

A GENOMIC SELECTION APPROACH TO PEANUT BREEDING

Davis Gimode, Jason Wallace, Corley Holbrook, Thomas G. Isleib, Ye Chu, Peggy Ozias-Akins

Introduction

Due to the continual reduction in cost of next generation sequencing (NGS), it is easier to develop high quality molecular markers and to use them in genotyping experiments¹. At the same time analytical techniques and statistical methods have improved sufficiently enough to facilitate their use in breeding methods such as genomic selection (GS) for crop improvement². GS takes advantage of high numbers of molecular markers developed at a relatively low cost. A GS experiment depends on a population for which previous phenotypic data has been collected. The same population is genotyped, typically using single nucleotide polymorphism (SNP) markers. This is called a training population (TP). A statistical model is then developed to associate the phenotype data with the genotype data and to make predictions of future phenotype scores, these predictions are the genomic estimated breeding values (GEBVs). Based on the derived GEBVs, selections can be made on a related population known as a breeding population (BP) that has been genotyped but not phenotyped. This way, breeders can skip a cycle of selection, hence saving on time and resources that would have been used for phenotyping. GS is particularly promising for studying hard-to-phenotype complex traits such as aflatoxin resistance in peanut. GS has been successfully applied to breeding in rice³, wheat⁴⁻⁶, maize⁷⁻⁹ and soybean².

Results

We applied GS to 3 peanut recombinant inbred line (RIL) populations. The first population, C1801 (Florida 07 X GP NC WS 16), had 383 RILs; the second population, C1802 (Florida 07 X C76-16 C1798), had 127 RILs; and the third population, C1798 (Tifruner X C76-16), had 152 RILs. All of C1801 had previously been genotyped using 999 SNPs and 78 SSR markers while phenotypic data on 8 traits had been collected for half of the lines (192) over a period of 1-4 years. These traits included Late Leaf Spot resistance (LLS; 4 years), grade (3 years), yield (3 years), Early Leaf Spot resistance (ELS; 2 years), double seed (2 years), *Tomato spotted wilt virus* resistance (TSWV; 1 year), double single count (1 year) and seed weight (1 year). Using subsets of the TP (the half that was genotyped and phenotyped), a statistical model was developed. GEBVs obtained for the traits were correlated with the actual trait values to determine how accurate the predictions were. The correlations ranged from 0.21 for TSWV to 0.79 for LLS indicating the robustness of the models for GS.

For the C1802 and C1798 RILs, all 625 lines in both population were genotyped again using the Affymetrix v2 SNP chip and GEBVs were calculated for aflatoxin resistance. The prediction model was based on the Affymetrix v1 SNP data available for the population (889 and 679 SNPs respectively) and phenotype data collected over 2 and 3 years for each population respectively. Due to the complex nature of the trait, relatively low correlation values were obtained (0.3 and 0.2 for C1802 and C1798, respectively). However, the values were comparable to the heritability (measure of the amount of variation explainable by genetic factors) of the trait in both populations: 0.2 and 0.1 for C1802 and C1798, respectively. This indicated that the GS model was able to give reasonable prediction values for the trait. The derived GEBVs were used to select 29 lines that had not been phenotyped before for experimental validation. Due to damage inflicted on the rainout shelters following hurricane Irma, it was not possible to do the field experiment in 2017.

References

1. Zargar, S. M. *et al.* Recent advances in molecular marker techniques: Insight into QTL mapping, GWAS and genomic selection in plants. *J. Crop Sci. Biotechnol.* **18**, 293–308 (2015).
2. Zhang, J., Song, Q., Cregan, P. B. & Jiang, G.-L. Genome-wide association study, genomic prediction and marker-assisted selection for seed weight in soybean (*Glycine max*). *Theor. Appl. Genet.* **129**, 117–130 (2016).
3. Spindel, J. *et al.* Genomic Selection and Association Mapping in Rice (*Oryza sativa*): Effect of Trait Genetic Architecture, Training Population Composition, Marker Number and Statistical Model on Accuracy of Rice Genomic Selection in Elite, Tropical Rice Breeding Lines. *PLoS Genet.* **11**, e1004982 (2015).
4. Battenfield, S. D. *et al.* Genomic Selection for Processing and End-Use Quality Traits in the CIMMYT Spring Bread Wheat Breeding Program. *Plant Genome* **9**, 0 (2016).
5. Rutkoski, J. *et al.* Genetic Gain from Phenotypic and Genomic Selection for Quantitative Resistance to Stem Rust of Wheat. *Plant Genome* **8**, 0 (2015).
6. Poland, J. *et al.* Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome J.* **5**, 103 (2012).
7. Bernardo, R. & Yu, J. Prospects for Genomewide Selection for Quantitative Traits in Maize. *Crop Sci.* **47**, 1082 (2007).
8. Albrecht, T. *et al.* Genome-based prediction of testcross values in maize. *Theor. Appl. Genet.* **123**, 339–50 (2011).
9. Technow, F., Bürger, A. & Melchinger, A. E. Genomic prediction of northern corn leaf blight resistance in maize with combined or separated training sets for heterotic groups. *G3 (Bethesda)*. **3**, 197–203 (2013).