Pigeonpea \([\textit{Cajanus cajan} \, (\text{L.}) \, \text{Millsp.}]\) is an important legume crop of tropical and sub-tropical regions of the world. Pigeonpea belongs to the sub-tribe Cajaninae of the agriculturally most important tribe Phaseoleae under sub-family Papilionoideae of the family Leguminosae (Fabaceae). \textit{Cajanus cajan} is the only domesticated species under Cajaninae. The true origin of pigeonpea is still disputable. However, various studies have indicated that pigeonpea originated from its closest wild relative \textit{Cajanus cajanifolius} most probably in India. Because of its multiple usages as food, fodder, fuel wood, nitrogen fixation, soil conservation and adaptability to diverse soil and climatic conditions, pigeonpea plays an important role in subsistence agriculture.

Pigeonpea germplasm represents a diverse set of landraces and heterogenous feral forms that are adapted to various agro-ecological settings. Despite extensive phenotypic diversity, molecular evidences have suggested very low genetic diversity within cultivated pigeonpea when compared to its wild relatives. The only means to broaden the genetic base of domesticated \textit{C. cajan} is to introgress genetic diversity from the wild gene pool and understanding how diversity is assorted among pigeonpea and its wild relatives has practical implications. The two main reasons behind reduced genetic diversity are the founder effect and inbreeding. The founder effect occurs when only a few individuals are used while domestication of a wild species; while introducing a crop into a new region or when breeders use only a few cultivars for all subsequent crop improvement. Domestication
creates genetic bottlenecks that can decrease genetic diversity, change allele frequencies, increase linkage disequilibrium (LD), and eliminate rare alleles in the resulting population. The process of domestication and divergence of different groups of cultivated germplasm is also influenced by gene flow from wild relatives.

The wild relatives of pigeonpea are crossable with cultivated pigeonpea and possess resistance to one or the other biotic or abiotic stresses in addition to having agronomic traits useful in crop improvement. For example, *Cajanus volubilis* has been reported as a resistance source against sterility mosaic disease. Similarly, *Cajanus acutifolius* has resistance to pod borer and pod fly and tolerance to drought and salinity and *Cajanus lineatus* has resistance to sterility mosaic disease, *Alternaria* blight and drought tolerance. However, the development of improved pigeonpea types through hybridization with wild relatives has limited success. *Fusarium* wilt is an important soil borne disease of pigeonpea which causes significant yield losses throughout pigeonpea growing areas. Based on several pathological and molecular studies, it has been concluded that the use of cultivars resistant to the disease is the effective means of wilt control. Hence, resistance breeding continues to be a major breeding objective in pigeonpea improvement programmes.

Genomic resources like molecular markers, gene tags, genetic maps and transcriptome or genome sequence data are prerequisites for undertaking molecular breeding in any crop. Microsatellites or simple sequence repeats (SSRs) are short tandem repeats of 1-6 nucleotides
evenly dispersed throughout the genome and are present in all eukaryotic
genomes. SSRs are becoming standard DNA markers for evolutionary and
population genetic studies and are being widely used in marker assisted
breeding. Microsatellites have also attracted scientific attention because they
have been shown to be part of or linked to some important genes of
agronomic interest. However, the conventional method of development of
SSR markers from genomic libraries is expensive, labour intensive and time
consuming. Hence, recent studies have used bioinformatics tools to detect
SSRs in sequence data generated from large scale genome sequencing
projects. With the establishment of sequencing projects for gene discovery
programs in several plant species, a wealth of DNA sequence information
has been generated and deposited in online databases. These sequences
form a good source for the identification of SSRs in such crop plants.

Genes conferring resistance to phytopathogenic bacteria, viruses,
fungi and nematodes have been cloned from several plant species.
Molecular markers tightly linked to the genes conferring resistance will be of
advantage in breeding for resistant varieties. Resistance gene analog
polymorphism (RGAP) technique has proven to be efficient in identification
of molecular markers for disease resistance. Resistance gene analogs
(RGAs) are potential genic regions that contain structural motifs such as
NBS (nucleotide binding site) and LRR (leucine rich repeat) that are common
to most of the cloned resistance genes. PCR based methodologies have
been developed to easily isolate resistance gene analogs from a wide variety
of plant species in which they used degenerate primers to amplify RGAs.
RGAs have the potential to serve as closely linked markers for marker assisted breeding or even as candidate resistance genes. So far the sources of wilt resistance in pigeonpea have been limited and there is a need to study the evolutionary reasons for presence of limited variation for this important trait in cultivated gene pool. The present study aims to relate genetic diversity to genes conferring wilt resistance and also attempts to identify the evolutionary bottlenecks induced by the process of domestication that might have played a role in occurrence of reduced levels of resistance in cultivated germplasm. The present study has been planned to carry out work in pigeonpea with the following objectives:

**Objectives**

1. Mining of public databases for microsatellites or Simple Sequence Repeat (SSR) markers and characterization of identified SSRs.
2. Isolation, identification and pure culturing of the wilt pathogen *Fusarium udum* from seed samples/germplasm of pigeonpea collected from diverse agro-climatic conditions in India.
3. Screening the cultivated pigeonpea germplasm for resistance and susceptibility against *Fusarium udum*.
4. Identification and characterization of resistance gene analogues (RGAs) responsible for determining resistance to *Fusarium* wilt in pigeonpea.

A brief account of results of the present study is given below:

**Results**
Fifty one pigeonpea accessions from different eco-geographic regions of India including 13 wilt resistant genotypes and seven wild species were used for the study.

**Characterization of Simple Sequence Repeats (SSRs) in pigeonpea**

1. **Abundance of SSRs in pigeonpea genome**

The DNA database of NCBI (National Centre for Biotechnology Information) was scanned for the identification of SSRs in the pigeonpea BAC (Bacterial Artificial Chromosome)-end sequences (BESs). The downloaded BESs were analysed for SSRs using the Tandem Repeat Finder (trf) program. The study involved identification of presence of repeat motifs in the pigeonpea DNA sequences, totalling about 88,066 from which 22,358 (25.39%) repeats were identified. Out of these 2773 (12.4%) had mono- to 10-mer motifs. The dinucleotide was the most abundant class of SSRs (1559), followed by trinucleotide (366) and hexanucleotide repeats (236). Tetra- and pentanucleotide repeats occurred at lower proportions. The identified di- and tri-nucleotide repeats with copy number \( \geq 5 \) were selected for primer designing in order to amplify the target SSR regions. A total of 139 primers were synthesized.

2. **Diversity analysis in pigeonpea germplasm**

Total genomic DNA was isolated from leaves of single plant per accession using CTAB method. Out of 139 SSR primer pairs screened, sixty-seven primer pairs producing very good, scorable and polymorphic amplification product were selected for diversity analysis in 58 pigeonpea accessions.
including seven wild relatives. A total of 251 alleles were detected in the genotypes analyzed with an average of 3.75 alleles per locus.

The SSR data was used for genetic diversity analysis using softwares FreeTree and NTSYSpc. The dendrogram generated based on Nei and Li’s genetic distances classified the 58 pigeonpea accessions into five clusters. The wild species formed a separate cluster. The value of Nei and Li’s genetic distance ranged from 0.132-0.411 for varieties from north and 0.189-0.486 for varieties from peninsular region. The value of matrix correlation coefficient (r) was 0.825 which indicated good fit and support for clustering.

3. Differentiation and partitioning of genetic variation

In order to test the degree of differentiation of the accessions belonging to various regions of India, analysis of molecular variance (AMOVA) was conducted using software Arlequin version 3.5.1.2. The total 51 genotypes were divided into two groups- released varieties and cultivated germplasm and into five sub-groups (north, west, peninsular, east and central) based on the source of material. Analysis of molecular variance indicated that variation among groups accounted for -2.36% of total variation and 14.44% of variation was contributed by differences among sub-groups within groups. The percentage of variation among individuals within sub-groups was 48.47% and variation within individuals contributed 39.45% of total variation.

A comparison of pairwise $F_{ST}$ among seven sub-groups indicated that the varieties from north had differentiated very little as compared to others and varieties from west were highly differentiated from germplasm from peninsular and east regions as well as varieties from peninsular region. The
pairwise coancestry coefficient of varieties from west with germplasm from east (0.332) and peninsular region (0.316) was higher than all other pairs of sub-groups. The Nei’s distance was higher for varieties from west region with germplasm from east and varieties and germplasm from peninsular region. The varieties from peninsular region observed highest average number of pairwise differences within sub-groups (22.242) followed by varieties from north (19.791). The average number of pairwise differences between sub-groups was highest (25.761) between varieties from peninsular region and varieties from west.

4. Allelic diversity and population sub-structure

The population diversity measures were calculated using POPGEN32 software. The highest number of alleles per locus was observed for the varieties from peninsular region (2.000±0.870) followed by germplasm from east (1.970±0.887) and varieties from north region (1.899±0.819) and germplasm from central region (1.899±0.873) while for germplasm from peninsular region, the lowest number of alleles was recorded (1.448±0.585). Similar results were observed for the estimates of Nei’s gene diversity, Shannon’s information index, number of polymorphic loci and percent polymorphism. Analysis of two groups indicated that the varieties representing accessions from north, west and peninsular regions showed the highest number of alleles per locus (2.239±0.971) followed by germplasm from central north, east and peninsular India (2.194±0.909). Similar trend was observed for the effective number of alleles, Shannon’s information index, Nei’s gene diversity and number of polymorphic loci.
Analysis of population substructure was conducted using the model-based clustering algorithm implemented in STRUCTURE v.2.1. Bar plots from the population sub-structure analyses indicated large scale sharing of alleles among varieties from north and west region and germplasm from northern region.

**Isolation of *Fusarium udum* from pigeonpea seeds**

For isolation of *Fusarium udum*, pigeonpea seeds were surface sterilized with 1% sodium hypochlorite for 1 min and then rinsed with sterile distilled water. Ten seeds of each accession were placed on moistened blotter papers and freshly prepared potato dextrose agar (PDA) medium with five seeds per plate. The plates were incubated at 25°C for 8-10 days and observed for growth of the pathogen *i.e.* *Fusarium udum*. The pathogen was isolated and subcultured on PDA medium for identification. The growth was observed after 7-8 days. The isolate produced brown colour pigmentation in agar medium. Colour of both the macroconidia and microconidia was hyaline. Shape of macroconidia was sickle shaped with blunt ends to elongated sickle shaped with pointed at both ends while shape of microconidia was oval to round.

**Screening of pigeonpea genotypes for Fusarium wilt**

The 51 pigeonpea accessions were screened for resistance and susceptibility by sowing seeds in wilt sick plots at University of Agricultural Sciences (UAS), Raichur, Karnataka. The wilted plants appeared only four weeks after sowing. The death occurred rapidly within susceptible cultivars but was very slow in resistant cultivars. The onset of wilting and number of
wilted plants was recorded weekly; final counts were taken four months after sowing. Wilt incidence was calculated for each line using the formula:

\[
\text{Wilt incidence} = \frac{\text{Number of plants wilted}}{\text{Total number of plants}} \times 100
\]

The overall mean wilt caused by this technique was 67%. The wilt screening on infested soil revealed that out of 51 accessions only 13 accessions (AWR74/15, BDN1, BWR377, C11, ICP8863, ICP9174, ICP87119, IPA38A, KPL43, KPL44, PI3974-30sel, Pusa84 and Pusa992) were found to be resistant.

**Characterization of Resistance Gene Analogs (RGAs) in pigeonpea**

1. **Phylogenetic relationship between wild and cultivated pigeonpea**

The RGA primers reported in previous literature designed from the conserved regions of N gene of tobacco, RPS2 gene of *Arabidopsis* and L6 gene of flax were used for the present study. Some primers were also designed from RGA sequences of soybean and *Phaseolus* in public domain. A total of 50 primers were screened for PCR amplification in four pigeonpea accessions. Out of 50, only seven primers producing single amplification product were selected for final analysis. The amplified products were purified and sequenced. The DNA sequence data was first edited for ambiguities using Ridom Trace Edit version 1.1.0 and then aligned using CLUSTAL X software.

For characterization of RGAs the 58 pigeonpea accessions including seven wild related species analysed in this study were classified into three groups namely, released varieties, cultivated germplasm (including landraces, farmer’s variety and obsolete varieties) and wild species.
Population genetic analysis of the aligned sequences was performed using softwares MEGA 5.1 and DnaSP version 5.0. The software MEGA was used to calculate various genetic divergence statistics like nucleotide composition (base frequencies for each sequence), nucleotide pair frequencies, codon usage, disparity index, codon based test of neutrality, Fisher’s exact test of neutrality, evolutionary divergence between sequences, base composition bias difference between sequences, pairwise distance within groups, among groups and overall averages. The wild species showed higher values for the number of base substitutions per site between sequences in comparison to variety and germplasm with all primers. The estimates of difference in base composition bias per site were also highest in wild species for all primers except one primer RGA6. The minimum evolution trees were generated by MEGA using RGA sequence data. The five and six resistant pigeonpea accessions were grouped together as in the trees obtained by primers GM15 and primer gi13111699 respectively.

2. Polymorphism within and between wild and cultivated pigeonpea
Several measures of sequence diversity were calculated using DnaSP, and these included the number of segregating sites, the number of mutations, number of haplotypes, haplotype diversity, nucleotide diversity, π (π), the expected heterozygosity per nucleotide site and θ (θ), the number of polymorphic sites in a genotypic sample corrected for sample size, average number of differences, the neutrality tests- Tajima’s D test and Fu and Li’s test, the frequency of recombination, gene flow and genetic differentiation.
For the seven RGAs analysed using software DnaSP, a total of 296 variable sites were found (248 singleton variable sites and 48 parsimony informative sites). In addition, 56 indel sites were also detected. Total number of mutations in these loci was 310. The wild *Cajanus* species had substantial diversity with higher values of $\pi$ (expected heterozygosity per nucleotide site) and $\theta$ (nucleotide diversity, the number of polymorphic sites in a genotypic sample corrected for sample size) than those of varieties and germplasm. The haplotype diversity data also showed a higher genetic diversity in wild species for all loci.

### 3. Domestication Bottleneck

The marked reduction in nucleotide diversity recorded in cultivated pigeonpea germplasm and varieties compared to that of wild species suggested that a domestication bottleneck effect has occurred. The nucleotide diversity retained by cultivars was 4.33% ($\pi$) and 5.42% ($\theta$) of its wild population while the germplasm retained only 3.33% ($\pi$) and 5.35% ($\theta$) of wild species diversity. The value of average number of pairwise differences ($k$) was higher in all primers for wild species as expected but for germplasm the value of $k$ was less as compared to varieties.

The values for Tajima’s D were negative for all three groups of pigeonpea with all primers except with primer RGA6 in varieties and germplasm and in wild species with primer RGA1FCG. The wild species showed a positive result of Fu’s Fs test with primers RGA6, gi28190624, gi28190626 and gi13111699. The varieties on the other hand showed a negative value for all primers except gi28190626. The germplasm showed
positive value of Fu’s Fs test for three primers (RGA6, gi28190624, gi28190626) and negative value for other three primers (RGA1FCC, GM15, gi13111699).

The values of ZnS (average of $R^2$ over all pairwise comparisons, $Za$ (the average of $R^2$ over all pairwise comparisons between adjacent polymorphic sites) and $ZZ$ (Za-ZnS) values decreased in germplasm and cultivars as compared to wild species in all loci except gi28190626. The same was observed in case of values of Wall’s B (proportion of pairs of adjacent segregating sites that are congruent) and Wall’s Q (measure of linkage disequilibrium between adjacent pairs of segregating sites).

4. Effect of selection on diversity

A comparison of the varieties and germplasm with wild species indicated an average reduction of nucleotide diversity in both varieties and germplasm. The values of $\pi$ and $\theta$ for wild species were higher as compared germplasm and varieties. While the $\pi$ and $\theta$ values of varieties were higher than that of germplasm. On an average the nucleotide diversity retained by cultivars was 4.33% ($\pi$) and 5.42% ($\theta$) of its wild population while the germplasm retained only 3.33% ($\pi$) and 5.35% ($\theta$) of wild species diversity. The average number of pairwise differences ($k$) within wild species was very high for RGA loci gi13111699 (36.20) and gi28190626 (38.95) while it was very low in varieties and germplasm. The haplotype diversity ($Hd$) also showed decreased values in germplasm when compared with the wild species.

From the present study the following conclusions are drawn:

Conclusions
1. Availability of insufficient number of polymorphic SSRs in pigeonpea necessitated identification of new markers and their characterization to facilitate use in gene tagging, genetic diversity analyses and molecular map construction. A total of 139 primers were designed from the flanking regions of SSRs in BAC-end sequences. The efforts resulted in identification of 67 new polymorphic SSRs and their characterization in addition to 22,358 SSRs identified in genomic sequences.

2. Analyses indicated that the dinucleotide repeat motifs are the most abundant class of SSRs in the pigeonpea BAC-end sequences.

3. The results of genetic diversity analyses revealed presence of no distinct relationships between the genotypes from various geographic areas of cultivation and the place of origin, but higher diversity was observed among the released varieties in comparison to the cultivated germplasm.

4. The results of partitioning of genetic diversity also indicated the need to sample more localities as well as the need to collect samples of higher sizes within a region for assembling diverse germplasm.

5. The genetic diversity estimates are influenced greatly by sample sizes; hence for sub-groups such as varieties from west and germplasm from peninsular region with small number of samples, the diversity values are not indicative.

6. *Fusarium* wilt is soil borne disease but the pathogen may be carried as a contaminant in seeds. The pathogen can be isolated from seeds.
In the present study Fusarium udum was isolated from pigeonpea seeds collected from plants grown on wilt sick plots.

7. The screening of wilt by sowing in infested soil is simple, reliable, easy to apply and cost effective. It allows easy differentiation of resistant and susceptible plants.

8. The RGA primers designed for complementarities to the conserved regions of resistance genes were used to characterize them in 58 wild and cultivated pigeonpea accessions. The RGA analysis revealed that wild Cajanus species have substantial diversity with higher values of π and θ than those of varieties and germplasm.

9. The minimum evolution trees based on nucleotide diversity in 58 pigeonpea accessions for the products from RGA primers GM15 and primer gi13111699 grouped 5-6 wilt resistant genotypes together. This suggests that these two RGAs may be having some role in imparting resistance to wilt in these pigeonpea cultivars although the evidences are not unequivocal. Further confirmation of the relationship between the loci and resistance needs to be ascertained through analyses of biparental mapping populations.

10. A marked reduction in nucleotide diversity recorded in cultivated pigeonpea germplasm and varieties compared to wild species indicates a severe bottleneck during pigeonpea domestication which may be responsible for presence of a fraction of these variations for RGAs in cultivated genepool.
11. The low genetic variability amongst cultivars when compared with the wild suggests that natural and artificial selection has contributed to the selection of specific alleles and to changes of allelic frequencies at specific loci.

12. Further, higher diversity was observed in released varieties as compared to cultivated germplasm; the main reason being representation of germplasm from only Indian subcontinent while in development of released varieties, the parental lines might have included germplasm from other countries. Such results indicate that cross-border germplasm movement has contributed significantly to broadening of the genetic base of pigeonpea cultivars.

13. The wild gene pool of *Cajanus* contains not only high genetic diversity but also unique and rare alleles for agronomically important traits (e.g. *C. scarabaeoides* for wilt resistance; *C. platycarpus* has shown to be the only source of resistance to the P3 race of *Phytophthora* blight disease). The only means to broaden the genetic base of domesticated *Cajanus cajan* is to introgress genetic diversity from the wild gene pool. The pigeonpea breeding and improvement programs would benefit from the continued and expanded use of the huge amount of genetic diversity prevailing in the wild gene pool. Hence, an increased number of wild species from around the world are needed to be analysed to identify more number of useful genes in order to enhance the possibility of their incorporation in cultivars.
14. Further, genome wide association studies can be carried out to identify specific genomic regions controlling wilt resistance. The inheritance of resistance to *Fusarium* wilt is to be investigated more deeply along with the variability in the disease causing pathogen.