

# Identification of QTL responsible for an economically valuable soybean seed oil trait using next-generation sequencing methods

Crystal Heim<sup>1</sup>, Jason D. Gillman<sup>2</sup>  
<sup>1</sup>University of Missouri-Columbia, Division of Plant Sciences  
 crystal.buerke@gmail.com  
<sup>2</sup>USDA-ARS, Midwest Area, Columbia, MO 65211

## Introduction

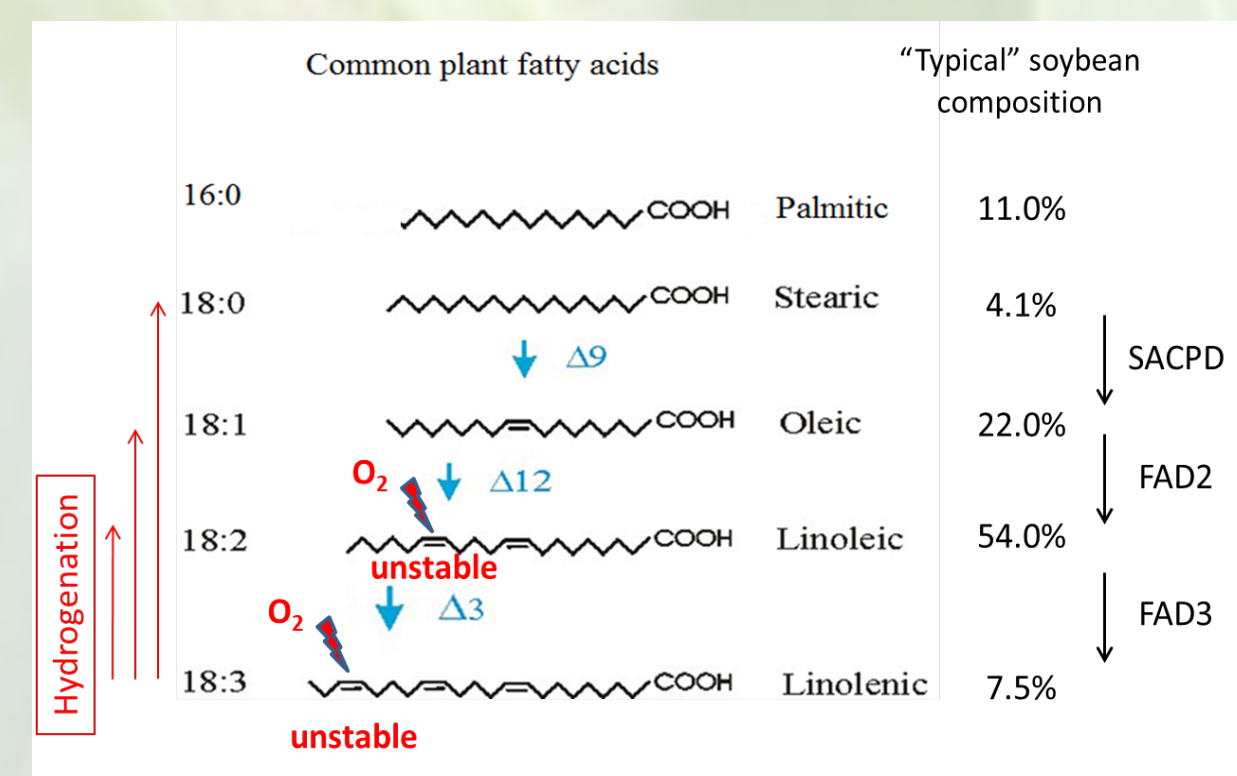


Figure 1. Soybean fatty acids: stability, typical composition

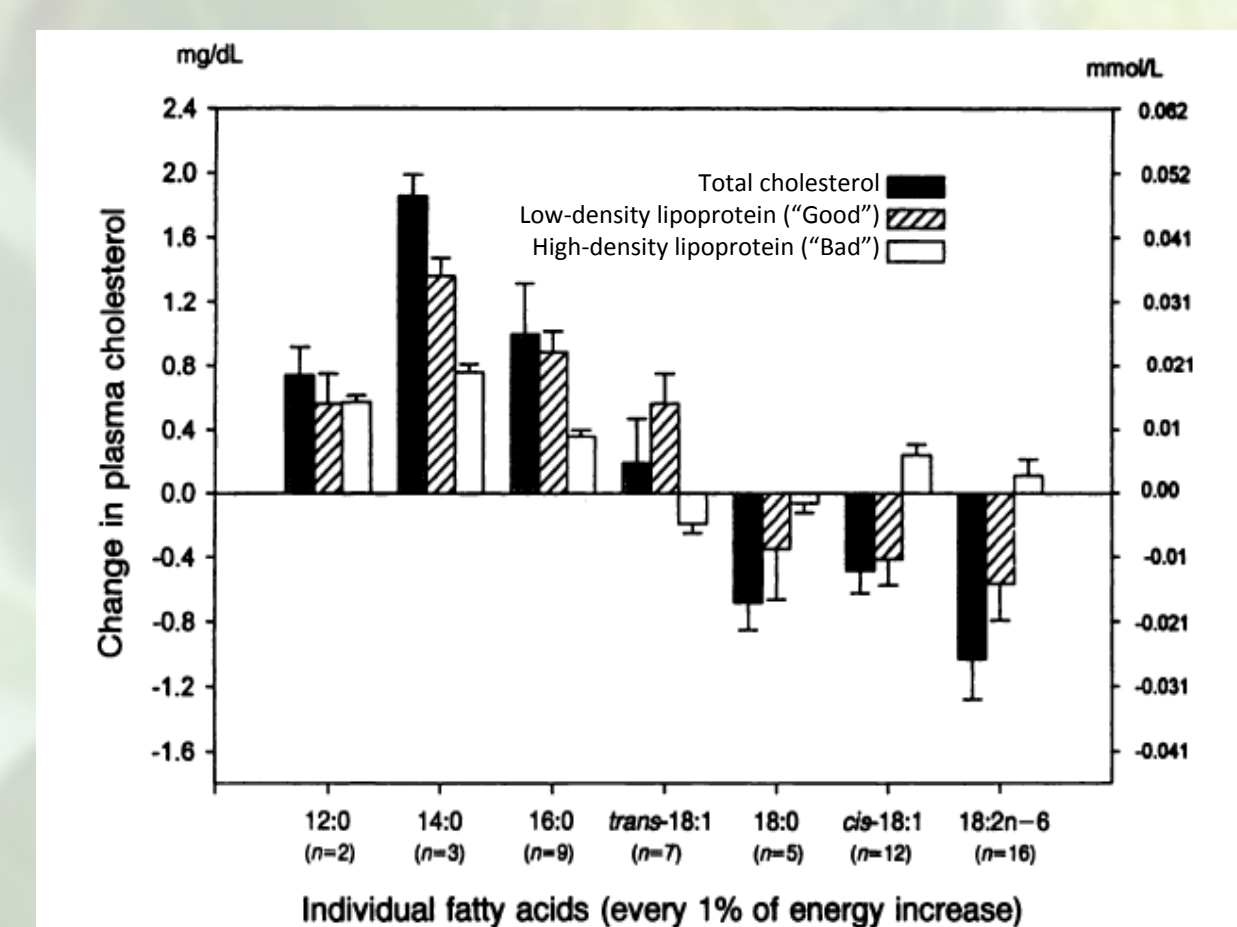


Figure 2. Individual fatty acids and their effects on cholesterol levels

## Stearic Acid

- Unsaturated fatty acids, such as linolenic (C18:3) and linoleic (C18:2), are oxidatively unstable, and tend to go rancid rapidly
  - These fatty acids must be hydrogenated to increase stability (Figure 1)
- The hydrogenation process results in the formation of *trans* fatty acids, which are tightly linked with heart problems in humans and "not generally regarded as safe" by the USDA
- Stearic acid (C18:0) is a long chain fully hydrogenated fatty acid that has been shown to have no negative effect on LDL levels (unlike other saturated fatty acids) (Figure 2)
- Soybean oil with elevated saturated fatty acid content is desirable for solid fat baking applications, and could potentially eliminate the need for expensive solid fats like cocoa butter and heart-unhealthy fats such as palm oil

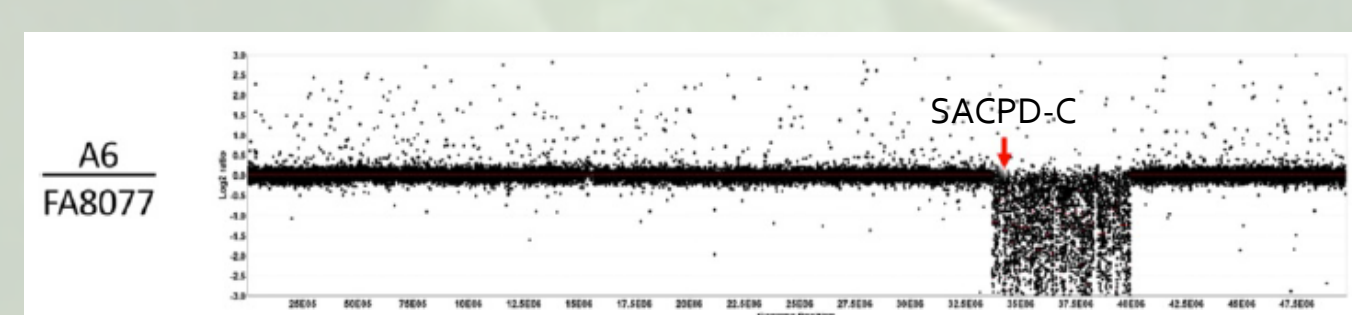


Figure 3. Comparative Genome Hybridization of A6 and its progenitor

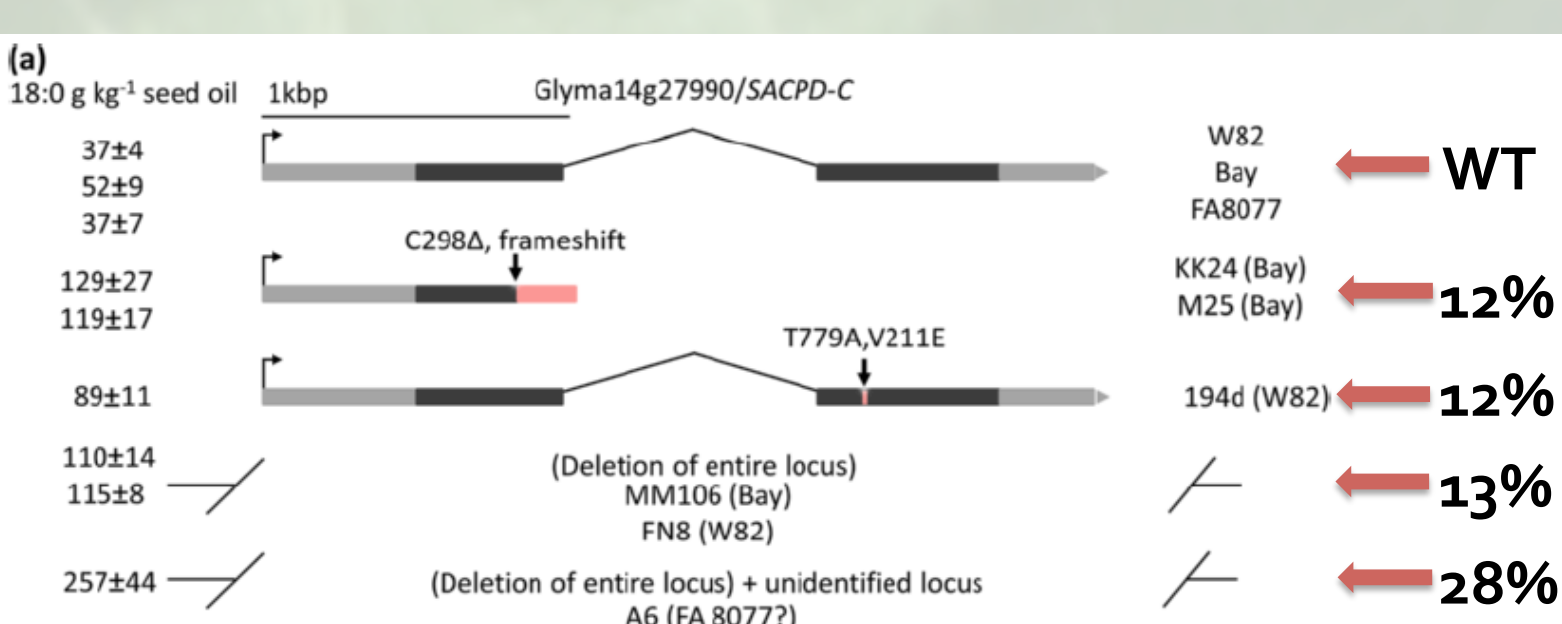


Figure 4. SACP-D in W82 and several stearic acid mutants and their % stearic acid

Line	Fatty acid					
	Palmitic	Stearic	Oleic	Linoleic	Linolenic	Arachidic
A6	8.0	28.1	19.8	35.5	6.6	2.0
FA 8077	8.4	4.4	42.8	36.7	7.6	<1.0

\*the exact provenance of A6 is contested

## A6 and 194D

- A6 is an M<sub>2</sub> plant selection from a sodium azide<sup>1</sup> induced mutant population with up to 30% stearic acid (Table 1)
  - Characterized by several large deletions, most notably a ~1/8 chromosome deletion on Gm14 (Figure 3)
    - Predicted to encompass 30 genes from the Glyma 1.0 high confidence gene set, including SACP-D-C
    - Exhibits poor germination and severe field defects such as reduced height and yield
- 194D is an EMS-induced mutant with moderate levels of stearic acid (~12%)
  - Due to a point mutation (V211E) in SACP-D (Figure 4)
- Zhang et al (2008) determined that the elevated stearic acid trait in soybean is due largely to alterations in SACP-D-C, a stearyl-acyl carrier protein-desaturase gene, responsible for hydrogenation of oleic acid to stearic acid

## Materials and Methods

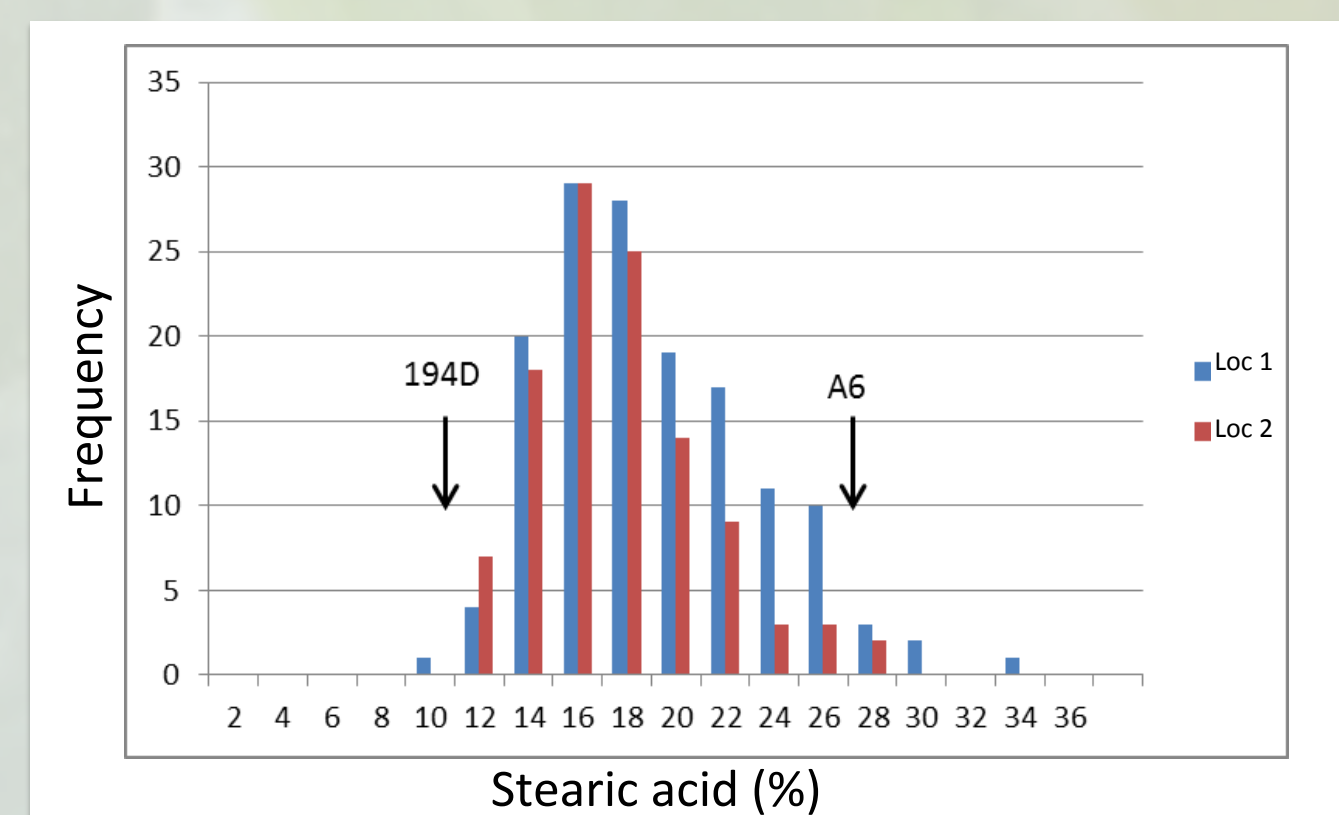


Figure 5. Distribution of stearic acid in RIL mapping population across two locations

**Formation of a RIL mapping population**  
 High stearic line A6 (~28%) × mid-stearic line 194D (~13%)  
 Grown in RCB design in three reps across two locations  
 Carry through to F<sub>5</sub>

**Phenotype F<sub>5</sub> seed for fatty acid composition using gas chromatography**  
 5 individuals seeds from each line

**Genotype population using Genotyping-by-Sequencing (GBS)**  
 176 individuals from F<sub>5</sub> population

**Generate a linkage map using R/qtl**

**Map reads to Glyma v2.0 and filter SNPs**  
 27,672 SNPs detected in TASSEL  
 Removed non-polymorphic and severely distorted markers  
 2,978 SNP markers useful for final analysis

**Perform a one-dimensional scan for QTL**  
 Using Haley-Knott regression in R/qtl  
 LOD thresholds determined by 1000 permutations

**Perform a two-dimensional scan to identify any interaction between QTL and any small-effect QTL**

**Determine the final model**

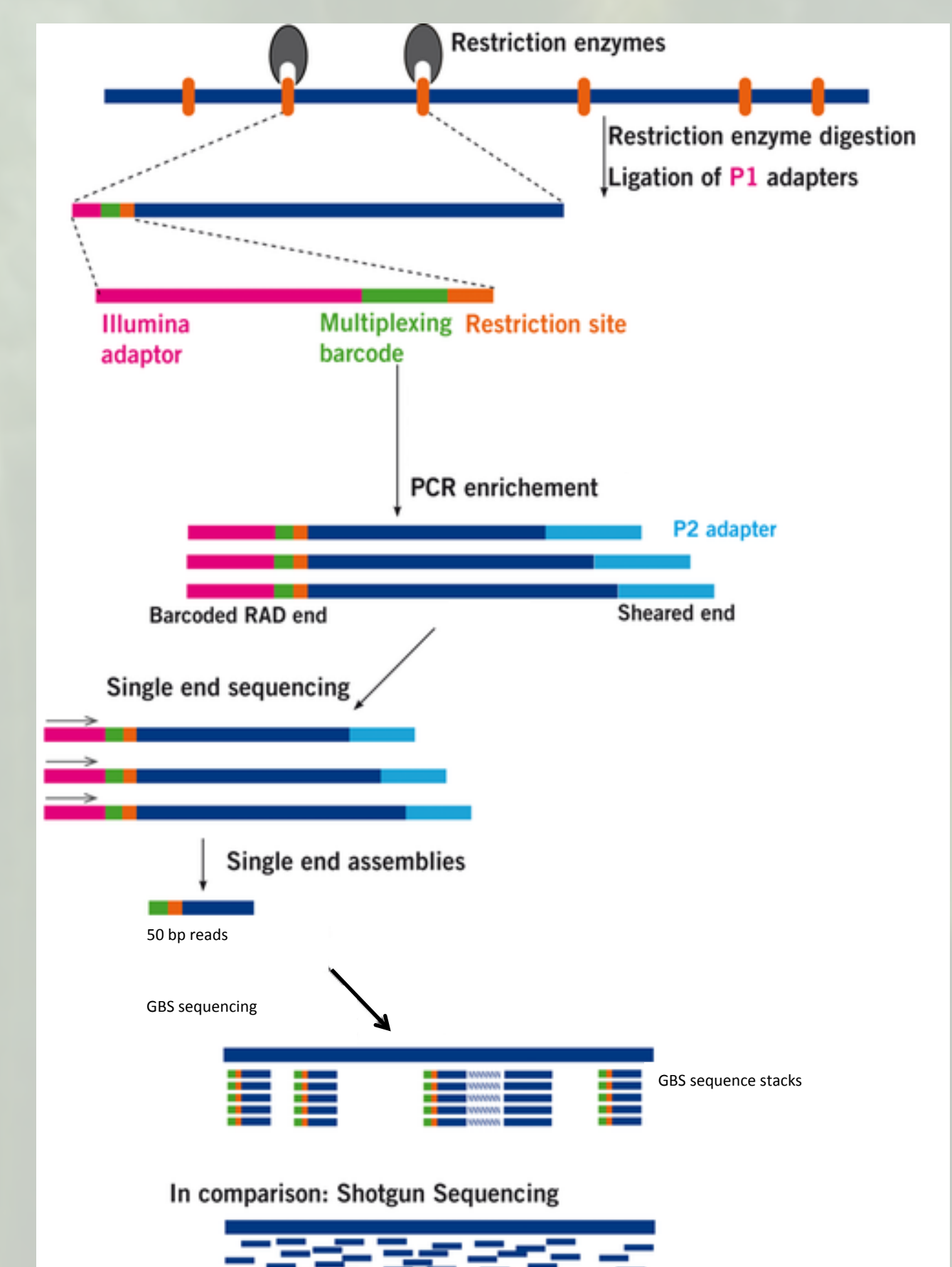


Figure 6. GBS workflow. Modified from <http://www.floragenex.com/rad-seq/>

## Results and Discussion

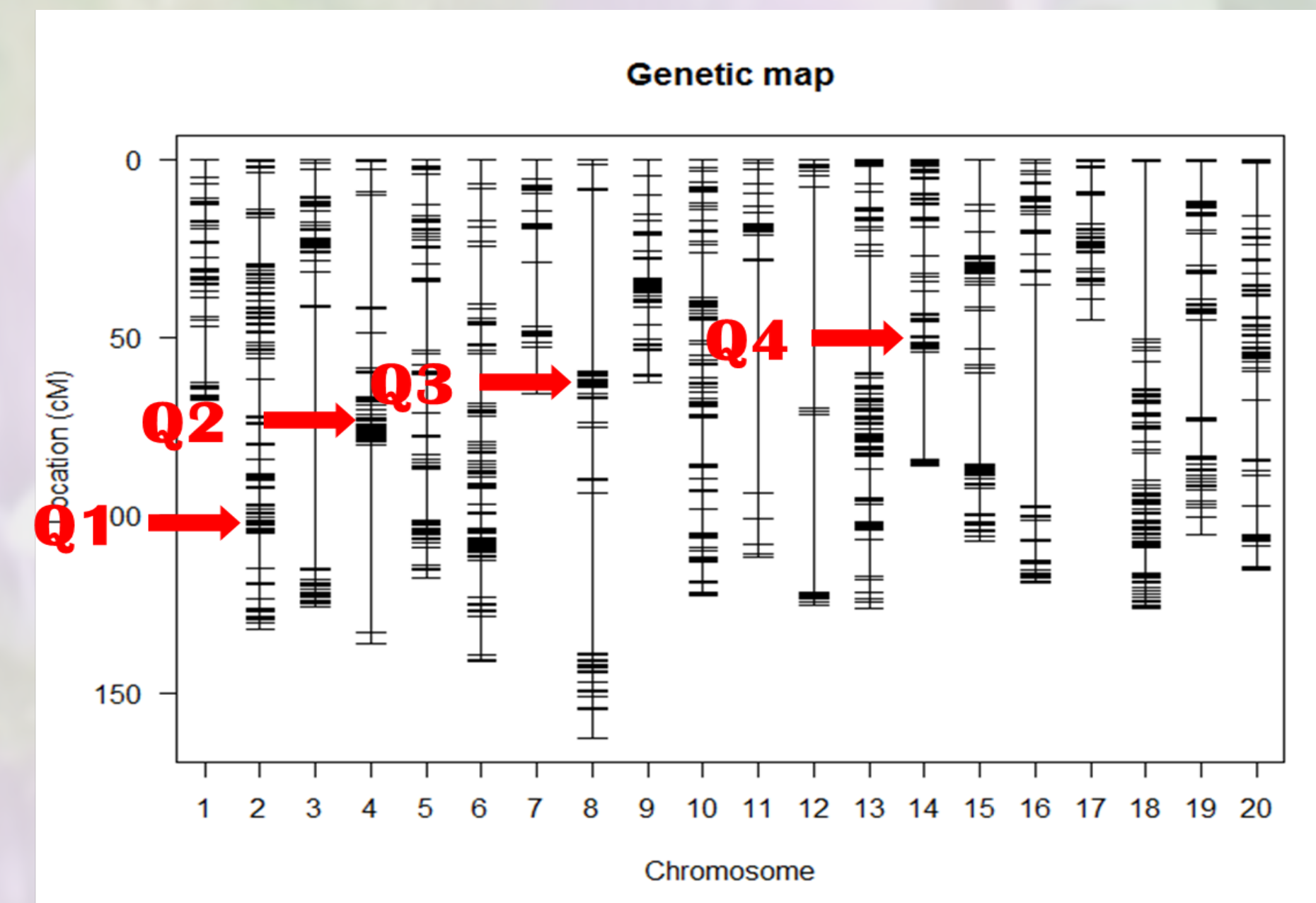


Figure 7. Linkage map constructed in R/qtl with four QTL

## QTL and LOD Scores

- QTL were detected on four different linkage groups (Figure 7)
  - These four QTL act additively to determine stearic acid content
- There is also evidence for interaction between Q3 and Q4, which explains ~8% of the variation in the trait in our population (Figure 10 and Table 2)
- One major QTL stands out (Q4), which explains roughly 59% of the variation in stearic acid content in our population (Table 2)
  - This is due to the SACP-D-C gene, which, as previously stated, is located on Gm14 (Figure 4)
- Final model:
  - Y<sub>ij</sub> = Q1 + Q2 + Q3 + Q4 + Q3:Q4

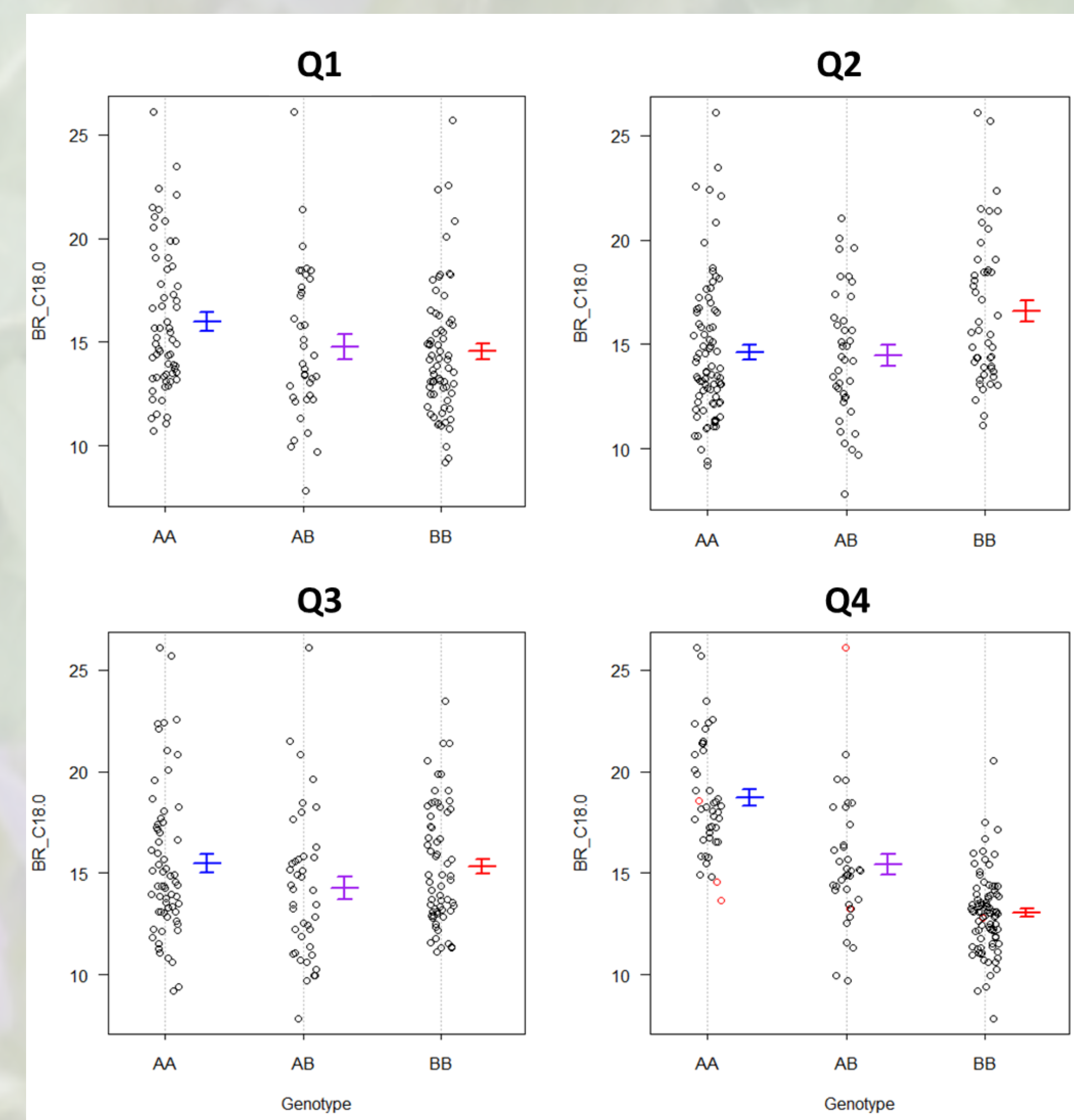


Figure 9. Phenotype vs. genotype values for all four QTL

A=allele from A6 B=allele from 194D

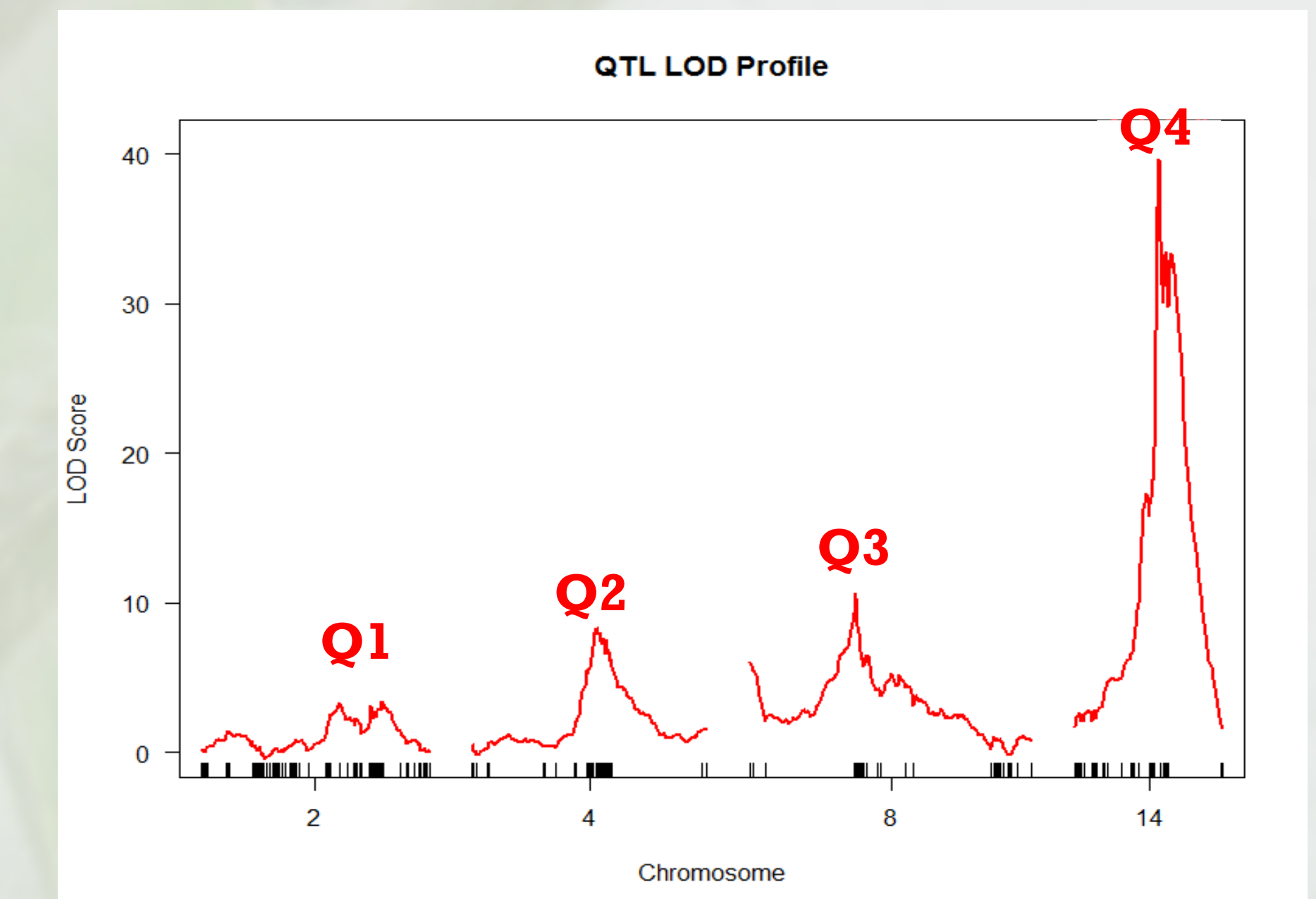


Figure 8. Four QTL and their respective LOD profiles

## Discussion

- Seed stearic acid in soybean is a complex trait
- Two of our QTL, Q1, and Q4, are in agreement with previous studies done on this trait
- Due to the very large effect of the QTL on Gm14 (Q4), it may be useful to analyze the data with Q4 as a covariate
- We hypothesize that the highly elevated stearic acid phenotype as well as yield defects seen in A6 may be due to genes encompassed by the 1/8 chromosome deletion on Gm14
  - Resequencing A6 will further elucidate the effects of this large deletion
- Further study of this economically valuable trait will be useful in restoring the soybean oil market lost to palm oil production and in replacing this heart-unhealthy fat in the human diet

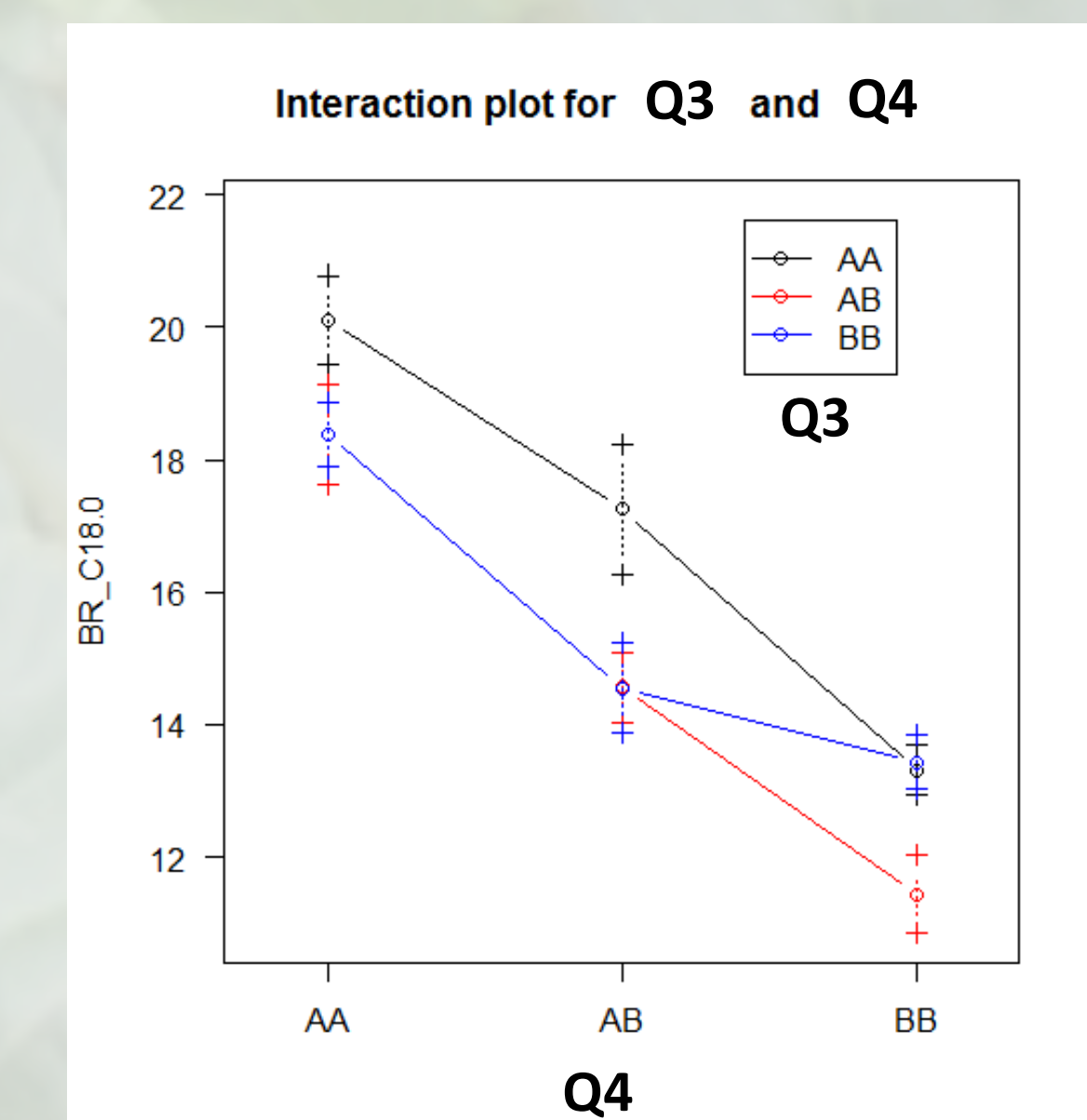


Figure 10. Effect plot showing the interaction between alleles at Q3 and Q4

QTL	df	SS	LOD	% var	F-value	P-value
Q1	2	57.91	3.363	2.902	7.486	<0.0001
Q2	2	152.69	8.279	7.653	19.74	<0.0001
Q3	6	203.5	10.666	10.2	8.77	<0.0001
Q4	6	1172.54	39.618	58.768	50.529	<0.0001
Q3*Q4	4	140.96	7.705	7.065	9.111	<0.0001

Source	df	SS	MS	LOD	% var
Model	12	1391.845	115.9871	43.89093	69.76008
Error	156	603.3434	3.867586		
Total	168	1995.188			

Table 2. ANOVA table with % variance in the trait explained by each QTL and overall model significance

## Future Plans

- Confirm results by re-running the analyses using data obtained from F<sub>6</sub> seed harvested in 2015
- Test for significant correlations between stearic acid content and other soybean traits
- Re-sequence high stearic line A6
- Develop an NIR calibration for stearic acid that will allow for non-invasive phenotyping of fatty acid composition.



## References

- Yu S, Derr J, Ethernan TD, Kris-Etherton PM: **Plasma cholesterol-predictive equations demonstrate that stearic acid is neutral and monounsaturated fatty acids are hypocholesterolemic.** *Am J Clin Nutr* 1995, 61(5):1129-1139.
- Food and Drug Administration. (2015). The FDA takes step to remove artificial trans fats in processed foods [Press release]. Retrieved from <http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm451237.htm>
- Zhang P, Burton JW, Upchurch RG, Whittle E, Shanklin J, Dewey RE: **Mutations in a Δ-Stearyl-ACP desaturase gene are associated with enhanced stearic acid levels in soybean seeds.** *Crop Sci* 2008, 48(6):2305-2313.
- Gillman D, Stacey MG, Cui Y, Berg H, Stacey G: **Deletions of the SACP-D locus elevate seed stearic acid levels but also result in fatty acid and morphological alterations in nitrogen fixing nodules.** *BMC Plant Bio* 2014, 14:143.
- Hammond EG, Fehr WR: **Registration of A6 germplasm line of soybean.** *Crop Sci* 1983, 23:192-193.



If you would like to view my poster online, scan here.