Evaluating Approaches To High-Throughput Phenotyping and Genotyping for Genomic Selection in Alfalfa

Kelly Robbins, Cornell University

OBJECTIVES & MOTIVATION

- Develop a genotyping approach to estimate genetic relationships between alfalfa populations, which are not genetically distinct individuals.
  - Population-level sequence-based genotyping to estimate population allele frequencies.
  - Use pairwise F_{st} to estimate additive covariance between populations.
- Incorporate aerial high-throughput phenotyping to predict performance and genetic merit of breeding materials.
  - Predict population performance using genetic relationships and vegetative indices.
  - Longitudinal random regression to estimate genotype-specific growth curves.

RESULTS

- F_{st} increased accuracy over dominant markers.
- Variable prediction with single NDVI time points.
- Growth curves incorporate many time points, highly related to forage yield and quality.

STUDY DESIGN & MATERIALS

- Diallel of 9 alfalfa germplasm sources (Segovia-Lerma et al. 2004).
  - 9 parental, 36 hybrid populations.
  - Forage yield 1997, 1998.
- Eight Cornell varieties and breeding populations.
  - Imaged (NDVI) every ~ 4:3 days.
  - 4 harvests across 2019 and 2020.

Table 1. Leave one out genomic prediction accuracy for genetic covariances estimated by pairwise F_{st} or covariance of allele frequencies (covAF) of eight Cornell populations evaluated in Geneva, NY.

Harvest	Year	F_{st}	covAF
2	2019	0.27	0.43
3	2019	0.75	0.73
1	2020	0.51	0.24
2	2020	0.94	0.97
Sum of 4 Harvests	-	0.90	0.79

Figure 1. Prediction accuracy using a leave one family out strategy for a diallel population with 9 parental populations, and 36 hybrid populations of alfalfa. For each of the nine parents, all entries with that parent were removed and predicted using the remaining eight families and the additive genetic covariance estimated using pedigrees, dominant markers (1544 AFLPs; Segovia-Lerma et al. 2003), or F_{st} statistics calculated from variant frequencies determined by whole-genome resequencing.

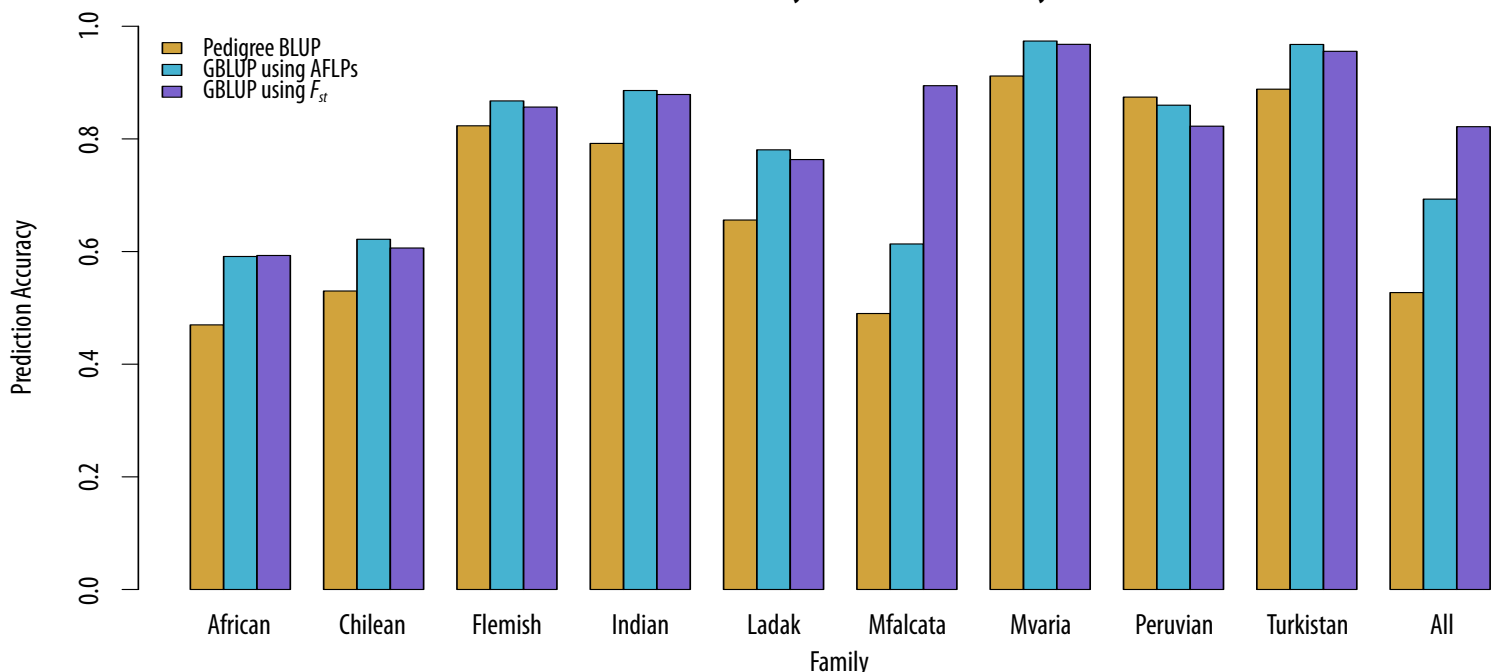Leave One Family Out Prediction Accuracy

Figure 2. Genetic growth curve deviations are the genetic differences from a mean genetic curve estimated in a longitudinal random regression using Legendre polynomials. Genetic growth curves have the mean growth curve added back in. The area under the genetic growth curve estimates the accumulation and partitioning of photosynthate to forage yield (FY) in two harvests in 2019 and 2020. Blue = high FY, Red = low FY.
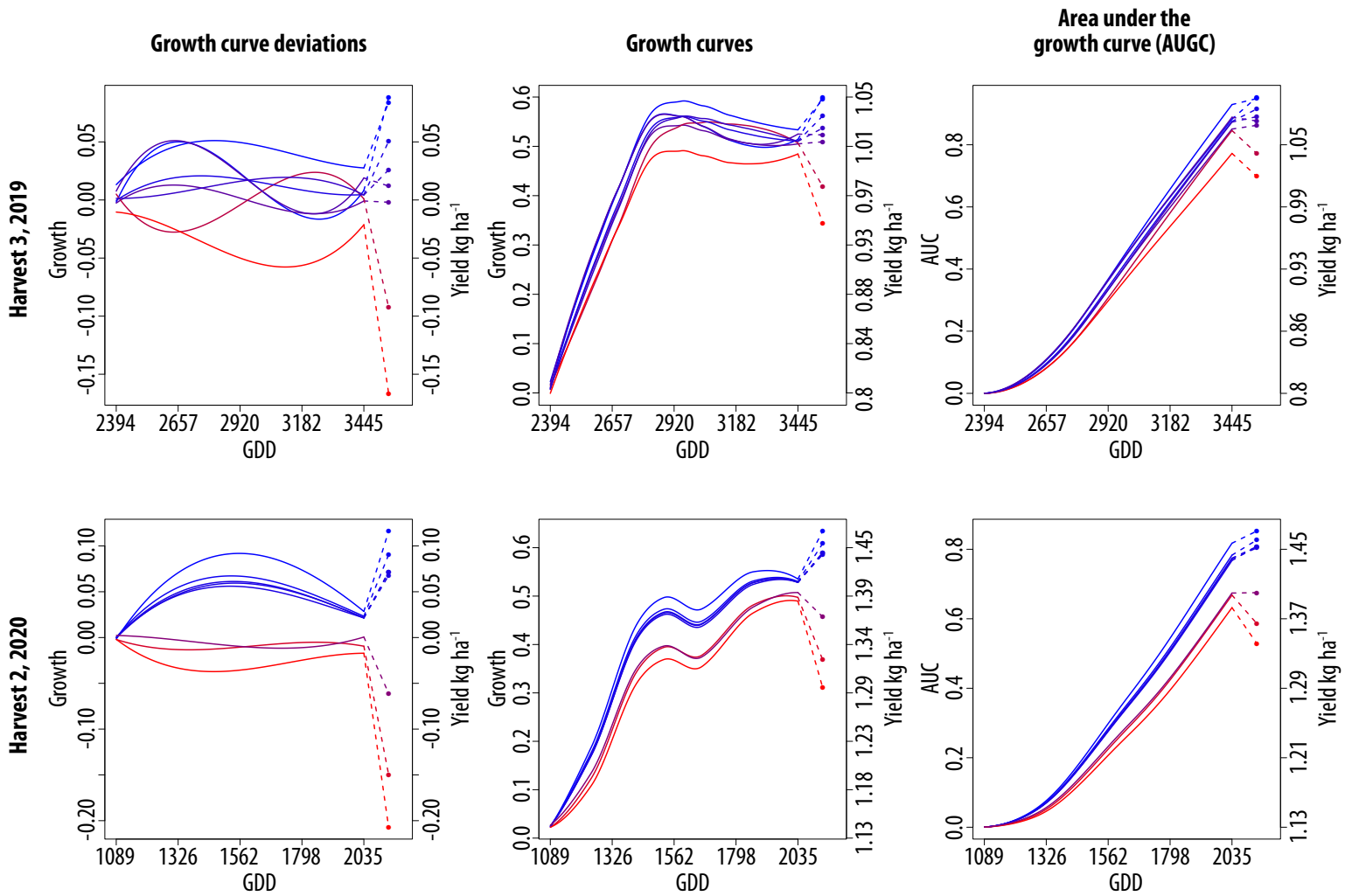
Table 2. Genetic correlations of Legendre polynomial parameters, L0, L1, L2, and L3 with forage yield (FY) in two harvests in 2019 and 2020. Intercepts (L0) were high correlated to forage yield (FY), while quadratic terms were highly negatively correlated with FY.

Forage Yield Harvest 3, 2019				
	L0	L1	L2	L3	FY
L0		0.04	-0.91	0.12	0.81
L1			-0.04	-0.98	-0.15
L2				-0.11	-0.95
L3					0.28

Forage Yield Harvest 2, 2020				
	L0	L1	L2	L3	FY
L0		0.54	-0.99	0.72	0.95
L1			-0.55	-0.09	0.36
L2				-0.66	-0.91
L3					0.87

CONCLUSIONS

- Pairwise F_{st} values serve as efficient estimates of genetic relatedness between populations.

- Genetic correlations between forage yield and vegetative indices are high, especially in first half of regrowth period.

- Vegetative indices were predictive of forage yield and quality, but including genetic co-variance was more important.

- Early growth tends to lead to higher forage yields, but with lower quality. Quality reduction likely related to maturity.

Data Availibility:
Images processed and stored on Imagebreed online database (Morales et al. 2020).

Alfalfa
NAFA
U.S. Alfalfa Farmer Research Initiative