	5	4	1	222	0	2	2	0.4	0.0043
72	Ca2S032873_1646_3001	A	5	4	1	694	0	4	3	0.7	0.0028
73	Ca2C34020	A	4	3	1	265	0	1	2	0.5	0.0021
74	Ca2C41295	A	4	3	1	278	0	1	2	0.5	0.002
75	Ca2S082575_1212_1008	A	5	4	1	363	0	1	2	0.4	0.0013
76	C2S217264_1973_0317	A	7	6	1	191	0	1	2	0.286	0.0021
77	Ca2S12857	A	5	4	1	188	0	3	3	0.7	0.0077
78	Ca2C11129	A	6	5	1	232	1	3	2	0.334	0.0057
79	Ca2S10894	A	7	6	1	223	0	1	2	0.286	0.0018
80	Ca2S38128	A	4	3	1	564	0	4	3	0.833	0.0039
81	Ca2S37998	A	5	4	1	276	0	4	2	0.4	0.007
82	Ca2S41121	A	5	4	1	415	0	4	2	0.4	0.0046
83	Ca2C37368	A	5	4	1	209	0	7	2	1.15	0.0161
84	Ca2C36626	A	4	3	1	314	0	2	2	0.5	0.0035
85	Ca2C36391	A	5	4	1	392	1	1	2	0.4	0.0012
86	Ca2C34583	A	7	6	1	274	0	2	2	0.286	0.003
87	Ca2C34413	A	5	4	1	233	0	6	4	0.9	0.0124
88	Ca2C3406	A	7	6	1	237	0	1	2	0.572	0.0017
89	Ca2C3233	A	4	3	1	919	0	8	3	0.833	0.0047
90	Ca2C32031	A	5	4	1	295	0	7	4	0.9	0.0114
91	Ca2C31438	A	7	6	1	816	0	1	2	0.286	0.0005
92	Ca2C15786	A	5	4	1	286	0	5	2	0.4	0.0084
93	Ca2C1614	A	5	4	1	296	1	1	2	0.4	0.0016
94	Ca2C41582	A	5	4	1	415	0	6	2	0.4	0.0069
95	Ca2C37037	A	5	4	1	295	0	4	2	0.4	0.0065
96	Ca2C7044	A	5	4	1	200	0	2	2	0.4	0.0048

97	Ca2C161	A	5	4	1	779	0	21	3	0.95	0.0129
98	Ca2C6341	A	5	4	1	450	1	3	2	0.4	0.0032
99	Ca2C17163	A	5	4	1	478	0	3	2	0.4	0.003
100	Ca2C23568	A	5	4	1	286	0	3	2	0.4	0.005
101	Ca2C26976	A	5	4	1	284	0	1	2	0.4	0.0017
102	ca2c33338	A	5	4	1	322	0	1	2	0.4	0.0015
103	ca2c42261	A	7	6	1	157	0	2	2	0.286	0.0052
104	Ca2C1043	A	4	3	1	694	0	2	3	0.833	0.0016
105	Ca2C36298	A	5	4	1	160	0	2	3	0.8	0.006
106	Ca2S039865_2799_2981	A	5	4	1	161	0	4	2	0.75	0.0119
107	chs	A	3	2	1	127	0	2	2	0.666	0.0105
108	ge13bg	A	3	2	1	512	0	4	2	0.666	0.0052
109	kpi1	A	3	2	1	259	0	3	2	0.666	0.0077
110	ntp	A	3	2	1	286	0	1	2	0.666	0.0023
111	rgr4	A	3	2	1	276	0	1	2	0.666	0.0024
112	SHMT	A	3	2	1	493	0	2	2	0.666	0.0027
113	tk	A	3	2	1	665	0	2	2	0.666	0.002
114	AGLC100	A	8	7	1	215	0	2	2	0.25	0.0036
115	AGLC198	A	8	7	1	458	0	2	2	0.25	0.0017
116	AGLC93	A	7	6	1	306	0	3	3	0.524	0.004
117	AGLC129	A	7	7	0	528	0	2	2	0.572	0.0015
118	AGLC136	A	7	6	1	436	0	5	2	0.762	0.0047
119	AGLC163	A	6	5	1	81	0	1	2	0.334	0.0054
120	AGLC8	A	7	6	1	469	0	3	2	0.286	0.0026
121	AGLC51	A	8	7	1	87	0	1	2	0.25	0.0044
122	AGLC57	A	8	7	1	596	0	1	2	0.25	0.0006
123	AGLC2	A	8	7	1	415	1	5	3	0.607	0.0046
124	AGLC7	A	8	7	1	262	0	1	2	0.25	0.0015
125	AGLC14	A	8	7	1	392	0	4	3	0.464	0.0039
126	AGLC15	A	8	7	1	465	0	5	2	0.25	0.0041
127	AGLC17	A	8	7	1	241	0	3	2	0.25	0.0048
128	AGLC19	A	8	7	1	271	0	2	2	0.25	0.0028
129	AGLC21	A	8	7	1	606	0	1	2	0.25	0.0006
130	AGLC22	A	8	7	1	815	0	3	2	0.25	0.0014
131	AGLC23	A	8	7	1	295	0	3	4	0.786	0.0039

132	AGLC28	A	8	7	1	259	0	1	2	0.25	0.0015
133	AGLC30	A	8	7	1	245	0	2	3	0.464	0.0031
134	AGLC44	A	8	7	1	686	0	2	2	0.25	0.0011
135	AGLC45	A	8	7	1	357	0	1	2	0.25	0.0011
136	AGLC61	A	8	7	1	392	1	2	2	0.25	0.002
137	AGLC72	A	8	7	1	110	0	3	3	0.464	0.0105
138	AGLC74	A	8	7	1	319	0	2	2	0.25	0.0024
139	AGLC76	A	8	7	1	364	0	2	2	0.25	0.0021
140	AGLC77	A	8	7	1	578	0	7	3	0.464	0.0047
141	AGLC78	A	8	7	1	291	0	4	4	0.822	0.0053
142	AGLC82	A	8	7	1	216	0	1	2	0.25	0.0018
143	AGLC84	A	8	7	1	361	0	1	2	0.25	0.0011
144	AGLC94	A	8	7	1	370	0	4	3	0.464	0.0042
145	AGLC111	A	8	7	1	327	0	2	2	0.536	0.0024
146	AGLC112	A	8	7	1	155	0	1	2	0.25	0.0025
147	AGLC126	A	8	7	1	592	0	4	3	0.679	0.0026
148	AGLC131	A	8	7	1	137	0	1	2	0.25	0.0028
149	AGLC137	A	8	7	1	340	0	4	2	0.25	0.0045
150	AGLC168	A	8	7	1	523	0	3	2	0.25	0.0022
151	AGLC178	A	8	7	1	279	0	3	2	0.25	0.0041
152	AGLC179	A	8	7	1	149	0	3	2	0.536	0.0078
153	AGLC193	A	8	7	1	689	0	6	4	0.643	0.0034
154	AGLC196	A	8	7	1	95	0	1	2	0.25	0.0041
155	AGLC202	A	8	7	1	605	0	1	2	0.25	0.0006
156	AGLC3	A	8	7	1	233	0	2	2	0.25	0.0033
157	AGLC171	A	8	7	1	140	0	2	3	0.464	0.0055
158	AGLC213	A	19	8	11	379	0	28	9	0.74	0.0211
159	AGLC216	A	12	7	5	393	0	6	3	0.319	0.0051
160	AGLC122	A	16	9	7	529	2	36	1	0.467	0.0205
161	AGLC212	A	18	9	9	696	1	62	5	0.929	0.0259
162	AGLC217	A	13	8	5	174	0	4	2	0.282	0.0074
163	AGLC214	A	16	9	7	229	0	14	4	1.038	0.0184
164	CaHa36	A	6	4	2	122	0	5	2	0.634	0.0179
165	CaHa61	A	6	4	2	197	0	2	3	0.6	0.0044
166	CaHa506	A	6	4	2	372	0	1	2	0.334	0.0012

1	TC81224	B	20	9	11	377	0	8	7	0.726	0.006
2	TC85165	B	20	9	11	276	0	9	6	0.586	0.0092
3	TC85414	B	19	9	10	189	0	6	7	0.667	0.0091
4	TC86258	B	20	9	11	168	0	9	6	0.684	0.0151
5	TC92821	B	17	9	8	164	0	7	5	0.518	0.0126
6	TC94373	B	19	9	10	262	1	19	6	0.547	0.0207
7	TC103928	B	20	9	11	271	0	9	5	0.576	0.0094
8	TC87800	B	15	8	7	364	3	22	1	1.067	0.0186
9	TC76606	B	16	9	7	311	0	21	1	1.062	0.0203
10	TC77515	B	19	9	10	725	2	79	3	1.047	0.0312
11	TC77624	B	19	9	10	157	0	11	6	0.468	0.02
12	TC77707	B	20	9	11	363	0	26	8	0.7	0.0202
13	TC101057	B	19	9	10	240	0	1	2	0.456	0.0012
14	TC87719	B	20	9	11	377	0	9	3	0.728	0.0067
15	TC87769	B	17	7	10	498	0	39	6	0.746	0.0232
16	TC87946	B	14	8	6	247	0	1	2	0.407	0.0013
17	TC95048	B	19	9	10	390	1	41	7	0.616	0.0301
18	TC96130	B	19	9	10	527	4	44	5	1.041	0.0239
19	Mt106141	B	3	2	1	212	0	3	3	1.001	0.0094
20	Mt106628	B	3	2	1	173	0	2	2	0.666	0.0077
21	Mt123162	B	3	2	1	656	0	1	2	0.666	0.001
22	Mt123479	B	3	2	1	287	0	5	2	0.666	0.0116
23	Mt124331	B	3	2	1	315	0	4	2	0.666	0.0085
24	Mt124935	B	3	2	1	148	0	1	2	0.666	0.0045
25	Mt125375	B	3	2	1	190	0	1	2	0.666	0.0035
26	Mt127721	B	3	2	1	420	0	10	2	1.167	0.0159
27	Mt133126	B	3	2	1	174	0	1	2	0.666	0.0038
28	Mt6799803	B	3	2	1	354	0	4	2	0.666	0.0075
29	Mt6802127	B	3	2	1	290	0	1	2	0.666	0.0023
30	Mt6803180	B	3	2	1	92	0	1	2	0.666	0.0072
31	Mt6811198	B	3	2	1	91	0	1	2	0.666	0.0073
32	Mt6815341	B	3	2	1	149	0	2	2	0.666	0.0089
33	Mt6817377	B	3	2	1	295	0	3	2	0.666	0.0068
34	Mt6836854	B	3	2	1	244	0	3	2	0.666	0.0082

35	Ms6943512	B	3	2	1	220	0	1	2	0.666	0.003
36	LG111	C	19	9	10	626	4	64	4	0.81	0.0293
37	LG91	C	18	8	10	534	2	39	10	0.912	0.0212
38	LG95	C	20	9	11	312	0	23	2	0.816	0.0208
39	LG99	C	19	9	10	520	2	82	1	1.056	0.0451
40	LG101	C	19	9	10	460	4	119	2	1.049	0.074
41	LG83	C	17	9	8	249	3	31	1	1.063	0.0368
42	LG90	C	20	9	11	320	1	25	8	0.656	0.022
43	LG73	C	20	9	11	391	2	33	6	1.037	0.0238
44	LG80	C	20	9	11	476	3	56	7	0.823	0.0332
45	LG87	C	17	8	9	530	0	47	6	0.912	0.0262
46	LG104	C	16	8	8	609	2	52	9	0.87	0.0257
47	LG105	C	19	8	11	330	0	23	9	1.011	0.0199
48	LG103	C	15	9	6	658	1	60	3	1.057	0.028
49	LUP51	D	13	9	4	302	1	14	2	1.07	0.0149
50	LUP211	D	20	9	11	265	0	17	9	0.653	0.0181
51	LUP246	D	16	9	6	371	0	20	7	0.892	0.0162
52	LUP255	D	13	5	8	737	1	43	4	1.039	0.0188
53	LUP276	D	9	7	2	224	0	1	2	0.223	0.0016
54	LUP240	D	20	9	11	250	0	15	8	0.747	0.0169
55	LUP241	D	20	9	11	297	2	25	8	1.024	0.0237
56	LUP94	D	20	9	11	186	0	13	10	0.8	0.0197
57	LUP120	D	18	8	10	325	1	55	6	1.039	0.0492
58	LUP235	D	19	9	10	268	2	16	7	1.036	0.0171
59	LUP302	D	16	8	8	176	0	14	9	0.829	0.024
60	LUP318	D	20	9	11	221	0	10	9	0.653	0.0128
61	LUP326	D	12	9	3	264	1	37	6	0.962	0.0464
62	Gm2077934	E	3	2	1	253	0	5	2	0.666	0.0132
63	Gm2084815	E	3	2	1	123	0	1	2	0.666	0.0054
64	Gm2091985	E	3	2	1	144	0	1	2	0.666	0.0046
65	Gm2096212	E	3	2	1	446	0	1	2	0.666	0.0015
66	Gm2099239	E	3	2	1	166	0	2	3	1.001	0.008
67	Gm2103135	E	3	2	1	492	0	2	2	0.666	0.0027
68	Gm2113547	E	3	2	1	111	0	1	2	0.666	0.006
69	Gm2117188	E	3	2	1	143	0	2	3	1.001	0.0093

70	Gm2118543	E	3	2	1	185	0	3	2	0.666	0.0108
71	Gm2120918	E	3	2	1	508	1	5	2	0.666	0.0066
72	Gm2123242	E	3	2	1	328	0	3	2	0.666	0.0061
73	Gm2124498	E	3	2	1	93	0	1	2	0.666	0.0072
74	Gm2125123	E	3	2	1	410	0	3	2	0.666	0.0049
75	Gm2127540	E	3	2	1	143	0	2	2	0.666	0.0093
76	Gm2129282	E	3	2	1	157	0	6	2	0.666	0.0255
77	Gm2131909	E	3	2	1	113	1	1	2	0.666	0.0059
78	Gm2132124	E	3	2	1	367	0	5	3	1.001	0.0091
79	Gm2139030	E	3	2	1	80	0	3	3	1.001	0.025
80	Gm2076345	E	3	2	1	74	0	3	3	1.001	0.027
81	Tp6849647	F	3	2	1	144	0	4	3	1.001	0.0185
82	Tp6849720	F	3	2	1	376	0	6	2	0.666	0.0106
83	Tp6850763	F	3	2	1	472	0	2	3	1.001	0.0028
84	Tp6854083	F	3	2	1	613	0	4	3	1.001	0.0044
85	Tp6857294	F	3	2	1	494	0	4	2	0.666	0.0054
86	Tp6857470	F	3	2	1	298	0	2	2	0.666	0.0045
87	Ah6928256	G	3	2	1	772	0	9	2	0.666	0.0078
88	Ct6874464	I	3	2	1	187	0	2	2	0.666	0.0071
89	Ct6875390	I	3	2	1	427	1	7	2	0.666	0.0109
90	Ct6875951	I	3	2	1	388	1	4	3	1.001	0.0069
91	Ct6876038	I	3	2	1	185	0	1	2	0.666	0.0036
92	Pc1722835	J	3	2	1	301	0	2	2	0.666	0.0044
93	Ps1768451	K	3	2	1	257	0	6	2	0.666	0.0156
94	Ps1770858	K	3	2	1	121	0	1	2	0.666	0.0055
95	Pv4540	L	3	2	1	578	0	28	2	0.666	0.0323
96	Pv4675	L	3	2	1	613	0	2	2	0.666	0.0022
97	Rp1788051	M	3	2	1	350	1	1	2	0.666	0.0019
98	Rp1788514	M	3	2	1	127	0	3	2	0.666	0.0157

***Gene Source**: A: Chickpea EST; B: *Medicago truncatula*; C: *Lotus japonicas*; D: *Lupinus spp.*; E: *Glycine max*; F: *Trifolium pratense*; G: *Arachis hypogaea*; H: *Crotolaria tenuifolia;* I: *Phaseolus coccineus*; J: *Pisum sativum*; K: *Phaseolus vulgaris*; L: *Robinia pseudoacacia*

Annexure 4

Details on primer ID and sequences of CISR markers

S. no.	Marker name	Primer ID	Forward primer sequence (5'- 3')	Reverse primer sequence (5'- 3')	Expected product size (bp)
1	CISR001	CAAC_JG11_356TF	AAATACTGGAGATACAAGCTGATGC	AGTTCCCCCAAGAAGCTCAT	616
2	CISR002	ICC4958_CAAA529TF	TTTGTAAAACCAACAAGGTTGC	TGTACCAATAGGTGAATTGCCA	459
3	CISR003	gi|169747732|gb|FE671906.1|FE671906 mh1_0011_E07 mh1_0011_E07	ATAATTACCGGGGGTAAGCG	GAACATACGAAAGCGTTGGA	946
4	CISR004	ABCDE_Contig_1444	TAATCAAAAATGGCTGTGCG	ATTGCCCCTTTTACGCTAGG	523
5	CISR005	ABCDE_Contig_1796	ATTGATGGTCAGATGTCGCA	CCAAACCAAAATTCCAATGC	647
6	CISR006	gi|169746952|gb|FE671390.1|FE671390 mh1_0005_B09 mh1_0005_B09	TCATGACTGGCTGAAGCAAC	TTCTGGAGAGTTCCTCGCAT	575
7	CISR007	ABCDE_Contig_1421	GCTGTCCTATGATGTCAGCCT	TTCTCATTTGCAGCCACAAG	369
8	CISR008	ABCDE_Contig_1752	CCTGTGAAAGCTTGACCTCC	AGCTGATATCCATAGAACATTGC	490
9	CISR009	ABCDE_Contig_1309	CGTCCTTGGTCATTCCATCT	AATTGTTGGTGGTTGGGACA	1075
10	CISR010	ABCDE_Contig_1309	TTCCACCTAGAGGAACTCCA	ACAACCTGCGACAGTCAAG	549
11	CISR011	ABCDE_Contig_1787	TTGTCATAGGTTGAGCTTTGTG	TTAGATCTCACCGAACGCAA	553
12	CISR012	ABCDE_Contig_2242	CCCTGGCGTTGTAAATCAGT	CTACAGTTATTGGGGGCGTC	1183
13	CISR013	ABCDE_Contig_2242	AATACCAAGACACCGCATCC	CGAAGATGATGAGCAGCAGA	513
14	CISR014	ABCDE_Contig_2242	GGTATAGGCTTCCCATCCCA	CAATTGCTCGAACAAAAGACA	1574
15	CISR015	ABCDE_Contig_705	GCAGCACCCGGTAACAAG	TGGGTCTGATGATGGGATTT	1585
16	CISR016	ABCDE_Contig_929	TGTGATGGAGTGGACAGTGG	AGATTCCATGGGGTGAGGTT	865
17	CISR017	CAAC_JG11_569TF	GGCACTAGAAAAGGTGGCTG	CATGTAAACAGGGGCCATTC	992
18	CISR018	CAAC_JG11_738TF	CCTGTTGTTTTGAAAGTTCCTG	GCCATGGTTGGATATAACGG	932
19	CISR019	CAAC_JG11_C61TF	AGGTTGGCTGCTCCTGATTA	CAATCTGGCATGGGAAAAGT	1239
20	CISR020	CAAC_JG11_K48TF	CGTTGATCAGCTTTCTGCTG	CCTTCTCCAACGGATAAGCA	1423

21	CISR021	CAAC_JG11_K48TF	CGTTGATCAGCTTTCTGCTG	ACAAAATTCCCTCAAGGGCT	1423
22	CISR022	CAAF_ICCV2_O37TF	GGTCAGTGAGGATGTTGCTG	ATCTGCCTTCCTTTCCCATT	871
23	CISR023	ICC1882_ICC1882_CD68_A04	CCGTAGCTGGTCCTTTGTTG	AGATTACTCCAGCCCAACCC	427
24	CISR024	ICC4958_CD107_D09	CAACACAGTAAATCCATTCCCC	GGGAGGAGATGCTTCTGTTA	1197
25	CISR025	gi|169746344|gb|FE671368.1|FE671368 mh1_0004_H10 mh1_0004_H10	TCCAATCAATGCTACAGCCA	CAGGCAATTGTGTCAGAGGA	795
26	CISR026	gi|169746344|gb|FE671368.1|FE671368 mh1_0004_H10 mh1_0004_H10	CAGCAGCAGAAAGAAATCTGTG	ATCCCCTGTGCTGTTGTTTC	605
27	CISR027	gi|169747248|gb|FE671794.1|FE671794 mh1_0010_B02 mh1_0010_B02	AAGCGGAACCCCAACTATTC	CCAACTTCCAAGAACCTCCA	716
28	CISR028	gi|169748192|gb|FE671994.1|FE671994 mh1_0013_A05 mh1_0013_A05	ATGGTCCAACATCTTCCTGC	TTTTCTTGGAGGCTGCATTT	818
29	CISR029	ABCDE_Contig_1056	ATGCCTGCATCACCTTATCC	GAAAACACCTCCAAGGCTCA	753
30	CISR030	ABCDE_Contig_1294	AATTGAAGGCAGTGTTTGGG	AGGCAGACTATGCTGTGAAGG	921
31	CISR031	ABCDE_Contig_1379	TCAAAGGAGCCAGAATCACC	CCTGGAAACCTTCCATCAGA	1327
32	CISR032	ABCDE_Contig_1741	CTAATTGGAGCAGGAGCAGC	AACGGAGCTTGTGGTGAATC	486
33	CISR033	ABCDE_Contig_1741	AGCCAACTGTGGATGAAACC	TAGAAGTCGGACTCGGGAAC	490
34	CISR034	ABCDE_Contig_2004	GGGTTCGTGGTGCTTACAGT	ATTTCTGTGGATGCTCCCTC	1041
35	CISR035	ABCDE_Contig_24	TTTGAAAGGATTGGTACGGC	CATATGCCATGGCTCTCTCA	530
36	CISR036	ABCDE_Contig_797	TTGTCAAAAATGAGAGGATTGG	GGTTCTGACCATTTGAGGGA	433
37	CISR037	ABCDE_Contig_971	TCTGCCAAGAGAAGGTCCAT	GAGACACAAGATATTCCAGATGC	1170
38	CISR038	CAAC_JG11_X71TF	GATGAGTACTTGCCCCAAGC	TCTTTCTCGAGTTGTCTGCG	427
39	CISR039	CAAF_ICCV2_M33TF	TTTGGTTCTTCTTGGAGTTGC	CCAATAGTCAAGAAATGGACCAA	869
40	CISR040	ICC1882_ICC1882_CD69_B09	GCATTAGGCCAAAGGAACAT	GTTCAAGATCCGAGGCAAAT	391
41	CISR041	ICC1882_ICC1882_CD69_B09	GCAAGATACAGCCATCACCA	TTCATTGAAATATGTACTCAGCG	483
42	CISR042	gi|169747276|gb|FE671822.1|FE671822 mh1_0010_E02 mh1_0010_E02	ATAAGGCAACCATTGGAGCG	ATCAGCACCGCGATAGAAAG	1123
43	CISR043	gi|169747300|gb|FE671846.1|FE671846 mh1_0010_G09 mh1_0010_G09	GAAGCTTCTAAAGAGGGCCG	AGTTTTACGAGCATGAGCAGG	542
44	CISR044	ABCDE_Contig_1017	ATCAAAGGCATGGAAGCAGA	CAATAAACAGAGCATGCATAACC	381
45	CISR045	ABCDE_Contig_1017	TGGTTCAGTTGATCCAGTTACAA	TTAAATTTGTTAAGAATCCACGCA	1104
46	CISR046	ABCDE_Contig_1017	TCCAGCAACGTTGTTCTGTT	TGTTTGTCAATTTGGCGAGA	874
47	CISR047	ABCDE_Contig_1298	CGTGATCAGATTCACCAAGC	GCTTCCTCTTCTCCAACAACC	538
48	CISR048	ABCDE_Contig_1360	CTTCAAGTCTTCCAAAGCGG	TGCAAACGAATTGTGAAGGA	413
49	CISR049	ABCDE_Contig_1563	CAGGAACCTGTGCTACCTGG	CTTCTTATTTCAGCCAGCCG	684
50	CISR050	ABCDE_Contig_1696	CAACTGCCCGTGCTATACG	TTTGCTGCAGTCTCAAAATCA	885
51	CISR051	ABCDE_Contig_1696	CCCCCTATTAGGAAACAAGGA	ATTTTGCGGAAGTCGTCAAT	629

52	CISR052	ABCDE_Contig_1696	AATCCGAATTCAGGCTTTGA	TTGTCCAGTCCATCAAACACA	1045
53	CISR053	ABCDE_Contig_1714	TGCAAAAATAACGTGCAAGC	CAAGCCTAGCCCTCTCCTCT	431
54	CISR054	ABCDE_Contig_1775	ATAGCATGCACCACAACCAG	TCAGGACTCTTTTATTTGCTCG	553
55	CISR055	ABCDE_Contig_1793	ATGGCGTCGAAACGTATCTT	GGGTGGCTTGAAGGGATAAT	1616
56	CISR056	ABCDE_Contig_1793	TCCAAACATCAACAGTAATGGC	TGCATTGAACAGCTAGGACG	517
57	CISR057	ABCDE_Contig_1850	CATATTGCTCCATGAAATCATCA	GGTGCCATGTATTCATTCCC	1255
58	CISR058	ABCDE_Contig_1892	TTGGTGGCACTGAAAGTACG	GACTCTCGATGAGGGCAGTC	826
59	CISR059	ABCDE_Contig_196	CTCAATGATAACCAACAGGAGC	TGAGAATGTAGCTGCTCCTGC	395
60	CISR060	ABCDE_Contig_197	TACTTTAAGGCTTGGAATAGCAATG	GCTTGGACTTGTATAACAACAACAT	1012
61	CISR061	ABCDE_Contig_2361	TCCCTTGTGGAGGCACTACT	TGAGAGAGTGGAGACCCAGC	831
62	CISR062	ABCDE_Contig_2361	TTCACGAAGCTGTTCAGTCG	TGTCTTGACCCCCGACAT	909
63	CISR063	ABCDE_Contig_2361	TCCCTTGTGGAGGCACTACT	TGAGAGAGTGGAGACCCAGC	754
64	CISR064	ABCDE_Contig_77	GCTGGAATTTGACCATAGCC	TTGCAAAGAAGTACCATGCC	421
65	CISR065	ABCDE_Contig_950	CTGACCCCCATTCTGTGTGT	TGGAACTTTTCTGGACAGGG	494
66	CISR066	ABCDE_Contig_951	AGGATTGGGTATAGTATGATTGCTG	CATTGAAGCAGGCATCGTAA	604
67	CISR067	ABCDE_Contig_951	GTCCTCACAAGATTCAAGGGA	TTTCTGGCCTTTTGGCTATC	792
68	CISR068	ABCDE_Contig_951	GTCCTCACAAGATTCAAGGGA	AATTCAGGGCTCGAAAGTCA	792
69	CISR069	ABCDE_Contig_952	CACTGTTGGTGATGGAGCTG	GTGTCCCACAAACCCAGATT	862
70	CISR070	ABCDE_Contig_952	TGGATTCCAGAGTTGAAGCA	AGCTGTGGTAATTGGAACGG	667
71	CISR071	ABCDE_Contig_958	GAGCCAGTATCCTCGATTGG	TGCATCAAGCTTTCCACTGT	1735
72	CISR072	CAAC_JG11_I40TF	ATGCGGAGACAGCAACTTTT	ACCCAATACTCAGGAGGTCG	701
73	CISR073	CAAC_JG11_I40TF	GGTTGTTCATGATGCTGGTG	AAAGTCCCTCAAAGAACCGC	557
74	CISR074	CAAC_JG11_X71TF	GATGAGTACTTGCCCCAAGC	TCTTTCTCGAGTTGTCTGCG	420
75	CISR075	CAAC_JG11_X71TF	TGAGGAAGAAGATTCCACTGC	CTCCCCGGTTAAACCAGAA	937
76	CISR076	CAAF_ICCV2_F36TF	GGTTTAGGGGGTAAAGTGCG	GAACAGATTCCAGACACGCA	572
77	CISR077	CAAF_ICCV2_H60TF	CAGGTCAGGACGTCCAACTTAC	TGTTCCTCGCTACCAAAACC	1238
78	CISR078	CAAF_ICCV2_H60TF	GGGACGAAGTACTGCTTGGA	TGCTTTCCTCTTTCCTGTGAA	812
79	CISR079	ICC1882_CAABB42TF	TCCGTGTACTTGCCTCATCA	TTCTCCATTTCCTGCTGGTT	752
80	CISR080	ICC4958_CAAAE88TF	GTTGCTAGATGGTAGCCGGA	AGCAGAAGACACGTCAGCAA	631
81	CISR081	ICC4958_CAAAV84TF	ATTCCTGAACAACCTGACGC	GTTTCAGGTTCTGGCAAAGG	946
82	CISR082	gi|169746954|gb|FE671392.1|FE671392 mh1_0005_B11 mh1_0005_B11	TTTGGGGAGGTTTGATTTTG	TCCCTTCCAGGTATTCTCCA	1720
83	CISR083	gi|169748034|gb|FE671488.1|FE671488 mh1_0006_C02 mh1_0006_C02	ACAACCTTGCCTTCCAGTTG	TCTGTGATGGTTTCTGTGGC	1411
84	CISR084	gi|169748259|gb|FE672025.1|FE672025 mh1_0013_C12 mh1_0013_C12	TCATATATTCTGACACTTTGACCAGA	TGGCTGGACACCATTACATC	1150

85	CISR085	gi|169748729|gb|FE672087.1|FE672087 mh1_0014_A03 mh1_0014_A03	TTCCTATGACAAGGCTTCGG	CCCAATTTGACTCCATGCTT	837
86	CISR086	gi|169748729|gb|FE672087.1|FE672087 mh1_0014_A03 mh1_0014_A03	ATGCAGTCATTGATGGGGAT	TTGTAAGAGTGCCTCCAGCC	569
87	CISR087	ABCDE_Contig_1526	AGAAACCAAGCATGGCAAAC	CTTTGGAAAACCAATCCCAA	751
88	CISR088	ABCDE_Contig_1526	ACTCTGCACCTCCCAAAAGA	CTAGATGGGCAATGCTTGGT	629
89	CISR089	ABCDE_Contig_1690	TCAACAGGGTCCCTAGGAGTT	GGGTGATCCCAAAGAAGTGA	914
90	CISR090	ABCDE_Contig_1983	AAGAGCAAACGCCAACCTAC	GCAATTCATCTAAAGCTTCTTTCA	611
91	CISR091	ABCDE_Contig_824	AAGGAGGCTGCTGAGATGAA	GAAACTTCCTTCACGATTTCATC	654
92	CISR092	CAAC_JG11_U04TF	TGTTACTGGCTCCTCTGTCG	AGCTCATCAAGGTAGCGCAT	672
93	CISR093	gi|169747614|gb|FE672316.1|FE672316 mh1_0016_E04 mh1_0016_E04	ATGCCCAGAAGCAGAAAATG	TCAGAGCTAGCATCATCCACA	469
94	CISR094	gi|169748367|gb|FE672397.1|FE672397 mh1_0017_D01 mh1_0017_D01	CAATCAGAGTTTGAAGATTCAGGA	ACTGCAGCTTGTGTGTGACC	672
95	CISR095	ABCDE_Contig_1102	CCGTTGCACACTACCACAAT	CCAGTGGATGGCATTTTTCT	445
96	CISR096	ABCDE_Contig_1102	TATTGCCACTCTGGCTGTTG	TGTAAAAATCTTCAAGGTATGCTTGT	479
97	CISR097	ABCDE_Contig_1102	TATTGCCACTCTGGCTGTTG	TGTAAAAATCTTCAAGGTATGCTTGT	537
98	CISR098	ABCDE_Contig_1499	TATTGTTGGATTGTGCTGCC	ATGCTTCCATGTTGTAGGGC	920
99	CISR099	ABCDE_Contig_1724	TCCTCCTAAGGGTGGTTTCA	TGTTGGGTGCCATATAATGAA	1128
100	CISR100	ABCDE_Contig_1757	ATTAACGAAAACCCAGGGGC	CAAGCTGTCCTTGCATCAAA	1228
101	CISR101	ABCDE_Contig_1757	AACGCCTAGGAACGAACTGA	TCAAAATGAATGCAAAAACGA	435
102	CISR102	ABCDE_Contig_1962	ATGTATGCCCTCAAGCTTCC	CAAATCTGAATGTGGAGGAGG	417
103	CISR103	ABCDE_Contig_1962	CAGCTTTGTTTCCCACCAGT	GCTTACTACCGTGGAGCCAT	365
104	CISR104	ABCDE_Contig_1962	AAAGAAGACTCGTCAGTAACGTCAT	TGATGGCAAGCGAATCAAAT	1122
105	CISR105	ABCDE_Contig_2202	AAATCTCCAAACTGCCAAGAAA	CTGTGATCCTTCTTTATCTTCCTTG	355
106	CISR106	ABCDE_Contig_2266	TTGAAGCAATCACTCTGAGCC	GAGAGGATCAAAGTCGGTGG	1401
107	CISR107	ABCDE_Contig_2428	TTGGGGTTTGTTCCTTTACG	AGTTGCCTGACACGGAGATT	730
108	CISR108	CAAC_JG11_938TF	GGATGAGTTCTCTGGAGCTGA	GCAAGGGTAAATGAGGACGA	1778
109	CISR109	CAAC_JG11_K59TF	TGGTTATCACACAAACATCAACA	ATGGGATCAGGTTCACTTGC	1465
110	CISR110	CAAC_JG11_T90TF	TGTATGTGCAGAGCATCTTGAA	CCTTAGGTGTGCCTGGTGTT	357
111	CISR111	gi|169747792|gb|FE671174.1|FE671174 mh1_0002_G09 mh1_0002_G09	TTCATCACTTCAAGCAACCG	GCACACTCTTTGTTGACCCA	554
112	CISR112	gi|169748372|gb|FE672402.1|FE672402 mh1_0017_D06 mh1_0017_D06	TGTTACATGTGGTTGGTGCC	CCTCTGCATTTTTCATGTTCC	466
113	CISR113	ABCDE_Contig_422	AGGTTCAGCGCACTATGCTT	TCAGGAAGGCCAAATTCAAC	481
114	CISR114	CAAC_JG11_C75TF	TTGATTTGTCAGTTAACATAACTCTGG	GACGTAACTTTCGCTGAGGC	524
115	CISR115	CAAC_JG11_J54TF	ATATGGCAGCCAATCCTTTC	CAATACTGTCAGGGCACCAA	672

116	CISR116	CAAC_JG11_U90TF	ATTGGGAATGTGTCTCGGTG	TCGGAGTTTGATTTTTCTTCG	866
117	CISR117	ICC4958_CAAA972TF	GCGATCGAACATGAAGATGA	AAATCATTCACTTTCTTTGCCA	900
118	CISR118	ICC4958_CAAA972TF	GGGACCTTATTCCACAAGCA	CTGCTCCTCCTGCATCATTT	490
119	CISR119	ICC4958_CAAAK76TF	TGCAACTGCAGCCATCTTAC	ATGAAGGGAGCATCAAGAGG	1743
120	CISR120	gi|169747568|gb|FE672270.1|FE672270 mh1_0016_A03 mh1_0016_A03	AATTGACTGCCATGAACAAGC	CCCTTGGGGGATGTTGACTA	670
121	CISR121	gi|169747568|gb|FE672270.1|FE672270 mh1_0016_A03 mh1_0016_A03	TTTTGTCCATTGCAGAAGAAA	CCCTAGTGGGTTTTAGTCCC	406