breedgenr: a non-parametric based breeding program simulator with applications on rice

Bartholomé Jérôme, Laval Jacquin, Tuong-Vi Cao, Nourollah Ahmadi

CIRAD, UMR AGAP, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Avenue Agropolis, 34398 Montpellier. Cedex 5, France.

Theme: Genetic improvement

Effective harnessing by crop breeding programs of genetic gains expected from the integration of genomic resources, high throughput phenotyping capabilities, and efficient statistical methods to model the complexity of empirical data requires adjustments of the breeding strategies. Tools that simulate breeding features are needed to explore the effects of those adjustments and their potential interactions. Simulations are useful because they allow rapid replicated testing of a wide range of hypotheses at low cost, for example, the initial feasibility of genomic selection or the impact of the reference population structure. Simulation of breeding strategies is well adapted to investigate long-term effects of selection, which are often infeasible using real experiments due to time and cost requirements. Different tools have been proposed to reflect the complexity of the studied population (genomes structure, genetic architecture of the traits, and relatedness among individuals). However, few of them are specifically designed to simulate breeding strategies in crop species. We have therefore developed *breedgenr*, a simulation tool dedicated to meet breeder's needs to rationalize the different steps of a breeding scheme. *breedgenr* is an R package that present two distinctive features compared to existing tools. First, *breedgenr* uses real genotypic and phenotypic data from breeding programs to generate the reference population. Thus, it limits the number of possible hypotheses and scenarios regarding the evolutionary history and the structure of the breeding populations. Second *breedgenr* is based on non-parametric approaches to calibrate genotype-phenotype relationship. Consequently, it does not rest on specific assumptions regarding the genetic model and the genetic architecture of the phenotypic traits considered. Robustness and functionality of *breedgenr* were evaluated using real datasets on rice and different breeding schemes. The objective was to design breeding schemes that integrate different genomic selection scenarios. The results confirmed the robustness of the calibration of genotype-phenotype relationships based on the chosen non-parametric kernel methods, and the capability of *breedgenr* to simulate breeding populations and different breeding schemes: pedigree breeding and recurrent selection. Further ascertaining of *breedgenr* robustness and adjustment-expansion of its functionalities are presented.