

***In silico* characterization of Guar genes homologous to soybean E1–E4 maturity genes**

Criollo L.*, Zamalutdinov A., Potokina E.

The Project Center for Agro Technologies, Skolkovo Institute of Science and Technology, Moscow, Russia

**Luisa.Criollo@skoltech.ru*

Key words: guar; photoperiod sensitivity; flowering time; homologous soybean genes

Motivation and Aim: Guar (*Cyamopsis tetragonoloba* (L.) Taub.) is an economically important short-day legume, whose seeds are the main source of Guar gum, a compound used as a thickening agent in the food industry and as a fracturing fluid during oil and gas extraction [1]. In Russia the main obstacle to its cultivation is the long photoperiod during summer which delays flowering in guar [2]. In other legumes such as Soybean, the flowering time is mostly influenced by four maturity genes, denominated as E1, E2, E3 and E4 [3]. However, in Guar these genes have not been characterized, which creates an issue for the implementation of genome editing technologies in these flowering-related genes. In this study, we used a bioinformatics approach to characterize the Guar genes homologous to soybean E1–E4 maturity genes.

Methods and Algorithms: The nucleotide sequences of the soybean E1–E4 maturity genes were retrieved from NCBI. Subsequently, BLAST was used to perform a sequence similarity search in the guar genome and identify homologous sequences. The genomic sequences with high percentage of homology were retrieved and gene prediction tools such as AUGUSTUS or FGENESH were used to predict the exon-intron structure of the candidate genes. MEGA 11 was used to carry out sequence alignment and identify conserved regions. Functional annotation was performed using InterProScan and I-TASSER server. The candidate genes were amplified through PCR in Guar accessions with different flowering time and sequenced using Sanger method. Additionally, transcriptome data was used to confirm the expression of the E1–E4 homologous genes in Guar.

Results: BLAST search against the Guar genome allowed the identification of homologous sequences to the soybean E1–E4 genes. Sequence alignment between Guar homologs and soybean genes revealed high sequence similarity and several conserved regions, indicative of the presence of conserved structural features. The gene structure prediction showed similar exon-intron structure between Guar candidate genes and their respective soybean counterparts. Functional annotation showed putative protein domains in the Guar homologs, providing a better understanding of their potential biological function. Sequence analysis of the Guar homologs in early and late flowering accessions provided a first insight into the presence of potential SNP associated with flowering time regulation. Finally, analysis of the transcriptome data confirmed the expression of the E1–E4 homologous genes in Guar.

Conclusion: The Guar genes homologous to soybean E1–E4 maturity genes were identified and characterized using several bioinformatic tools. Analysis of the sequence of Guar homologs in early and late flowering accessions provided a first insight into the presence of potential SNP associated with flowering time regulation. The *in silico* characterization of these Guar genes lays the foundation for further functional studies

and genome editing experiments targeting these genes to analyze the genetic network that underlies the flowering time regulation in Guar.

Funding: The study is supported by Russian Science Foundation grant No. 24-26-00073.

References

1. Hasan A.M., Abdel-Raouf M.E. Applications of guar gum and its derivatives in petroleum industry: A review. *Egypt J Pet.* 2018;27(4):1043-1050. doi 10.1016/j.ejpe.2018.03.005
2. Grigoreva E., Tkachenko A., Arkhimandritova S. et al. Identification of Key Metabolic Pathways and Biomarkers Underlying Flowering Time of Guar (*Cyamopsis tetragonoloba* (L.) Taub.) via Integrated Transcriptome-Metabolome Analysis. *Genes.* 2021;12(7):952. doi 10.3390/genes12070952
3. Liu L., Song W., Wang L. et al. Allele combinations of maturity genes E1-E4 affect adaptation of soybean to diverse geographic regions and farming systems in China. *PLoS One.* 2020;15(7):e0235397. doi 10.1371/journal.pone.0235397