Evaluating Approaches to High-Throughput Phenotyping & Genotyping for Genomic Selection in Alfalfa
Cornell University - Robbins

Project Award: $36,811

Justification:
- Genetic gain in alfalfa has approached stagnation in the past few decades, limiting benefits to alfalfa farmers. Adoption of new breeding technologies has also lagged due to the complexity of the genetics, a high phenotypic burden and a paucity of public funds for a crop that is one degree of separation too far from the consumer's mouth and interest. Evaluation of breeding material requires multiple harvests per year for multiple years, limiting the size and number of field trials. The low heritability of forage yield also demands extensive replication, further limiting the number of breeding populations that can be evaluated. The ability to screen more material will lead to higher effective selection intensities, and increase the frequency of developing populations that outperform current varieties.

High throughput phenotyping (HTP) technologies could drastically reduce the phenotypic burden in alfalfa by replacing a plot harvester with a multi-spectral imaging drone and a swather for some harvests, locations, and/or replications. Quantitative genetic models can be built to accurately predict forage yields from spectral imaging, especially given that the harvested product is imaged directly. Images taken throughout the production years of a stand can also provide insight into genotype by environment interactions (GE), in which varieties have differential growth responses under different conditions. Understanding the genetic signal in differential growth response will allow for identication of breeding targets and optimal population change for some set of predictable environmental conditions.

Inclusion of genome-wide markers can improve these types of prediction models by enabling related material to share information. These genomic prediction models can allow for reduced replication, sparse testing and even prediction of unobserved populations. Estimating realized genetic relationships in alfalfa is complicated by the fact that varieties are not genetically distinct individuals. As an obligate outcrosser, alfalfa must always be bred on a population level, where varieties are released as synthetics to avoid inbreeding and take advantage of population-level heterosis. This has limited implementation of marker-based selection because large numbers of individuals must be genotyped and intermated to avoid inbreeding in future generations. Single individuals are not representative of a variety as a whole, and genotyping many individuals from each variety is costly and restrictive.

We are currently in the process of evaluating a new genotyping strategy for alfalfa, where DNA from many (100) individuals is bulked in a given breeding population or variety for genotyping. Because much of quantitative genetics and selection theory hinges on population level parameters, the current machinery can be easily adapted to breeding on a population level. Allele counts within each variety or breeding population, as opposed to allele counts within each individual, will be used to estimate genetic relationships between populations. This genotyping strategy will allow for prediction of additive effects for genetic gain, as well as dominance effects to exploit population level heterosis.

Genomic selection (GS), in which genome-wide markers are used to predict performance and choose parents, has been shown to be a promising method for population improvement, but is in its infancy for implementation into plant breeding programs. For a forage breeding program, we envision the use of numerical optimization approaches to define optimal allele contributions and select parental

populations. Parental breeding populations would then be intermated in optimal proportions to produce idealized allele frequencies in the offspring population using an insect pollinator. Prediction models would be updated on a yearly basis from both high- and low-throughput phenotypes of material in the field.

An affordable, population-based genotyping method would need the ability to count alleles in a given sample, a task well suited for sequence-based methods. However, we chose not to pursue a GBS type approach due to the costs associated with the patent on the method (KeyGene, U.S. patent 8,815,512 B2), the inability to scale up economically due to the high sequencing depth required, and the reduced representation of the genome. Therefore, a publicly available, cost effective, high-throughput, high-density, sequence-based genotyping platform is desperately needed for alfalfa. Amplicon sequencing-based methods could scale up economically if well designed. We have been in contact with the Breeding Insight initiative (USDA, 2019) to use all sequence data collected from this project to help build this kind of marker platform for alfalfa.

This proposal is in concert with another proposal as part of an effort to gather supporting information for further federal funding from the USDA. The other proposal is requesting funds to sequence diverse materials from the nine alfalfa germplasm sources (Barnes et al. 1977) that have been previously phenotyped in Las Cruces, NM (Segovia-Lerma et al. 2004). That material will be used as a proof of concept for population level genomic prediction, as well as provide a pseudo-pan-genome of alfalfa that can be used to help construct a high-throughput marker platform. The sequence requested in this proposal is that of Cornell breeding material that will allow us to validate the HTP prediction models, as well as contribute toward the effort of an affordable marker platform. Integration of population level genetic relationships with high-throughput multi-spectral imaging has the potential to revolutionize the modern forage breeding program in the age of Digital Ag.

Objectives:
- The objectives of this project are to 1) Estimate the genetic correlations of multi-spectral indices with forage yield and quality using population-level genomic relationships; 2) Determine efficacy of phenotype reduction using spectral indices; 3) Evaluate the efficacy of using a population bulk genotyping approach to predict yield performance in elite germplasm; and 4) Fit population- specific curves for each harvest using genomic relationships and spectral indices.