NON-HIERARCHICAL CLUSTERING OF CASSAVA GERMOPLASM BASED ON QUANTITATIVE TRAITS

Oliveira, EJ[1], Aud, F[1], Morales, CFG[1], Santos, VS[1], Kotwiski, F[2], Costa, V[2]

eder.oliveira@embrapa.br

[1]*Embrapa Cassava & Tropical Fruits, Cruz das Almas (BA), Brazil*

[2]*Bahiamido Agroindustrial Services, Laje (BA), Brazil*

The knowledge of the phenotypic variation of cassava (*Manihot esculenta* Crantz) germplasm allows estimating the genetic variability to support the selection of contrasting genitors. Therefore, the aim of this work was to define homogeneous groups of cassava germplasm based on yield traits, disease resistance and rot quality using *k*-means as a non-hierarchical method. Breeding values estimated by Best Linear Unbiased Predictor (BLUP) were used for the cluster analysis. The number of groups was defined according to the stabilization of the smallest within sum of squares. Seventeen clusters were defined to represent the diversity of the germplasm, whose number of accessions ranged from 7 (Group 15) to 69 (Group 9). In general, accessions belonging to Groups 1, 4, 7, 12, 15 and 16, showed good agronomic traits such as high fresh rot yield and starch yield (> 60.7 t.ha^{-1} and 18.6 t.ha^{-1}, respectively). In contrast, compared to other groups, the differences in disease severity in the groups mentioned previously were low, except for Group 15 (low bacteriosis severity). The groups obtained showed strong differences, as evidenced by the within values of sums of squares, which ranged from 215.1 (Group 15) to 2338.3 (Group 8). The *k*-means algorithm allowed the formation of consistent groups based on yield traits, disease resistance and rot quality. Therefore, the *k*-means algorithm was efficient in the formation of groups with the lowest within genotypic variation, especially concerning large amounts of data, such as in cassava germplasm banks.