

## The genes determining synthesis of pigments in cotton

Mikhailova A.S.<sup>1,2\*</sup>, Strygina K.V.<sup>2,3</sup>, Khlestkina E.K.<sup>2,3</sup>

<sup>1</sup> St. Petersburg State University, St. Petersburg, Russia

<sup>2</sup> N.I. Vavilov All-Russian Institute of Plant Genetic Resources, St. Petersburg, Russia

<sup>3</sup> Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

\* e-mail: a.mikhailova@vir.nw.ru

**Key words:** anthocyanins, bHLH, flavonoids, MBW complex, Myb, proanthocyanidins

**Motivation and Aim:** Flavonoid compounds proanthocyanidins (PAs) and anthocyanins in cotton fibres and stems provide the brown and red colour, respectively [1]. Naturally coloured cotton fibres are hypoallergenic and do not lead to environmental pollution in the textile production. Stems contained flavonoid compounds are insensitive to various environmental stress factors as ultraviolet radiation, anomaly temperature changes and fungal infections [2]. The anthocyanins and PAs biosynthesis is under control of R2R3-Myb, bHLH-Myc and WD40 transcription factors (TF) forming MBW regulatory complex [3]. The aim of the present study was the identification and characterization of homeologous and paralogous R2R3-Myb and bHLH-Myc genes copies in genomes of diploid and allotetraploid cotton species (*G. arboretum* A<sub>2</sub>, *G. raimondii* D<sub>5</sub>, *G. hirsutum* AD<sub>1</sub> and *G. barbadense* AD<sub>2</sub>).

**Methods and Algorithms:** The search of homologous sequences was made in databases for not annotated cotton sequences using BLAST. The cluster analysis using the MEGA software was based on NJ algorithm. The number of non-synonymous substitutions per non-synonymous sites (*Ka*), the number of synonymous substitutions per synonymous sites (*Ks*) and the *Ka/Ks* ratio for *Myb* and *bHLH* genes were calculated. The obtained value was used for the calculation of divergence time of duplicated copies.

**Results:** The information about known *R2R3-Myb* (orthologs of *AtTT2*) and *bHLH-Myc* (orthologs of *AtTT8*) genes was used to identify duplicated genes by BLAST search in genome of cotton. Due to this analysis, we identified three groups of *Myb* and one group of *Myc* genes. The results of investigation of phylogenetic relationship and the analysis of structural organization of duplicated regulatory genes are represented in current report. The calculation of the divergence time revealed that the duplication of *Myb* genes occurs on early stages of cotton genome development. Genotype-specific features are revealed for some duplicated genes.

**Conclusion:** The duplications of the *R2R3-Myb* genes occurred several times in the course of cotton evolution. All identified sequences have conserved R2R3 and bHLH domains. These findings can be used further for the development of diagnostic PCR- and CAPS-markers in marker-assisted selection of cottons with naturally coloured fibres.

### References

1. Feng H., Li Y., Wang S., Zhang L., Liu Y., Xue F. Molecular analysis of proanthocyanidins related to pigmentation in brown cotton fibre (*Gossypium hirsutum* L.). *J. Experimental Botany*. 2014;65(20):5759-5769.
2. Khlestkina E.K. Genes determining the coloration of different organs in wheat. *Russian J. Genetics: Applied Research*. 2013;3(1):54-65.
3. Gong W., Du X., Jia Y., Pan Z. Color cotton and its utilization in China. In cotton fiber: physics, chemistry and biology. – Cham: Springer International Publishing. 2018;117-132.