Accuracy of genomic selection in a rice synthetic population developed for recurrent selection.

Cécile Grenier[1,2], Tuong-vi Cao[2], Yolima Ospina[1], Joe Tohme[1], Brigitte Courtois[2], Nourollah Ahmadi[2]

1: CIAT, A.A. 6713, Cali, Colombia
2: CIRAD, UMR AGAP, F-34398 Montpellier, France

Genomic selection is a promising breeding strategy for rapid improvement of complex traits. We investigated the accuracy of genomic estimate breeding values (GEBV) through cross validation in a synthetic population (SP) of broad genetic diversity developed for upland rice breeding for Latin America, through recurrent selection (RS). Three hundred and three S4 lines extracted from a SP that has undergone several RS cycles were phenotyped for flowering time (FL), plant height (PH) and grain weight per plant (GW), and genotyped with an average density of one marker per 22.9 kb, using genotyping by sequencing. Cross validation was performed hundred times, for four ratios of training and validation population sizes (TP/VP), using ridge regression best linear unbiased prediction (RR-BLUP), least absolute shrinkage and selection operator (LASSO) and Bayesian linear regression methods (BL). Accuracy was estimated as the correlation between GEBV and true breeding value.

Linkage disequilibrium was high (average $r^2>0.5$ at 20 kb; $r^2>0.25$ at 1.5 Mb distance). Accuracy of GEBV increased with the size of the training population, regardless of the model. For FL and TP/VP varying from 150/153 to 270/33, the GEBV accuracy ranged from r=0.20 (0.05) to r=0.22 (0.15) under RR-BLUP, from r=0.19 (0.09) to r=0.28 (0.17) under LASSO, and from r=0.28 (0.03) to r=0.32 (0.17) under BL, before marker selection based on alleles frequency. GEBV accuracy was more sensitive to markers' allele frequency than to the total number of markers.

The best selected S4 will be recombined and their progenies tested for GEBV accuracy across generations.