MultiGWAS: An integrative tool for Genome Wide Association Studies (GWAS) in tetraploid organisms

Luis Garreta[1], Ivania Cerón-Souza[1], Manfred Palacio[1], and Paula Reyes-Herrera[1]

[1]AGROSAVIA

September 11, 2020

Abstract

The Genome-Wide Association Studies (GWAS) are essential to determine the genetic bases of either ecological or economic phenotypic variation across individuals within populations of model and non-model organisms. For this research question, current practice is the replication of the GWAS testing different parameters and models to validate the reproducibility of results. However, straightforward methodologies that manage both replication and tetraploid data are still missing. To solve this problem, we designed the MultiGWAS, a tool that does GWAS for diploid and tetraploid organisms by executing in parallel four software, two for polyploid data (GWASpoly and SHEsis) and two for diploids data (PLINK and TASSEL). MultiGWAS has several advantages. It runs either in the command line or in an interface. It manages different genotype formats, including VCF. It executes both the full and naïve models using several quality filters. Besides, it calculates a score to choose the best gene action model across GWASPoly and TASSEL. Finally, it generates several reports that facilitate the identification of false associations from both the significant and the best-ranked association SNP among the four software. We tested MultiGWAS with tetraploid potato data. The execution demonstrated that the Venn diagram and the other companion reports (i.e., Manhattan and QQ plots, heatmaps for associated SNP profiles, and chord diagrams to trace associated SNP by chromosomes) were useful to identify associated SNP shared among different models and parameters. Therefore, we confirmed that MultiGWAS is a suitable wrapping tool that successfully handles GWAS replication in both diploid and tetraploid organisms.

Hosted file

`MultiGWASv14.pdf` available at https://authorea.com/users/358323/articles/480510-multigwas-an-integrative-tool-for-genome-wide-association-studies-gwas-in-tetraploid-organisms