Genomic prediction using low density marker panels in *Coffea canephora*: Predictive ability using random forest algorithm

Ithalo Coelho de Sousa[1], Cynthia Aparecida Valiati Barreto[2], Matheus Massariol Suela[2] Talita Estéfani Zunino Santana[3], Jackson Tavela da Silva[1], Emilly Ruas Alkimim[4], Letícia de Faria Silva[1], Eveline Teixeira Caixeta[5], Moysés Nascimento[1]

Abstract

Coffea canephora is one of the most important coffee species and is responsible for approximately 40% of Brazil's coffee production. Among the traits used to evaluate coffee production, vegetative vigor has good potential. This study aimed to evaluate the trade-off between density marker panels size and the predictive ability using random forest algorithm to predict the genetic merit for vegetative vigor in *Coffea canephora*. The data used composes the breeding program of the Empresa de Pesquisa Agropecuária de Minas Gerais in partnership with Universidade Federal de Viçosa and Empresa Brasileira de Pesquisa Agropecuária-Café. This data consisted of 165 genotypes of *Coffea canephora* and 14387 markers. To evaluate the dimensionality reduction by predictive ability (PA), we evaluated different density marker panels (from 25 to 14387), using Random Forest (RF) to select markers and to predict the genetic merit of the genotypes. The average of the PA varied from 0.28 (14387 markers) to 0.51 (200 markers). In general, as the number of SNPs increased, the PA decreases, which can be explained by the fact that the tree-building process is *greedy*. The results show that the SNP's reduction can improve the PA for vegetative vigor in *Coffea canephora*.

Keywords: Coffee; Random Forest; Predictive ability; Regression Tree.

[1] Department of Statistics, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil
[2] Department of General Biology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil
[3] Deparment of Animal Science, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil
[4] Federal University of Triangulo Mineiro, Iturama, Minas Gerais, Brazil
[5] Brazilian Agricultural Research Corporation, Embrapa Coffee, Brasília, DF-Brazil