Selection for Yield and Seed Composition Traits with Lines from a Recombinant Inbred Line Population

Jay Gillenwater

9/7/2021

Core ideas here.

Author list + affiliations

Abbreviations (alphabetical order: Abbreviation, definition; next abbreviation, next definition; …)

CAS, Caswell research farm; CLA, central crops research station; PLY, Tidewater Research Station; Pop 201, oil mapping population 201; Pop 202, oil mapping population 202; RIL, recombinant inbred line; SDWT, one hundred seed weight; SSD, single seed descent;

# Abstract

# Introduction

* Importance of yield.
* Importance of oil, protein coupled with oil. Stats on worldwide meal consumption, value contributed by high protein.
* Difficulty in simultaneously increasing yield, protein.
* Need for diverse germplasm with high protein, yield, good agronomic qualities.

# Materials and Methods

## Phenotypic data collection

* List each relevant phenotype and how it was measured
  + Yield, protein/oil, lodging, agronomic score, seed weight.

Genotypes were evaluated for both agronomic and seed composition traits. Agronomic data collected included seed yield, hundred seed weight (SDWT), lodging, height, maturity date, and a composite agronomic score. The agronomic score was a visual metric which sought to quantify traits not directly covered by the other phenotypes like pod load and plot uniformity to provide an estimate of a genotypes general desirability. Height was measured in centimeters from the soil to the top of the plant. Lodging was scored on a scale of 1 to 5 where a score of 1 indicates that all plants are erect and a score of 5 indicates that all plants in a plot are on the ground(I. Fehr, 1987). Agronomic score was also measured on a scale of 1 to 5 where a score of 1 indicates lines that are the most desirable in a population while a score of 5 indicates that a genotype is among the least desirable. Maturity was recorded at the R8 maturity date(W. R. Fehr & Caviness, 1977). Following harvest seed yield, SDWT, seed oil content, and seed protein content were measured after seed was allowed to dry to approximately 7% moisture content in a greenhouse. Seed yield and SDWT were measured in grams after seed had been sifted to remove debris and cracked seed. Seed protein and oil contents were measured with a Perten DA 7250 NIR instrument®.

## Pedigree

Lines used in this study were selected from two F5-derived recombinant inbred line populations that were segregating for seed protein and seed oil contents, denoted as populations 201 and 202 (Pop 201 and Pop 202). The first population, population 201 consisted of 273 recombinant inbred lines (RILs) derived from a cross between the high protein breeding line LMN09-119 and the high oil breeding line N09-09 in 2015. Population 202 consisted of 197 RILs derived from a cross between the high protein breeding line LMN09-19 and the high oil breeding line N13-47, also made in 2015. RILs were advanced to the F5 generation by single seed descent (SSD) at which point a single plant was harvested to represent each F5 derived RIL.

## Population development

* Criteria for selection from the starting RIL populations.
  + Two trials per RIL population.
* How were lines selected for progression to the 2020 season.
  + Two trials in total.
* In 2019, 80 RILs with the highest bulk weight, seed protein, and seed oil content with non-extreme maturity dates were selected from Pop 201 and Pop 202. The selections were used to make two yield trials per population, named Test 1 - Test 4 where Test 1 and Test 2 consisted of RILs from Pop 201 and tests 3 and 4 consisted of RILs from Pop 202.

## Plot techniques

* Plot techniques: How many rows, row length, spacing, seeding rate (from lead sheets)
* How data was converted between three-row and four-row trials

## Statistical analysis

* Statistical model used to analyze data, fixed vs random effects what software was used.
  + Model used.
  + ANOVA results for model term significance (mainly genotype) both within environment and in all environments.
  + Software used to fit the model
  + How were genotype means compared (LSD?)

# Results and Discussion

* Results of the ANOVA analysis
  + Were G and GE effects significant for traits in the individual and combined environments?
  + Present ANOVA tables for the combined analysis for each trait in main manuscript body (space allowing) and the individual environments in the supplementary section
* Pearson correlation coefficients between phenotype LSMeans for individual and combined analyses.
* Report agronomic performance of well performing lines.
  + Well performing lines are those with yield within a LSD of the check average.
  + Report yield, maturity, lodging, heinght, seed weight, seed protein, seed oil, and meal protein.
    - Meal protein maybe from the SPROC program(?) **look into how to calculate this**

# Acknowledgements

# References

Fehr, I. (1987). *Soybeans: Improvement, production and uses,”* [Journal Article].

Fehr, W. R., & Caviness, C. E. (1977). *Stages of soybean development* [Journal Article].