

Diversity panel provides accurate genomic estimate of breeding values for complex traits, within the progenies of biparental crosses involving members of the panel.

Ben Hassen M¹, Cao TV², Jaquin L², Colombi C³, Orasen G¹. *, Rakotomalala J⁴, Razafinimpiasa L⁴, Bertone C¹, Volante A⁵, Biselli C⁵, Vale G⁵, Ahmadi N^{2*}

1: Department of Agriculture, University of Milan, Via Giovanni Celoria, 16, 20133 Milano, Italy

2: Cirad, UMR AGAP, Avenue Agropolis, -34398 Montpellier Cedex 5, France

3: Fondazione Parco Tecnologico Padano, Via Einstein, Loc. Cascina Codazza, Lodi, 26900 Italy

4: FOFIFA, Antananarivo, Madagascar

5: CREA- Council for Agricultural Research and Economics, Rice research unit, S. S. 11 to Torino Km 2.5, Vercelli, 13100, Italy

* Corresponding autor.

Abstract

Recently, the conjunction of high-throughput marker technologies and new statistical methods has given birth to a new breeding strategy called genomic selection (GS). It refers to methods that use genome-wide dense marker genotyping for the prediction of genetic values with enough accuracy to allow selection based on that prediction alone. It consists of two steps: estimation of SNP effects within a training population (TP), and prediction of genomic estimated breeding values (GEBV) in the candidate population (CP) based on SNP genotypes.

We present here GS for water use efficiency in rice, in the framework of a pedigree breeding scheme. The TP was composed of 284 accessions belonging to the temperate *japonica* and tropical *japonica* rice groups. The CP was composed of 99 F5-F7 lines derived from 36 crosses involving 32 accessions of TP. The two populations were genotyped with an average marker density of 4.8 per kb, with minimum minor allele frequency of 2.5. Phenotypic traits considered included flowering time (FL), grain yield (GY) and nitrogen balance index (NI) under conventional irrigation (CI) and aerobic system (AS). Phenotypes were modeled using two statistical regression methods: genomic best linear unbiased prediction (GBLUP) and reproducing kernel Hilbert Space (RKHS). The models were tested with three incidence matrixes corresponding to densities of 4.8, 9.5 and 13.8 marker per kb, and to linkage disequilibrium (LD) thresholds of $r^2 \leq 1$, $r^2 < 0.98$ and $r^2 < 0.81$. For each combination of phenotypic trait, incidence matrix and prediction method, accuracy of GEBV was investigated with 100 replicates using 100 random drawings of 2/3 of accessions of TP to train the prediction model. The accuracies of GEBV prediction were reasonably high for GY (0.41; Sd=0.03) and for NI (0.36; Sd=0.04), low for FL (0.26; Sd = 0.07). They were not significantly different under CI and AS. Likewise, the accuracies of prediction were not significantly different between GBLUP and RKHS methods. The size of the incidence matrix affected negatively the accuracy of prediction for GY and NI, suggesting negative effect of long-distance LD. These are the very first results of inter-population prediction in rice breeding. GEBV predictions need to be optimized by exploring potential effect of population structure within both TP and CP, and the minimum number of progenies for individual crosses.