National Alfalfa & Forage Alliance – 2017 Alfalfa Check-off Research Grant
Final Report
2/4/2018

Title: Maximizing alfalfa's yield potential

PI	E. Charles Brummer, University of California, Davis
Address:	1234 PES Building, Plant Sciences Dept.,
	Univ. of California, Davis, One Shields Ave., Davis, CA 95616
Phone:	530-574-6133
Email:	ecbrummer@ucdavis.edu

Abstract: Alfalfa yield has changed little over the past 25 years, but improving yield is essential if alfalfa will remain in crop rotations, and ultimately dairy or livestock rations, and not be displaced by other crops, such as corn silage. Improving yields through breeding could be facilitated by some combination of improved selection methodology, use of new, unexplored germplasm, and the deployment of sensors to enable routine indirect estimation of yield in breeding trials. In the experiments described here, we address all three routes. First, using genomic selection, we showed that we could use markers to develop populations with high vs. low yield, although "progress" toward lower yield was more successful than increased yield. Some reasons for these results are discussed, but the potential for genomic selection to work remains. We attempted to discriminate among germplasm accessions using frequency-based molecular markers, but for reasons that are not clear, our experiment failed and known relationships were not observed. This experiment is being repeated. Finally, we used a multispectral camera to estimate plant height and yield in five and 25-plant plots. The sensor data correlated with yield data in the five plant plots very well (r = 0.90) but less well with height. However, height was accurately measured using drone-based sensors on larger 25-plant plots. The potential of sensor data to estimate forage yield and height repeatedly throughout the season in breeding nurseries looks promising. Collectively and despite some setbacks, we are optimistic that yield improvement in alfalfa is feasible and that new technologies can help us get there.

Keywords: alfalfa, yield, genomic selection, germplasm, sensors

Introduction:
High yield of highly nutritious forage is critical for profitable alfalfa production. Yield advances are being made in other crops, in particular, corn silage, so the need to boost alfalfa's yield potential just to remain competitive with other crops is critical. Unfortunately, most evidence shows that alfalfa yield has not improved over the past several decades, with stagnation in most areas of the country (Brummer & Casler, 2014). Improvements in alfalfa have occurred, as undoubtedly, resistant cultivars will have higher yield, particularly later in stand life, under disease or insect pressure (Lamb et al., 2006). However, these yield gains are effectively defensive – while they protect yield, they are not advancing yield potential. University variety trials in California show clear indications that even across years, yields have not increased since the 1970s (Fig.1).

My research group is investigating three interlinked objectives related to alfalfa yield – heterosis, yield selection, and dormancy. Heterosis can be defined as the superiority of a hybrid (formed from crossing inbreds or populations) to perform better than the midparent or the high parent value, with the latter being the more interesting from a commercialization standpoint. Although alfalfa cultivars are typically

marketed as synthetic varieties, semi-hybrid cultivars theoretically could be produced and they could express heterosis for yield (Brummer, 1999). We have evaluated the potential for yield heterosis among alfalfa germplasm, exploring both yellow-flowered falcata (Riday and Brummer, 2002; 2005) and non-dormant germplasm (Sakiroglu and Brummer, 2007), complementing conceptually similar work done by Ian Ray's group in the early 2000s (e.g., Segovia-Lerma et al., 2004). The substantial amount of underused germplasm from the USDA-National Plant Germplasm System and other genebanks worldwide could be used to create genetically distinct populations useful to form heterotic groups *de novo* based on breeding methods equivalent to reciprocal recurrent selection.

Second, we have applied genetic markers to assess yield and yield-related traits (e.g., Robins et al., 2007; Li, Alarcon-Zuniga, et al., 2016). Recently, we have used high-throughput array-based SNP markers (Li et al., 2014a) and genotyping-by-sequencing (GBS) (Li et al., 2014b) to quickly and cost effectively cover the entire genome. We used GBS markers and yield data from a clonal phenotypic selection program (Li et al., 2015) and also from two Italian half-sib family selection programs (Annicchiarico et al., 2015) to create genetic prediction models. These models suggest that we can accelerate yield improvement using GBS marker-only genomic selection (GS). These markers can also be used for genome-wide association studies (GWAS) (Sakiroglu and Brummer, 2016).

More recently, we have been using GBS markers to assess population distinctiveness. We evaluated populations of CUF101 that had been divergently selected for autumn dormancy – taller or shorter plants in autumn – using GBS markers. Marker loci whose allele frequencies shifted as a result of selection may be associated with dormancy. We genotyped four independent 24 plant bulks of each population. The results show that the four replicate bulks tightly cluster but the three populations are clearly differentiated (Fig. 2). The O and H populations are not different phenotypically for plant height in autumn (the selection criterion), yet they can be clearly separated by markers nonetheless. The importance of this experiment is that we can easily separate populations based on the marker allele frequencies in bulked samples. Thus, this opens the door to apply markers more widely in alfalfa breeding where all germplasm is based on populations, not individual genotypes as in inbred crops.

Both conventional and marker-assisted selection depends on accurate phenotypes of the traits under selection. For traits like yield and even plant height, repeatedly measuring large nurseries throughout the year is both arduous and expensive. Automated phenotyping for plant size, growth and development stage, and plant health and stress level has been demonstrated in a number of field crops (e.g., white clover [Inostroza et al., 2016]; cotton [Andrade-Sanchez et al., 2014]; and wheat and soybean [Bai et al., 2016]). Importantly, sensor-based phenotyping has been used to accurately measure alfalfa height (Pittman et al., 2015).

The objectives for this project included:
Objective 1. Evaluate the yield gain possible from genomic selection,
Objective 2. Classify germplasm by genotyping populations that could form heterotic pools
Objective 3. Characterize alfalfa growth using proximal and/or remote sensing

Materials and Methods:
Objective 1. For this experiment, we evaluated ten germplasms in replicated field trials in Tulelake, CA and Ithaca, NY (see Fig below). These germplasms derived from a population created by Dr. Don Viands at Cornell Univ. in the 1990s. From that population, we conducted phenotypic and genomic selection. We have previously described this population, the NE-1010 clonal selection population, in an experiment

using SSR markers for association analysis (Li et al., 2011a). Briefly, NY0358 was formed by intercrossing three elite, semi-dormant cultivars and recombining the resulting population twice.

The NY0358 population was selected for yield for two cycles using clonal evaluations at multiple locations. At each location, about 200 individual plants were included at the beginning of each cycle. These plants were clonally propagated using stem cuttings, and three replications were planted to the field. In each replication, three clones were included in a plot; thus, each individual plant was replicated nine times at each location in each cycle. Yield data were collected on individual plots, bulking the biomass of the three clones within the plot. In the first cycle, data were obtained from Ithaca, New York; Ste.-Foy, Québec; and Ames, Iowa (PI Brummer participated in the first cycle of selection when he was at Iowa State Univ.). The second cycle included locations near Ithaca, New York; Ste.-Foy, Québec; and Lethbridge, Alberta. After each cycle, the top yielding 10% of genotypes based on an across location analysis of total annual yield were recombined to form the next cycle. The second cycle selections were intercrossed in 2012. Separate populations specific to New York were also developed in an analogous manner. Co-PIs Dr. Don Viands and Dr. Julie Hansen at Cornell University have overseen the clonal selection program since its inception and have increased seed from all cycles of selection. These populations are being evaluated in trials planted in 2014, separate from the experiment described here.

The clonal selection experiment offered an excellent opportunity to attempt genomic selection in alfalfa. Fortuitously, we had DNA remaining from our previous experiment (Li et al., 2011a) of the initial set of 190 C0 plants evaluated in the clonal nursery. Subsequently, we extracted DNA of the 200 individual plants from Cycle 1 which were evaluated in the field and from which the parents of Cycle 2 were selected. Thus, we were able to conduct DNA marker analysis using nearly 400 individual plants for which yield data had been collected.

We used GBS to score SNP markers throughout the genome. We used the *Ape*KI protocol developed by (Elshire et al., 2011) to generate markers, and multiplexed approximately 100 genotypes per lane. Genotyping was done using Illumina HiSeq 2000 sequencing systems at the University of Texas Genome Sequencing and Analysis Facility.

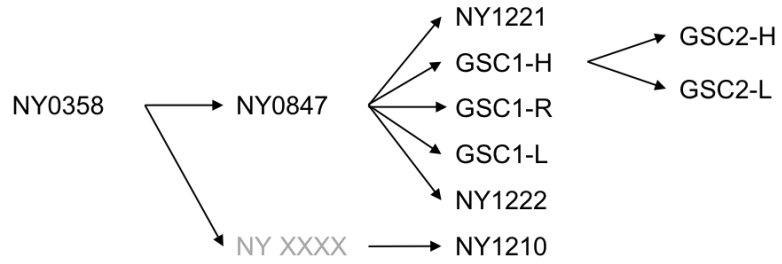
The phenotypic and genotypic data from the two sets of ~200 individuals were used to develop GS prediction models; for our purposes, we used these populations as training populations for our marker-only selection program. Numerous GS models have been developed (Heslot et al., 2012), but based on results from Heslot et al. (2012), we decided that Ridge Regression-Best Linear Unbiased Prediction (RR-BLUP) would be an appropriate choice for biomass yield, and therefore, used the computer package RR-BLUP (Endelman, 2011). The genomic selection model development was described in detail in Li et al. (2015).

After evaluating models from various combinations of locations, we decided to focus on the model developed from the C1 NY total yield data because these data were more robust than the other locations or than any individual harvest. We then grew and genotyped 400 individual seedlings from NY0847 using GBS. Our goal was to select from a population also used for phenotypic selection. We conducted GBS as described previously, multiplexing 100 genotypes in a single lane of the HiSeq 2000. Following SNP scoring and imputation, we computed GEBVs for each individual plant based on the score summed across all markers and then selected the top 20 genotypes based on GEBVs. We restricted selections to no more than four individuals from any given maternal half-sib family to maintain variation in the population. These 20 individuals were intermated in the greenhouse by hand without emasculation to form the GSC1-H

population. An analogous population based on the lowest 20 GEBVs was also formed (GSC1-L). In addition, a random selection of 20 plants from the 400 plant population was intermated as a control (GSC1-R).

For the second cycle of selection, seeds of each maternal half-sib family derived from the intercross resulting in GSC1-H were germinated and DNA from 20 plants from each of the 20 families was isolated for marker analysis. We selected the top and bottom 20 individuals as done for Cycle 1 to form GSC2-H and GSC2-L, respectively. We then increased seed of all populations shown in the figure in 2016.

To evaluate our progress from selection, we planted trials in spring 2017 at three locations in the northern US, the regions to which this germplasm is most adapted. These standard alfalfa cultivar trials were planted in Ithaca, New York with Drs. Viands and Hansen, Arlington, WI with Dr. Riday, and in Tulelake, CA in conjunction with co-PIs Orloff and Putnam. The trial in Arlington failed to establish due to heavy rain after planting. We collected yield data in the establishment year (2017) and first full production year (2018) in Tulelake and in 2018 only in Ithaca.



Schematic of ten populations evaluated in the Genomic Selection trial. NY0358 is the base population for the other nine. The derived NY populations were developed using phenotypic selection based on clones (the differences being the number of locations of data used to select plants). The GS populations were developed using marker-only selection. “H” designates selection based on the genomic model for high yield and “L” for low yield. The “R” population was generated by randomly choosing individuals for intercrossing.

Objective 2. Classify germplasm by genotyping populations that could form heterotic pools

We evaluated ~200 germplasms, including UC breeding germplasm, fall dormancy check cultivars, and NPGS accessions, mostly of non-dormant germplasm. The methodology we used for SNP marker genotyping, bioinformatics, and population diversity analysis were identical to those we have been using in our lab for the past several years (Li et al., 2014b; Li et al., 2015; Munjal and Brummer, 2018). We collected tissue samples from populations in Spring 2018, as bulked samples of 25 plants each for each of the populations to be assayed. We constructed GBS libraries based on the protocol of Elshire et al. (2011) as modified by Li et al. (2014b). We are now using the GBS-SNP-CROP pipeline for (SNP) discovery (Melo et al., 2016) because its flexibility results in a larger number of markers being identified.

We compared allele frequencies at each SNP locus among the populations. Sample specific allele frequencies were calculated as their maximum-likelihood estimates given by the number of reads representing an allele at a given marker in a given sample divided by the total number of reads representing that marker within that sample. We then compared populations based on allele frequencies by computing genetic distances among populations and conducting a principal components analysis as we have described previously (Annicchiarico, et al., 2016).

Objective 3. Characterize alfalfa growth using proximal and/or remote sensing

Our goal was to assess the ability of drone-based sensors on predicting plant height and/or biomass of alfalfa breeding plots or fall dormancy standard check plots. For the experiment described here, we evaluated either half-sib family height and yield based on the five plant plots or fall dormancy standard check trial plots of 25 plants per plot.

To capture sensor data from our field experiments, we used the following drone platform and sensing equipment. A Red-Edge-M™ multispectral camera made buy MicaSense® was mounted to a DJI Matrice 100 quadcopter platform. The field was flown at an altitude of 15 meters, with drone route, speed and image capture timing set to capture images with about 50% overlap. All images taken with the camera perpendicular to the plane of the field. No other angles were captured.

The cloud computing service DroneDeploy was used to stitch together the images taken into a orthographic image (orthomosaic) of the field. This service also creates a height map for each pixel in the image based on GPS data from the drone platform, and photogrammetry techniques. We obtained one of these maps from each of the cameras on the multispectral camera system.

We aligned each orthomosaic in QGIS, a free GIS mapping tool, using the drone GPS data. Using QGIS, we created outlines of the plot space to record data for each plot. Calculated NDVI was used to create a mask separating the plant material from the soil. From this we obtained the total canopy surface area in square meters. Because our plots are in leveled fields, it is expected that the ground height measurement is fairly consistent within a plot, and so the median height of the ground was used to control for any major outliers due to soil surface disturbances. We subtracted this value from the height of each pixel in the plant area, and recorded the mean and the maximum value of height above soil level in meters. By multiplying canopy area by height, we obtain a rough estimate of plant volume.

Results and Discussion

Objective 1.

The plots in Tulelake (photo below) and Ithaca established well but the trial in Arlington failed to establish. We harvested twice in the seeding year (2017) in Tulelake but no yield harvests were taken in the establishment year in Ithaca. The first full production year was 2018, and yields were recorded at both locations, with the data shown in Table 1. Our genomic prediction model was based on data collected in Ithaca, so that location would be expected to be the most likely to show improvements. Although Tulelake is roughly at the same latitude as Ithaca, the vastly different production system – irrigated vs rainfed – would suggest that yield gains in Tulelake could be less than in Ithaca.

First year yields show that the two locations gave largely similar results (Table 1). When comparing the first cycle genomic selection populations, GSC1-H had higher yield than GSC1-L. This indicated that our markers were able to discriminate between high and low yielding plants. However, GSC1-H was not higher than NY1221, the population selected phenotypically for higher yield;

unfortunately, neither showed an improvement over the base population. Thus, genomic selection seems quite successful at identifying low yielding plants, but less valuable at increasing yield in this experiment. The GSC1-R population would be expected to fall between the high and low populations and the results show that it does. The cycle 2 GS populations perform poorly, suggesting that the transferability of our GS model from one cycle to the next was poor. We may also be seeing an impact of inbreeding depression in these results. We also have anecdotal evidence that the fall dormancy of the GSC2 populations has declined, perhaps by genetic drift during our GS for higher yield. The inadvertent dormancy effect may counter the positive effect of selection using our GS model. In any case, the robustness of GS models for more than one cycle is poor in this experiment, at least.

Table 1. Total biomass dry matter yield across 3 (Ithaca) or 4 (Tulelake) harvests in 2018, the first full production year. Trials planted in spring 2017.

Entry	Ithaca	Tulelake	Overall
	----------T/acre----------		
NY0358	7.18	8.83	8.00
NY1210	7.22	8.74	7.98
GSC1-H	7.17	8.77	7.96
NY0847	6.90	8.88	7.89
NY1222	6.80	8.98	7.89
NY1221	6.85	8.82	7.83
GSC2-L	6.77	8.82	7.80
GSC1-R	6.82	8.60	7.71
GSC2-H	6.46	8.56	7.51
GSC1-L	6.49	8.40	7.45
Mean	6.87	8.74	7.80
LSD (5%)	0.34	0.35	0.25

There are several reasons why our GS model may not have been more successful. First, the model was based on spaced plant yield data and not on data generated from swards. Although many breeding programs are based on spaced plants, the correlation of spaced plant data with sward performance can be low, as shown in recent years in tall fescue (Waldron et al.) and switchgrass (ref). Thus, one way to improve GS modeling would be to develop models based on sward yields. Of course, the same argument can be used to improve traditional phenotypic selection and our current NAFA project aims to assess this possibility.

In addition, our GS model was a rather simple model that did not account for the full complexities of autotetraploid genetics – namely, it did not differentiate among the three heterozygote genotypic classes (for example, AAAT, AATT, and ATTT). A more refined model, possible using statistical frameworks developed since we did this selection, could result in a better prediction of yield. Finally, we need to do a better job at keeping fall dormancy constant across cycles, as this change alone could result in yield differences. This, together with a minimization of potential inbreeding depression, when applied to better phenotypic data, might improve our ability to predict yield.

We are conducting a second production year of data in 2019 to get a full picture of the productivity of these populations.

Objective 2.

Our goal with this objective was to use population-based allele frequencies to assess relatedness of our UC breeding germplasm together with alfalfa accessions from the National Plant Germplasm System's alfalfa collection. We very successfully differentiated populations separated from one another by selection for more or less dormancy (Munjal and Brummer, 2018), suggesting that we could similarly identify relationships among many diverse germplasms. Unfortunately, our experiment failed. We successfully generated SNP markers for all accessions, but no meaningful clustering occurred when we analyzed the allele frequencies. We are not certain what went wrong with the procedure, but several options present themselves. First, we could simply have had DNA mix-ups. DNA from some of the accessions were extracted by a visiting scientist who was not well-versed in lab technique; however, the lack of clustering was not isolated to those samples. We think this reason is unlikely to be the case.

More likely, we had a procedural issue with the GBS itself. Perhaps we simply did not get enough sequencing, which was less than that for the previous experiment. But more likely we simply had a series of related problems, from old reagents to poor quality DNA, that all conspired to give the ambiguous results. We are currently redoing the experiment using new reagents, new DNA extractions in some cases, and some altered methodology that has worked better for other labs on campus.

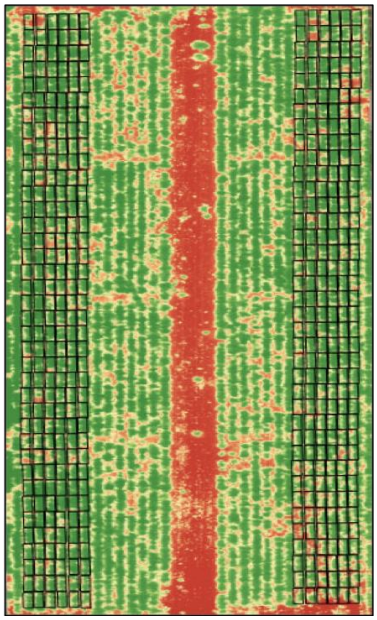
Objective 3.

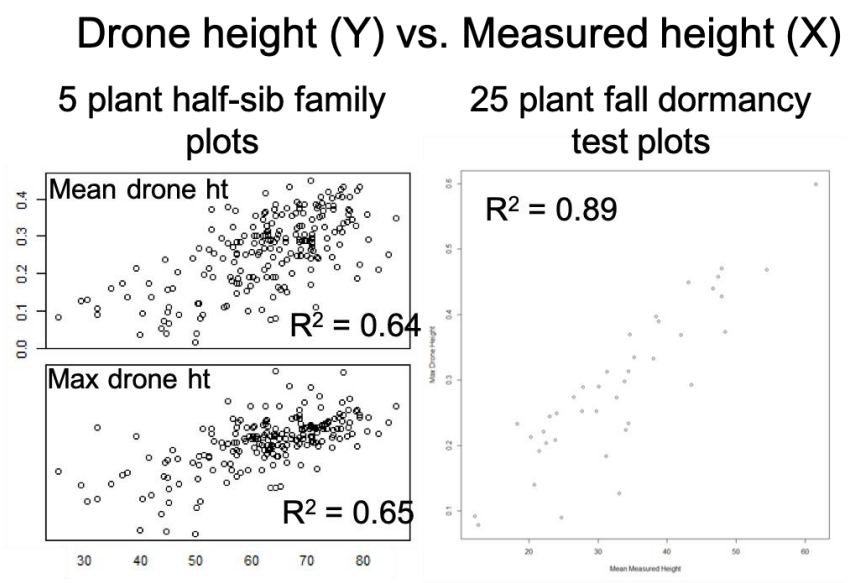
We obtained sensor data from drone flights over two trials: a set of half-sib families planted in 5-plant plots and a fall dormancy standard test using 25 plant plots. The figure below shows two sets of the half-sib plots on either side of a larger breeding nursery. We measured plant height on these plots and measured total wet and dry weights on a plot basis.

We computed correlations between the volume and height measured or computed from sensor data and the measured height, dry weight, or wet weight of the plots. Our initial attempt at correlations showed no relationship between drone data and the measured values, but this was due solely to an error in specifying plot boundaries and numbering. Once that was corrected, we found the following:

Measured Height on Drone Mean Height:	0.64
Measured Height on Drone Max Height:	0.64
Wet Weight on Volume:	0.91
Dry Weight on Volume:	0.90

Thus, the drone based heights were reasonably good at measuring height in five-plant plots, but the 0.64 was too low for selection or predicting dormancy level, for example. More interesting was the NDVI and height correlation of 0.90! This is very favorable. If confirmed by further evaluations this year, this would have substantial value for measuring yield and selection for increased productivity. Turning to the dormancy trial data, we found a better correlation between drone height and the 25-plant standard test plot, nearing 0.90. These results suggest that drone based phenotyping of yield and height is feasible, although how broadly we can extend these results across yield trials or different experiments isn't clear. We likely will need to recalibrate depending on the experiment, and possibly the harvest. Further results will be coming in 2019.





Acknowledgements: Funding for this study was provided by the U.S. Alfalfa Farmer Research Initiative of the National Alfalfa & Forage Alliance.