Maximizing alfalfa's yield potential

E. Charles Brummer
1234 PES Building; Plant Sciences Dept.; Univ. of California, Davis; Davis, CA 95616
530-574-6133
ecbrummer@ucdavis.edu

Abstract:
Although alfalfa is a major crop in the United States, yield improvement has been limited in the past several decades. One way to potentially improve yield is to apply modern genomic and phenotyping technologies to breeding programs that will enable faster cycle time and/or faster and easier data collection, which could improve genetic progress. In this project, we evaluate progress from our initial genomic selection experiment, which showed that yield increases are possible using this method, although further work is needed to optimize the application of the technology. The second objective is to use half-sib progeny evaluations to select for yield and to determine if yields can be indirectly estimated using drone-based sensors. Our initial data suggests that multispectral sensors can effectively predict yield, which would make the evaluation of large yield trials easier, potentially improving yield. Future work will attempt to combine genomic selection, high-throughput yield phenotyping, and half-sib selection in solid-seeded swards to develop higher yielding cultivars.

Introduction:
Yield is the bottom line for crop breeding programs. Even if one considers other traits of importance, such as nutritive value, the need for maximizing production given the constraints of the other trait(s) remains. Oddly, and somewhat inexplicably, yield improvement in alfalfa has stalled for the past ~25 years across the nation (Brummer and Casler, 2014). Based on a variety of data from university trials to USDA statistics, we can conclude that overall biomass production per acre is more-or-less stagnant, with gains relative to older varieties likely to be due more to improved disease/pest resistance than to pushing the yield ceiling *per se*.

Does it matter that yields are not increasing, or at least, not increasing by much? I think the answer is yes. Higher production per unit area is necessary to accommodate the needs of a growing population desiring more animal products, to overcome the loss of agricultural land to urbanization, degradation, or other uses, and importantly, to remain competitive with other crops, such as corn silage, whose productivity is increasing. In addition to providing desirable nutritive value to ruminants, alfalfa also provides numerous positive environmental services to an agroecosystem (nitrogen fixation, perennial soil cover, improved soil organic carbon, to name three) (Olmstead and Brummer, 2008). Thus, ensuring it remains in cropping systems has both direct animal feeding benefits as well as indirect benefits to other crops in rotation.

The lack of yield progress can be ascribed to the perennial nature of alfalfa, to the harvesting of the entire plant (and hence, the inability to make gains in harvest index as in grain crops), and to the conflicting demands of long term persistence, higher nutritive value, better disease resistance, and incorporation of new traits all of which divert attention from direct yield selection. The first two reasons undoubtedly pose limitations relative to grain or oilseed crops,

but there is no obvious physiological reason why yield improvement could not be made. However, the latter reason is a red herring, as every other crop faces similar conflicts among multiple traits – wheat, for instance, must maintain strict milling quality standards and yet wheat yields have been increasing despite that limitation. Although there are plausible reasons why alfalfa yield improvement could lag other crops, there's no reason that yield improvement (over and above defensive disease/insect resistances) shouldn't be possible.

This research project posits that selecting explicitly for yield in solid seeded sward plots, by using genomic selection to minimize cycle time, and by using remote sensing platforms to indirectly assess yield throughout the stand life of an alfalfa breeding line, we can improve biomass yield *per se*.

Materials and Methods:

Objective 1. Genomic selection in alfalfa

Using the GS model developed by Li et al. (2015), we conducted two cycles of GS (Fig. 1) using RR-BLUP and random forest imputation for missing values to compute genomic estimated breeding values. In the first GS selection cycle, we grew seedlings of NY0847 and

Figure 1. Populations developed by Genomic (GS) or Phenotypic (PS) Selection. (C1 = Cycle 1; C2 = Cycle 2; H = High and L = Low GEBV; R = Random).

genotyped 384 plants using GBS (methods as in the paper). Using the GS model constructed from NY data only, we computed GEBV and selected 20 plants with high GEBV (GSC1-H), 20 with low GEBV (GSC1-L), and also chose 20 plants randomly for a control population (GSC1-R).
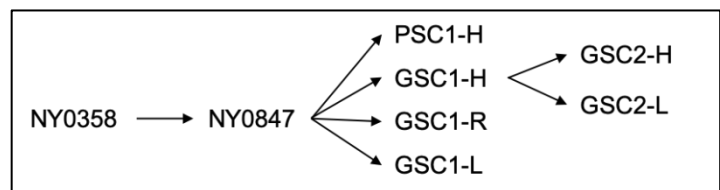
The 20 plants from each population were intercrossed to form the cycle 1 populations. Subsequently, we conducted a second cycle from the GSC1-H population to produce the GSC2-H and GSC2-L populations. Seed of all populations in Fig. 1 were increased in cage isolations to produce sufficient seed for plot evaluations (Fig. 2). Field trials were sown in Ithaca, NY and Tulelake, CA (Fig. 2) in spring 2017. Two full production years were harvested in 2018 and 2019 across four harvests per year in Tulelake and three in Ithaca.
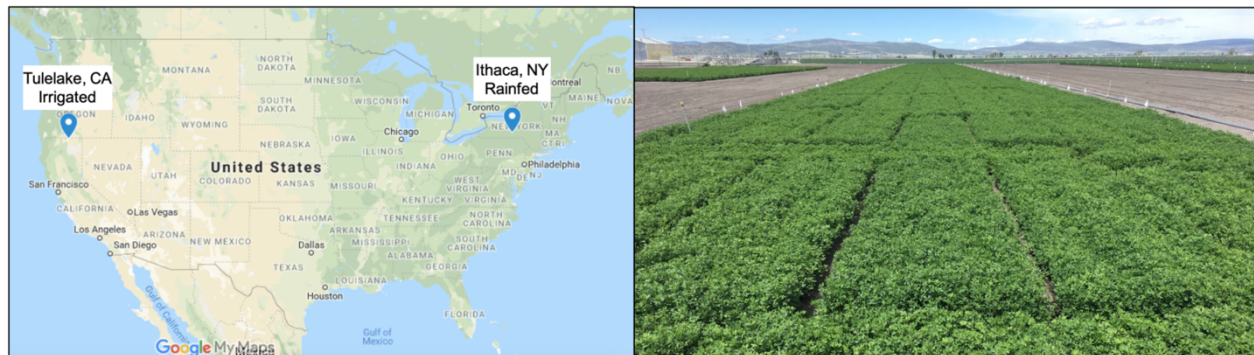

Figure 2. Locations of the genomic selection (GS) yield evaluation trials GS trials. The Tulelake trial is shown above left.

Objective 2: Develop populations and conduct yield evaluation trials.
We have developed half-sib families of four UC Davis breeding populations. Population A was planted in late 2018 in Davis in two replicates of 80 half-sib families sown into sward plots of 1.25 m x 2.5 m. Population B, consisting of 63 families, was planted in three replications into high and low salinity sites at the UC Westside Research and Extension Center near Five Points, CA in June 2019. Due to extremely limited seed, in each rep, seven plants of each entry were sown in rows with one foot between plants within rows and three feet between plots.

Finally, two closely related populations, Populations C and D, with about 100 half-sib families each, were sown on the UC Plant Sciences Farm in May 2020. These populations were intended to be sown in October 2019, but a seed production failure at our contract seed grower in Idaho resulted in us producing seen on backup populations and delaying sowing until 2020. Covid-19 delayed our planting by about six weeks. Trials were planted at two locations, one in Yolo County, which will be managed with deficit irrigation and the other in Solano County managed with full irrigation. These plots were transplanted with 24 plants spaced in four rows of six plants with 8" between plants within rows and between rows and 18" between plots side to side and 3' end to end. Yield data of Pop A were collected approximately monthly throughout 2019 and are continuing in 2020 using a self-propelled forage harvester with an electronic weigh system. Yield data of Pop B were collected for two harvests at the end of 2019 and will continue in 2020. Yield data for Pops C and D will be taken beginning in August 2020 and continue throughout 2021.

Tissue for all populations either has been collected or will be collected this summer (2020). We had anticipated using genotyping-by-sequencing to generate marker profiles for these families. However, with the rapidly changing marker landscape, we have decided to delay genotyping so that we can implement a newer amplicon-based sequencing system by LCG technologies. We are hoping this system will be both economical and rapid, and importantly, provide robust marker data for use in GS with far less missing data that often seen with GBS. We will test it during 2020.

Objective 3: Evaluate sensors to estimate yield.
In 2019, we imaged the alfalfa breeding population sown in 2018 prior to each harvest. We collected imagery with a multispectral Micasense RedEdge-M. Raw images were made into orthomosaics and digital surface models (DSMs) in Pix4Dmapper, and subsequently processed in QGIS 2.18. A grid of shapefiles was added over the plots in QGIS. In the imagery, I separated the alfalfa from soil + mowed weeds using an NDVI threshold of about 0.8. This makes a raster "classification layer" of alfalfa vs. not alfalfa. The average height and area of alfalfa in each plot was then extracted and used to calculate canopy volume of each plot. We also had a service lab at UC Davis fly plots with a drone fixed with a LiDAR system. Although we have been attempting to get data from them, we still do not have those data in hand, frustratingly. The results from our multispec camera, shown below, may be sufficient for our purposes, however.

Project Objectives and Corresponding Results:

Project Objectives	Project Results
1. Complete a second year of data collection on the existing GS selection evaluation experiment.	**1.** Genomic selection shows promise to improve yield over phenotypic selection, though further research is needed.
2. Develop, genotype, and begin phenotypic evaluation of half-sib breeding populations.	**2.** We have developed half-sib breeding populations and are collecting yield data in the field. We have delayed genotyping in order to use a better marker platform.
3. Evaluate remote and/or proximal sensors to determine biomass yield in breeding populations.	**3.** Using a multispectral camera, we have identified a strong correlation (r = 0.80) between sensor predicted volume and biomass yield that works across harvests. Confirmation is needed, but this certainly looks promising currently.

Results and Discussion:

Obj. 1. Across two production years (2018 and 2019), the high and low GS populations differed in total biomass yield in the target location, Ithaca, and averaged across locations, while the "random" population fell intermediate (Fig. 3). No differences were noted in the non-target Tulelake, CA location across two years (Fig. 8), although the differences seen in Ithaca were present in Tulelake in 2018 only (data not shown). Across the two years, the GSC1-H population produced more yield than the PSC1-H population (P=0.08) in Ithaca but not in Tulelake, the non-target environment.
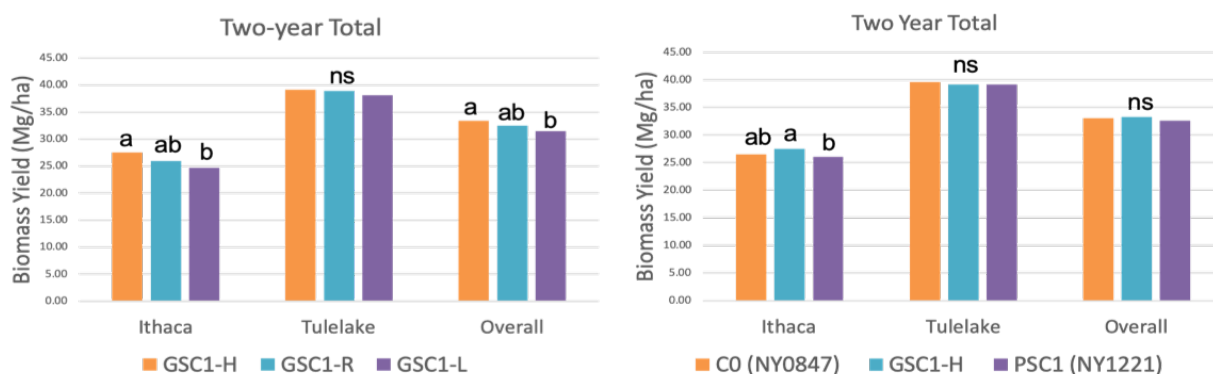

Figure 3. (Left) The performance of GS populations selected for high yield (GSC1-H), for low yield (GSC1-L), or at random (GSC1-R) at two locations across two years. (Right) The performance of GS vs. phenotypic selection compared to the original base population (C0).

A second cycle of selection was ineffective, with the GSC2-H population actually lower yielding (Fig. 4). One reason for the lack of progress in Cycle 2 could have been because of an inadvertent increase in fall dormancy (decrease in fall dormancy rating), which we documented in a separate fall dormancy trial at Davis (Table 1). Dormancy has a well-known negative
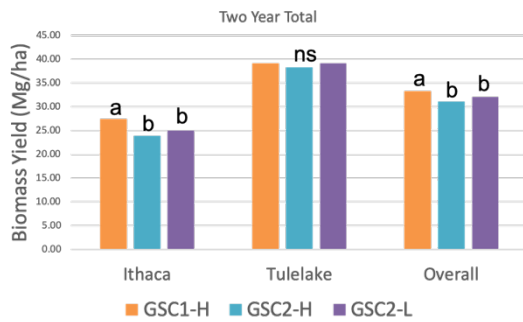
Figure 4. Yields of GS cycle 2 populations compared to GSC1-H, the progenitor of both. Fall dormancy of the C0 and C1 populations based on FD from two years in Davis.

Table 1. Fall dormancy ratings for four experimental populations evaluated in two years at Davis, CA in 2018 and 2019.

Population	Fall Dormancy Rating
C0 (NY0847)	3.2
GSC1-H	3.2
GSC2-H	1.7
GSC2-L	2.2

relationship with yield. Further, we may have seen negative effects of inbreeding depression due to small population sizes or the model might simply have broken down due to allele frequency changes resulting from selection and/or genetic drift.

The original GS model used for selection was based on phenotypic data from a clonal selection experiment, in which plants were grown as three plant plots in rows, an arrangement different from actual production conditions where plants are grown in swards as shown in the plot picture above (Fig. 7). Also, the phenotypic data represented genotypic rather than breeding values; hence, the model may not accurately represent breeding values (although we did show that predictions across cycles of selection were reasonably good (r = 0.4; Li et al., 2015).

The results of this project suggest that GS is a promising method to improve forage yield in alfalfa. Further research is necessary to optimize the system and to understand the issues with the second cycle of selection.

Obj. 2 and 3. Preliminary data on yield and plant height prediction in alfalfa
We have four yield trials on-going at the current time (June 2020) (Fig. 5). These trials all have one to three years to go before they are concluded. The yield trial of Pop A was harvested on 7/22/2019, 8/26/2019, and 10/2/2019. At each harvest period, six plant height measurements were taken per plot, the plot area was imaged using a multispec camera mounted on a drone, and yield was measured using a forage harvester (Fig. 5).

Based on the drone-based measurements of plot area and plant height, we estimated biomass yield. The fitted values have a strong relationship with biomass yield (R^2 = 0.76; Fig. 6). More interestingly, we used developed a prediction equation from the October harvest and used that to predict yield in July and August harvests. The sum of predicted July, predicted August, and actual October yield had a very strong correlation to the sum of actual yields across the three harvests (Fig. 6, right). This is the most important result, suggesting we can skip harvesting altogether some harvests and predict the yield based on sensor data. By cutting down the number of plots that need to be harvested throughout the year, we can increase plot numbers and decrease time needed to harvest.

Figure 5. Yield trials in Davis (2018, left), Five Points (2019, center top), and Davis (2020, bottom center). Sward trial being imaged using a multispec camera mounted on a drone (top right) and harvested for biomass (bottom right) at Davis in 2019.
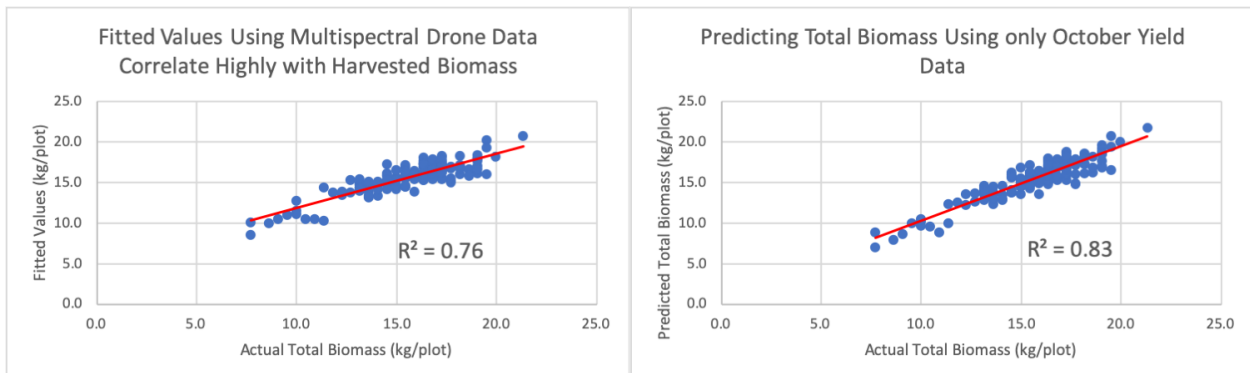

Figure 6. Relationships between predicted yields and actual yields in Davis, 2019.

Future research.
Obviously the experiments in objectives 2 and 3 are continuing. We will collect yield and sensor data throughout 2020 on all trials; this will be funded through the current 2020 NAFA grant. Future research will hopefully be funded through a pending NIFA-AFRI grant proposal that will cover genomic selection and sensor evaluation on these plots after 2020. Selections will be made from all trials based on the yield performance and used to develop new populations with (hopefully) higher yield.

References:

Brummer, E.C. and M.D. Casler. 2014. Cool-season forages. In S. Smith, B. Diers, J. Specht, and B. Carver (eds.). Genetic gain in major U.S. field crops. CSSA Spec. Publ. 33. ASA, CSSA, and SSSA. Madison, WI. doi:0.2135/cssaspecpub33.c3.

Li, X., Y. Wei, A. Acharya, J.L. Hansen, J.L. Crawford, D.R. Viands, R. Michaud, A. Claessens, and E.C. Brummer. 2015. Genomic prediction of biomass yield in two selection cycles of a tetraploid alfalfa breeding population. Plant Genome 8. doi:10.3835/plantgenome2014.12.0090.

Olmstead J., Brummer E.C. (2008) Benefits and barriers to perennial forage crops in Iowa corn and soybean rotations. Renewable Agric. Food Syst. 23:97-107.

Keywords:

Alfalfa, sensors, biomass yield, genetic markers, genomic selection