1.0 INTRODUCTION

1.1. The genus *Gossypium*

*Gossypium* is the genus of cotton which belongs to the mallow family, Malvaceae. Cotton is native to tropical and subtropical areas of both old and new world. The name of the genus *Gossypium* is derived from the Arabic word *goz*, which means a soft substance. The genus *Gossypium* contains the largest number of species in the tribe *Gossypioideae* and comprises about 50 species divided into eight genome types, with distinct species and evolutionary histories in Asia, Africa, Arabia, the Americas, and Australia. The genus *Gossypium* was evolved ~ 5-10 millions year ago. The tetraploid cotton (AADD genome) was evolved ~ 1.5-2 million years ago by hybridization event in between its progenitor diploid cotton (A genome and D genome). The five tetraploid cotton species are *G. hirsutum*, *G. tomentosum*, *G. mustelinum*, *G. barbadense*, and *G. darwinii*. *Gossypium* is a shrub and the species of this genus are very much diverse in morphology and adaptation, ranging from perennials in Australia to trees in Mexico. The plant height in present cropping system ranges from 1-2 meter. The leaves are broader and contain three to five lobes. The fruit of the cotton is known as ‘boll’. The length of cotton fiber varies from 2.2 to 3.0 cm with 11 to 22 mm diameter but in *G. barbadense* it reaches to length of 6.0 cm. The cotton boll (ovary) contains about 25–30 seeds (ovules). Each seed bears 13,000-21,000 single celled fiber (Kim *et al.*, 2001). Two types of fiber are found, namely lint and fuzz. The longer fiber that undergoes spinning into yarn is known as lint and the shorter one is fuzz. The fibers are found naturally in colours of white, green and brown. *Gossypium hirsutum* (>90% of world production), *G. barbadense* (3-4%), *G. arboreum* and *G. herbaceum* (2%) are the most widely cultivated commercial species of cotton. A number of cotton varieties have been developed by cross breeding of these species for different desirable traits such as drought tolerant, insect resistance and fiber quality. Cotton fiber is economically important.

1.2. Economic potential of *Gossypium* spp. fiber

Cotton (*Gossypium* spp.) is one of the most important premium sources of textile fiber, edible oil in the world and has been used for clothing, fine papers and
other needs. More than 80 countries grow the economically important cotton including Australia, China, Africa, India, Pakistan, the USA and Uzbekistan. China occupies top position with regards to cotton production. India covers top position with the area under cotton production but it ranks second after China in production. Cotton fibers can be used for producing innumerable commodities, ranging from textile fabrics and computer screens to automobile brakes. Import and export of the cotton are carried out by more than 150 countries. Economic impact is estimated to be approx. $500 billion year worldwide (National Cotton Council 2006, http://www.cotton.org/). Cotton is a cash crop and lint fiber is used in textile industries to make clothes. The oil from seed is used for the manufacture of vanaspathi. Cotton flour is used to make bread and biscuit. Cotton seed cakes are used as organic manure. Cotton is used in stuffing pillows, cushions, mattresses etc. It is used in generating rubber tyres, carpet, blankets and cordages. Fatty acids obtained from the oil are used in the manufacture of insecticide, fungicides and plastics. Cellulose industries use cotton as a raw material. Cotton plays a vital role in the country’s economic growth.

The second largest contributor of employment after agriculture is the textile sector.

The cotton cultivation sector not only engages around 6 million farmers, but also involved another about 40 to 50 million people relating to cotton cultivation, cotton trade and its processing. Cotton fiber undergoes four overlapping developmental stages namely fiber initiation, elongation, secondary cell wall biosynthesis and maturation. The desirable characteristics of cotton fiber are length, strength and fineness. Fiber length is the elongation of an epidermal cell of the ovule. Fiber strength represents the crystalline cellular component of the secondary wall. Cellulose deposition in fiber produces microfibrils in a spiral arrangement and their binding gives strength.

1.3. Distribution of Gossypium spp.

Gossypium spp. is distributed worldwide. Distribution of 50 species of cotton divided into eight groups (A, B, C, D, E, F, G and K) is well documented by Fryxell (1979, 1992). The genome group A, which contains two species, is geographically distributed in Africa and possibly Asia. The genome group B, which contains three
species, is geographically distributed in Africa and Cape Verde Islands. The genome group C, which contains two species, is geographically distributed in Australia. The genome group D, which contains thirteen species, is geographically distributed in primarily Mexico also Peru, Galapagos Islands and Arizona. The genome group E, which contains more than seven species, is geographically distributed in Arabian Peninsula, Northeast Africa and Southwest Asia. The genome group F, which contains only one species, is geographically distributed in East Africa. The genome group G, which contains three species, is geographically distributed in Australia. The genome group K, which contains twelve species, is geographically distributed in NW Australia. The AD genome that contains five numbers of species is geographically distributed in new world tropics and subtropics including Hawaii. In India, the planting season of cotton normally occurs from March to September and the harvesting season is from October to February. India contains three cotton producing zones such as northern zone, central zone and southern zone. The northern zone, which is also known as *G. hirsutum* and *G. arboreum* zone, comprises Punjab, Haryana and Rajasthan. The central zone, which is also known as *G. hirsutum*, *G. arboreum*, *G. herbaceum* and hybrid zones, comprises Maharashtra, Madhya Pradesh and Gujarat. The southern zone, which is also known as *G. hirsutum*, *G. arboreum*, *G. herbaceum*, *G. barbadense* and hybrid zones, comprises Andhra Pradesh, Karnataka and Tamil Nadu. Gujarat, Maharashtra and Andhra Pradesh are the leading state in cotton production in India. They produce about 70% cotton of total produced in India. The short and medium cotton is produced by the northern zone, long fiber is normally produced by southern zone and long and medium fiber is produced by central zone.

1.4. Molecular advancement in improving *Gossypium spp.* fiber

There is a great need for improvement of the cotton fiber using molecular techniques. The molecular basis of cotton fiber development is still unanswered. Recent techniques used in understanding the cotton fiber development are gene cloning, expressed sequence tags (EST) characterization, microarrays, phytohormonal regulation and molecular markers. The followings are some reports which have been studied in cotton fiber development.

Earlier studies have reported a set of genes that may control the fiber initiation, elongation and secondary cell wall biosynthesis (SCW). The *GhMyb25 like*
transcription factor was reported to play key role in fiber initiation (Walford et al. 2011). Similarly, the GbPDF1 was reported as a main regulator of cotton fiber initiation (Deng et al. 2011). Some other known fiber initiation specific genes are GaMYB2 (Wang et al., 2004), GaRDL1 (Wang et al. 2004), GhHD1 (Wu et al. 2007). Many genes also have been reported to play role in both fiber initiation and elongation such as GhMYB109 (Pu et al. 2008), GhFLA1 (Li et al. 2010), GhGA20ox1-3 (Xiao et al. 2010). Likewise, genes related to fiber elongation such as GhKCS13/CER6 (Qin et al. 2007), GhAPX1 (Li et al., 2007; Qin et al., 2008), GhACT1 (Li et al., 2005), GhTUB1 (Li et al., 2002) and GhPFN1 (Wang et al. 2005) have been identified.

GhGlcAT1 (Wu et al., 2005, 2007) and GhRLK1 (Li et al., 2005) genes are identified for elongation and secondary cell wall biosynthesis stage.

Apart from these several plant hormones also have been validated in cotton fiber development. Auxin is known to play important role in fiber initiation (Zhang et al., 2011). Brassinosteroid play role in fiber elongation (Sun et al., 2005). Similarly, Ethylene also regulates fiber elongation (Shi et al., 2006). Abscisic acid negatively regulates fiber growth (Beasley and Ting, 1974). Cytokinin accelerates fiber initiation before anthesis and inhibits fiber growth after this (Chen et al., 1997).

Cotton fiber is an excellent model to study the cell elongation and cellulose biosynthesis. It involves four overlapping developmental stages namely fiber initiation, elongation, SCW and maturation. Cotton fiber initiation is key point which determines the number of fiber initials on each seed coat. Fiber elongation is the duration which determines the final length. Fiber strength and fineness is controlled by SCW. Therefore, genomic approaches are mostly focused on identification and functional validation of potential candidate genes for improving cotton fiber yield and quality.

In our previous study, microarray was performed in between five contrasting genotypes of G. hirsutum (JKC725, JKC777, JKC703, JKC737 and JKC783) grouped into two group namely superior (JKC725 and JKC777) and inferior (JKC703, JKC737 and JKC783) genotypes (Nigam et al., 2013). We identified several potential genes including transcription factors that may play important role in fiber development. One such transcription factor SQUAMOSA promoter binding like transcription factor (SPL) was identified expressing highly in superior genotype in compare to inferior
genotype. The SPL encode plant specific transcription factor and share a highly conserved DNA binding (SBP) domain (Yamasaki et al., 2004). They promote both flowering and vegetative growth by binding to the GTAC consensus nucleotide sequence of its target gene promoter (Wu and Poethig, 2006). In *Arabidopsis* this gene is negatively regulated by miR156 (Wu et al., 2009).

The present study aim at functionally validating SPL gene designated as *GhSPL5D* in cotton fibre development. The interaction of this gene with miRNA156 was also validated in cotton fiber. Further, overexpression and suppression (RNAi) strategies were used to study the function of *GhSPL5D* and miR156 in cotton fiber development. The present study also aims to improve the cotton fiber quality and yield by high level expression of candidate gene, *GhSPL5D*, in cotton fiber. This thesis demonstrates that high level expression of *GhSPL5D* in cotton fiber improves the fiber initiation and boll biomass.

### 1.5 Importance of the study

The importance of the study here is to explore the role of SPL transcription factor in cotton fiber development. The improved cotton fiber quality namely fiber length, strength and fineness could be achieved by expressional alteration of *SPL* gene in transgenic lines. Increase in the yield of cotton may be achieved by overexpression and knockdown of *SPL* gene. Interaction analysis of miR156, miR172 with *SPL* transcription factor in cotton fiber development could be validated.

### 1.6 objectives

There are four major objectives of my present study:-

1. Identification of candidate gene *SQUAMOSA* promoter binding like transcription factor (SPL) based on expression profiling in different developmental stages of cotton fibre.
2. Functional characterization of protein by making overexpression and knockdown expression (RNAi) lines of cotton.
3. Interaction analysis of SPL with miR156 and miR172.
4. Molecular characterization of SPL gene and elucidation of its role in cotton fibre development.