Validation and use of SNP markers for genetic compliance of commercial varieties in the coffee seed and nursery sectors

Barrera Santos[1]*, Berny Mier Y Teran Jorge C[1]., Umaña Emilia[1], Alvarado Julio[1], Cruz Jesslin[1], More Jimmy[1], Castro Elly[1], Zhang Dapeng[2]

[1]World Coffee Research; [2]United States Department of Agriculture, Agricultural Research Service *santos@worldcoffeeresearch.org

Introduction

➤ Nurseries play a crucial role in providing millions of plants to coffee farmers, so it is essential that nurseries offer pure and genetically traceable plants to farmers.

➤ Relying entirely on phenotypic characterization to ensure variety compliance is risky due to low phenotypic diversity.

➤ Using DNA markers to monitor genetic compliance is reliable and cost-efficient. KASP SNP markers are a preferred option due to their accuracy, stability, reproducibility, and high-throughput capabilities.

➤ The main objective of this project was to validate SNP markers at a large-scale sampling and establish an accurate fingerprint database to identify varieties of coffee arabica in seed and nursery sectors.

Materials/Methods

➤ The fingerprint database was developed using 1,424 Arabica coffee samples collected from five Latin American countries. The validation of this database included 30,716 samples from Guatemala, Honduras, El Salvador, Nicaragua, and Peru.

➤ We utilized 45 KASP-SNPs markers, originally published by Zhang et al. (2021). SNP genotyping results were acquired through outsourcing from Intertek.

➤ SNPs data were used to develop a fingerprint database that determines the genetic makeup of the 23 varieties of coffee arabica.
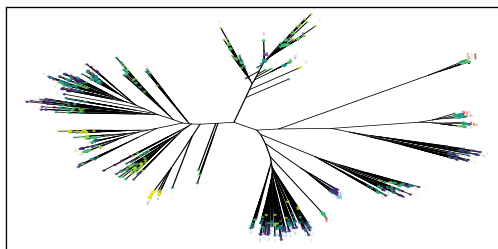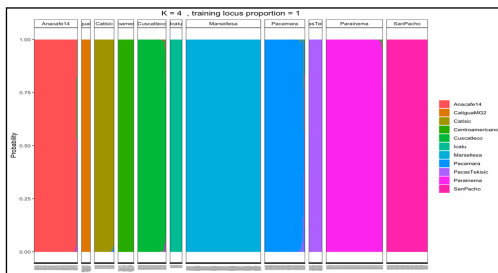

Figure 1: Phylogenetic tree of references

Figure 2: Assignment test probabilities of references using a K-fold cross validation.
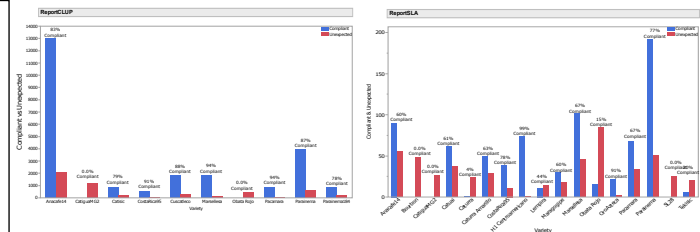

Figure 3: Genetic compliance of varieties evaluated in 2023 in 5 Latin American countries

Results/Discussion

The results indicated variation across varieties, as well as within varieties. This variation was observed in certain cases, likely because many coffee varieties are released at early generation stages (F4-F7), which lead to the presence of residual heterozygosity. Capturing this variation has allowed us to better identify the set of markers to use as a fingerprint for variety identification and nursery and seed production purity assessment. When we utilized the database to assess the genetic compliance of the seed and nursery sectors, we found that for certain varieties, the degree of compliance varied significantly, ranging from 0% to as high as 99%, when comparing the reported varieties to the expected varieties.

Conclusion/Perspectives

The fingerprinting database created from published SNP markers and implemented by WCR has proven to be a useful tool in identifying the authenticity of coffee varieties in the seed and nursery sector. Further research is needed to explore the potential of molecular characterization in developing new coffee varieties and preserving the genetic diversity of coffee.

References:

• Zhang, D., Vega, F.E., Solano, W. et al. Selecting a core set of nuclear SNP markers for molecular characterization of Arabica coffee (Coffea arabica L.) genetic resources. Conservation Genetic Resources 13, 329–335 (2021). https://doi.org/10.1007/s12686-021-01201-y

• World Coffee Research. (2023). Coffea arabica KASP genetic markers for low-density genotyping. [Dataset]. World Coffee Research. https://doi.org/10.3886/ICPSR36966.v1.