NAME OF THE SCHOLAR: SHWETA
NAME OF THE SUPERVISOR: PROF. JAWAID AHMAD KHAN
DEPARTMENT: DEPARTMENT OF BIOSCIENCES
TITLE OF THE THESIS: *In silico* search of miRNA and siRNA targets against begomoviruses in cotton (*Gossypium hirsutum*) host plant

ABSTRACT

Cotton leaf curl virus (CLCuV) is a monopartite begomovirus, belongs to the family *Geminiviridae* and transmitted by whitefly (*Bemisia tabaci*). CLCuV genome encodes six genes in virion and complementary-sense separated by a non-coding intergenic region (IR), namely *AC1* (Replicase, Rep), *AC2* (Transcriptional Activator Protein, TrAP), *AC3* (Replication Enhancer protein, REn), pathogenicity enhancer protein- *AC4*, *AV1* (Coat Protein, CP) and *AV2* (pre-CP) gene. These genes are vital for viral replication and disease development in the host plant. IR contains the invariant TAATAT/AC motif responsible for the initiation of rolling circle DNA replication. CLCuV genomes are associated with betasatellites and alphasatellites. Both satellite DNA genomes are approximately half the size of CLCuV DNA genomes. Betasatellites encodes β*C1* genes, essential for induction of typical disease symptoms. Alphasatellites encodes α*Rep* gene and have no known contributions to CLCuV driven pathogenesis. In India, CLCuD was appeared in an epidemic form during the year 1997 in the state of Rajasthan affecting an area of 0.1 million hactares. This disease caused significant losses to cotton crop estimating total value of > $ 5 billion from 1992 to 1997 in Pakistan.

In plants, microRNAs (miRNAs) and small interfering RNAs (siRNAs) are the major contributors of the small RNA species. They play critical roles in multiple cellular processes through post-transcriptional regulation of RNA targets. They are known to provide resistance against plant-infecting viruses, via binding to their cognate complementary viral nucleotide sequences. Small RNA - mRNA pairs undergo RNA interference (RNAi) by target cleavage, translational inhibition or both. Here a computational approach was applied to identify the targets of miRNAs and siRNAs against CLCuV species such as *Cotton leaf curl Alabad virus* (CLCuAV), *Cotton leaf curl Allahabad virus* (CLCuAlV), *Cotton leaf curl Burewala virus* (CLCuBV), *Cotton leaf curl Bangalore virus* (CLCuBaV), *Cotton leaf curl Gezira virus* (CLCuGV), *Cotton leaf curl Kokhran virus* (CLCuKV), *Cotton leaf curl Multan virus* (CLCuMV) and *Cotton leaf curl Rajasthan virus* (CLCuRV)). On the basis of

complementarity between small RNAs (miRNAs and siRNAs) and viral mRNA targets, a large numbers of putative viral targets were identified.

In Chapter 3, I have predicted cotton (*Gossypium hirsutum*) encoded seventy nine miRNA targets in the genome of CLCuV. Forty one most potential miRNA families targeting their complementary sequences, with perfect or nearly perfect complementarity at multiple loci of CLCuV genome were identified. Artificial synthesis of the identified miRNAs and their expression in cotton could be an effective strategy to combat CLCuD infection.

In Chapter 4, conserved miRNAs targeting *AC1*, *AC2*, *AC3*, *AC4*, *AV1*, *AV2*, βC1 and αRep genes of CLCuV species were predicted. Members of twenty miRNA families were found to possess conserved targets among eight CLCuV species. These conserved miRNAs could serve as potential substrate for generating therapeutic intervention to generate CLCuD resistant plants.

In Chapter 5, CLCuV encoded miRNAs and their targets against CLCuV species were identified. Six miRNA families were identified to be encoded by CLCuV species. For *AC1*, *AC2*, *AC3*, *AC4*, *AV1*, *AV2*, βC1 and *αRep* genes 24, 13, 10, 26, 35, 36, 9 and 6, numbers of miRNA targets were identified in the genome of CLCuV species. Identification of these CLCuV encoded unique miRNA will improve the understanding of host- pathogen interaction pathways and the function of viral miRNAs in cotton plants.

In Chapter 6, cotton (*G. hirsutum*) encoded five hundred sixty nine siRNAs targets in the promoter of CLCuV were predicted. In CLCuV_AC1 promoter region and CLCuV_AV1 promoter region, a total number of 229 and 174, respectively siRNA targets were found. These siRNAs can be used to mimic ligands of CLCuV encoded proteins that will bind to active sites of proteins and leads to their functional inactivation.

In Chapter 7, a user friendly web interface was designed to demonstrate expression level of cotton encoded miRNAs at various developmental stages. The web interface was tentatively named as "Cotton miRNA Expression database" (COMED). This user friendly graphical interface will provide glimpse of miRNA expression variation at various flowering and ovule stages of cotton.

The results presented in this thesis are novel in its own kind by reporting small RNA targets against CLCuV encoded transcripts for the first time. This study would serve as milestone to understand complex interaction between plant and viruses and unfolds some of the leading facts about this complicate intricate relationship. It will complement the current knowledge of the function and evolution of cotton small RNAs.