Factors affecting the accuracy of genomic predictions in testcrosses of maize biparental population

Faktori koji utječu na efikasnost genomskih predviđanja u testrižancima biparentalne populacije

Galić, V., Mazur, M., Brkić, A., Volenik, M., Jambrović, A., Zdunić, Z., Šimić, D.

FACTORS AFFECTING THE ACCURACY OF GENOMIC PREDICTIONS IN TESTCROSSES OF MAIZE BIPARENTAL POPULATION

Galić, V.[1], Mazur, M.[1], Brkić, A.[1], Volenik, M.[1], Jambrović, A.[1, 2], Zdunić, Z.[1, 2], Šimić, D.[1, 2]

SUMMARY

*Genomic prediction accuracy (r_{MP}) is affected by many factors, such as the trait heritability, training population size and structure, and the number of markers. This study's objective was to investigate the factors associated with r_{MP} for the ear height and the plant height in two planting densities in testcrosses of maize (*Zea mays* L.) IBM population. Genetic correlations between the training and validation populations were calculated. The high heritability estimates and correlations between the traits were observed. The non-zero estimates of r_{MP} for all trait-density combinations implied an efficiency of genomic selection. The lower than expected values of genetic correlations were observed between the training and validation populations. However, a strong correlation was observed between a genetic correlation of training and the validation population and r_{MP} in all three sizes of training populations assessed (20-40%, 40-60%, and 60-80%), suggesting that the size of the training population can be kept low by an appropriate selection while maintaining a high r_{MP} Further studies of relationships between the training and validation populations with larger effective population sizes are suggested, as reducing the size of training population while maintaining a high r_{MP} can facilitate a more effective allocation of resources in a maize breeding program.*

Keywords: genomic selection, genomic prediction accuracy, training population size, planting density, plant architecture

INTRODUCTION

Maize is one of the most important crops in the world used as food, feed and for industrial processing. Architecture of maize plants is very important, especially in high planting densities, as the intensity of light attenuates (Maddonni et al., 2006), although optimal planting density that maximizes grain yield varies with the environment (Assefa et al., 2017). Lobell et al. (2014) found that higher planting densities may lead to increase in maize grain yield, but also imply higher sensitivity to high vapor pressure deficit. High planting densities can also negatively affect photosynthesis and grain yield on single plant basis (Franic et al., 2015). Morphological consequences of increased density are the delay in differentiation of ears resulting in bareness, flowering asynchrony and kernel abortion.

Key parameters of plant architecture are plant and ear height. Plant height (PH) is a trait positively correlated with grain and forage yield and cultivars with short stalk usually yield less than the higher ones (Freeman et al., 2007). Ear height (EH) exhibits positive genetic and phenotypic correlations with PH, but this correlation is not always favorable. Higher positioned ears can change the stalk balance and increase the risk of stalk lodging and ear loss. Higher planting densities are known to increase ear and plant height posing risk for further losses in yield (Li et al., 2018). Stalk strength is

(1) Vlatko Galić, Ph. D., Maja Mazur, Ph. D. (maja.mazur@poljinos.hr), Andrija Brkić, Ph. D., Mirna Volenik, Antun Jambrović, Ph. D., Prof. Dr. Zvonimir Zdunić, Assoc. Prof. Domagoj Šimić - Agricultural Institute Osijek, Department of Maize Breeding and Genetics, Južno predgrađe 17, Osijek, Croatia, (2) Centre of Excellence for Biodiversity and Molecular Plant Breeding, Svetošimunska 25, Zagreb, Croatia

tightly correlated to lower ear placement, so breeding efforts for stronger stalk often result in lower placed ears (Flint-Garcia et al., 2003) confounding the effects of the two. Considerable portion of variation for breeding progress remains as EH and PH in maize are very complex, highly heritable traits controlled by many small effect loci (Peiffer et al., 2014).

Genomic predictions allow the breeders to fit the marker data in a model for the prediction of marker effects on quantitative traits. Since the marker data are collinear and the collinearity issue deepens as the number of markers increases, the penalized regression models were suggested (Meuwissen et al., 2001). The ridge regression best linear unbiased prediction (rrBLUP) model provides a precise and fast tool to predict the marker effects in breeding populations in a mixed-model equation style (Endelman, 2011). The accuracy of genomic predictions depends on many factors, some of the most important ones being the properties of the examined trait, the number of markers used for genotyping, heritability (Combs and Bernardo, 2013), and the training population size (Zhang et al., 2017). However, when building the crop prediction models, environmental variables also seem to leverage for many of the effects, and considerable efforts must be invested to improve the accuracy of genomic predictions (Rincent et al., 2017). Another significant factor affecting the prediction accuracy is the training population structure. If the population selected for calculation of the marker effects does not represent the population structure of the prediction population well, the accuracy of the predictions is expected to shrink (Isidro et al., 2015). The non-zero marker effects are calculated for the marker data in the rrBLUP model (Endelman, 2011), and a variation captured by marker effects is equivalent to that captured in the QTL analysis (Bernardo, 2010), the only difference being fitting of all the markers to the phenotype in the rrBLUP model, compared to fitting only the significant markers in the QTL models. This paper's objective was to analyze the factors determining the accuracy of the genomic predictions of the two high-heritability traits, the EH and PH, respectively, in two planting densities.

MATERIAL AND METHODS

Plant material and experimental design

The 191 intermated recombinant inbred lines (IRIL) of maize IBM*Syn4* (Intermated B73 x Mo17) population (Lee et al., 2002), along with both the IBM parental lines (B73 and Mo17), were testcrossed to a common tester inbred line. The inbred line 84-28A of Iodent genetic background, an Agricultural Institute Osijek's proprietary, was used to produce the testcrosses. The produced testcrosses, along with 11 checks (216 hybrids totally), were sown in a completely randomized block design with two replicates in two planting densities: a low and a high

one. According to Zhu et al. (2013), a planting density with ~56000 plants/ha was considered low, while the density of ~95000 plants/ha was considered high. The plot size amounted to 8.4 m^2, and the total of 864 plots were evaluated. The experiment was sown on April 23, 2014 on the Experimental Field of the Agricultural Institute Osijek in Osijek, Croatia (45°32'N, 18°44'E), with soybean as a preceding crop in a three-year crop rotation (soybean, maize, wheat/barley). The soil type was an eutric cambisol. Standard agricultural practices were applied.

Phenotyping and genotyping

The EH and PH phenotyping was performed 10 days subsequent to the pollination with a 4-m long ruler. Five adjacent plants were measured in each row, comprising 10 measurements per plot. The ear height was measured as a distance between a node bearing the topmost ear on the stalk and the soil surface, while the plant height was measured as a distance between the tassel tip and the soil surface. Mean of all measurements from a plot was considered a single observation. The IBM Neighbours 2 linkage map, anchored with 2178 genetically mapped markers, mostly the SNPs and SSRs ones, was used to conduct this study. The length of the genetic map amounted to 7090 cM, and the average distance between the markers amounted to 3.2 cM.

Statistical analysis

All statistical analyses were conducted in the R programming language (R core team, 2016). The variance components and their respective 95-percent confidence intervals were calculated with a linear mixed model solver, as implemented in the *lme4* package (Bates et al., 2014). In the mixed model, genotype, density, and genotype x density interaction were treated as the random effects, while a block and a replicate were considered as the fixed effects. The variance components' confidence intervals were calculated based on 1000 bootstraps. Heritability was estimated on an entry-mean basis (Hallauer et al., 2010), as follows: $H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_{GxD}^2 + \sigma_e^2}$, where σ_G^2 represents a genotypic variance, σ_{GxD}^2 a genotype x density interaction variance, and σ_e^2 a residual variance. A genetic correlation coefficient was approximated according to Ziyomo and Bernardo (2013), i.e., by calculating the Pearson's product-moment correlation of the rrBLUP genome-wide marker effects between the traits and between the densities for each trait.

Genome-wide marker effects

Genomic prediction models were set to test the predictive ability of the marker data for the plant height traits in testcrosses in different planting densities. The models were set using the ridge regression mixed model solver with the markers as the random effects implemented in the rrBLUP R library (Endelman, 2011). The markers with 50% or more of the missing data were

not used for predictions, while the markers with less than 50% of the missing data were imputed with a position mean value. The phenotypes used for each trait-density combination were the means of two replicates. A cross-validation with the random sampling of 60% of the individuals was run 500 times for each cycle. The 60% of the sampled individuals comprised the training population, in which a model for the prediction of the marker-effect best linear unbiased predictions (BLUPS) was calculated (trained), while the rest of the population was used in each cycle to validate the model. The correlation coefficients were calculated for each trait in each cycle between the predicted and the observed values (r_{MP}). To investigate the selection effects in training population on the genomic prediction accuracy, a sampling algorithm was conceived. This algorithm's objective was to test the effects of the training population size and the genetic correlation between the training and validation populations on the genomic prediction accuracy. The algorithm calculated the BLUPS of the marker effects in the training data set, than calculated the BLUPS in the validation data set, and eventually calculated a correlation coefficient between the obtained marker effects, approximating the genetic correlation between the sets (Ziyomo and Bernardo, 2013). The sampling started with 20% of the data comprising the training set (39 individuals), incrementing by five in each progressive step, up to the point in which ~80% of the data comprised the training set (154 individuals), and ~20% was a validation set. In each step, the genetic correlation coefficients between the two populations were calculated, as were the correlations between the genetic correlations between the sets and the accuracies of the genomic predictions. The samplings and the calculations in each step were repeated 500 times for each trait, finally comprising 48000 data points.

RESULTS AND DISCUSSION

The mean value for EH in a high planting density was higher than in a low planting density. The mean value for PH was also slightly higher in the high one when compared to the low planting density (Table 1). The testcrosses of B73 parental inbred line have demonstrated a higher EH and PH in both planting densities when compared to the testcrosses of Mo17 parental inbred line. The high proportions of genetic variances were observed for both traits, which explains the very high calculated heritabilities of 0.85 for the EH and 0.84 for the PH (Table 1). The obtained levels of heritability estimates are comparable to the heritabilities for EH and PH reported in other studies conducted using the testcrosses (Li et al., 2016; Li et al., 2017). The heritability estimates for the EH in the inbred lines are often very high, exceeding the 0.90 value (Peiffer et al., 2014; Li et al., 2017), but a genetic component of variance in testcrosses is expected to be lower (Smith, 1986). The extent of genotype x density was also reported to be relatively low for these traits (Li et al., 2016; Li et al., 2018). The 95-percent confidence intervals have demonstrated the relatively narrow estimates of genetic and error variance components, implying a low error margin of the estimates.

Table 1. The means and standard errors of mean for the ear and plant heights of the maize IBM population and its parental lines testcrosses, along with the variance components and heritabilities

Tablica 1. Srednje vrijednosti i standardne greške za visinu klipa i visinu biljke u IBM populaciji kukuruza i test-križancima roditeljskih linija, zajedno s komponentama varijance i heritabilnostima

Trait *Svojstvo*	Ear height (cm) *Visina klipa (cm)*		Plant height (cm) *Visina biljke (cm)*	
Density *Gustoća*	Low *Niska*	High *Visoka*	Low *Niska*	High *Visoka*
Mean ± SE	121.7 ± 0.54	128.4 ± 0.59	300.6 ± 0.63	301.1 ± 0.69
B73	132.5 ± 3.75	138.5 ± 0.64	309.5 ± 5.12	313.7 ± 2.92
Mo17	125.7 ± 1.79	125.2 ± 0.48	303.5 ± 1.89	308.5 ± 2.46
σ^2_G	76.87 (63.99–102.11)		107.74 (85.01–135.51)	
σ^2_D	25.57 (0.00–100.74)		0.01 (0.00–2.14)	
σ^2_{GXD}	0.01 (0.00–6.27)		0.00 (0.00–8.73)	
σ^2_e	56.79 (48.88–62.31)		80.11 (68.90–87.74)	
H²	0.85 (0.77–0.89)		0.84 (0.76-0.89)	

σ^2_G – genotypic variance/*genotipska varijanca*, σ^2_D – variance explained by the planting density treatment/*varijanca tretmana gustoće sjetve*, σ^2_{GXD} – variance explained by the genotype x planting density interaction/*varijanca interakcije genotip x gustoća sjetve*, σ^2_e – model residual variance/*varijanca ostatka*, *H²* – estimate of entry-mean heritability/*procjena heritabilnosti*

A moderate phenotypic correlation was observed between the EH and the PH (Table 2). A genetic correlation between the EH and the PH was strong, indicating a possibility for the presence of linkage or pleiotropy, and the high levels of phenotypic and genetic correlations for each trait between the densities were observed. The EH and PH traits are known to be highly correlated, both phenotypically and genetically (Peiffer et al., 2014; Liu et al., 2016; Li et al., 2017; Vanous et al., 2018). Much of this correlation structure can be attributed to a high proportion of the shared loci with the pleiotropic effects (Liu et al., 2016; Li et al., 2017; Vanous et al., 2018). The reason for this stabile reaction of testcrosses could be a density of 9.56 plants/m^2, which was considered high in the current study, but in the favorable growing conditions it can be considered an optimal planting density (Li et al., 2018).

Table 2. The coefficients of phenotypic (r_p) and genetic (r_g) correlations between the ear height (EH) and the plant height (PH), EH-PH, and between the two plant densities for each trait. All correlations are significant at the $p<0.01$ level

Tablica 2. Koeficijenti fenotipskih (r_p) i genetskih korelacija (r_g) između visine klipa (VK) i visine biljke (VB) VK-VB te između dviju gustoća za svako svojstvo. Sve korelacije su značajne na razini p<0,01

Correlation *Korelacija*	EH-PH *VK-VB*	EH *VK*	PH *VB*
r_p	0.597	0.743	0.713
r_g	0.729	0.746	0.717

The testing of the training population's optimal size and a genetic correlation between the training and the validation population was carried out in the current study. The significantly lowest r_{MP} value of 0.235 was observed for all algorithm iterations in all four trait-density combinations (Figure 1), when ~20-40% of the data were used as the training population, namely 39–79 out of 191 individuals. A significantly higher r_{MP} was observed (0.305) when ~40–60% of the data comprised the training population (79–114 individuals), while the highest mean r_{MP} of 0.393 was obtained when the ~60-80% of the data were used (114–154 individuals). All r_{MP} values were significantly different from zero at $P=<0.001$, suggesting the effectiveness of genomic selection. The non-zero efficiencies of rrBLUP genomic prediction model are expected for the complex traits with a high heritability, such as the plant and ear height (Combs and Bernardo, 2013; Lian et al., 2014). A highly strong correlation was observed between the genetic correlation between the training and the validation populations and the r_{MP} values. Generally low values of genetic correlations between the training and the validation populations were observed. The highest genetic correlation between the populations was obser-

ved when ~40-60% of the data comprised the training population (0.172), compared to the other two groups (0.155 and 0.152).
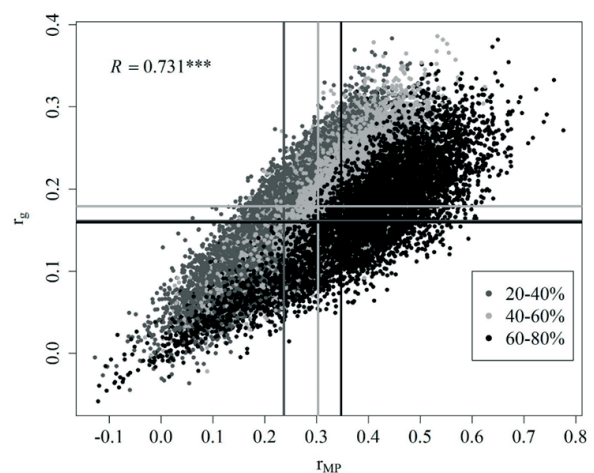

Figure 1. A linear model of genetic correlation (r_g) between the training and validation sets and the genomic selection accuracies r_{MG} across all four trait-density combinations. Dot colors represent the size of a training set. The vertical bars represent the grand mean of a genomic selection accuracy for the group, and the horizontal lines represent the grand mean of genetic correlations between the training and the validation populations for a group

Slika 1. Linearni model genetske korelacije (r_g) između radnoga i validacijskog seta i efikasnosti genomske selekcije r_{MG} kroz sve četiri kombinacije svojstvo – gustoća. Boje točaka predstavljaju veličinu radnoga seta. Okomite linije predstavljaju srednju vrijednost efikasnosti genomske selekcije za grupu, a vodoravne linije srednju vrijednost genetskih korelacija između radne i validacijske populacije za grupu

The wide ranges of r_{MP} values observed for all trait-density combinations indicate a potential for the optimization of training population in the genomic predictions of the EH and the PH. The size of the training population is known to influence the genomic prediction accuracy, and the training sets that capture more variation of the genetic structure in validation populations are expected to demonstrate a higher prediction accuracy (Lorenz and Smith, 2015). Biparental populations are widely used in breeding of cross-pollinated species, and genomic selection is extensively employed in the commercial breeding programs. The main goal of the genomic selection is to maximize a genetic gain per cycle of breeding and to reduce the costs of finding the best progenies and crosses, so the breeders could benefit from phenotyping of the smaller training populations while maintaining the accuracy of predictions.

The low observed values of genetic correlations between the training and the validation populations indicate a low efficacy of representing the loci affecting the

traits of interest in both populations. The better mutual genetic populational representation may lead to a considerable increase in the r_{MP}, but the larger effective population sizes are necessary to derive the more robust conclusions. The higher genetic correlations between the training and the validation populations apparently leverage the r_{MP} values in smaller than in the large training populations (Fig. 1) more strongly, implying that the appropriate selection of individuals in the training population can maintain a high r_{MP} value. Apparently, in the smallest and in the largest training populations assessed, the genetic correlations between the training and the validation populations was lower if compared to that when the size of the training population amounted to ~40–60% of the original dataset. A plausible reason for that is, to a certain point, that the genetic variation of the whole data set becomes better represented by a larger training population. Furthermore, the calculation of marker effects on small effective populations leads to the poor estimates of true marker effects (Zhang et al., 2017), leading to the underestimates of r_{MP} as well as to that of the r_g in the current study. A strong correlation between the r_g and the r_{MP} suggests that, by an appropriate representation of the target population by the training population, the r_{MP} can be considerably shifted towards the right side of the distribution assessed across the cycles of the cross-validation.

CONCLUSION

The nonzero mean r_{MP} values indicated the effectiveness of genomic selection for the high heritability traits in all three population size groups assessed. A strong correlation between a genetic correlation of the training and validation populations and the r_{MP} indicate that the training population can be fairly small, with an appropriate selection of individuals for phenotyping, while maintaining the high levels of r_{MP}. Further studies, with larger effective populations, are necessary to comprehensively inspect the genetic factors facilitating a high r_{MP} in the small training populations. The reduction of training population size while maintaining a high r_{MP} could facilitate a more effective allocation of resources in the breeding program.

ACKNOWLEDGEMENT

REFERENCES

1.   Assefa, Y., Prasad, P. V. V., Carter, P., Hinds, M., Bhalla, G., Schon, R., … Ciampitti, I. A. (2017). A new insight into corn yield: Trends from 1987 through 2015. *Crop Science 57*(5), 2799–2811. https://doi.org/10.2135/cropsci2017.01.0066

2.   Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software 67*(1). https://doi.org/10.18637/jss.v067.i01

3.   Bernardo, R. (2010). Breeding for Quantitative Traits in Plants (2nd ed.). Woodbury, MN: Stemma Press.

4.   Combs, E., & Bernardo, R. (2013). Accuracy of Genomewide Selection for Different Traits with Constant Population Size, Heritability, and Number of Markers. *The Plant Genome 6*(1), 0. https://doi.org/10.3835/plantgenome2012.11.0030

5.   Endelman, J. B. (2011). Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome Journal 4*(3), 250. https://doi.org/10.3835/plantgenome2011.08.0024

6.   Flint-Garcia, S. S. A., Jampatong, C., Darrah, L. L., & McMullen, M. D. (2003). Quantitative Trait Locus Analysis of Stalk Strength in Four Maize Populations. *Crop Science 43*(1), 23. https://doi.org/10.2135/cropsci2003.0023

7.   Franic, M., Mazur, M., Volenik, M., Brkic, J., Brkic, A., & Simic, D. (2015). Effect of Plant Density on Agronomic Traits and Photosynthetic Performance in the Maize IBM Population. *Poljoprivreda 21*(2), 36–40.

8.   Freeman, K. W., Girma, K., Arnall, D. B., Mullen, R. W., Martin, K. L., Teal, R. K., & Raun, W. R. (2007). By-plant prediction of corn forage biomass and nitrogen uptake at various growth stages using remote sensing and plant height. *Agronomy Journal 99*(2), 530–536. https://doi.org/10.2134/agronj2006.0135

9.   Hallauer, A. R., Carena, M. J., & Filho, J. B. M. (2010). Quantitative Genetics in Maize Breeding. https://doi.org/10.1007/978-1-4419-0766-0

10. Isidro, J., Jannink, J. L., Akdemir, D., Poland, J., Heslot, N., & Sorrells, M. E. (2015). Training set optimization under population structure in genomic selection. *TAG. Theoretical and Applied Genetics 128*(1), 145-158. https://doi.org/10.1007/s00122-014-2418-4

11. Lee, M., Sharopova, N., Beavis, W. D., Grant, D., Katt, M., Blair, D., & Hallauer, A. (2002). Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. *Plant Molecular Biology 48*(5-6), 453-461. https://doi.org/10.1023/A:1014893521186

12. Li, H., Yang, Q., Fan, N., Zhang, M., Zhai, H., Ni, Z., & Zhang, Y. (2017). Quantitative trait locus analysis of heterosis for plant height and ear height in an elite maize hybrid zhengdan 958 by design III. *BMC Genetics 18*(1), 1-10. https://doi.org/10.1186/s12863-017-0503-9

13. Li, J., Xie, R. Z., Wang, K. R., Hou, P., Ming, B., Zhang, G. Q., … Li, S. K. (2018). Response of canopy structure, light interception and grain yield to plant density in maize. *The Journal of Agricultural Science* (1997), 1-10. https://doi.org/10.1017/S0021859618000692

14. Li, X., Zhou, Z., Ding, J., Wu, Y., Zhou, B., Wang, R., … Wu, J. (2016). Combined Linkage and Association Mapping Reveals QTL and Candidate Genes for Plant and

Ear Height in Maize. *Frontiers in Plant Science 7*(June), 1-11.
https://doi.org/10.3389/fpls.2016.00833

15. Lian, L., Jacobson, A., Zhong, S., & Bernardo, R. (2014). Genomewide prediction accuracy within 969 maize biparental populations. *Crop Science 54*(4), 1514-1522.
https://doi.org/10.2135/cropsci2013.12.0856

16. Liu, Y., Hou, X., Xiao, Q., Yi, Q., Bian, S., Hu, Y., … Huang, Y. (2016). Genetic Analysis in Maize Foundation Parents with Mapping Population and Testcross Population: Ye478 Carried More Favorable Alleles and Using QTL Information Could Improve Foundation Parents. *Frontiers in Plant Science 7*(September), 1-13.
https://doi.org/10.3389/fpls.2016.01417

17. Lobell, D. B., Roberts, M. J., Schlenker, W., Braun, N., Little, B. B., Rejesus, R. M., & Hammer, G. L. (2014). Greater sensitivity to drought accompanies maize yield increase in the U.S. Midwest. *Science 344*(6183), 516-519.
https://doi.org/10.1126/science.1251423

18. Lorenz, A. J., & Smith, K. P. (2015). Adding genetically distant individuals to training populations reduces genomic prediction accuracy in Barley. *Crop Science 55*(6), 2657–2667.
https://doi.org/10.2135/cropsci2014.12.0827

19. Maddonni, G. A., Cirilo, A. G., & Otegui, M. E. (2006). Row width and maize grain yield. *Agronomy Journal 98*(6), 1532-1543.
https://doi.org/10.2134/agronj2006.0038

20. Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Meuwissen_etal2001Genetics1819.pdf.
https://doi.org/11290733

21. Peiffer, J. A., Romay, M. C., Gore, M. A., Flint-Garcia, S. A., Zhang, Z., Millard, M. J., … Buckler, E. S. (2014). The genetic architecture of maize height. *Genetics 196*(4), 1337-1356.
https://doi.org/10.1534/genetics.113.159152

22. R Core Team (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

23. Rincent, R., Kuhn, E., Monod, H., Oury, F. X., Rousset, M., Allard, V., & Le Gouis, J. (2017). Optimization of multi-environment trials for genomic selection based on crop models. *Theoretical and Applied Genetics 130*(8), 1735-1752. https://doi.org/10.1007/s00122-017-2922-4

24. Smith, O. S. (1986). Covariance between line per se and testcross performance. *Crop Science 26*(3), 540-543.
https://doi.org/10.2135/cropsci1986.0011183X002600030023x

25. Vanous, A., Gardner, C., Blanco, M., Martin-Schwarze, A., Lipka, A. E., Flint-Garcia, S., … Lübberstedt, T. (2018). Association Mapping of Flowering and Height Traits in Germplasm Enhancement of Maize Doubled Haploid (GEM-DH) Lines. *The Plant Genome 0*(0), 0.
https://doi.org/10.3835/plantgenome2017.09.0083

26. Zhang, A., Wang, H., Beyene, Y., Semagn, K., Liu, Y., Cao, S., … Zhang, X. (2017). Effect of Trait Heritability, Training Population Size and Marker Density on Genomic Prediction Accuracy Estimation in 22 bi-parental Tropical Maize Populations. *Frontiers in Plant Science 8*(November), 1-12.
https://doi.org/10.3389/fpls.2017.01916

27. Zhu, L. Y., Chen, J. T., Li, D., Zhang, J. H., Huang, Y. Q., Zhao, Y. F., … Liu, Z. Z. (2013). QTL Mapping for Stalk Related Traits in Maize (Zea mays L.) Under Different Densities. *Journal of Integrative Agriculture 12*(2), 218-228.
https://doi.org/10.1016/S2095-3119(13)60221-X

28. Ziyomo, C., & Bernardo, R. (2013). Drought tolerance in maize: Indirect selection through secondary traits versus genomewide selection. *Crop Science 53*(4), 1269–1275.
https://doi.org/10.2135/cropsci2012.11.0651

FAKTORI KOJI UTJEČU NA EFIKASNOST GENOMSKIH PREDVIĐANJA U TESTKRIŽANCIMA BIPARENTALNE POPULACIJE

SAŽETAK

Mnogi faktori, kao što su heritabilnost svojstva, veličina i struktura radne populacije i broj markera, utječu na efikasnost genomskih predviđanja (r_{MP}). Cilj ovoga rada bio je ispitati faktore povezane s r_{MP} za visinu klipa i visinu biljke u dvije gustoće sjetve kod test-križanaca IBM populacije kukuruza (Zea mays L.). Izračunane su genetske korelacije između radne i validacijske populacije. Dobivene su visoke procjene heritabilnosti i korelacije između svojstava. Procjene r_{MP} bile su različite od nule za sve kombinacije svojstvo -gustoća. što ukazuje na efikasnost genomske selekcije. Vrijednosti genetskih korelacija između radne i validacijske populacije bile su niže od očekivanih. Međutim, jaka korelacija dobivena je između genetske korelacije radne i validacijske populacije i r_{MP} za sve tri veličine radne populacije (20-40%, 40-60% i 60-80%). što ukazuje da se uz odgovarajuću selekciju veličina radne populacije može održati niskom zadržavajući visoku vrijednost r_{MP}-a. Potrebna su daljnja istraživanja odnosa između radne i validacijske populacije s većom efektivnom veličinom populacije jer smanjivanje veličine radne populacije, uz održavanje visoke vrijednosti r_{MP}-a može omogućiti učinkovitiju alokaciju resursa u oplemenjivačkome programu kukuruza.

Ključne riječi: genomska selekcija, efikasnost genomskih predviđanja, veličina radne populacije, gustoća sjetve, arhitektura biljke