

GWAS-Assisted Genomic Prediction for Forage Nutritional Quality Traits in Elite Alfalfa (*Medicago Sativa* L.) Germplasm

Harpreet Kaur, New Mexico State University

Cesar Medina Culma, USDA-ARS

Christopher Pierce, New Mexico State University

Arshdeep Singh Gill, New Mexico State University

Jon Reich, Alforex Seeds

Steve Damon, Corteva Agriscience

Long-Xi Yu, USDA-ARS

Ian Ray, New Mexico State University

Alfalfa (*Medicago sativa* L.) is a highly-valued forage legume because of its nutritive quality. The goals of this study were to identify genomic regions, and to evaluate genomic prediction (GP) models, associated with forage nutritional quality traits in an elite New Mexico Genome Selection (NMGS) alfalfa population grown near Las Cruces, NM. Forage samples for nutritive value assessment were collected from seeded plots of half-sib families derived from 215 semi- and non-dormant NMGS plants immediately prior to the third forage harvest in 2018 and 2019. Among the maternal parents of the 215 NMGS half-sib families evaluated, 12,884 genome-wide single nucleotide polymorphism (SNP) markers were identified using a *Medicago truncatula* reference genome assembly. These SNP markers, and best linear unbiased estimates for 15 nutritional traits, were used to conduct a genome-wide association study (GWAS) based on six polyploid gene action models as provided by the R package, GWASpoly. Twenty-seven markers were significantly associated with 12 different forage quality traits (e.g., lignin, neutral detergent fiber digestibility, *in-vitro* true dry matter digestibility, etc.) in two years. All SNP markers were also used for GP analyses of the 15 nutritional quality traits using ridge regression BLUP (rrBLUP), genomic BLUP (GBLUP), support vector machine (SVM)-linear, SVMgaussian, random forest, Bayes A, Bayes B, Bayes C, and Bayesian Lasso methods. For all GP methods, the prediction accuracies (i.e., coefficient of correlation, r) between predicted and observed values were low to moderate. For instance, average r ranged between -0.04 (magnesium in 2018) to 0.26 (lignin in 2018) for the rrBLUP model over both years. GWASassisted GP using a subset of SNP markers possessing GWASpoly scores >1 greatly enhanced prediction accuracies for all methods (e.g., r for rrBLUP ranged from 0.74 [ash in 2018] to 0.87 [neutral detergent fiber digestibility in 2019]). Additionally, moderate to high prediction accuracies ranging from 0.57 (magnesium) to 0.88 (lignin) were obtained using a weighted GBLUP method with a complete set of markers weighted according to their GWASpoly scores. The data were also analyzed using a recently developed *Medicago sativa* reference genome assembly and similar results were obtained.