2. REVIEW OF LITERATURE

2.1 Disease scenario

Pigeonpea crop suffers from a large number of diseases caused by fungi, bacteria, viruses and nematodes etc. Some of the important occurring fungal diseases and their pathogens are leaf spot (*Alternaria alternata*), leaf spot (*Alternaria tenuissina*), leaf spot (*Cercospora cajanî*), leaf spot (*Cercospora cajanica*), leaf and stem spot (*Cercospora indica*), leaf spot (*Cercospora instabilis*), decay of tender leaves and shoot (*Choanephora cucurbitarum*), leaf spot (*Cladosporium cladosporioides*), leaf spot (*Cochilobulus lunatus*), anthracnose (*Colletotrichum cajani*), stem cancer and die back (*Colletotrichum capsici*), stem rot (*Diplodia cajî*), wilt (*Fusarium udum*), wilt (*Fusarium oxysporum*), wilt (*Fusarium solanî*), powdery mildews (*Levellilia taurica*), stem rot (*Macrophauma cajanica*), powdery mildew (*Oidium* sp.), root and stem rot (*Sclerotium rolfsii*) and leaf gall (*Synchitrium phaseoli-radiata*) (Richardson, 1990; Kumar *et al.* 2001; Jamaluddin *et al.* 2004). Among these diseases *Fusarium* wilt has been found to be the most destructive all over the world and the present work was proposed to be carried out on the same.

2.2 Wilt disease in pigeonpea

*Fusarium* wilt (*Fusarium udum Butler*) is an important soil borne disease of pigeonpea, which causes significant yield losses in susceptible cultivars throughout the pigeonpea growing areas. The disease was first recorded by Butler (1906) in India. Cent per cent grain losses has been reported due to
Fusarium wilt when occurred at pre-pod stage, 67 per cent at pod maturity stage and 29.5 per cent at pre harvest stage (Kannaiyan and Nene, 1981). The existence of variants/races of *F. udum* has been reported and has been cited as a major drawback in the development of pigeonpea varieties resistant to *Fusarium* wilt (Okior and Kimani, 1997). Being a soil-borne pathogen, *Fusarium udum*, the fungus enters the host vascular system at root tips through wounds leading to progressive chlorosis of leaves, branches, wilting and collapse of the root system (Jain and Reddy, 1995). Partial wilting of the plant as if there is water shortage even though the soil may have adequate moisture distinguishes this disease from termite damage, drought, and phytophthora blight that all kill the whole plant. Partial wilting is associated with lateral root infection, while total wilt is due to tap root infection (Nene, 1980; Reddy *et al.* 1993). The most initial characteristic internal symptom is a purple band extending upwards from the base of the main stem. The xylem develops black streaks and this results in brown band or dark purple bands on the stem surface of partially wilted plants extending upwards from the base visible when the main stem or primary branches are split open (Reddy *et al.* 1990; Reddy *et al.* 1993).

*Fusarium* wilt is soil borne but the pathogen may be carried as a contaminant of pigeonpea seed (Upadhyay and Rai, 1983). Pigeonpea has traditionally been screened for wilt resistance in wilt-infested fields (Butler, 1908; Deshpande *et al.* 1963). Several screening techniques have been reported for *Fusarium* wilt out of which the best results were obtained in seeds sown in infested soil (Haware and Nene, 1994; Okior, 1998). Five techniques, namely,
sowing seeds or transplanting seedlings into infested soil, dipping roots or soaking seed in a spore suspension, and stem injection were tested under glasshouse conditions on four cultivars of pigeonpea with different levels of resistance (Okior, 1998). Sowing seed in infested soils gave the highest mortality and allowed for easy differentiation of resistant and susceptible plants. The stem injection induced very low wilting and required significant labour to inject the plants. The other techniques either gave severe wilting, inconsistent results or low wilting, or were considered unreliable. This study recommended that sowing seeds in infested soil in a glasshouse can be adopted as a standard procedure for scoring wilt (Okior, 1998). The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT, Patancheru, India) also uses field screening but has reported cases of inconsistent results (Nene et al. 1981).

The inheritance of resistance to Fusarium wilt is not fully understood. Conflicting reports have been made on the inheritance of resistance to Fusarium wilt in pigeonpea. Pal (1934) reported that resistance to pigeonpea wilt was controlled by multiple factors while Shaw (1936) suggested that wilt resistance was conditioned by two complementary genes. Resistance to Fusarium wilt has been reported to be under the control of two complementary genes (Parmita et al. 2005), single dominant gene (Pawar and Mayee, 1986; Pandey et al. 1996; Singh et al. 1998; Karimi et al. 2010), two genes (Okior, 2002), major genes (Sharma, 1986; Parmita et al. 2005), and a single recessive gene (Jain and Reddy, 1995). Parmita et al. (2005) reported significant role of a single dominant gene and dominant epistatic gene interaction in controlling resistance to wilt. In a
cross between one resistant (ICP 8863) and two susceptible (ICP 2376 and LRG 3C) lines, resistance was found to be controlled by a single recessive gene. The gene was designated pwr1 (Jain and Reddy, 1995). Recently real-time PCR based detection assay was developed for *Fusarium udum*. The qPCR assay specifically differentiated the *F. udum* from closely related species of *Fusarium*, other test microbes and environmental samples (Mesapogu *et al.* 2011).

Mahesh *et al.* (2009) studied the morphological and cultural variability of six isolates of *Fusarium udum* Butler collected from Bangalore, Kolar, Hoskote, Ramanagar, Anekal and Jagalur. All the isolates showed the significant variations with respect to morphological characters *viz.*, the size of macro conidia and micro conidia varied from 10.51-18.70 × 1.27-3.10μm and 3.62-8.12 × 0.96-1.80μm respectively. Number of septa of macro conidia and micro conidia varied from 2.12-2.93 and 0-0.61 respectively. Colour of both the macro conidia and micro conidia was hyaline. Shape of macro conidia was sickle shaped with blunt ends to elongated sickle shaped with pointed at both ends while shape of micro conidia was oval to round. Among the media used to study the growth of *Fusarium udum* isolates, all the six isolates produced maximum growth on Richard’s agar medium (84.33 mm). Czapeck’s agar medium was found to be best for sporulation of *Fusarium udum* isolates except for Kolar isolate, Richard’s agar medium was found to be best for sporulation.

The major portion of pigeonpea improvement is being carried out at International Crop Research Institute for semi-arid tropics (ICRISAT-Patancheru), India. Since sterility mosaic and *Fusarium* wilt are the major pigeonpea diseases, breeding
varieties with dual resistance was given priority, of which Asha (ICPL 87119) 15, Laxmi (ICPL 85063) and Maruti (ICP 8863) are good examples of resistant varieties developed by ICRISAT. A variety developed at Indian Institute of Pulses Research (IIPR), IPA 204, derived from a cross (Bahar x AC 314-314) is also tolerant to wilt. Use of cultivars resistant to the disease is the only effective means of wilt control.

2.3 Genetic diversity in pigeonpea

Pigeonpea is an important grain legume of the Indian subcontinent, South-East Asia and East Africa. The availability of limited genomic resources and low levels of genetic diversity in the primary gene pool have constrained genetic improvement of pigeonpea. The wild relatives of cultivated species can be an important source of genetic variability for desired agronomic traits, including resistance to various biotic and abiotic stresses and seed quality. Until a couple of years ago pigeonpea was considered an orphan legume crop but now substantial amount of genomic resources have been generated. Availability of genome sequence will accelerate the utilization of pigeonpea germplasm resources in breeding. Gepts (1999) discussed the use of molecular markers for improving the efficiency of plant breeding programs because at the molecular level recognizing the presence or absence of a particular gene is independent of plant part or plant age. Also, in contrast to morphological traits, molecular markers are not influenced by various pleiotropic and epistatic interactions. The first step in molecular breeding, therefore, is to establish linkage between a gene and its marker locus. Subsequently specific DNA diagnostic tests can be applied
to assist plant breeders in selection. The identification of useful breeding lines with the help of linked molecular markers is popularly known as marker assisted selection (MAS).

2.4 Simple sequence repeats (SSRs) evolution in pigeonpea

Simple sequence repeats (SSRs) or microsatellites are becoming standard DNA markers for plant genome analysis and are being used as markers in marker assisted breeding. De novo generation of microsatellite markers through laboratory based screening of SSR enriched genomic libraries is highly time consuming and expensive. The first set of SSR markers was developed in pigeonpea by Burns et al. (2001). An alternative is to screen the public databases of related model species where abundant sequence data is already available. Recently many genomic programs are underway leading to the accumulation of voluminous genomic and expressed sequence tag (EST) sequences in public databases. Microsatellites have also attracted scientific attention because they have been shown to be part of or linked to some genes of agronomic interest (Yu et al. 2000). Together, limited genomic resources and low levels of genetic diversity in the primary gene pool have constrained genetic improvement of pigeonpea. To accelerate the application of genomics to improve yield and quality the first draft genome sequence of a popular pigeonpea variety ‘Asha’ was generated. Eleven pigeonpea chromosomes showed low but significant synteny with the twenty chromosomes of soybean. The genome sequence was used to identify large number of hypervariable ‘Arhar’ simple sequence repeat (HASSR) markers, 437 of which were experimentally validated
for PCR amplification and high rate of polymorphism among pigeonpea varieties. These markers will be useful for fingerprinting and diversity analysis of pigeonpea germplasm and molecular breeding applications (Singh et al. 2011; Varshney et al. 2011).

2.4.1 Abundance of microsatellites plant genome

All the genomic sequences of *Medicago* from the public domain database were searched and analysed of di, tri, and tetra nucleotide repeats. Of the total of about 1,56,000 sequences which were searched, 7,325 sequences were found to contain repeat motif and may yield SSR which will yield product sizes of around 200 bp. Of these the most abundantly found repeats were the tri-nucleotide (5,210) group (Mahalakshmi et al. 2002). In a similar study a total of 875 EST-SSRs were identified from 772 SSR containing ESTs. The dinucleotide repeats were the most abundant and accounted for 50.9% of the Eucalyptus genome (Yashoda et al. 2008).

Raju et al. (2010) constructed 16 cDNA libraries from four pigeonpea genotypes that are resistant and susceptible to *Fusarium* wilt (ICPL 20102 and ICP 2376) and sterility mosaic disease (ICP 7035 and TTB 7) and a total of 9,888 ESTs were generated and deposited in dbEST of GenBank. Clustering and assembly analyses of these ESTs resulted into 4,557 unigenes. 3,583 SSR motifs were identified in 1365 unigenes and 383 primer pairs were designed. Assessment of a set of 84 primer pairs on 40 elite pigeonpea lines showed polymorphism with 15 markers with an average of four alleles per marker and an average polymorphism information content value of 0.40.
The pigeonpea genomics work gained momentum with the development of a set of 88,860 BAC (bacterial artificial chromosome)-end sequences (BESs) after constructing two BAC libraries by using HindIII (34560 clones) and BamHI (34560 clones) restriction enzymes. Analysis of BESs for microsatellites identified 18,149 SSRs, from which a set of 6,212 SSRs were selected for further analysis. A total of 3072 novel SSR primer pairs were synthesized and tested for length polymorphism on a set of 22 parental genotypes of 13 mapping populations segregating for traits of interest. Based on these markers, the first SSR-based genetic map comprising of 239 loci was developed (Bohra et al. 2011).

Odeny et al. (2007) developed microsatellite markers and evaluated their potential use in pigeonpea genetics and breeding. About 208 microsatellites were isolated by screening a non-enriched partial genomic library. AT and TG class of repeats were the most abundant dinucleotide repeats while TAA and GAA were the most abundant trinucleotide repeats. The diversity analysis readily distinguished all wild relatives from each other and from the cultivated germplasm.

2.4.2 Diversity analysis in pigeonpea

Genetic divergence in forty early maturing genotypes of pigeonpea (Cajanus cajan) from different geographic regions was analyzed based on morphological traits. All genotypes were grouped into three clusters. The analyses revealed that genetic diversity was independent of geographical origin (Murthy and Dorairaj, 1990). Genetic divergence among 49 genotypes of pigeonpea belonging to
different eco-geographic regions was studied by using Mahalanobis $D^2$ statistics (Rekha et al. 2011). They were grouped into 6 clusters but the clustering pattern of genotypes did not follow geographical origin, suggesting that geographical isolation may not be the only factor causing genetic diversity. It was concluded that the selection of parents for hybridization should be more based on genetic diversity rather than geographic diversity.

Twenty two SSR markers of different crop species origin were used to assess polymorphism through their SSR fingerprinting of 16 cultivated pigeonpea genotypes. Four hundred twenty five bands were amplified in all the sixteen genotypes. A total of 46 SSR fragments were amplified. Eight primers showed 100% polymorphism. Based on dendrogram constructed using the similarity coefficient values, 16 genotypes were grouped into two distinct clusters. Cluster I comprises mostly late duration genotypes while cluster II comprises medium duration genotypes except CO-6 and Bahar. Both the clusters and sub-cluster in the dendrogram were supported by high bootstrap values, thus indicating that the SSR could be a good choice to classify the genotypes (Singh et al. 2008).

Dutta et al. (2011) developed 550 validated genic-SSR markers in pigeonpea using deep transcriptome sequencing. Genetic diversity analysis was done on 22 pigeonpea varieties and eight wild species using 20 highly polymorphic genic-SSR markers. The number of alleles at these loci ranged from 4-10 and the polymorphism information content values ranged from 0.46 to 0.72. Neighbour-joining dendrogram showed distinct separation of the different groups of pigeonpea cultivars and wild species.
A total of 24 pigeonpea (*Cajanus cajan*) cultivars representing different maturity groups were evaluated for genetic diversity analysis using 10 pigeonpea specific and 66 cross-genera microsatellite markers. Of the cross-genera microsatellite markers, only 12 showed amplification. A total of 45 alleles were amplified by the 22 markers. Nine markers showed 100% polymorphism. SSR primers from pigeonpea were found to be more polymorphic (37%) as compared to common bean and lentil markers (Datta *et al.* 2010).

Saxena *et al.* (2010) isolated 36 microsatellite loci from a SSR-enriched genomic library of pigeonpea genotype ‘Asha’. Primer pairs were designed for 23 SSR loci, of which 16 yielded amplicons of expected size. Thirteen SSR markers were polymorphic amongst 32 cultivated and eight wild pigeonpea genotypes representing six *Cajanus* species. These markers amplified a total of 72 alleles ranging from two to eight alleles with an average of 5.5 alleles per locus. The polymorphic information content for these markers ranged from 0.05 to 0.55 with an average of 0.32 per marker. Phenetic analysis clearly distinguished all wild species genotypes from each other and from the cultivated pigeonpea genotypes.

Upadhyaya *et al.* (2008) with the objective of enhancing the utilization of pigeonpea germplasm in breeding and genomic research developed a composite collection of 1000 accessions and profiled using 20 SSR markers. Aruna *et al.* (2009) quantified diversity in a collection of *Cajanus* species selected from a wide geographic range using amplified fragment length polymorphism (AFLP), simple sequence repeats (SSRs) and restriction fragment length polymorphism (RFLP).
Polymorphism was higher among the wild accessions than among the cultivated genotypes. Low level of genetic diversity was also revealed in cultivated pigeonpea as compared to its wild relatives using diversity arrays technology (DArT). Most of the diversity was among the wild relatives of pigeonpea or between wild species and cultivated *Cajanus cajan* (Yang et al. 2006).

Odeny et al. (2009) developed microsatellite markers from an enriched library of pigeonpea and also tested the transferability of soybean microsatellites in pigeonpea. Primers were designed for 113 pigeonpea genomic SSRs, 73 of which amplified interpretable bands. Thirty-five of the primers revealed polymorphism among 24 pigeonpea breeding lines. The number of alleles detected ranged from 2 to 6 with a total of 110 alleles and an average of 3.1 alleles per locus. GT/CA and GAA class of repeats were the most abundant dinucleotide and tri-nucleotide repeats respectively. Additionally, 220 soybean primers were tested in pigeonpea, 39 of which amplified interpretable bands. But due to lack of polymorphism they were not of much use. Despite the observed morphological diversity, a little genetic diversity was detected within cultivated pigeonpea as revealed by the developed microsatellites. Besides, a mapping population (F6 RILs) was developed for resistance to *Fusarium* wilt. Nine markers showed easily scoreable differences between parents.

Songok et al. (2010) studied genetic relationships among 88 pigeonpea accessions from a presumed centre of origin and diversity, India and a presumed secondary centre of diversity in East Africa using six microsatellite markers. Forty-seven alleles were detected in the populations studied, with a mean of
eight alleles per locus. Populations were defined by region (India and East Africa) and sub-populations by country in the case of East Africa and state in case of India. Substantial differentiation among regions was evident from Roger’s modified distance and Wright’s $F$ statistic. Greatest genetic diversity in terms of number of alleles, number of rare alleles and Nei’s unbiased estimate of gene diversity ($H$) was found in India as opposed to East Africa. This supported the hypothesis that India is the centre of diversity and East Africa is a secondary centre of diversity. Within East Africa, germplasm from Tanzania had the highest diversity according to Nei’s unbiased estimate of gene diversity, followed by Kenya and Uganda. Germplasm from Kenya and Tanzania were more closely related than that of Uganda according to Roger’s modified distance. Within India, results did not indicate a clear centre of diversity. Values of genetic distance indicated that genetic relationships followed geographical proximity (Songok et al. 2010).

Dubey et al. (2011) reported generation of large scale genomic resources for pigeonpea. FLX/454 sequencing carried out on a normalized cDNA pool prepared from 31 tissues produced 4,94,353 short transcript reads (STRs). The comparison of pigeonpea transcriptome assembly showed similarity to soybean gene models. Additionally, illumina 1G sequencing was performed on *Fusarium* wilt and sterility mosaic disease challenged root tissues of 10 resistant and susceptible genotypes. A large set of markers including 8,137 simple sequence repeats, 12,141 single nucleotide polymorphisms and 5845 intron-spanning regions were identified.
A comprehensive transcriptome assembly for pigeonpea has been developed using Sanger and Second generation sequencing platforms (Kudapa et al. 2012). The resultant transcriptome assembly, referred to as CcTA v2, comprised 21,434 transcript assembly contigs (TACs) with an N50 of 1,510 bp, the largest one being 8 kb. Of the 21,434 TACs, 16,622 (77.5%) could be mapped on to the soybean genome build 1.0.9 under fairly stringent alignment parameters. Based on knowledge of intron junctions, 10,009 primer pairs were designed from 5,033 TACs for amplifying intron spanning regions (ISRs). By using in silico mapping of BAC-end-derived SSR loci of pigeonpea on the soybean genome as a reference, putative mapping positions at the chromosome level were predicted for 6,284 ISR markers, covering all 11 pigeonpea chromosomes. A subset of 128 ISR markers was analyzed on a set of eight genotypes. While 116 markers were validated, 70 markers showed one to three alleles, with an average of 0.16 polymorphism information content (PIC) value.

Amplified Fragment Length Polymorphism (AFLP) analysis in pigeonpea revealed close relationship of cultivated genotypes with some of its wild relatives. A total of 561 AFLP loci were to study genetic diversity of wild and cultivated genotypes of pigeonpea. Analysis of molecular variance (AMOVA) revealed significant strong population structure when genotypes were structured according to continent of origin ($F_{ST}=0.22$) also when structured into cultivated and wild genotypes ($F_{ST}=0.16$). Maximum polymorphic loci were observed in cultivated species which is due to more number of genotypes used. Clustering analysis revealed most cultivated genotypes grouped into one major cluster while, the wild
genotypes grouped into many clusters revealing greater diversity within wild species as compared to cultivated genotypes (Ganapathy et al. 2010).

Genomic relationships among 11 species in the genus *Cajanus* was revealed by seed protein (albumin and globulin) polymorphisms. SDS-PAGE analysis of seed albumins and globulins from two pigeonpea, *Cajanus cajan*, cultivars (DSLR-17 and BDN-2) and ten wild species, including *C. cajanifolius*, *C. lineatus*, *C. sericeus*, *C. acutifolius*, *C. lanceolatus*, *C. reticulates*, *C. albicans*, *C. scarabaeoides*, *C. volubilis* and *C. platycarpus*, resulted in 34 albumin and 27 globulin polypeptides. Proximity matrix analysis based on electrophoretic banding patterns of albumins and globulins jointly revealed *C. cajanifolius* to be closest to *C. cajan* having similarity coefficients of 0.595 and 0.676, respectively. Cluster analysis also exhibited the grouping of *C. cajanifolius* with *C. cajan* in one cluster (Panigrahi et al. 2007).

The phylogenetic relationship of pigeonpea [*Cajanus cajan* (L.) Millsp.] and its wild relatives was reported based on seed protein profiles. A considerable variation was detected among the protein profiles of different accessions of *C. cajan* while those of wild species were very specific and distinctly different from each other. The clustering of 10 wild species and *C. cajan* more or less agrees with their sectional classification and available data based on morphological characteristics, crossability, genome pairing in hybrids and nuclear RFLPs (Jha and Ohri, 1996).

Diversity in 28 accessions representing 12 species of the genus, *Cajanus* arranged in 6 sections including 5 accessions of the cultivated species, *C. cajan*,
and 4 species of the genus *Rhyncosia* available in the germplasm collection at ICRISAT was assessed using RFLP with maize mtDNA probes. Cluster analysis of the Southern blot hybridization data with 3 restriction enzymes – 3 probe combinations placed the genus *Rhyncosia* in a major group well separated from all the species belonging to the genus *Cajanus*. Within the genus *Cajanus*, the 4 accessions of *C. platycarpus* belonging to section *Rhynchosoides* formed a separate group in contrast to those in other sections of pigeonpea. In the section, *Cajanus* all the 5 accessions of *C. cajan* were grouped together and *C. cajanifolius* belonging to the same section was in a subgroup by itself closer to the main group. The four accessions of *C. scarabaeoides*, were together and the other species belonging to section *Cantharospermum* were in different subgroups. The intra-specific variation was seen even within accessions of certain pigeonpea wild species such as *C. scarabaeoides*, *C. platycarpus*, *C. acutifolius*, and even the cultivated species of *C. cajan* (Sivaramakrishnan et al. 2002).

Wasike et al. (2005) used AFLP to study genetic variability and relatedness between Asian and African pigeonpea cultivars. Forty-one samples, 32 African and 9 Asian varieties were subjected to the analyses. Phenetic analysis revealed no major clusters and indicated limited genetic variability among the samples. Analysis of molecular variance (AMOVA) at continent wide hierarchical level, revealed a significantly weak population structure ($F_{ST} = 0.05$, $P= 0.001$) and Fishers’ exact tests ($P<0.05$) provided no support for population differentiation. AMOVA based on treating the cultivars as samples from a
panmictic population revealed a stronger genetic structure ($F_{ST} = 0.09$, $P=0.001$). This study suggested that East Africa pigeonpeas were closely related but less genetically diverse than Indian cultivars.

Kassa et al. (2012) reported genetic patterns of domestication in pigeonpea and wild *Cajanus* relatives using 752 single nucleotide polymorphisms (SNPs) derived from 670 low copy orthologous genes. Among all species analyzed *Cajanus cajanifolius* was found to be the most probable progenitor of cultivated pigeonpea. Multiple lines of evidence suggested recent gene flow between cultivated and non-cultivated forms, as well as historical gene flow between diverged but sympatric species. Evidence supported that primary domestication occurred in India, with a second and more recent nested population bottleneck focused in tropical regions. Abundant allelic variation and genetic diversity was found among the wild relatives, with the exception of wild species from Australia for which a third bottleneck unrelated to domestication within India was reported. Domesticated *C. cajan* possessed 75% less allelic diversity than the progenitor clade of wild Indian species, indicating a severe ‘bottleneck’ during pigeonpea domestication.

Genetic diversity was analyzed among 77 pigeonpea genotypes adapted to South American regions based on microsatellite markers and their transferability was evaluated in *Phaseolus vulgaris* and *Vigna unguiculata* species (Barbosa de Sousa et al 2011). The number of alleles per locus ranged from 2 to 12, with an average of 5.1 alleles. The PIC values ranged from 0.11 to 0.80 (average 0.49) and the D values from 0.23 to 0.91 (average 0.58). The
averages of observed and expected heterozygosity were 0.25 and 0.47, respectively, showing a deficit in heterozygosity. A model-based Bayesian approach implemented in the software STRUCTURE was used to assign genotypes into clusters. A dendrogram was constructed based on the modified Roger’s genetic distances using a neighbor-joining method (NJ). A total of four clusters were assembled by STRUCTURE and a strong tendency of correspondence between the Bayesian clusters in the NJ tree was observed. The genetic distance ranged from 0.09 to 0.62 (average 0.37), showing a low genetic diversity in the pigeonpea genotypes. Transferability of pigeonpea-specific microsatellites revealed a cross-amplification and the presence of polymorphic alleles in *P. vulgaris* and *V. unguiculata*.

### 2.5 Resistance gene analogs (RGAs) in crop plants

All known resistance (R) genes can be grouped into a few classes based on their sequence structure and functional domain/motifs. Most of these R-genes belong to the nucleotide binding site (NBS)-leucine rich repeat (LRR) type (Martin, 1999). R-genes of this class share conserved domains and structural similarities even though they are from diverse taxonomic groups (monocots and dicots) and confer resistance to viral, fungal or bacterial pathogens. More than 50 R-genes have been cloned so far from a variety of plant species (Wenkai *et al.* 2006). These genes confer resistance to a diversity of pathogens including bacteria, fungi, oomycetes, viruses, insects and nematodes (Martin *et al.* 2003). Leister *et al.* (1996) developed a PCR based method to easily isolate resistance gene analogues, from a wide variety of plant species, in which they used degenerate
primers that amplify between the kinase 1a motif of the NB-ARC domain and the GLPL motif that lies about 160 amino acids further downstream. RGAs have been successfully isolated using the PCR based approach from a wide range of plants including potato (Leister et al. 1996), soybean (Kanazin et al. 1996), *Arabidopsis* (Speulman et al. 1998), maize (Collins et al. 1998), rice (Leister et al. 1998), wheat and barley (Seah et al. 1998), tomato (Ohmori et al. 1998), lettuce (Shen et al. 1998), bean (Rivkin et al. 1999), citrus (Deng et al. 2000), coffee (Noir et al. 2001), chickpea (Huettel et al. 2002), barrel medic (Zhu et al. 2002), grapevine (Di Gaspero and Cipriani, 2002), peanut (Bertioli et al. 2003), cotton (Tan et al. 2003), pine (Liu and Ekramoddoullah, 2003), strawberry (Martinez et al. 2004), oat (Irigoyen et al. 2006), buffalo grass and *Argostis* species (Budak et al. 2006). Many of these RGAs map in close proximity to known resistance genes. Traditional breeding methods are very time consuming and development of resistant cultivars may take up to 15-20 years. RGAs can be utilized as a marker to be applied in marker assisted selection for early release of disease resistant varieties. The RGA fragments were used as molecular markers for tagging the disease resistance loci in wheat (Chen et al. 1998), rice (Ilag et al. 2000), melon (Mas et al. 2001), cowpea (Gowda et al. 2002), *Lycopersicon* (Zhang et al. 2002), common bean (Lopez et al. 2003), cocoa (Lanaud et al. 2004). Hence, this technique is useful in identifying potential disease resistance loci and help breeders to fish out resistance gene over species and genera.
2.5.1 Association of RGAs with resistance genes

Kanazin et al. (1996) used primers designed for conserved sequences from coding regions of disease resistance genes N (tobacco), RPS2 (Arabidopsis) and L6 (flax) to amplify similar sequences from soybean [Glycine max (L.) Merr.]. Nine classes of RGAs were detected. Genetic mapping of members of these classes located them to eight different linkage groups. Several RGA loci mapped near known resistance genes. Clustering and sequence similarity of members of RGA classes suggested a common process in their evolution.

The degenerate primers designed from conserved NBS-LRR regions of known disease resistance genes were used to amplify different resistance gene like (RGL) DNA fragments from Arabidopsis thaliana accessions Landsberg erecta and Columbia. Almost all cloned DNA fragments were genetically closely linked with known disease resistance loci. Most RGL fragments were found in a clustered or dispersed multi-copy sequence organization, supporting the supposed correlation of RGL sequences and disease resistance loci (Aarts et al. 1998).

Degenerate primers based on conserved NBS of resistance genes of Arabidopsis, flax and tobacco were used to amplify resistance gene analogs of 500bp in length in rice. The fragments were cloned and analyzed based on southern blot analysis. Fourteen clones, each representing one of the 14 categories of RGAs were mapped onto the rice genetic map using a Nipponbare (japonica) x Kasalath (indica) mapping population consisting of 182 F2 lines. Of the 14 clones representing each class, 12 could be mapped onto five
chromosomes of rice with major cluster of 8 RGAs on chromosome 11 (Mago et al. 1999).

Degenerate oligonucleotides designed to recognize conserved coding regions within the nucleotide binding site (NBS) and hydrophobic region of known resistance (R) genes from various plant species were used to target PCR to amplify resistance gene analogs (RGAs) from a cowpea (Vigna unguiculata L.Walp.) cultivar resistant to Striga gesnerioides. PCR products consisted of a group of fragments approximately 500 bp in length that migrated as a single band during agarose gel electrophoresis. The nucleotide sequence of fifty different cloned fragments was determined and their predicted amino acid sequences compared to each other and to the amino acid sequence encoded by known resistance genes, and RGAs from other plant species. Cluster analysis identified five different classes of RGAs in cowpea. Gel blot analysis revealed that each class recognized a different subset of loci in the cowpea genome. Several of the RGAs were associated with restriction fragment length polymorphisms, which allowed them to be placed on the cowpea genomic map (Gowda et al. 2002).

Comparative sequence analysis of the resistance gene analog marker locus aACT/CAA (originally found to be tightly linked to the multiallelic barley Mla cluster) from genomes of barley, wheat and rye revealed a high level of relatedness among one another and showed high similarity to a various number of NBS-LRR disease resistance proteins (Mohler et al. 2002). Using the sequence-specific polymerase chain reaction, RGA marker aACT/CAA was mapped on group 1S chromosomes of the Triticeae and was associated with
disease resistance loci. In barley and rye, the marker showed linkage to orthologous powdery mildew resistance genes *Mla1* and *Pm17*, respectively, while in wheat linkage with a QTL against *Fusarium* head blight disease was determined.

In chickpea, using the RGA primers, which are designed based on the conserved motifs present in characterized R-genes, Bulk Segregant Analysis (BSA) was performed on a resistant bulk and a susceptible bulk along with parents for ascochyta blight resistance. Of all available RGAs and their 48 different combinations, only one RGA showed polymorphism during BSA. This marker was evaluated in an F7:8 population of 142 RILs from an interspecific cross of *C. arietinum* (FLIP 84-92C) × *C. reticulatum* (PI 599072) and was mapped to *Cicer* linkage map (Rajesh et al. 2002).

Eight resistance gene analogs were isolated from wild rice, *Zizania latifolia* by degenerate primers designed according to conserved motifs or around the nucleotide binding site of known NBS-containing plant resistance genes. Eight RGAs were classified into 6 distinct groups based on their deduced amino acid sequences similarity of 60%. Eight *Zizania* RGAs belong to the non-TIR NBS-LRR subgroup (Chen et al. 2006).

Degenerate primers designed based on known resistant genes and resistance gene analogs were used in combinations to elucidate RGAs from *Sorghum bicolor*, cultivar M35-1. Most of the previously tried primer combinations resulted in amplicons of expected 500–600 bp sizes in sorghum along with few novel combinations. Restriction analysis of PCR amplicons of expected size
revealed a group of fragments present in a single band indicating the heterogeneous nature of the amplicon. Many of these were cloned and sequenced and their predicted amino acid sequences compared to each other and to the amino acid sequences of known R-genes revealed significant sequence similarity. A cluster analysis based on neighbor-joining method was carried out using sorghum RGAs (SRGAs) together with several analogous known R-genes resulting in two major groups; cluster-I comprising only SRGAs and cluster-II comprised of known R-gene sequences along with three SRGAs. Further analysis clearly indicated similarity of SRGAs in overall sense with already known ones from other crop plants (Totad et al 2005).

Basak et al (2007) cloned and sequenced one yellow mosaic virus-resistance linked R gene homolog (RGH) from Vigna mungo, line VM-1, GenBank accession number AY297425. Later, two other RGHs from YMV resistant lines, V. mungo WBU 108 and V. radiata Pusa 9072 were selectively amplified using R gene targeted degenerate primers and were cloned subsequently (AY301991 and trIQ7XZT9, respectively). Characterization of these three RGHs and analysis of a total of 221 R-genes and RGHs raised the question of the evolution and distribution of the R-genes/RGHs in the family Fabaceae, to which the primary hosts of the YMV belong. The phylogenetic analyses indicated that two-third of the sequences are of the TIR-NBS type, while about one third represent the Non-TIR subfamily. Simultaneous presence of the TIR and the Non-TIR domains within the Fabaceae indicated divergent evolution and heterogeneity within the NBS domain. The finding reflected that the
successful introgression of the functional R gene could be possible to the disease susceptible cultivars within the tribes Phaseoleae and Trifoleae.

A PCR approach with degenerate primers designed from conserved NBS–LRR (nucleotide binding site – leucine rich repeat) regions of known disease resistance genes was used to amplify and clone homologous sequences from 5 faba bean (Vicia faba) lines and 2 chickpea (Cicer arietinum) accessions. Sixty-nine sequenced clones showed homologies to various R-genes deposited in the GenBank database. The presence of internal kinase-2 and kinase-3a motifs in all the sequences isolated confirmed that these clones correspond to NBS-containing genes. Using an amino-acid sequence identitiy of 70% as a threshold value, the clones were grouped into 10 classes of resistance gene analogs (RGA01 to RGA10). A phylogenetic tree based on the deduced amino-acid sequences of 12 representative clones from the 10 RGA classes and the NBS domains of 6 known R genes (I2 and Prf from tomato, RPP13 from Arabidopsis, Gro1–4 from potato, N from tobacco, L6 from flax), clearly indicated the separation between TIR (Toll/interleukin-1 receptor homology: Gro1–4, L6, N, RGA05 to RGA10) and non-TIR (I2, Prf, RPP13, RGA01 to RGA04) type NBS–LRR sequences (Palomino et al. 2006).

The resistance gene analog polymorphism (RGAP) has been used to identify tightly linked markers for disease resistance genes and to enrich the genetic map with a different class of markers in crops, including barley (Chen et al. 1998; Toojinda et al. 2000), tomato (Sanjukta et al. 2007) and wheat (Xie et al. 2008). The previous studies have shown that the RGAs might be the part of
resistant genes, or link tightly to it, or have no association with it (Leister et al. 1996; Collins et al. 1998). Wherever studied, PCR cloning of disease resistance analogs had been a promising approach to obtain disease resistance gene candidates and to develop molecular markers (Feuillet et al. 1997; Chen et al. 1998). Resistance gene analogues of *Cicer* were isolated by different PCR approaches and mapped in an inter-specific cross segregating for *Fusarium* wilt by Restriction Fragment Length Polymorphism (RFLP) and Cleaved Amplified Polymorphic Site (CAPS) analysis (Huettel et al. 2002). Mutlu et al. (2006), developed resistance gene analog polymorphism (RGAP) markers for common bean (*Phaseolus vulgaris* L.), which co-localize with disease resistance gene and QTL in common bean.

Oligonucleotides already designed from sequence motifs conserved between resistance genes *N* of tobacco and *RPS2* of *Arabidopsis thaliana* were used as PCR primers (AS1/S2) to scan the rice blast disease resistant Moroberekan genomic DNA. The fragment amplified by the primer AS1/S2 was cloned and sequenced. The PCR products for the other three primers were sequenced directly. Homology search of the resultant nucleotide sequences and deduced amino acid sequences with the reported sequences available in public data bases of NCBI BLASTn and PSI blast indicated the presence of resistance protein-like gene in BRGA-1(blast resistant gene analogue-1), putative retro-elements and putative retro-transposons proteins in BRGA-2, mitochondrial DNA in BRGA-3 and NBS-LRR type resistance protein and NB-ARC domain containing expressed protein of *Oryza sativa* in BRGA-4 (Selvaraj et al. 2011).
2.5.2 Characterization of RGAs in crop plants

Isolation, cloning and characterization of resistance gene analogs were reported in pearl millet by Ramachandra et al. (2011). Using specific primers designed from the conserved NBS regions, 22 RGAs were cloned and sequenced from pearl millet (Pennisetum glaucum L. Br.). Phylogenetic analysis of the predicted amino acid sequences grouped the RGAs into nine distinct classes. GenBank database searches with the consensus protein sequences of each of the nine classes revealed their conserved NBS domains and similarity to other known R-genes of various crop species. One RGA 213 was mapped onto LG1 and LG7 in the pearl millet linkage map.

Joshi et al. (2011) used bioinformatic tools to detect and characterize NBS type R-genes from Curcuma longa transcriptome. Insilico characterization of EST database resulted in the detection of 28 NBS types R-gene sequences in Curcuma longa. All the 28 sequences represented the NB-ARC domain, 21 of which were found to have highly conserved motif characteristics and categorized as regular NBS genes. Most alignment occurred with monocots (67.8%) with emphasis on Oryza sativa and Zingiber sequences. All best alignments with dicots occurred with Arabidopsis thaliana, Populus trichocarpa and Medicago sativa.

A PCR strategy was used to amplify resistance gene analogues in Vigna spp. using degenerative primers designed at the conserved motif of cloned plant NBS–LRR R-genes. Out of nine RGA fragments amplified, five sequences showed homologies to various R-genes deposited in the GenBank database.
Phylogenetic analysis of *Vigna* RGAs showed all RGAs belonged to TIR-NBSLRR class of R-genes. The amino acid identity of *Vigna* RGAs to various *R* genes ranged from 3% to 37%. In addition, eight AFLP-RGA primer combinations were used to amplify 11 AFLP-RGA fragments showing 45% polymorphic markers between two genotypes namely, MYMV resistant TNAU Red and susceptible VRMGg 1. Through AFLP-RGA analysis it was found that about 18% of ricebean alleles were introgressed into resistant RIL F9 (Mahadeo, 2009).

An attempt was made to understand the genetic difference between mungbean and ricebean for the presence/absence and expression pattern of the homologue of a resistant gene namely *N*-gene of tobacco. PCR analysis using degenerate primers designed from conserved regions of *N*-gene of tobacco from NCBI Genbank database, revealed the presence of *N*-gene homologue in all the accessions of both mungbean and ricebean. Agroinoculation studies confirmed the resistance of ricebean accession TNAU-red and susceptibility of mungbean variety VRM 1 against mungbean yellow mosaic virus (MYMV). Semi-quantitative RT-PCR analysis revealed the down-regulation of the *N*-gene homologue in mungbean and up-regulation in ricebean upon MYMV infection. This differential expression of the homologue of *N*-gene of tobacco upon MYMV infection may play a crucial role in conferring resistance against MYMV (Rajgopal, 2008).

Genomic sequences sharing homology with NBS region of known resistance genes were isolated and characterized from under-exploited plant species (*Pongamia glabra, Adenanthera pavonina, Clitoria ternatea, Solanum trilobatum*) using PCR approach with primers designed from conserved regions.
of NBS domain and all the four RGAs isolated had high level of identity with NBS-LRR family of RGAs deposited in the GenBank. The extent of identity between the sequences at NBS region varied from 29% (P. glabra and S. trilobatum) to 78% (A. pavonina and C. ternatea), which indicates the diversity among the RGAs (Thirumalaiandi et al. 2008).

The resistance gene analog approach was used to analyze genetic diversity among the 40 sugarcane cultivars that vary in their resistance to red rot disease. About 29 RGA primers designed from the conserved domains of resistance proteins were used. The genetic similarity values ranged from 58.4 - 90% with the mean genetic similarity of 74.2%. Cluster analysis resulted in a dendrogram with 3 major clusters and a clear distinction of resistant and susceptible varieties was observed. A total of 25 specific fragments amplified by 14 primers were identified to be associated with resistance and 8 specific fragments amplified by 8 primers were associated with susceptibility. Amplification of the red rot resistant variety Bo 91 and the red rot susceptible variety CoC 67 with the twenty nine RGA primers, followed by sequencing and homology analysis revealed significant homologies with the RGA’s of rice, maize and sugarcane (Jayashree et al. 2010).

Genomic DNA sequences sharing homology with NBS region of resistance gene analogs were isolated and characterized from resistant genotypes of finger millet using PCR based approach with primers designed from conserved regions of NBS domain. Attempts were made to identify molecular markers linked to the resistance gene and to differentiate the resistant bulk from
the susceptible bulk. A total of 9 NBS-LRR and 11 EST-SSR markers generated 75.6% and 73.5% polymorphism respectively amongst 73 finger millet genotypes. NBS-5, NBS-9, NBS-3 and EST-SSR-04 markers showed a clear polymorphism which differentiated resistant genotypes from susceptible genotypes. By comparing the banding pattern of different resistant and susceptible genotypes, five DNA amplifications of NBS and EST-SSR primers (NBS-05504, NBS-09711, NBS-07688, NBS-03509 and EST-SSR-04241) were identified as markers for the blast resistance in resistant genotypes. Principal coordinate plot and UPGMA analysis formed similar groups of the genotypes and placed most of the resistant genotypes together showing a high level of genetic relatedness and the susceptible genotypes were placed in different groups on the basis of differential disease score (Panwar et al. 2010).

Degenerate primers designed based on known resistance genes were used in combinations to elucidate resistance gene analogs from Curcuma longa cultivar Surama. The three primers resulted in amplicons with expected sizes of 450-600 bp. The nucleotide sequence of these amplicons was obtained through sequencing; their predicted amino acid sequences compared to each other and to the amino acid sequences of known R-genes revealed significant sequence similarity. The finding of conserved domains, viz., kinase-1a, kinase-2 and hydrophobic motif, provided evidence that the sequences belong to the NBS-LRR class gene family. The presence of tryptophan as the last residue of kinase-2 motif further qualified them to be in the non-TIR-NBS-LRR subfamily of resistance genes. A cluster analysis based on the neighbor-joining method was
carried out using *Curcuma* NBS analogs together with several resistance gene 
analogs and known R-genes, which classified them into two distinct subclasses, 
corresponding to clades N3 and N4 of non-TIR-NBS sequences described in 
plants (Joshi *et al*. 2010).