



P435

Linkage disequilibrium and population structure in grain sorghum inbred lines using SNP markers

Karla Silva¹, Cicero Beserra de Menezes², Claudia Teixeira Guimarães², Leonardo Pimentel¹, Maria Marta Pastina², Robert Eugene Schaffert², Aluizio Oliveira¹, Jurandir Vieira de Magalhães²

¹Universidade Federal de Viçosa, Viçosa, Brazil. ²Embrapa Milho e Sorgo, Sete Lagoas, Brazil.

Studies of genetic diversity are essential for breeding programs to define breeding strategies and to identify superior inbred lines for the development of hybrids and to determine possible heterotic groups. The purpose of this study was to estimate the genetic diversity of sorghum inbred lines using linkage disequilibrium, population structure and clustering analyses. One hundred and sixty sorghum lines, including 109 restorer lines (R-lines) and 51 maintainer lines (B-lines), were genotyped with SNP markers using Genotyping by sequencing technique. The pairwise Linkage Disequilibrium (LD) between the SNPs markers was calculated based on the allelic frequency correlation (r^2) using the TASSEL software. The genetic structure of the population was estimated using the model-based Bayesian clustering method implemented in STRUCTURE software. Allele frequencies in each of the K clusters (from 1 to 10) were estimated, and for each line, the percentage of its genome derived from each cluster was estimated. A total of 86.342 single nucleotide polymorphisms (SNPs) were generated, remaining 29.649 SNPs with $MAF \geq 5\%$ and less than 20% of missing data. The study revealed a slow LD decay ($r^2 = 0.2$) of ~ 800 kb throughout the genome, probable due to the small population size or genetic drift. The population structure organized the inbred lines in two subpopulations, represented by restorer (R) and maintainer (B) lines, what was confirmed by the clustering analysis. The linkage disequilibrium analysis and the population structure were effective to investigate the genome of the breeding line population, and it was possible to quantify the degree of association between alleles of different loci of the population and confirm the possible existing subpopulations within the group of lines studied.