

GWAS using the IIIVmrMLM model provides new genomic associations with agronomic traits in Chickpea

Duk M.^{1, 2*}, Gurkina M.V.³, Nekrasov A.Y.⁴

¹ *Peter the Great St.Petersburg Polytechnic University, St.Petersburg, Russia*

² *Ioffe Institute, RAS, St. Petersburg, Russia*

³ *N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Astrakhan Experiment Station of VIR, Astrakhan, Russia*

⁴ *N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Kuban Experiment Station of VIR, Krasnodar, Russia*

* duk@mail.ioffe.ru

Key words: chickpea; *Cicer arietinum*; GWAS; IIIVmrMLM; favorable alleles

Motivation and Aim: Chickpea is one of the most important crops and is cultivated both for food (kabuli type) and animal feed (desi type) [1]. Although it is grown in many countries and used in many traditional Asian dishes, the genetic diversity of this crop has been relatively reduced due to domestication. Due to this, chickpea plants are susceptible to various abiotic stresses [2] and diseases. In this regard, the search for new variants associated with different agronomic traits is important for breeding programs. In addition, agronomic traits in chickpea vary greatly under different environmental conditions, so traditional genome-wide association searching methods yield different variants in different environments that may not be meaningful in other environments. The new IIIVmrMLM genomic variant search method [3] allows it to be used in multiple environments as well as in each environment separately. In addition, GWAS methods based on the MLM algorithm have a Bonferroni correction for multiple testing that is too stringent to find associations with complex traits in crop [4]. The IIIVmrMLM method solves this problem by estimating effects using a single multilocus model and then by estimating non-zero effects using the likelihood ratio test.

Methods and Algorithms: The collection of the All-Russian Institute of Plant Genetic Resources (VIR), which we used in our study, consisted of 171 chickpea accessions, of which 147 accessions were landraces. We considered 12 phenotypic traits measured in two regions of Russia, Astrakhan (46°06'N, 48°04'E, altitude 24 m) and Kuban (45°18'N, 40°52'E, altitude 138.9 m) in 2022. We used the IIIVmrMLM package [5] in Multi_env mode to search in two environments and in Single_env mode to search for variants in each environment separately. Phenotypic data were quantile normalized for each trait. We considered phenotypic traits such as: days from emergence to start of flowering, flowering duration in days, days from emergence to full maturity, plant height, height to first pod, weight of all pods per plant, number of all pods per plant, weight of plant with pods (biological yield), weight of one hundred seeds, number of primary and secondary branches and leaf size.

Results: We found a total of 233 QTNs and 204 QEIs (QTN-environment interactions) for 12 traits using IIIVmrMLM in Multi_env mode and 455 associations in total for each of locations using a Single_env mode. Associations were located in all 8 chromosomes. QTNs found for each of environments explained 36.56 % to 55.5 % of the variation in Astrakhan and 36.32 % to 53.26 % of the variation in Kuban. In the Multi_env mode,

QTN explained about 27.5 % of the variation, whereas QEI explained on average 38.8 % of the variation, ranging from 24 % to 52.4 % for the different traits. The 29 QTNs found using Single_env model at each location were previously found by the FarmCPU, Blink, SUPER, MLM, and MLM methods using the GAPIT package [6, 7], whereas only 8 QTNs and 5 QEIs using Multi_env model were found previously by other methods. The 43 QTNs and 40 QEIs found with Multi_env model, as well as 63 QTNs predicted with Single_env model, were located within or around genes. There were only 9 overlaps between the lists of these genes.

In addition, we found favorable alleles for such traits as days from emergence to maturity, weight of the plant with pods, weight and number of pods per plant. The IIIVmrMLM method allowed us to estimate the additive effect of each variant in each location, which made it possible to find accessions with a good combination of favorable alleles for Astrakhan and Kuban. As a result, we found 5 landraces with good combinations of QTN and QEI for Kuban and two landraces with good combinations of QTN and QEI for Astrakhan.

Conclusion: When IIIVmrMLM was applied to data from single or multiple environments, the results obtained differed significantly. The Single_env model applied separately to data from each location had more overlap with GWAS results found with GAPIT. However, such models did not take into account G×E interaction and may result in associations which cannot be reproduced in another climatic conditions. The Multi-env model implemented in IIIVmrMLM estimates effects of genes and all types of interactions between alleles, genes and environments. The environment-specific alleles of the QEI loci are of particular interest, as they can be used by breeders to obtain varieties with improved adaptation to specific climatic conditions.

Funding: The study is supported No. 22-46-02004.

References

1. Varshney R.K. et.al. Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nat Genet.* 2019;51(5):857-864. doi 10.1038/s41588-019-0401-3
2. Jha U.C., Kole P.C., Singh N.P. QTL mapping for heat stress tolerance in chickpea (*Cicer arietinum* L.). *Legume Res.* 2019;44(4):382-387. doi 10.18805/LR-4121
3. Zhang Y.-W., Tamba C.L., Wen Y.-J. et al. MrMLM v4.0.2: An R Platform for Multi-Locus Genome-Wide Association Studies. *Genomics Proteomics Bioinformatics.* 2020;18(4):481-487. doi 10.1016/j.gpb.2020.06.006
4. Zhang Y.-M., Jia Z., Dunwell J.M. Editorial: The Applications of New Multi-Locus GWAS Methodologies in the Genetic Dissection of Complex Traits. *Front Plant Sci.* 2019;10:100. doi 10.3389/fpls.2019.00100
5. Li M., Zhang Y.-W., Xiang Y., Liu M.-H., Zhang Y.-M. IIIVmrMLM: The R and C++ Tools Associated with 3VmrMLM, a Comprehensive GWAS Method for Dissecting Quantitative Traits. *Mol Plant.* 2022;15:1251-1253. doi 10.1016/j.molp.2022.06.002
6. Wang J., Zhang Z. GAPIT version 3: Boosting power and accuracy for genomic association and prediction. *Genomics Proteomics Bioinformatics.* 2021;19(4):629-640. doi 10.1016/j.gpb.2021.08.005
7. Duk M.A., Kanapin A.A., Bankin M.P., Vishnyakova M.A., Bulyntsev S.V., Samsonova M.G. Genome-Wide Association Analysis in Chickpea Landraces and Cultivars. *Biophysics.* 2023;68(6):952-963