Review of Literature
CHARACTERISTICS OF THE CHICKPEA CROP

Classification

Chickpea (*Cicer arietinum* L.) is a self pollinating diploid species (2n=2x=16) with a genome size of 7.38 x 10^8 bp (Arumuganathan and Earle, 1991). The genus *Cicer* belongs to family Fabaceae, subfamily Papilionaceae and forms its own tribe, Cicerea Alef. It comprises more than 43 species, nine of which (including the cultigen, *Cicer arietinum*) are annual; 33 are perennial and one unspecified (van der Maesen, 1987). All the species are grouped into four sections based on morphological characteristics, life cycle and geographical distribution: *Monocicer*, *Chamaecicer*, *Polycicer* and *Acanthocicer* (van der Maesen, 1987). All annual species except *C. chorassanicum* (*Chamaecicer*) belong to the section *Monocicer*. Annual species have been subdivided into four groups on the basis of crossability (Ladizinsky and Adler, 1976a), karyotype (Ocampo et al., 1992; Abbo et al., 1994), isozyme polymorphisms (Tuwafe et al., 1988; Gaur and Slinkard 1990a, 1990b; Labdi et al., 1996) and seed storage protein characteristics (Ahmad and Slinkard, 1992). The first group contains the cultigen, its presumable ancestor *C. reticulatum* (Ladizinsky, 1975) and *C. echinospermum*. Both wild species produce fertile hybrids in crosses with the cultigen, though fertility barriers between chickpea and *C. echinospermum* do exist (Singh and Ocampo, 1997). One perennial species, *C. anatolicum*, first grouped into section *Polycicer*, has also been placed into the first group based on isozyme similarities (Kazan and Muehlbauer, 1991). The second group comprises *C. bijugam*, *C. pinnatifidum*, *C. judaicum* and *C. yamashitae*. The remaining two species, *C. cuneatum* and *C. chorassanicum*, can be crossed neither with each other nor with any other species and, therefore, make up the third and fourth group respectively (Kazan and Muehlbauer, 1991).

Distribution and yield

Chickpea ranks third among the world's food legumes or pulse crops (FAO, 1999). It is grown in an area of about 11.1 million ha, with a total annual production of 9.1 million tons. The Indian subcontinent (India, Pakistan, Myanmar, Bangladesh and
Nepal) accounts for about 80% of the global production while the rest is produced in eastern Africa, Mediterranean and Near East countries, Australia, southern Europe, and North and South America.

**Ecology**

Chickpea is a self-pollinated crop. Cross-pollination is rare; only 0-1% is reported (Singh, 1987; Smithson et al., 1985). It is grown usually as a rain-fed cool-weather crop or as a dry climate crop in semi-arid regions. Optimum conditions include 18-26°C day and 21-29°C night temperatures and annual rainfall of 600-1000 mm (Duke, 1981; Smithson et al., 1985). Frost, hailstones, and excessive rains damage the crop. Though sensitive to cold, some cultivars can tolerate temperatures as low as -9.5°C in early stages or under snow cover. Daily temperature fluctuations are desired with cold nights with dewfall. Relative humidity of 21-41% is optimum for seed setting.

**Origin**

van der Maesen (1972) believed that the *Cicer* species originated in the southern Caucasus and northern Persia. However, Ladizinsky (1975) reported the center of origin to be southeastern Turkey. van der Maesen (1987) recognized the southeastern part of Turkey adjoining Syria as the possible center of origin of chickpea based on the presence of the closely related annual species, *C. reticulatum* Ladizinsky and *C. echinospermum* P.H. Davis. The earliest excavated chickpea remains have been dated to the Pre-Pottery Neolithic B period of a number of Near-East sites and the earliest remains of chickpea seeds were unearthed from archaeological digs within or near the known distribution range of *C. reticulatum* (Zohary and Hopf, 1993). Due to a very limited distribution of the wild progenitor, the common view is that chickpea was domesticated somewhere in the west arch of the Fertile Crescent alongside the rest of founder crops of the Near-East Neolithic agriculture (Zohary and Hopf, 1993; Lev-Yudan et al., 2000). It is worth noting that the range of *C. reticulatum* is the only region in the Fertile Crescent where all the wild progenitors of the founder crops of the Near-East Neolithic agriculture grow together. This includes the wild species of
Review of literature
diploid and tetraploid wheat, barley, lentil, pea and flax as well as wild rye (Lev Yadun et al., 2000). In India, the earliest occurrence of chickpea dates back to 2000 BC at Atranjikhera in Uttar Pradesh, although it may have been introduced independently to the southern parts of the country by sea (Chowdhury et al., 1971; van der Maesen, 1987).

Among the eight wild annual *Cicer* species, two that are native to eastern Turkey are closely related to the cultigen. The first, *Cicer echinospermum* P.H.Davis (echinate seed coat) grows in steppe plant formation on soils of basaltic origin. The second closely related species is *Cicer reticulatum* Landiz. (reticulate seed coat), which is found in oak shrub formations on hilly limestone bedrock (Ladizinsky, 1975). Based on meiotic chromosome pairing data *C. reticulatum* was suggested as the immediate wild progenitor of domesticated chickpea (Ladizinsky and Adler, 1976a, 1976b). This early identification is also supported by seed storage protein profiles (Ladizinsky and Adler, 1975) and by more recent morphological comparisons (De Leonardis et al., 1996) as well as by DNA marker analysis (Patil et al., 1995).

*C. reticulatum* was first collected and described in 1975 (Ladizinsky, 1975). Ever since only 10 populations have been located in southeast Turkey (Ladizinsky, 1995). So far, not all the ten accessions have been utilized in genetic analysis (Gaur and Slinkard, 1990a, 1990b; Singh and Ocampo, 1997).

Vavilov (1950) suggested two primary centers of chick diversity, Southwest Asia and the Mediterranean center, and designated Ethiopia as a secondary center. He observed that large seeded varieties were cultivated in the Mediterranean basin and progressively small seeded (*desi*) varieties abounded eastward in the Indian subcontinent. It is believed that large seeded chickpea was introduced into India through Kabul, Afghanistan (therefore named *kabuli*) in the mid-to-late 17th century. The spread of chickpea to tropical Africa, North and South America, and Australia has occurred in more recent times. Wild species are most abundant in Turkey, Iran, Afghanistan, and Central Asia (Duke, 1981).
Composition

Chickpea seed has 38-59% carbohydrate, 3% fiber, 4.8-5.5% oil, 3% ash, 0.2% calcium, and 0.3% phosphorus. Digestibility of protein varies from 76-78% and its carbohydrate from 57-60% (Hulse, 1991, Huisman and van der Poel, 1994). Raw whole seeds contain per 100 g: 357 calories, 4.5-15.69% moisture, 14.9-24.6 g protein, 0.8-6.4 % fat, 2.1-11.7 g fiber, 2.4-8 g ash, 140-440 mg Ca, 190-382 mg P, 5-23, 9 mg Fe, 0-225 μg β-carotene equivalent, 0.21-1.1 mg thiamin, 0.12-0.33 mg riboflavin, and 1.3-2.9 mg niacin (Duke, 1981; Huisman and van der Poel, 1994). Sprouting is said to increase the proportionate amounts of ascorbic acid, niacin, available iron, choline, tocopherol, pantothenic acid, biotin, pyridoxine, inositol, and vitamin K. The leaves of the plants exude malic and oxalic acid. Wild species often have similar glandular secretions (Duke, 1981). Percent fatty acid compositions are: Desi: oleic 52.1, linoleic 38.0, myristic 2.74, palmitic 5.11, and stearic 2.05; Kabuli: oleic 50.3, linoleic 40.0, myristic 2.28, palmitic 5.74, stearic 1.61, and arachidic 0.07%. The leaves contain 4-8% protein (Duke, 1981).

Uses

Chickpea is grown in tropical, sub-tropical and temperate regions. Kabuli type is grown in temperate regions while the desi type chickpea is grown in the semi-arid tropics (Muehlbauer and Singh, 1987; Malhotra et al., 1987). Chickpea is valued for its nutritive seeds with high protein content, 25.3-28.9 %, after dehulling (Hulse, 1991). Chickpea seeds are eaten fresh as green vegetables, parched, fried, roasted, and boiled; as snack food, sweet and condiments; seeds are ground and the flour can be used as soup, dal, to make bread and as a nutritive supplement added to wheat flour; prepared with pepper, salt and lemon and served as a side dish. Dal is the split chickpea without its seedcoat, dried and cooked into a thick soup or ground into flour for snacks and sweetmeats (Hulse, 1991). Sprouted seeds are eaten as a vegetable or added to salads. Young plants and green pods are eaten like spinach. A small proportion of canned chickpea is also used in Turkey and Latin America. Animal feed is another use of chickpea in many developing countries. An adhesive may also be prepared; although not water-resistant, it is suitable for plywood. Gram husks, and
green or dried stems and leaves are used for stock feed; whole seeds may be milled directly for feed. Leaves are said to yield indigo like dye. In Chile, a cooked chickpea-milk (4:1) mixture is good for feeding infants, to effectively control diarrhea. Chickpeas yield 21% starch suitable for textile sizing, giving a light finish to silk, wool, and cotton cloth (Duke, 1981). Among the food legumes, chickpea is the most hypocholesteremic agent; germinated chickpea was reported to be effective in controlling cholesterol level in rats (Geervani, 1991). Medicinal applications include use for aphrodisiac, bronchitis, catarrh, cutamenia, cholera, constipation, diarrhea, dyspepsia, flatulence, snakebite, sunstroke, and warts. Chickpea seeds are considered antibilious (Duke, 1981).

**Constraints on productivity of chickpea**

Chickpea is predominantly grown under rain-fed conditions in a post-rainy season, on marginal lands, often without monetary inputs. The crop therefore is vulnerable to various biotic and abiotic constraints, which occur under these conditions. The main fungi that affect chickpea are *Fusarium oxysporum* Schlechtend.:Fr. f. sp. *ciceris* (Padwick) Matuo & K. Sato, causing the plant to wilt and *Ascochyta* blight caused by *Ascochyta rabiei*. Ascochyta blight is the most serious disease in North India, Pakistan, the U.S. and the Middle East (sometimes causing 100% losses) (Smithson et al., 1985). Blight causes brown spots on leaves, stems, pods and seeds (Kaiser, 1992). Other fungi known to attack chickpea include leaf spot (*Alternaria* sp.), *Ascochyta pisi*, rust (*Uromyces ciceris-orientini*), gray mould (*Botrytis cinera*), powdery mildew (*Levillula 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 (Duke, 1981; Smithson et al., 1985; Kaiser, 1988; van Emden et al., 1988). Pod borer (*Helicoverpa armigera*), the most important pest, feeds on leaves and developing seeds (Smithson et al., 1985). Cutworms (*Agrotis* sp.), lesser armyworms (*Spodoptera exigua*), leaf minor, groundnut aphid (*Aphis craccivora*), pea aphid (*Acythosiphon pisum*), cowpea bean seed beetle
Review of literature

(Callosobruchus maculatus), and adzuki bean seed beetle (C. chinensis) are also important. Many storage insects specifically Bruchid sp. are a serious pest of stored chickpea. Callosobruchus chinensis lowers seed viability. 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 (van Emden et al., 1988).

Among the abiotic factors, drought stands to be the number one problem in major chickpea growing regions because the crop is grown on residual moisture and is eventually exposed to terminal drought (Johansen et al., 1994). In west Asia and North African countries, low temperature causing freezing injury or death or delayed onset of podding reduces yield tremendously (Singh, 1987). Heat and salinity problems are relatively important following drought and cold stresses (Singh et al., 1994).

MOLECULAR MARKERS FOR PLANT GENOME ANALYSIS

Molecular (DNA) markers reveal sites of variation at the DNA sequence level. Unlike morphological markers, these variations do not always show themselves in the phenotype, and each might be nothing more than a single nucleotide difference in a gene or a piece of repetitive DNA. DNA markers are rapidly being adopted by crop improvement researchers globally as an effective and appropriate tool for basic and applied studies addressing biological components in agricultural production systems.

Origin of DNA markers

Before the advent of DNA markers, protein markers, including seed storage proteins, structural proteins, and isozymes (Smith and Smith, 1992) were exploited for genetic diversity assessment and genetic linkage map development. Protein markers are the basis for a newly emerging research area called proteomics. The major limitations of these markers are that much of the genome (including much of the most polymorphic portions of it that are less subject to evolutionary restrictions) does not code for genes, different biochemical procedures are required to visualize allelic differences for enzymes having different functions, and, many proteins are several post transcriptional steps away from underlying DNA sequence polymorphism and
thus can mask variation present at that level (e.g., differences in tri-nucleotide sequences coding for the same amino acid, intron sequences that are post transcriptionally removed from the mRNA, and post translational modifications can all contribute to reduced polymorphism expression at the protein level compared to that at the DNA level).

In 1980, a landmark paper outlined the basic principle, which would represent the foundation of an explosion in molecular genetic analysis of plants and animals (Botstein et al., 1980). It was realized that the utility of restriction enzymes, in providing physical landmarks in the DNA, extended far beyond molecular cloning. The concept was that mutations in restriction sites, or mutations which altered the distance between adjacent restriction sites, could be visualized as "DNA markers." This technique referred to, as restriction fragment length polymorphism (RFLP) has been the basis of an explosion of genetic mapping activity.

**Molecular Basis of DNA Markers**

DNA markers arise as a result of several different classes of mutations. The simplest event is substitution of as little as a single nucleotide, differentiating two genotypes. When a base substitution eliminates a restriction site, it changes the length of DNA fragments detected by the relevant assay, and thus represents a discrete marker which is directly representative of an individual's genotype. In assays based on the polymerase chain reaction, base substitution within the region to which a PCR primer would (otherwise) bind has exactly the same effect.

Alternatively, rearrangements in the DNA intervening between two restriction sites, or two priming sites, can generate DNA markers. Such rearrangements might include insertion or excision of mobile DNA elements, or error in replication of arrays of tandemly repeated DNAs. Rearrangements tend to create DNA markers that can be detected by several different restriction enzymes, or different PCR primers flanking the same region, while base substitutions are specific to particular restriction enzymes or PCR primers.
Types of DNA markers

Most DNA markers represent anonymous DNA polymorphism [(e.g., restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and microsatellite markers)] and do not correspond to any gene of known function. However, some markers (including cDNA and expressed sequence tags (ESTs), as well as the protein markers) do pinpoint individual genes. DNA markers are visualized either by use of radioactivity (autoradiography), fluorescence, or by direct chemical staining of the DNA itself. DNA markers are generated by a wide variety of techniques, differing greatly in their reliability (repeatability and robustness), difficulty of operation, expense, and the nature of the polymorphism that they detect. Because of these differences, they also vary greatly in their suitability for various uses. They may be hybridization based (e.g., RFLP), or polymerase chain reaction (PCR) based (e.g., RAPD and AFLP); they may detect single locus or multiple locus differences; and the markers detected may be inherited in a presence/absence, dominant, or co-dominant manner. A number of widely used DNA marker techniques are AFLP (Amplified fragment length polymorphism), AP-PCR (Arbitrarily primed PCR), DAF (DNA amplification fingerprinting), ESTs (Expressed sequence tags), ISSR (Inter-SSR amplification), RAPD (Random amplified polymorphic DNA), RFLP (Restriction fragment length polymorphism), retrotransposon-based markers, SCAR (Sequence characterized amplified region), STMS (Sequence tagged microsatellite sites) etc.

Amplified fragment length polymorphism (AFLP)

Although RFLP remains a widely used technique, two basic limitations have motivated development of several alternative technologies. The first limitation is the quantity of DNA required. RFLP analysis typically requires 50-200 micrograms of DNA per individual, to generate a DNA fingerprint of the entire genome. Large-scale extraction of DNA is tedious and laborious whereas use of smaller amount of DNA would increase sample throughput. In contrast to RFLPs, PCR based techniques require only about 10% of this amount of DNA, as template. RAPDs have been reported to be an easy, inexpensive and fast PCR technique. Yet, the reproducibility of
RAPDs has been reported to be problematic, probably due to the unpredictable behaviour of short primers (10-mer) in PCR reactions (He et al., 1994).

AFLP, first developed by Vos et al. (1995) is reproducible, high multiplex assay with the ability to generate a large number of polymorphic genetic loci. The multiplex assay refers to the number of different genetic loci that may be simultaneously analyzed per experiment. AFLP is a DNA fingerprinting technique that combines the advantages of the classical hybridization based fingerprinting (Restriction Fragment Length Polymorphism) and PCR based fingerprinting (Rapid Amplified Polymorphic DNA).

The power of AFLP is based upon the molecular genetic variations that exist between closely related species, varieties or even cultivars. These variations in DNA sequence are exploited by this technique such that “fingerprints” of particular genotype can be routinely generated. Fingerprints can be produced without prior sequence knowledge using a limited set of generic primers. The primers consist of a longer fixed portion (~15 base pairs) and a short (2-4 base pairs) random (selective) portion. The fixed portion gives the primer stability (and hence repeatability) and the random portion allows it to detect many loci. The number of fragments detected in a single reaction can be “tuned” by selecting specific primer sets. The AFLP technique is robust and reliable because stringent reaction conditions are used for primer annealing. Polymorphism is detected as band presence/absence (so it is usually interpreted as dominantly inherited, although reports of co-dominant inheritance have also been made). AFLP markers are often inherited as tightly linked clusters in centromeric and telomeric regions of chromosomes, but randomly distributed AFLP markers also occur outside these clusters.

The major steps involved in the technique are (i) generation of AFLP libraries by restriction endonuclease digestion of the DNA followed by ligation of adapters, (ii) amplification of the restriction fragments and (iii) gel analysis of the amplified fragments (Fig. 1).
Microsatellite markers

In 1989, mammalian geneticists described a new technique based upon rapidly evolving regions of the genome (Weber and May, 1989). It had been previously known that arrays of short DNA elements repeated in tandem tend to be imprecisely replicated during DNA synthesis (replication slippage), and frequently generate new alleles with different numbers of repeating units (Tautz, 1989). Repeating units consisting of 1-5 bases are referred to as microsatellites and also known as simple sequence, simple sequence repeats (SSRs), simple repetitive DNA sequences, short tandem repeats (STRs), and simple sequence motifs (SSMs). In order to exploit microsatellites to detect variations, several groups introduced microsatellite amplification by locus-specific PCR (Litt and Luty, 1989; Smeets et al., 1989; Tautz, 1989; Weber and May, 1989). This strategy, was called STMS analysis (sequence tagged microsatellite sites, Beckmann and Soller, 1990). The technique involves using the repeated motif as a probe against genomic or cDNA libraries to identify a clone in which it is present. These clones are then end-sequenced and primers are designed against the unique DNA flanking the repeated sequence. These primers are then used to amplify the repeated sequence. Polymorphism is usually due to differences in length of the amplified product (Fig. 2). Many alleles are available for many of these marker loci (co-dominance). The method is highly repeatable, identifies a single locus, and targets hypervariable regions of the genome.

Several lines of evidence suggest that repetitive sequences may have functional significance in eukaryotic genomes (Flavell, 1986; Ting, 1995). In plants, it has been found that cell and nuclear size as well as duration of mitosis and meiosis are correlated with the nuclear DNA content. Tandemly repetitive sequences are frequently associated with chromosomal landmarks such as centromeres and telomeres (Bedbrook et al., 1980; Miklos, 1985). This suggests that repetitive sequences might play a role in chromosome pairing, meiotic recombination, or speciation (Brutlag, 1980; Flavell, 1982). However, some repetitive elements appear to have no functional significance, and simply accumulate in the genome because they do not produce any phenotypic disadvantage.
Microsatellites have been used for understanding genome evolution, analysis of phylogenetic relationships among related taxa, genetic mapping, DNA fingerprinting, physical mapping and gene cloning.

Advantages of DNA markers

DNA markers offer numerous advantages over the conventional (morphological) marker systems. DNA markers provide a choice of an unlimited number of visible traits, they are independent from environmental factors, there is no effect of temporal and spatial expression, a decreased number of breeding generations are required, they provide a uniform method of scoring, there is no need to use phenotypic scoring until the end, they are informative of the percentage of genome inherited from each parent, and they identify specific chromosomal regions inherited from each parent.

Desirable properties of molecular markers

The basic attributes of a reliable molecular marker include a high level of polymorphism, co-dominant inheritance, unambiguous designation of alleles, frequent occurrence in the genome, even distribution throughout the genome, selectively neutral behavior (no pleotropic effect), easy access (no cloning), easy and fast assay (by procedure amenable to automation, high reproducibility, easy exchange of data between laboratories, and development at reasonable cost) (Weising et al., 1998).

Applications of DNA markers

DNA markers have made a potential impact on various crop improvement programs. These include assessment of genetic diversity and derivation of evolutionary relationships, mapping complex traits relevant to crop improvement, marker assisted selection etc. in plant breeding programs, map-based cloning of important genes and cultivar identification and plant variety protection.
Detection of Genetic diversity in plants using molecular markers

The diversity of species in an environment has been shown to contribute to the sustainability and productivity of the ecosystem. Reduction of biological diversity may result in increased losses of nutrients from the soil and lower the productivity of a plant community. Further, plant genetic resources provide the foundation for the maintenance and improvement of crop agriculture. Throughout the course of history, plant genetic resources have been acquired, exchanged, selected, improved and preserved.

Without a continued source of variability, the ability of plant breeders to improve agronomic performance that is based on complex genetic combinations could decline. Diversity can allow farmers the opportunity to hedge genetically based risks with respect to annually unforeseen difficulties that accrue from environmental, pest and disease pressures (Smith and Smith, 1992). Because it is impossible to directly survey plant populations for all possible gene and genotypic combinations that may contribute to valuable traits, data from DNA markers can be used as a substitute with the assumption that the diversity of marker loci directly reflects the diversity of useful genes (Schoen and Brown, 1993).

The analysis of genetic variation or diversity in plants has been traditionally assessed by analysis of morphological or biochemical traits. The assessment of phenotype may not be a reliable measure of genetic difference because of the influence of environment on gene expression. The analysis of the plant genome using DNA markers allows the direct assessment of variation in genotype.

Most genetic resource collections are being expanded to incorporate wild and weedy relatives of crop species. Genes introduced from wild relatives constitute an important source of variation for the improvement of domesticated species. Conservation of these wild and weedy accessions is, and will be, particularly challenging because these entries are often highly heterogeneous and heterozygous. The application of DNA marker based strategies for characterization can help maximize the genetic diversity of wild and weedy material selected for inclusion in collections and in core subsets of collection’s genes (Schoen and Brown, 1993). DNA markers are further being used in molecular phylogenetics in determining evolutionary
relationships between and within various species. The basic principle underlying this is that if genomes evolve by the gradual accumulation of mutations, then the amount of nucleotide sequence difference between a pair of genomes should indicate how recently those two genomes shared a common ancestor. Two genomes that diverged in the recent past would be expected to have fewer differences than a pair of genomes whose common ancestor is more ancient. This means that by comparing three or more genomes with each other it should be possible to work out the evolutionary relationships between them.

**Origin and maintenance of genetic variation in plant populations**

Genetic variation in plant populations may be caused and maintained by a variety of mechanisms including mutation, recombination, migration and gene flow, genetic drift and genetic selection.

Mutation: Single base mutations can arise by simple base substitution or by insertion or deletion of a nucleotide. A single base substitution may alter a single amino acid in a protein or may be silent, depending on the position within the codon. The insertion or deletion of a single base can cause a frame shift and can result in a major change in the protein encoded and the resultant phenotype. Changes in large numbers of nucleotides may result from the duplication of segments, their inversion or deletion. All of these changes may occur at random and have dramatic effects or result in no change, depending on the presence of important genes in the DNA segment undergoing mutation.

Recombination: Mating results in the recombination of alleles in individuals and allows the spread of new alleles arising by mutations throughout the population. Plants that reproduce sexually may be inbreeding (selfing) or outbreeding (outcrossing) or may be intermediate. Outbreeding may be forced by self-incompatibility. Other plants reproduce asexually, eliminating recombination as a mechanism for the generation of genetic variation.

Migration: Genes may move from one population to another by migration. This may often occur by dispersal of seed.
Genetic Drift: Small populations allow random or chance changes, resulting in the loss of rare alleles.

Selection: The frequency of alleles in a population may change as a result of natural selection acting on alleles that confer either an advantage or disadvantage to the survival and reproduction of the individual in which they occur.

Types of molecular markers used for genetic diversity analysis

The complete array of techniques used for analysis of DNA can be applied to the assessment of genetic variation in plants. These techniques may be divided into those that involve molecular marker methods and those based upon the comparison of gene sequences at specific loci. The DNA marker methods may also be used for plant identification. The analysis of differences in sequences at specific loci may be used to develop a specific diagnostic test for identification purposes but is more commonly applied to the analysis of relative genetic distances in phylogenetic analysis. In many poorly described systems an analysis of genetic variation must precede the development of diagnostic methods to distinguish the genetic groups present. The type of molecular method used to measure genetic distances in plants will vary depending upon the magnitude of the genetic differences being assessed.

The application of the various types of markers in the assessment of diversity among genus, species, accessions within a species, or parental varieties from breeding programs will vary according to a number of criteria. The choice of the most appropriate type(s) of marker system depends on various factors which include the type of diversity information needed, the level of variation expected or indicated, the accessibility of probes and primer sets, the time constraints of the specific project, and the level of operational and financial investment available.

Phylogenetic trees

In phylogenetic studies, the evolutionary relationships among a group of organisms are illustrated by means of a phylogenetic tree. A phylogenetic tree is a graph composed of nodes and branches, in which only one branch connects any two adjacent nodes. The nodes represent the taxonomic units, and the branches define the
relationships among the units in terms of descent and ancestry. The branching pattern of a tree is called topology. The branch length usually represents the number of changes that have occurred in that branch. The taxonomic units represented by the nodes can be species, populations, individuals, or genes.

When dealing with phylogenetic trees, it is required to distinguish between external nodes and internal nodes. External nodes represent the extant taxonomic units under comparison and are referred to as operational taxonomic units (OTUs). Internal nodes represent ancestral units.

Phylogenetic trees can be either rooted or unrooted. In a rooted tree there exists a particular node, called root, from which a unique path leads to any other node. The direction of each path corresponds to evolutionary time, and the root is the common ancestor of all the OTUs under study. An unrooted tree is a tree that only specifies the relationships among the OTUs and does not define the evolutionary path.

Recent genetic diversity studies in plants using AFLP

Cereals

Rice: In the case of rice, the biodiversity of 57 rice accessions originating from South and Southeast Asia has been studied using 4 primer pair combinations (Zhu et al., 1998). A total of 410 bands were generated of which 44% were polymorphic. These markers could distinguish all the 57 accessions. Markers from only 5 combinations covered most areas of rice chromosomes. These AFLP markers appeared not to be confined to any particular regions of rice chromosomes. They concluded that AFLP markers have a wide coverage throughout the rice genome.

Wheat: The level of genetic diversity in eleven German and Austrian winter wheat cultivars has been estimated using 16 AFLP primer combinations along with RFLPs and SSR (Bohn et al., 1999). The marker index has been found to be low for SSR and RFLPs but high for AFLP, which is attributed to the large number of loci detected per AFLP combination. It was possible to identify each wheat cultivar with the aid of a single AFLP assay. In this study, 11 primer combinations yielded a total of 599 fragments of which 117 were polymorphic across all eleven winter cultivars.
Barley: Three separate groups have analyzed the extent of genetic diversity in various barley accessions. Ellis et al. (1997) have studied 41 barley cultivars using 14 AFLP primer combinations that revealed a total of 300 polymorphic amplification products. Most of the primer combinations uniquely identified in excess of 90% of the genotypes assayed and 2 primer combinations uniquely identified all the genotypes. Russell et al. (1997) have used 6 AFLP primer combinations on 18 cultivated barley accessions. A total of 297 AFLP bands were scored and only 46.8% were polymorphic. In this case nearly all primer combinations were able to discriminate between the 18 accessions used. In a study by Qi and Lindhout (1997), 96 primer combinations have been used on 2 barley lines to check the efficiency of primer combinations to detect polymorphism. They observed that out of 96, 48 combinations were able to detect sufficient polymorphism. These 48 combinations were then used on 16 barley lines, which represented a wide range of genetic variation in barley. One hundred and six AFLP bands were observed among these 16 lines of which 67% were polymorphic. They concluded that AFLP is an extremely efficient technique for marker generation in barley.

Tef: Tef (Eragrostis tef (Zucc)) is a major cereal crop of Ethiopia and is also grown as forage crop in several other countries including USA. Forty-seven accessions of Tef and two of their wild relatives were analyzed using 11 AFLP primer combinations (Bai et al., 1999). A total of 316 polymorphic DNA fragments were detected. The level of polymorphism was found to be low (18%). In tef, AFLP was able to detect an average of 83 DNA fragments per reaction, which is 13 times more than that detected by RAPD and at least 20 times more than RFLP.

Foxtail millet: Foxtail millet (Setaria italica) is a staple food crop of China and has been cultivated in Northern China since Neolithic times (8000 years ago). A total of four AFLP primer combinations were used to assess genetic diversity and patterns of geographic variation among 39 accessions of foxtail millet and 22 accessions of green foxtail millet. One hundred and sixteen polymorphic bands were obtained (Ennequin et al., 2000). Data from this study revealed that no specific geographical structure could be extracted.
Legumes

Soybean: A total of 759 AFLP fragments were detected using 5 primer combinations in a sample of 23 accessions of wild and cultivated soybean (Maughan et al., 1996). Of the total fragments 36% were polymorphic. Within the G. max group, adapted soybean cultivars were tightly clustered, illustrating the relatively low genetic diversity present in the cultivated soybean.

Lentils: In the case of Lens culinaris, four AFLP primer combinations were used to evaluate and study the diversity and phylogeny of 54 lentil accessions representing 6 populations of cultivated lentil and its wild relatives (Sharma et al., 1996). The four combinations yielded 23, 25, 52, and 48 polymorphic bands respectively. The use of 148 AFLPs arising from 4 primer combinations was able to discriminate between genotypes, which could not be distinguished using 88 RAPDs. The level of variation detected within the cultivated lentil with AFLP analysis indicated that it is a more efficient marker technology than RAPD analysis for lentils.

Pea: Three primer combinations detected a total of 106 bands of which 47.8% were polymorphic (Lu et al., 1996). They have compared DNA based RFLP and various PCR based techniques regarding their informativeness and applicability for genetic diversity analysis. The result show that PCR based techniques could replace RFLP in the estimation of genetic diversity.

Wild Bean: The genetic structure between and within gene pools of a core collection of wild Phaseolus vulgaris L. using the technique of AFLP was determined (Tohme et al., 1996). Two primer combinations produced 203 bands of which over 90% were polymorphic. Major groups or gene pools have been recognized according to their geographical distribution.

Azuki: Yee et al. (1999) have studied genetic similarity among selected azuki (Vigna angularis) accessions from distinct production regions of Asia. They have used 19 AFLP primer combinations to analyze 58 azuki accessions. These generated 1158 scorable bands of which 18% were polymorphic. This study showed that AFLP is more efficient than RAPD in detecting polymorphism in azuki. Cluster analysis revealed groupings among accessions; however, no apparent correlation with putative geographical origins was detected.
Medicinal Plants

Neem: *Azadirachta indica* A. Juss., or neem, is a multipurpose tree which is commercially important for its medicinal and biopesticidal properties. Genetic diversity was estimated in 37 neem accessions from different eco-geographical regions of India and four exotic lines from Thailand using AFLP markers (Singh et al., 1999). Seven primer combinations generated 422 amplification products of which 69.8% were polymorphic. Cluster analysis separated 37 Indian genotypes from 4 Thai lines and indicated that the neem germplasm within India constitutes a broad genetic base.

*Withania*: *Withania* is an important medicinal plant due to its anticancerous properties. Seven AFLP primer combinations were used to investigate the inter- and intra-specific genetic variation present in 35 individuals of *W. somnifera* and 5 individuals of *W. coagulans* (Negi et al., 2000). A total of 520 bands were obtained of which an average of 52% bands were polymorphic among *W. somnifera* individuals. However, this value rose to 82% when *W. coagulans* was included. An important finding was the presence of three types namely Kashmiri, Nagori and an intermediate type in *W. somnifera*.  

Model plant *Arabidopsis*:

Two separate studies on assessment of genetic diversity within and between *Arabidopsis thaliana* ecotypes have been reported. The first of these is by Breyne et al. in 1999 who have assessed the degree of genetic diversity within and between 21 *Arabidopsis thaliana* ecotypes using two different enzymes combinations (EcoRI + MseI and SacI + MseI) and a total of 13 AFLP primer combinations. Their findings indicate that the choice of enzymes may influence the final result. The other study by Erschadi et al. (2000) subjected 19 different *Arabidopsis* ecotypes to analysis using 15 AFLP primer combinations. A total of 471 polymorphic bands were detected. The data obtained revealed that within this small set of ecotypes a group of ecotypes and a further single ecotype exhibit considerable genetic diversity in comparison to others.

Genetic diversity studies have been done on many other plants using AFLP. These include carrot (Shim and Jorgensen, 2000), grapevine (Cervera et al., 1998), cotton (Pillay and Myers, 1999), hop (Hartl and Seefelder, 1998), *Alstromeria* (Han et al., 2000), potato (Milbourne et al., 1997), willows (Barker et al., 1999), chicory
Review of literature

(Kiers et al., 2000), strawberry (Graham et al., 1996), lettuce (Hill et al., 1996),
coconut (Teulat et al., 2000), tea (Paul et al., 1997), loliwm (Cresswell et al., 2001),
banana (Loh et al., 2000), invasive weed Rubus alcefolius (Amsellem et al., 2000),
Morus (Sharma et al., 2000) and cassava (Elias et al., 2000).

Genetic Diversity studies in Cicer

The genetic diversity studies involving Cicer have mainly been focused on
either detecting diversity among various species of Cicer or within various accessions
of the cultigen C. arietinum (chickpea).

Attempts to introgress useful genes from wild related species to the cultivated
chickpea have been hampered by a limited understanding of species relationships
(Tayyar and Waines, 1996) as well as by strong post fertilization barriers (Ahmad et
al., 1988). In the past, the question of species' relationships have been investigated by
plant morphology (Robertson et al., 1997), karyotype analysis (Ocampo et al., 1992;
Tayyar et al., 1994), crossability studies (Ladizinsky and Adler, 1976a; Ahmad et al.,
1987; Singh and Ocampo, 1993), seed storage protein fractionation (Ahmad and
Slinkard, 1992) and isozyme analysis (Kazan and Muehlbauer, 1991; Ahmad et al.,

A total of 8 RAPD primers were used to distinguish 9 different Cicer taxa
representing the cultivated chickpea and 8 other related annual wild species (Ahmad,
1999). These 8 primers produced 115 RAPD markers. Four distinct clusters were
observed and reported. The first cluster comprised C. arietinum, C. reticulatum, C.
echinospermum followed by C. chorasanicum and C. yamashitae which comprised the
second cluster. C. pinnatifidum, C. judaicum and C. bijugam formed the third cluster.
C. cuneatum did not cluster with any of the species and was most distantly placed
from the cultivated species. The results are mostly concordant with previous studies.
Certain species-diagnostic amplification products specific to each of the nine species
were also identified.

In a recent study, 90 STMS markers were used to derive a phylogenetic tree
among 39 accessions of 8 annual and one accession of a perennial species of the genus
Cicer (Choumane et al., 2000). The phylogenetic tree revealed that chickpea and the
other members of its crossability group were more closely related to the perennial *C. anatolicum* than to other annual species of genus. The data has also been used to estimate the age of section *Monocicer* to be about 100,000 years.

**Genetic diversity within *C. arietinum***

The second aspect involves the analysis of genetic diversity present among various accessions of the cultigen *C. arietinum*. It has been shown that isozymes (Kazan and Meulhbauer, 1991; Labdi et al., 1996), and conventional RFLPs and RAPDs (Udupa et al., 1993) are inefficient in detecting polymorphism among chickpea accessions. Sant et al. (1999) have shown that oligonucleotides representing microsatellites can detect high level of polymorphism. They used 25 microsatellite-enzyme combinations to detect genetic diversity among 29 elite chickpea accessions. Further, in 1999, Udupa et al. have shown that there is a high degree of intraspecific polymorphism at microsatellite loci of chickpea. Twelve microsatellite STMS markers were used to distinguish between 78 chickpea accessions. A total of 169 alleles, including null alleles, were detected among the 78 accessions. The phylogenetic analysis showed that there is no relationship between accession and geographic origin, as indicated by clustering of accessions from different geographic locations. Recently, *Ty1-copia* retrotransposon-like elements have been identified in chickpea (Sant et al., 2000) and RFLP probes derived from these sequences have detected genetic variability among four *C. arietinum* accessions.

**Genetic mapping using molecular markers***

Mapping implies putting molecular markers in order, indicating the relative genetic distances between them, and assigning them to their linkage groups on the basis of the recombination values from all their pairwise combinations. Genetic mapping is made possible by the fact that the nuclear genome of higher organisms is organized and transmitted as linear units, called chromosomes. "Genetic linkage", or co-transmission from parent to progeny of genetic markers, which are close together on the same chromosome, provides a means for determining the order of DNA markers along the chromosome. By extending such analysis to hundreds, or even
thousands of markers, one can build up a "linkage map" describing relationship among the markers.

Genetic mapping permits study of any morphological, physiological, or developmental process in which genetic variants exist, with minimum prior information (Paterson and Wing, 1993). It provides a direct means for investigating the number of genes influencing a trait. It also provides the information needed to implement "DNA marker assisted selection", an approach of growing importance in plant and animal improvement. Using genetic maps, one can evaluate the correspondence of genes regulating phenotype in different populations, or even different taxa. Finally, genetic mapping is the first step towards "map-based cloning" of genes responsible for gene function, as well as new opportunities for engineering novel traits.

**Comparative mapping**

Over the past two decades, molecular genetics has revealed fundamental similarity in gene repertoire between taxa (species, genera or higher divisions in the evolutionary tree) that have been reproductively isolated for millions of years. Similarity of genes in different taxa thus provides a means to evaluate similarities and differences in the organization of genes along the chromosome in taxa, which cannot be subjected to genetic linkage analysis. Systemic relationship has been established extensively in cereals. It has been observed that there is a of conservation single copy genes and even the order of most of these genes is conserved among different cereals. This conservation of genes and gene orders is irrespective of haploid chromosome number (x) of the species studied. These vary from 5 (maize), 7 (wheat, barley and rye), 9 (foxtail millet), and 10 (sugarcane and sorghum) to 12 (rice) (Moore et al., 1995). It is also irrespective of the haploid DNA content (1C value) which vary 40 fold, rice being the smallest (400 megabytes) and wheat being the largest (16000 megabytes).

The commonality among organisms as a result of conservation of genes affords the opportunity to use genes as "DNA markers" (mileposts) common to different organisms. Thus, comparative mapping offers tremendous potential efficiencies in
genome analysis because new information about both location and function of gene in one organism is likely to be relevant to other related organisms.

**Markers for genetic mapping**

Any marker technique can be used for development of a map. However, co-dominant markers (e.g., RFLP and microsatellite) provide more information from F2 and backcross generations than markers producing predominantly presence/absence or dominantly inherited polymorphism. However, when mapping homozygous populations of random inbred lines, co-dominant markers offer little advantage over presence/absence and other dominantly inherited markers. For comparative mapping within and across species, use of RFLP markers as map anchor loci appears to be the best choice as the polymorphism they detect, appears to be evolutionary conserved in a more predictable manner than that of loci detected by hyper-variable SSR and AFLP markers. AFLP, and other highly polymorphic markers, can then be used to fill gaps in RFLP based genetic linkage maps, following bulk segregant analysis approaches.

**AFLP and SSR molecular marker maps of plant genomes**

Molecular marker maps are now available for most important species. Cervera et al. (2001) have constructed dense genetic linkage maps of three *Populus* species based on AFLP and microsatellite markers. Their map covers 86% of the genome. A map of barley covering 925.6 cM of the genome has been constructed by Mano et al. (2001). Gedil et al. (2001) have constructed a linkage map of cultivated sunflower using both RFLP and AFLP markers covering 1326 cM of the genome. An expanded and integrated map of cucumber is now made available by Bradeen et al. (2001). Liu et al. (2000) have constructed the genetic map of soybean using various markers including AFLPs, which covers 3713.5 cM of the genome. An anchored AFLP and retrotransposon-based map of diploid *Avena* has been constructed, which includes 372 AFLP markers by Yu and Wise (2000). A high resolution linkage map of the rice submergence tolerance locus *SubI* has been constructed by Xu et al. (2000). The linkage maps of *Pinus* (Don et al., 1999) and *Prunus* (Dirlewanger et al., 1998) have also been constructed. Rajore et al. (2001) have constructed a linkage map of
microsatellite markers in spruce. Paglia et al. (1998) have constructed STS linkage map of conifer (Norway spruce), covering a length of 2198.3 cM which represents 77.4% of the estimated genome length. A high-density molecular map of chromosome region harboring strip-rust resistance gene \( YrH52 \) and \( Yr15 \) derived from wild emmer wheat has been constructed by Peng et al. (2000). In their map 45 marker loci spanned a total map length of 107.7 cM. Ponce et al. (1999) have performed a high throughput genetic mapping in \textit{Arabidopsis thaliana} using microsatellites. A microsatellite map of wheat has been constructed by Roder et al. (1998). Ma et al. (2001) have performed molecular linkage mapping in rye. The map contains 184 markers present in all seven linkage groups covering 727.3 cM. A core genetic map of \textit{Hordeum chilense} has been constructed by Hernandez et al. (2001) and has been compared with maps of barley and wheat. Temnykh et al. (2000) have performed the mapping and genome organization of microsatellite sequence in rice.

**Genetic mapping in chickpea**

An STMS marker map of chickpea was constructed by Winter et al. (1999) that consisted of 120 markers mapped to 11 linkage groups covering a span of 613 cM. A total of 90 RILs from an inter-species cross between \textit{C. reticulatum} and the chickpea cultivar ICC4958 were used for mapping. The map displayed both clustering as well as a random distribution of loci. The majority of loci were located in three distinct regions of the genome. More recently, Winter et al. (2000) have reported an integrated molecular marker map of the chickpea genome which was established using 130 RILs from a wide cross between \textit{C. arietinum} and \textit{C. reticulatum}. A total of 354 markers were mapped on the RILs including 118 STMS, 96 DAFs, 70 AFLPs, 37 ISSRs, 17 RAPDs, 8 isozymes, 3 cDNAs, 2 SCARs and three loci that confer resistance against different races of Fusarium wilt. Of these, 303 markers cover 2077.9 cM in 8 large and 8 small linkage groups. Fifty-one (14.4%) markers were unlinked. A clustering of markers in central regions of linkage groups was observed. Markers of the same class, except for ISSR and RAPD markers, tended to generate subclusters.
Mapping of Quantitative Trait Loci (QTLs) using molecular markers

Many plant phenotypes reflect the effects of numerous independent genes, acting at different times during growth and development. Such multigenic or polygenic traits are especially cumbersome to manipulate by classical breeding methods. Having a complete genetic map is important in locating genes, because one can be confident that all regions of all chromosomes are searched. The number of genes accounting for variation in a phenotype can vary widely, in different pedigrees, taxa, or even environments. Phenotypic variation that is due to segregation at a single genetic locus is simple to map, in fact, if such variation is discrete, the trait itself is a "visible marker". By contrast, if the effects of the locus are obscured by environmental factors (such as differences in temperature, moisture, or nutrition), genetic mapping becomes difficult, requiring a mathematical approach which accommodates the fact that phenotype is only a quantitative (rather than qualitative) measure of genotype. Finally, if a phenotype is influenced not just by one such "quantitative trait locus" (QTL), but by many such loci segregating in the same population, genetic mapping becomes even more complicated.

Such complex traits, influenced by many genes, require large populations to map. Using "complete" genetic maps, together with populations typically including 200 or more individuals, such QTLs can be assigned to "likelihood intervals", typically spanning 10-30 cM (Paterson et al., 1988; Lander and Botstein, 1989). A large number of phenotypes important to agriculture, evolution, and medicine are influenced by quantitative trait loci, and "QTL mapping" has become a powerful new research tool across the life sciences. Most commonly, individual QTLs are described by their chromosomal location, the magnitude of their phenotypic effects, the effect of gene dosage at the locus and their interactions with other QTLs or unlinked genetic loci.

A large body of data from QTL mapping have supported the concept that relatively few genes explain large portions of the phenotypic variance in a trait, with increasing numbers of genes explaining progressively smaller fractions of phenotypic variance (Paterson et al., 1991).
Recent plant QTLs mapped using AFLP and SSR

Lerceteau et al. (2000) have detected various QTLs using AFLP for economically important traits in *Pinus sylvestris*. The female map spans 796 cM and the male map spans 1335 cM. Twelve QTLs for tree height, trunk diameter and volume were detected. Thirty-six QTLs for grain yield and grain related traits of maize have been identified using an AFLP map by Marsan et al. (2001). Nandi et al. (1997) have mapped QTLs for submergence tolerance in rice by AFLP analysis covering 1756 cM distributed over 12 chromosomes. DNA markers for fusarium head blight resistance QTL have been identified in two wheat populations. Sripongpangkul et al. (2000) have identified several genes/QTLs that control plant elongation and flood tolerance in rice. QTLs for aluminum tolerance in rice have been identified by Wu et al. (2000). Rouppve van der Voort et al. (1998) have identified a QTL for broad-spectrum resistance to cyst nematode species in potato. Varshney et al. (2000) have identified a microsatellite marker associated with QTL for grain weight in bread wheat. A microsatellite marker associated with a QTL for grain protein content of bread wheat was identified by Prasad et al. (1999). Bohn et al. (2000) have identified QTL for resistance against the European corn borer of maize. QTLs for resistance to insects in soybean have been identified by Terry et al. (2000). Lanceras et al. (2000) have identified QTLs for amylose content, gel consistency and gelatinizing temperature in Thai jasmine rice. Sixty-eight loci for 12 agronomically important traits in wild rice have been identified by Xiao et al. (1998).

Applications of molecular markers in plant breeding programs

A common objective in plant breeding is to transfer a specific gene (trait) into an otherwise desirable genotype, which has a remedied defect. Because the cultivated gene pool of many crops represents only a small portion of the total genetic diversity within a taxon, many genes for disease resistance or other traits are found only in wild or feral plants. The various applications of molecular markers in breeding programs are marker assisted selection, confirming the identity of breeding lines, establishing hybrid identity, testing for purity of breeding lines, evaluation of somaclonal variation and predicting hybrid performance. Shen et al. (2001) have evaluated near-isogenic
Review of literature

lines of rice introgressed with QTLs for root depth through marker-aided selection. Liu et al. (1999) have reported molecular identification and marker-assisted selection of *Pm21* gene conferring resistance to powdery mildew in wheat. Moury et al. (2000) have identified a CAPS marker to assist selection of tomato spotted wilt virus resistance in pepper.

**Map-based cloning**

Map based cloning provides a means to isolate naturally occurring mutations, even when obscured by the presence of multiple mutations, or by the vagaries of environment. Genetic linkage analysis is employed to locate a gene to a region of perhaps 0.1% or less of a genome, then physical mapping is used to clone a contiguous stretch of DNA including the gene. This technique is being widely used to clone agronomically important genes. Some of the recent reports include chromosome landing at the bacterial blight resistance gene *Xa4* locus using a deep coverage rice BAC library (Wang et al., 2001), cloning of the *Arabidopsis RSF1* gene using a mapping strategy reported by Spiegelman et al. (2000), isolation of LEUNIG gene, a putative transcriptional repressor that regulates AGAMOUS expression during flower development (Conner and Liu, 2000) and genetic and physical mapping of *xa13*, a recessive bacterial blight resistance gene in rice (Sanchez et al., 1999).

**BULKED SEGREGANT ANALYSIS**

Various breeding populations are used for identification of molecular markers such as cytogenetic stocks (e.g. nullitetrasomic lines), near isogenic lines etc. These lines require many backcrosses to develop and are therefore time-consuming to generate. An alternative method, bulked segregant analysis (BSA) was proposed by Michelmore et al. (1991) to circumvent these problems and allow rapid mapping of loci that do not segregate in the original populations used to develop the genetic map. The principle of bulked segregant analysis is the grouping of informative individuals together so that a particular genomic region can be studied against a randomized genetic background of unlinked loci. The method involves comparing two pooled
DNA samples of individuals from a segregating population originating from a single cross. Within each pool, or bulk, the individuals are identical for the trait or gene of interest but are arbitrary for all other genes. Two pools contrasting for a trait are analyzed to identify markers that distinguish them. Markers that are polymorphic between the pools will be genetically linked to the loci determining the trait used to construct the pool (Fig. 3).

**Recent studies using bulked segregant analysis**

Wang et al. (2000) have developed PCR based RAPD markers linked to the fertility restorer gene for polima cytoplasmic male sterility in rape seed (*Brassica napus*) using bulked segregant analysis. AFLP markers of restoring gene of the wild-abortive hybrid rice have been identified using bulked segregant analysis by He et al. (2000). Khan et al. (2000) have identified RFLP markers for resistance to wheat spindle streak mosaic bymovirus disease using bulked segregant analysis. Molnar et al. (2000) studied the inheritance and RAPD tagging of multiple genes for resistance to net blotch in barley using bulked segregant analysis. A CAPS marker to assist selection of tomato spotted wilt virus resistance in pepper has been identified by Moury et al. (2000). Chague et al. (1999) have isolated microsatellite and RAPD markers flanking the *Yr15* gene of wheat using near isogenic lines and bulked segregant analysis. Parker et al. (1999) have mapped loci associated with milling yield in wheat using bulked segregant analysis. Genetic mapping of a fusarium wilt resistance gene in melon has been performed by Wang et al. (2000). Chantret et al. (2000) have mapped powdery mildew resistance gene *MIRE* in wheat and detected a resistance QTL by bulked segregant analysis using microsatellies. A new major QTL associated with resistance to soybean cyst nematode has been identified by Schuster et al. (2001) using bulked segregant analysis. Monna et al. (1995) have identified RAPD markers linked to photoperiod sensitivity gene in rice using bulked segregant analysis.
DROUGHT AS A MAJOR CONSTRAINT OF CHICKPEA

Drought is a major constraint to increased productivity, as nearly 90% of the world's chickpea is grown rain-fed (Kumar et al., 1996). It is estimated that if moisture stress is alleviated, up to 50% increase in chickpea production could be achieved, with a present value of ca. US$900 million (Ryan, 1997). One way to escape end-of-season drought is to develop varieties with early growth vigor, early flowering, and early maturity (Calcagno and Gallo, 1993; Johansen et al., 1997).

In drought prone environments such as those in the tropics normally a strong positive association exists between water transpired by the crop and biomass formation (Sinclair et al., 1984). Therefore, rapid early growth of the crop is desirable. This will also ensure early attainment of full crop canopy and prevent soil-surface evaporation. Often end-of-season drought is associated with increasing temperature (Calcagno and Gallo, 1993; Singh, 1997). Sedgley et al. (1990) suggested that early pod set should be a prime strategy for avoiding drought stress in environments prone to end-of-season moisture stress. Thus development of early maturing varieties may help drought-escape and result in increased productivity and extending this crop to even more drought-prone areas (Kumar et al., 1996). Early maturing varieties will also allow increased population per unit area and consequently help maximize yield in drought-prone environment.

Early flowering and podding restricts vegetative growth in indeterminate plants like chickpea (Saxena et al., 1997). In subtropical environments, winter rains may induce excessive vegetative growth leading to dense canopy and high humidity. Such conditions are conducive for the development of foliar diseases. Thus restricted vegetative growth can help avoid seed yield losses in these environments. Therefore development of early flowering and podding cultivars should be a major objective for chickpea improvement.
EVOLUTION OF PHOTOPERIOD RESPONSE IN CHICKPEA

The Greek botanist Theophrastus (1997, translation) and the Roman historian Pliny (1971, translation) have described chickpea as a summer crop (sown in March/April and harvested in June/July). Such a crop begins and completes its life cycle under increasing photoperiod and rising temperature and depends mainly on stored moisture (Khanna-Chopra and Sinha, 1987; Kostrinski, 1974). The crop cycle of chickpea during initial stages of domestication is still unclear. Nevertheless, the present day crop cycle of the wild ancestor *C. reticulatum* is entirely different from the one described in the ancient reports. In the wild, *C. reticulatum* germinates after the autumn rains and develops vegetatively during the rainy winter accompanied by shortening photoperiod and cool temperatures. Flowering and reproduction occurs in the late spring of the ensuing year when mean temperatures are high and the days are long. It has been observed that spring sown wild *C. reticulatum* yield less than 1/5th biomass and seed produced. The possible reason to compromise on seed yield is that the earlier farmers knew of the devastating effects of the blight disease caused by the fungus *Didymella rabiei* (Kovacevski) v. Arx [(anamorph: *Ascochyta rabiei* (Pass.) Labr)]. In the Near East, the climatic conditions favoring spread of the disease occur from early February until early April. An Ascochyta epidemic would destroy the crop completely because the autumn sown crop would have a fully developed closed canopy by this time. Indeed, ascochyta blight is a major biotic constraint for chickpea production in the Mediterranean basin even today. Since the disease is not a serious problem in spring-sown chickpea, it is considered as the prime reason for the ancient practice of chickpea spring sowing. This shifting from the natural plant cycle to spring sowing was accompanied with selection in the direction of increased daylength sensitivity, which might turn into an adaptive advantage. This is because it might allow the plant to enter reproduction early enough in the season regardless of its developmental stage. Timing of flowering independent of daylength usually means that plant would enter reproduction upon accumulation of a certain biomass value typical to the genotype (Sachs, 1999). Following spring sowing, a delay of flowering until a relatively large number of nodes have developed might imply that the plant
would enter reproduction when soil moisture is nearly depleted and only a meager seed yield (if any) might be expected. On the other hand, following spring sowing, increase daylength sensitivity might turn into a major adaptive advantage. This is because it might allow the plant to enter reproduction early enough in the season regardless of its developmental stage. In such a way, seed set and pod development will take place before the onset of the summer drought and the grain yield (although modest) will be assured. The long-term consequence of selection of spring sowing was the nearly complete fixation of the relatively high daylength sensitivity in the Mediterranean kabuli germplasm (Kumar and Abbo, 2001).

Despite the Near East origin of the crop, currently about 80% of its global production takes place in the Indian subcontinent. Such remarkable adaptive success in an environment so different from its native origin is dependent on the presence of such allelic variation in major adaptive loci in the introduced material and adequate agrotechniques. The allelic variations and agrotechniques ensure crop establishment and correct timing of flowering.

In India, chickpea is mostly sown in October/November and in Ethiopia from August/September onward to January (van der Maesen, 1972). In both regions, the growing season is characterized with shortening photoperiod due to which the introduction of chickpea would have accompanied with problems related to poor adaptation, especially incorrect timing of flowering. It is difficult to assume that repeated sowing of a nonadapted crop was done until daylength insensitive types gradually occurred in the seed stock. Therefore, it has been suggested that the spread of chickpea into its Indian and East African growing areas must have required adequate allelic variation in flowering time genes to be present in the founder seed stocks (Kumar and Abbo, 2001). The seasonal daylength variation in the low latitude chickpea growing areas of India and Africa suggests that insensitive alleles at photoperiod response loci had a central role in the successful spread of chickpea into these regions. Such a variation might have included alleles at both major and minor photoperiod loci and perhaps temperature response loci as well. These off-types, which require reduced daylength, gave rise to Indian and African chickpea genepools.
THE PHENOMENON OF FLOWERING

The timing of the transition to flowering is determined by the interaction of the endogenous developmental processes of a plant with environmental cues that signal the onset of conditions favorable for reproductive success. Recently, genetics has been used to study the mechanism of flower initiation by analysis of genetic variation in species such as *Arabidopsis* and pea (Simpson et al., 1999). Since the flowering behavior of *Arabidopsis* mirrors that of many other plants, it is likely to be a model of broad utility. Many plants respond to changes in daylength and extended periods of cold temperature, since these are both predictable and reliable indicators of seasonal progression. *Arabidopsis* flowering is accelerated as the day length increases, an environmental condition that signals the onset of spring and summer in the higher latitudes. This process is known as photoperiod response. *Arabidopsis* flowering is also accelerated following an extensive period of cold treatment, an environmental condition that signals the end of winter and onset of spring. This process is known as vernalization response. In addition, flowering in *Arabidopsis* is stress responsive; it exhibits precocious flowering under conditions that are unfavorable for its survival. For example, light quality changes that accompany shading by near neighbors' enrichment in the far-red wavelength promote flowering.

Genes that affect flowering time in *Arabidopsis* have been identified through two complementary strategies. The first has utilized the variation present in naturally occurring ecotypes of *Arabidopsis* that vary in flowering time. The second approach has utilized a diverse spectrum of induced mutations that result in either early or late flowering. Taken together, there are currently about 80 genes and loci in *Arabidopsis* known to affect flowering time (Simpson et al., 1999).

**Multiple pathways regulating the floral transition**

The study of how flowering time mutants respond to environmental treatments, such as vernalization and photoperiod combined with the genetic analysis of epistasis, has established the existence of multiple pathways that control flowering time in *Arabidopsis* (Alonso-Blanco et al., 1998a; Koornneef et al., 1998a, 1998b). Two of
Review of literature

these pathways clearly mediate signals from the environment. The photoperiod promotion pathway integrates daylength into the flowering decision through a series of genes that sense and respond to the regular day-to-night transition. The genes PHYB, CRY2, CRY1, PHYA, ELF3, CCA1, and LFY and others constitute the photoperiod promotion pathway. The vernalization promotion pathway promotes flowering in many late-flowering ecotypes in response to an extended period of cold temperature. Analysis of VRN genes should identify the molecular processes important in vernalization and may clarify the connection between vernalization and light quality perception. In contrast to these environmentally responsive pathways, there is a collection of genes that promote flowering in a manner independent of environmental signals. Plants carrying mutations in these genes flower late in both long-day plants (LDs) and short-day plants (SDs), and these genes have been formed into a separate class known as the autonomous promotion pathway. Members of this epistatic class probably monitor or respond to endogenous cues such as an internal developmental clock. The genes FCA, FY, FPA, FVE, LD, and FLD constitute the autonomous pathway.

The phytohormone gibberellic acid (GA) affects multiple aspects of plant development. In Arabidopsis, mutants defective in GA biosynthesis and signal transduction pathways, flower early or late. The analysis of the epistatic relationships between GA mutants and other flowering time mutants suggests that they form a promotive pathway distinct from the others. The genes GAI, GA1, SPY, FPFL and others constitute the GA pathway.

In addition to these established pathways, another group of late-flowering mutants has been separated into a distinct section referred to as the FT subgroup. Members of this pathway flower late in LDs but show properties that distinguish them from other members of the photoperiod pathway. The genes FWA, FT, FD, FE and others constitute the FT pathway. Finally, the fact that so many early flowering mutants have been identified indicates that the floral transition is normally actively repressed. Multiple genes encoding repressors of the floral transition have been identified, some of which oppose the activity of specific promotive pathways (FRI, FLC etc) (Simpson et al., 1999).
Floral Transition-Meristem fate changes

The multiple pathways that regulate the timing of the floral transition act directly or indirectly on a group of genes that switch the fate of the meristem from vegetative to floral. *Arabidopsis thaliana* has a distinct vegetative phase during which the apical meristem produces lateral meristems developing into leaves subtending an axillary bud. The nodes do not elongate, resulting in the formation of a rosette. Floral transition is marked by the establishment of a floral fate in these meristems and by the suppression of leaf production. Although the appearance of flowers is the final result of the phase change, other changes occur earlier. These are somewhat gradual and can be observed in leaf morphology (Matzke et al., 1996) and the gradual appearance of trichomes at the abaxial side of the leaves and their gradual disappearance at the adaxial side (Chien and Sussex, 1996; Telfer et al., 1997). It has been proposed that phase changes involve a decrease of a floral repressor (Sung et al., 1992), called a controller of phase switching (COPS), which at critical low levels leads to the activation of the floral initiation process (FLIP) (Schultz and Haughn, 1993). The latter is controlled by the so-called Floral Meristem Identity or FLIP genes, such as *LEAFY (LFY)*, *APETALAI and 2 (AP1, AP2)*, *CAULIFLOWER (CAL)*, and *UNUSUAL FLORAL ORGANS (UFO)* (Haughn et al., 1995).

The floral meristem identity genes in turn, control a second set of genes that direct the formation of various flower parts (termed organ identity genes), which in *Arabidopsis* include *AG, AP3, PISTILLATA (PI)*, and *SUP*.

How the different pathways integrate to activate the floral meristem identity genes is an area of intense current investigation. Models designed to integrate the genetic and molecular data have been proposed and are constantly being updated (Schultz and Haughn, 1993; Martinez-Zapater et al., 1994; Coupland, 1995; Yang et al., 1995; Koornneef et al., 1998b and Levy and Dean, 1998). The models reflect the multifactorial nature of flowering time control as first proposed by Bernier (1988). Fig. 4 shows the model proposed by Simpson et al. (1999). The transition to flowering is perhaps not a single switch but a progressively acquired condition resulting in expression of genes regulating meristem identity. Redundancy, potentiation, and
additivity have been built into the system to ensure that when it occurs the transition is sharp.

**Recent studies of mapping of QTL(s) associated with flowering time and photoperiod response**

A large number of QTLs, associated with time-to-flowering and photoperiod response in diverse plant species, have been mapped. These include *Arabidopsis* (Kowalski et al., 1994; Clarke et al., 1995) in which five QTLs affecting flowering time were identified, one of these loci is coincident with *FRI* locus identified as the major determinant for late flowering and vernalization responsiveness in the ecotype Stockholm. Kuittinen et al. (1997) identified one major and six minor QTLs associated with flowering time in *Arabidopsis*. They conclude that major QTLs for adaptive traits can be detected in non-domesticated species. Alonso-Blanco et al. (1998a) studied the flowering behavior of two early *Arabidopsis* ecotypes *Ler* and *Cvi* and have attributed the flowering behavior differences between them to four loci referred to as *ED1, FLF, FLG* and *FLH*. Stratton (1998) has identified seven QTLs affecting flowering date and eight QTLs for rosette leaf number in *Arabidopsis* and has concluded that genotype-environment interactions for flowering time are controlled by many minor genes. Camargo and Osborn (1998); Bohuon et al. (1998), and Rae et al. (1999) have identified QTLs linked to flowering time in *Brassica oleracea*. In the case of rice, Maheswaran et al. (2000) have identified 15 QTLs associated with days to flowering of which four were identified as influencing the response to photoperiod. A major QTL controlling response to photoperiod was identified by Yano et al. (2000) and was shown to be closely related to the *Arabidopsis* flowering time gene *CONSTANS*. In sunflower, five QTLs associated with days to flowering were identified by Leon et al. (2000). Further, six QTLs associated with growing degree days to flowering and photoperiod response in an elite sunflower population were identified by Leon et al. (2001). Using a linkage map constructed of AFLP and SSR markers, three QTLs controlling bud set and six QTLs controlling bud flush were detected in *Populus* by Frewen et al. (2000). Additionally, five candidate genes believed to be involved in perception of photoperiod were placed on the QTL map. Vladutu et al. (1999)
confirmed the presence of two QTLs in maize that are involved in the transition of apical meristem from vegetative to generative structure. In wheat, two QTLs affecting heading time and photoperiod response were identified by Sourdille et al. (2000). Yamanaka et al. (2001) have isolated four QTLs for flowering time using a linkage map of soybean.

THE FLOWERING GENES OF CHICKPEA

In chickpea, information on the genetic control of flowering time is only beginning to accumulate. This is despite the fact that early flowering mediated by photoperiod insensitivity was suggested as a means to increase chickpea adaptability as early as 1971 (Sandhu and Hodges, 1971). Regrettably, no genetic studies followed until recent years (Kumar and van Rheenen, 2000; Or et al., 1999).

The flowering time of chickpea genotypes varies with latitude and temperature variations. ICRISAT, Patancheru, India conducted trials of breeding lines at three locations: Patancheru (18°N), Gwalior (26°N) and Hisar (29°N). The range for 25 genotypes tested in these locations did not overlap. The mean number of days to 50% flowering were 51, 76, and 96 respectively for these three locations. Thus the genes controlling flowering time are sensitive to temperature and day length.

The existence of wide genetic variation for flowering time was documented by Pundir et al. (1988), who evaluated the world chickpea germplasm maintained at ICRISAT and listed 43 accessions that flowered in less than 39 days at Patancheru (18°N). Most of these lines originated in tropical India (Maharashtra and Karnataka), a few in Ethiopia, two in Mexico and five in Iran (>30°N). This might indicate that mutations for early flowering genes also survived in subtropical environments. They probably out-yielded the traditional long duration varieties under severe drought conditions. Lack of knowledge on the genetic control of flowering time did not prevent Kumar et al. (1985) from developing extra-early chickpea ICCV2 as a transgressive segregant from a cross of five chickpea lines. However, further manipulation of these genes is difficult without understanding individual effects of
other genes governing this trait, interaction among them, and their response to variations in temperature and daylength.

**Genetic control of flowering time in chickpea**

A major recessive gene "efl-1", for early flowering has been identified in a cross between the extra-early variety ICCV2 and the medium-duration variety JG 62 (Kumar and van Rheenen, 2000). This gene is responsible for about 3 weeks' difference in flowering time between the two parents at ICRISAT, Patancheru, India. A super early chickpea segregant, ICCV96029, was selected from the F6 generation from a cross of two extra early varieties, ICCV2 and ICCV93929. ICCV96029 flowers about a week earlier than either of the parents (Kumar and Rao, 1996). The allele efl-1 is common between the two parents. Thus other complementary genes with smaller effects exist between these two extra early parents. Complementary gene action for flowering time was also evident in crosses between chickpea genotypes ICC4958 (India) and Guamuchil (Mexico), two of the five parents of ICCV2 (Kumar et al., 1985). Thus at least two different loci control flowering time in ICCV2. This observation was further corroborated by a diallel analysis among three extra early lines, ICCV2, ICCV93929 and Harigantars (ICC5810) that produced three different types of F1s, indicating that more than two complementing genes operate flowering time in chickpea. In these studies one of the three F1s (ICCV2 x ICCV93929) flowered earlier than the mid-parent, the second at the same time as the mid-parent and the third flowered later than the mid-parent.

Or et al. (1999) studied chickpea flowering time in a cross between an early line ICC5810 and a late flowering Israeli cultivar (Hadas) at Rehovot (32°N), Israel. The cross was designed to analyze the flowering syndrome of the Mediterranean chickpea stocks, hence the choice of the modern relatively late flowering cv. Hadas. The early parent ICC5810 was chosen based on the screening of Roberts et al. (1985) who characterized it as day-neutral type. The 3:1 segregation of late: early individuals among the F2 progeny was interpreted as an evidence to a major gene action affecting flowering time through determination of photoperiod response (PPD). In this cross, the late condition (photoperiod responsive allele) was dominant.
At present it is unclear whether the \textit{efl-1} gene described by Kumar and van Rheenen (2000) and the PPD gene reported by Or et al. (1999) differ from one other. However, there are indications that the major recessive allele for earliness in ICC5810 is located at the same locus as the \textit{efl-1} gene in ICCV2 (Kumar and Abbo, 2001). In the absence of DNA markers linked to the chickpea flowering genes it is not possible to relate either of the two genes to their counterparts among the well-defined pea or \textit{Arabidopsis} flowering genes.