

I. MSMC project title

MSMC 15-380: Genetic Mapping of a Unique Morphological Trait in Soybean and Evaluation of the Correlations with Yield Potential and Seed Composition

II. Time period covered

January 1st, 2017 – December 31st, 2017

III. Project leaders and co-primary investigators

PI: Andrew Scaboo, University of Missouri

Co-PI: Jason Gillman, USDA-ARS

IV. Layman's summary (limit one page)

The soybean breeding program at the University of Missouri grew 10 F₂ hybrid populations derived from 10 unique *Glycine max* x *Glycine soja* (wild soybean) bi-parental crosses in Columbia, MO in 2012. A single F₂ plant, among over 100,000 different plants, exhibited a unique growth habit. This plant stood upright, with productive lateral branches upright and equal to the entire length of the main stem. The F₃ seed were harvested from this single plant during the fall of 2012, and the F₃ plants were grown in Columbia, MO during the summer of 2013. During the growing season of 2013, some of the F₃ plants exhibited immediate and substantial growth of the lateral branches at the cotyledon node (the first node on the plant) and these branches grew at the same rate as the main stem throughout the lifecycle of the plants. Approximately 300 single F₃ plants were harvested during the fall of 2013 and F_{3:4} plant rows were grown in Columbia, MO during the summer of 2014. The unique cotyledon node branching pattern was again observed in some of the F₄ plants, suggesting this trait is stable, heritable, and potentially largely genetically controlled. Furthermore, this branching pattern was also observed in the *Glycine soja* plant introduction used as a parent to develop the initial hybrid population.

To understand the genetic control of this trait we developed an F₄-derived population in 2014, increased the seed in 2015, and developed a multi-location field experiment for 2016 and 2017. The field experiment has been largely successful over the last two growing seasons, although we discuss only 2016 results in this report. We will complete the entire analysis, including 2017 data, during the spring of 2018.

Using the data we have collected thus far, we have identified two genomic regions that are associated with average and total branch length (one location), two regions that are associated with seed protein (across 2 locations), two regions that are associated with seed oil (across 2 locations), one region associated with plant height (across 2 locations), and one region associated with seed yield (across 2 locations). Our results are novel for the unique branching trait and seed yield, and in agreement with previous studies for

the other agronomic and seed composition traits we measured in this study. Although we are currently unable to identify specific genes controlling these traits, we can use this information to develop new plant material and experiments to eventually understand the genetic architecture of these traits.

We also evaluated the relationship (phenotypic correlation) between all of the traits we measured in this experiment including the unique branching trait, seed protein and oil, seed yield, lodging, plant height, and maturity. Surprisingly, we found no correlation between the unique branching pattern and seed yield. This indicates that selecting for this trait, as we have identified and measured it, as a means to increase yield potential will likely not be effective in a breeding program. We were successful in finding positive correlations between the unique branching trait and plant height, maturity, and lodging. These results suggest that the unique branching trait is much more complex than we originally thought, and arises from numerous genetic and environmental factors that contribute to this unique branching pattern in the population we evaluated.

V. State your objectives in question form and discuss how your results answer these objectives

The objectives of this project were to answer two questions:

- Can we identify the genomic region(s) controlling the unique branching trait derived from wild soybean?
- What is the yield potential, seed composition, and correlation between traits in experimental lines exhibiting the unique branching trait?

Can we identify the genomic region(s) controlling the unique branching trait derived from wild soybean?

We found two genomic regions, or quantitative trait loci (QTL), in 2016 at Albany, MO that were significantly associated with the unique branching trait we have observed and evaluated in this study. One QTL was found on chromosome 14 to be significantly associated with average branch length, and one QTL on chromosome 12 was found to be associated with total branch length (Table 1). We are currently collecting and analyzing the data from two locations in 2017 to confirm our 2016 results. This will be complete by the spring of 2018. Additionally, we identified several genomic regions associated with seed protein and oil, maturity, plant height, lodging, and seed yield (Table 1).

What is the yield potential, seed composition, and correlation between traits in experimental lines exhibiting the unique branching trait?

The yield potential of lines in this population was relatively low compared to elite soybean varieties grown by farmers, although at the 2016 Novelty, MO location we found one experimental line that yielded significantly higher (71.25 bushels/arc) than

the yield of Osage and Ellis (44.65 bushels/arc and 66.1 bushels/arc, respectively), and similar to the yield of IA4005 (72 bushels/arc). At the Albany, MO location in the same year, the yield of this experimental line (70.4 bushels/arc) was similar to that of Osage (70.4 bushels/arc). This population, which was genetically 50% wild soybean, had a lower range of oil and higher range of protein (136.5-184.2 g kg⁻¹ and 391.3-481.1 g kg⁻¹, respectively) than those in the elite soybean cultivar checks (175.9-216.4 g kg⁻¹ and 371.5-205.5 g kg⁻¹, respectively).

Table 1. Identification of genomic regions (QTL) associated with traits measured in this study during 2016 at two locations in Missouri

	Chr#	Position (cM)	LOD score	% variance explained	P-value	QTL interval (cM)	AA>BB effect*
Albany, MO							
Avg. Branch length (cm)	14	161.0	5.86	10.46	1.38E-06	156.0-165.0	-4.2
Height (cm)	5	7.8	6.47	10.59	3.39E-07	2.0-13.0	8.2
Height (cm)	11	63.6	5.39	9.54	4.07E-06	60.3-69.0	-7.1
Height (cm)	16	9.3	7.27	15.48	5.37E-08	7.0-12.2	9.0
Oil (%)	8	36.3	11.8	14.45	1.58E-12	34.1-41.0	0.6
Oil (%)	20	70.3	19.1	29.36	7.94E-20	69.3-72.6	0.8
Protein (%)	14	60.3	13.9	14.98	1.26E-14	55.8-61.8	-1.2
Protein (%)	20	72.0	30.9	42.43	1.26E-31	69.5-75.4	-2.0
Total branch length (cm)	12	0	5.51	10.35	3.09E-06	0.0-6.0	-48.2
Yield	2	145.8	5.99	8.42	1.02E-06	143.0-148.7	7.1
Yield	8	27.2	5.86	14.05	1.38E-06	22.0-32.3	8.3
Yield	16	122.8	8.44	13.38	3.63E-09	117.0-123.4	8.9
Novelty, MO							
Height (cm)	5	20.0	5.33	5.05	4.68E-06	12.5-21.0	4.4
Height (cm)	12	21.2	7.98	14.72	1.04E-08	16.0-24.7	-8.6
Oil (%)	8	38.9	14.5	17.20	3.16E-15	37.0-41.0	0.7
Oil (%)	20	71.5	29.7	38.60	2.00E-30	69.5-72.6	1.1
Protein (%)	10	107.5	5.36	4.39	4.36E-06	103.7-110.7	-0.7
Protein (%)	14	59.0	9.37	10.80	4.27E-10	55.0-61.8	-1.0
Protein (%)	20	70.3	35.18	47.85	6.61E-36	68.2-75.0	-2.3
Yield	6	23.1	8.14	15.73	7.24E-09	20.0-23.8	5.9
Yield	8	39.1	6.52	12.61	3.02E-07	37.0-44.0	6.8
Yield	18	91.1	5.86	9.14	1.38E-06	88.9-93.4	-5.8

Although we were fully successful in carrying out our field and genetic experiments, we unexpectedly found no correlation between the unique branching pattern and seed yield. This indicates that selecting for increased branching, as we have identified and measured it, as a means to increase yield will likely not be an effective strategy in a breeding program. We found positive correlations between the unique branching trait and plant height, maturity, and lodging. These results suggest that the unique branching trait is much more complex than we originally thought, and that many genetic and environmental factors contribute to this unique branching pattern.

VI. Please answer the following:

A. How do your results benefit Missouri soybeans growers?

We found QTL's with positive allelic variation and effect in *Glycine soja* and expect to leverage this information to improve the overall yield potential and seed composition of the cultivated soybean. In addition, the utilization of wild soybean in breeding programs will help to improve the poor genetic diversity of soybean cultivars in North American, and prevent future yield limitations due to lack of genetic diversity.

B. Estimate financial return for the average Missouri soybean producer.

The financial return for Missouri farmers from this project will be in the long-term increases in profitability and productivity due to improved soybean genetic potential.

C. Do your results benefit the environment?

Our results do not directly benefit the environment; however, increasing the yield potential of soybean would lead to a reduced environmental footprint per unit area of soybean production, due to increased yield potential.

D. What products or processes can be commercialized from this research?

Two main avenues can commercialize genetic gains realized as soybean germplasm and varieties: a) licensing to private companies for sale as varieties and b) licensing to private companies for use in their respective breeding programs.

1. List disclosure(s) of inventions or plant varieties submitted to the MU Tech Transfer Office

None at present.

2. Identify potential disclosure(s) of inventions or plant varieties.

Soybean germplasm from this project has the potential to be developed as disclosed inventions in the future, as well as parents for future inventions/plant varieties.

E. How would you commercialize these products or processes?

Two main avenues can commercialize soybean germplasm and varieties: a) licensing to private companies for sale as varieties and b) licensing to private companies for use in their respective breeding programs.

F. If no specific products or processes were produced, how do you plan to make your results available to producers or industry?

The results of our research will be released in the form of peer-reviewed publications, as well as communicated in the form of presentations at scientific and farming related conferences.

G. Is additional time or research required before your results can be used by producers and industry?

Yes, some of our 2017 analyses are still in process, and can be fully completed with the funds already utilized. Moreover, we plan to continue this research with a new proposal that expands the research on the seed composition specific portions of the project, for which we were most successful in terms of identifying genetic architecture.

VII. List publications by type (popular press, thesis, journals, other) written or planned.

We plan to publish the results of this project in a scientific journal during 2018.

VIII. List cost of original project and actual expenditures. *The U.S. Department of Agriculture requires that we ask for budget information, including the number of hours spent on the project, the number of dollars remaining on*

account, as well as a breakdown of expenses. You are required to provide this information in your report. Please also include names and titles/positions of those whose time has been charged to this project.

Scaboo

As of January 4th, 2018, the account balance is \$471.21. \$19,906.84 has been spent on salary and wages, \$46,500.00 is subcontracted to USDA-ARS, \$9,833.95 has been spent on operating expenses, and \$1000.00 has been spent on land rent and equipment maintenance. The encumbered dollar amount is \$288.00. Thang La is the PhD Graduate Research Assistant working on this project and part of his assistantship was paid from this project.

Gillman

As of January 4th, 2018, the current account balance is \$0. \$19,096.67 has been spent on salaries and operating expenses and \$27,403.33 is encumbered. Two part-time student workers (Research Assistant at < 8 hours/week) are supported.

IX. List equipment purchased with MSMC funds, identifying inventory, and serial number.

None