Genetics of Pasting Properties of High-Quality Cassava Flour for Baking in Selected Cassava (Manihot Esculenta Crantz) Genotypes in Tanzania

Abubakar Mikidadi

Abstract

Partial substitution of high-quality cassava flour (HQCF) in baking bread saves wheat importation and enhances production of cassava through cassava commercialization. Moreover, the integration aspect provides nutritional benefits to the composite cassava-wheat flour-based products. However, information on pasting properties of HQCF is key for composite cassava-wheat based flour products. The pasting properties of cassava varieties in Tanzania are not known and the best time to harvest them for making composite cassava-wheat flour is also not known. Increased utilization of HQCF for partial substitution also requires cost-effective high throughput phenotyping technologies such as SCiO and rapid development of baked suited cassava varieties using Marker assisted selection. There is limited information on loci that are linked to pasting properties of HQCF. The use of SCiO for pasting properties determination has also not been explored. The objectives of this study were to i) identify the most appropriate cassava genotypes and best harvesting time for pasting properties for baking ii) explore the possibility of using SCiO™ for screening large number of genotypes for pasting properties and iii) to identify genomic regions that are linked to the key pasting properties of HQCF through genome-wide association studies (GWAS). In order to identify appropriate genotypes and harvest times for pasting properties, eleven genotypes were planted at TARI-Ukiriguru in 2020/2021 and 2021/2022 seasons. A diverse set of 236 cassava genotypes were planted using augmented design at TARI-Ukiriguru and Chambezi experimental stations during the 2020/2021 and 2021/2022 seasons for GWAS and SCiO exploration. Harvesting was done at 10 and 12 months after planting (MAP). The rapid visco analyzer (RVA) was utilized to analyze the pasting properties. SCiO was used for generating spectral data. R statistical software was used to evaluate genotype and harvest time variability using general linear model. Three machine-learning algorithms: Gradient Boosting Machines (*gbm*), Linear Model (*lm*), and Lasso and Elastic-Net Regularized Generalized Linear Models (*glmnet*) were used for model calibration, optimization and validation. GWAS was performed in R using appropriate packages. The pasting temperature for all tested varieties was below 74 °C, the minimum pasting temperature for baking. Lowest setback viscosity was recorded for Kipusa both at 10 and 12 MAP indicating that this is the best variety for baking products offering smother texture with reduced tendency of becoming overly firm upon cooling. Highest set back viscosity was recorded for Hombolo at all the harvesting times implying that this is the best variety for producing wheat-cassava based products such as noodles requiring a more solid structure after cooling. SCiO results indicated that the best validation and calibration accuracies of R^2=0.51 and

R^2=0.56 and R^2=0.66 and R^2 =0.66 were obtained for peak viscosity and breakdown respectively using *glmnet* with first derivative pretreatment only. However, the best calibration and validation accuracies for pasting temperature were achieved through stacked several pre-treatments suggesting that stacking more pre-treatments does not always result in a robust calibration model. These calibrations and validations accuracies suggest that, NIR is a reliable low-cost innovative tool to deploy to predict cassava root pasting properties. A total of 48 SNPs and 301 candidate genes were significantly associated with the pasting properties. These portions of the genome can be tracked to ensure that improved cassava varieties carry the advantageous segments for pasting properties improvements. The results of this study can enhance cassava utilization and reduce wheat imports in Tanzania.