The plant-specific transcription factors play a key role in regulation of gene expression by interacting with conserved sequences of the promoters region. Dof (DNA binding with one finger) is one of the most widely studied transcription factor involved with multifarious roles that are unique to plants. The proteins of Dof family comprise of 200-400 amino acids having a highly conserved domain (Dof domain) of 50-52 amino acids including a C2C2-type zinc-finger motif at its N-terminal region and binds specifically to DNA sequences with a T/AAAAG core. There exists great diversity in terms of number of Dof genes in different crops and only few of them have been characterized in cereals namely rice, wheat, barley and maize.

The present thesis entitled "Molecular cloning and in silico studies of Dof (DNA binding with one finger) transcription factor genes and domains of different cereals and millets" reports PCR based amplification of conserved Dof domain of cereals and millets, cloning of Dof genes of rice (Oryza sativa), wheat (Triticum aestivum), barley (Hordeum vulgare), sorghum (Sorghum bicolor), maize (Zea mays) and finger millet (Eleusine coracana), extensive in silico characterization of cloned genes and PCR amplified Dof domains. Further genome wide identification and in silico characterization of Dof gene family of recently sequenced sorghum genome has been studied. The significant findings and summary of the present investigation are mentioned below:

A total of 35 sets of primers specific for conserved Dof domain and Dof genes were designed based on available Dof sequences of cereals viz. rice, wheat, barley and maize. The PCR amplification with Dof gene-specific primers resulted in multiple band formation showing considerable variability among cereals and millets in terms of number and sizes of bands reflected the inherent diversity of Dof genes in the genomes of different crops. The phylogenetic tree constructed based on the banding patterns resulting from PCR amplification with Dof domain and gene-specific primers revealed existence of six cereals and six millets into two major clusters A and B, which were further bifurcated into two sub-clusters. The DNA polymorphism observed with these sets of primers for cereals and millets indicate a possibility of using these as functional markers for diversity study of agronomically important cereal crops after extensive validation.

The PCR amplified conserved Dof domains of cereals (rice, wheat, oat, sorghum, barley) and millets (finger millet, barnyard millet, proso millet, little millet, kodo millet, foxtail millet) were gel eluted and sequenced using domain-specific primers. The PCR amplified putative Dof genes from finger millet, rice, wheat, maize and barley were cloned in pGEM-T Easy vector
while sorghum *Dof* genes (*SbDof₁, SbDof₁₉, SbDof₂₃* and *SbDof₂₄*) were cloned in pBSK vector and subsequently sequenced using M13 universal primers. A total of 32 sequences representing *Dof* genes/domains of different cereals and millets were submitted to GenBank and assigned accession numbers EU586262 to EU586269, EU760631 to EU760640, FJ854501 to FJ854504, GQ924919 to GQ924921, GQ352370 to GQ352372 and HQ540084 to HQ540087.

These sequences representing *Dof* genes and *Dof* domains of cereals and millets were further subjected to *in silico* characterization for homology search, multiple sequence alignment, phylogenetic tree construction and motif analysis. Multiple sequence alignment of sequenced *Dof* domain and genes of cereals and millets revealed four conserved cysteine residues associated with the typical zinc-finger like structure. In few of the sequences slight alterations were also observed based on the partial sequence of the genes/domains resulting from the sequencing by gene and domain-specific primers. The phylogenetic tree constructed based on translated protein sequences of cloned *Dof* genes resulted in to distinct clusters for prolamine-box binding factor (PBF) *Dof* genes of monocots and dicots. The cloned *Dof* genes of *Eleusine coracana, Hordeum vulgare, Triticum aestivum* and *Zea mays* revealed their identity to PBF *Dof* based on the presence of motifs related to regulation of endosperm-specific seed storage protein encoding genes. The PBF *Dof* gene of finger millet is reported here for the first time based on *in silico* studies though it has already been reported in rice, wheat, barley and maize.

A total of 28 copies of *Dof* genes were *in silico* predicted from sorghum genome based on comparative genomic approach. The *in silico* predicted *Dof* genes were assigned third party accession number (TPA: BK006983-BK007006 and TPA: BK007079-BK007082) and were extensively characterized for chromosomal location, gene structure determination, phylogenetic tree construction and *cis*-regulatory element analysis using different bioinformatics tools. The predicted 28 *SbDof* genes were distributed on nine out of ten chromosomes of sorghum and most of these genes lacked introns. Chromosomes 1 and 3 exhibited a maximum of 7 and 6 *Dof* genes respectively, while chromosome 10 did not contain any *Dof* gene. The comparative phylogenetic analysis of *SbDof* proteins along with 30 rice and 36 Arabidopsis *Dof* proteins revealed six major groups (A-F). Motif analysis revealed the presence of conserved 50-52 amino acids *Dof* domain uniformly distributed across all *SbDof* proteins. The *cis*-regulatory element analysis of these predicted *Dof* genes suggested their diverse putative functions such as regulation of genes associated with seed storage proteins, photoperiodic control, growth hormones, abiotic and biotic stress. The *SbDof* members representing major group-A showed similarity with Arabidopsis...
cycling Dof factor (CDF) proteins associated with regulation of photoperiodic control of flowering. The \textit{SbDof19} and \textit{SbDof24} showed similarity with rice PBF \textit{Dof} in group-D.

The structural analysis of protein sequences of four cloned \textit{Dof} genes viz. HQ540084 (\textit{SbDof1}), HQ540085 (\textit{SbDof19}), HQ540086 (\textit{SbDof23}) and HQ540087 (\textit{SbDof24}) of sorghum was attempted. Secondary structure prediction of these proteins showed the presence of conserved Dof domain region having turns as predominant features along with sheets, coil and helix region. Three-dimensional structures for these four \textit{SbDof} proteins were \textit{in silico} predicted by multiple threading and iterative structural assembly simulations from online available I-TASSER server. The final models were submitted to PMDB database and assigned PMDB IDs i.e. PM0077395, PM0077396, PM0077397, PM0077398 and PM0076448 for SbDof1, SbDof19, SbDof23, SbDof24 and Dof domain, respectively. Superposition of Dof domain with predicted models of four \textit{SbDof} proteins revealed domain region to be structurally highly conserved. Active site identification and metal detection analysis suggested the involvement of cysteine residues in the coordination of metal.

The PCR amplification, cloning, sequencing of \textit{Dof} genes and Dof domain of cereals and millets followed by extensive \textit{in silico} characterization needs to be validated by wet lab experiments by performing expression profiling of the respective genes. The genome wide identification and \textit{in silico} characterization of \textit{Dof} gene family provides an opportunity to clone all the genes and analyze the functional diversity in sorghum crop. The utilization of transcription factor like Dof for crop improvement has been reported in many crops and can also be extended to sorghum based on the expression profiling of predicted \textit{Dof} genes.