

Automated phenotyping of 371 *Sorghum bicolor* recombinant inbred lines for QTL mapping



Bryce Askey¹, Qinnan Yang², Andy Benson², James Schnable³
¹Horticultural Science Department, University of Florida, Gainesville, FL, ²Department of Food Science, University of Nebraska-Lincoln, Lincoln, NE, ³Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE



Abstract

Sorghum, a drought resistant cereal crop, is an important human food source, biofuel, and animal feed component. To improve the production efficiency of sorghum, a quantitative trait loci (QTL) study was conducted to identify sections of the sorghum genome which control for economically valuable traits. Traditionally a labor-intensive process when done manually, phenotype measurement was automated by applying a computer vision method to RGB digital images of plants. With genotype data for the 371 BTx623 × IS3620C recombinant inbred lines, a genetic map was constructed, QTL identified, and their phenotypic effects estimated.

Background

Agricultural importance of sorghum

- 5th most grown cereal crop in the world
- Drought resistant, and able to survive on degraded soils
- Efficient at converting water and nutrient inputs into starch and protein



Human consumption - sorghum grain is rich in phytochemicals, and a staple food for many living in hot and arid regions

Biofuel production – due to the sugars in their stalks, sweet sorghum varieties have great potential for ethanol generation

Animal feed – newer varieties of sorghum can provide a feeding value comparable to maize while being cheaper/easier to grow

Figure 1. Uses of sorghum (Ledbetter, 2016) include human consumption (Brink, 2012), biofuel production (National Renewable Energy Lab, 2017), and animal feed (K-State Research and Extension, 2016).

Improving production & survivability

- Identifying genes controlling variation in traits allows for more efficient selective breeding, and faster improvements in yield and stress tolerance
- Traits of interest:
 - **Stalk height** – plant biomass and lodging resistance
 - **Panicle exertion** – ease of harvesting & pest resistance
 - **Number of leaves** – energy production & plant biomass
 - **Leaf angle** – maximum planting density

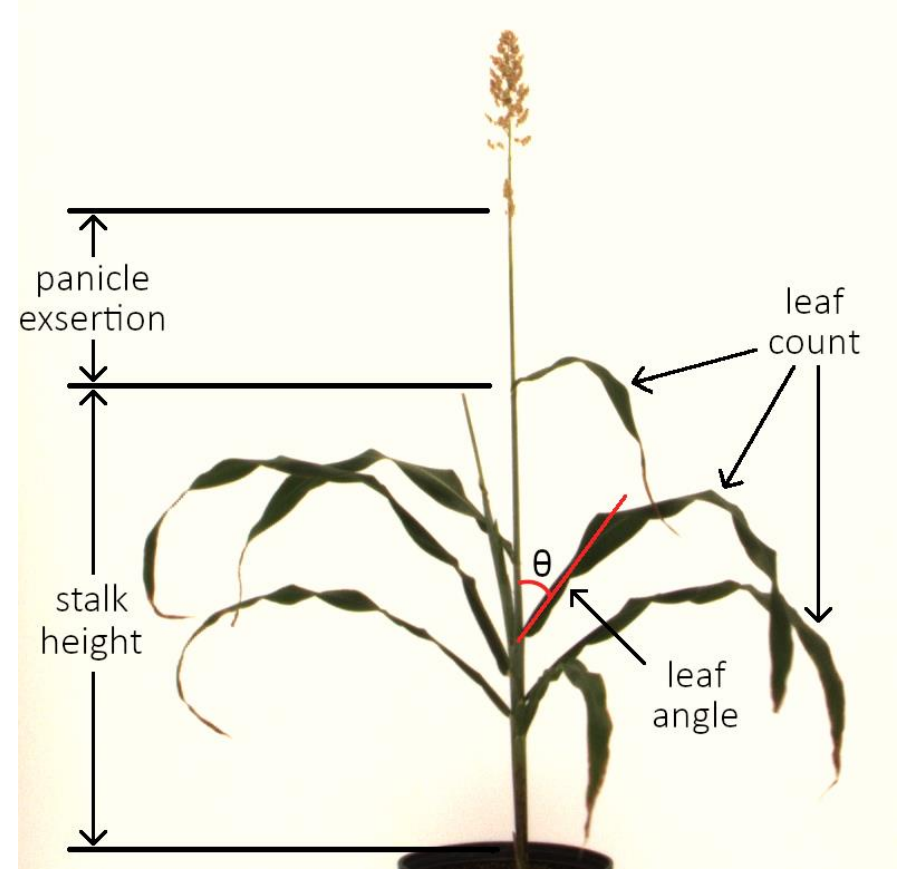


Figure 2. Sorghum plant with traits of interest identified

Automated phenotyping for QTL mapping

- QTL (Quantitative Trait Locus) mapping identifies genomic regions that control traits by correlating sequencing data with phenotype measurements
- Traditionally, phenotype measurements are taken by hand, which can be time-consuming, labor-intensive, and subject to inter-observer variability
- Automated phenotyping eliminates the need for manual measurements, and has potential to provide more objective data

Objectives

1. Develop a computer vision-based method to automate the measurement of four phenotypic traits (stalk height, panicle exertion, number of leaves, and leaf angle) from RGB images of sorghum plants
2. Correlate measured phenotype data with genotype data with a multiple QTL search to identify regions in the sorghum genome controlling variation in the traits

Methods

Plant growth and imaging

- 438 *Sorghum bicolor* plants were grown in controlled greenhouse conditions
 - 417 from a BTx623 × IS3620C family of RILs (371 unique lines used) (Kong et al, 2018), and 21 of genotype BTx623, a parent of the RIL family
- 130 days from sowing, RGB digital images of plants were taken

Computer vision

- Approach – segment plant into primary organs, then measure phenotypic traits

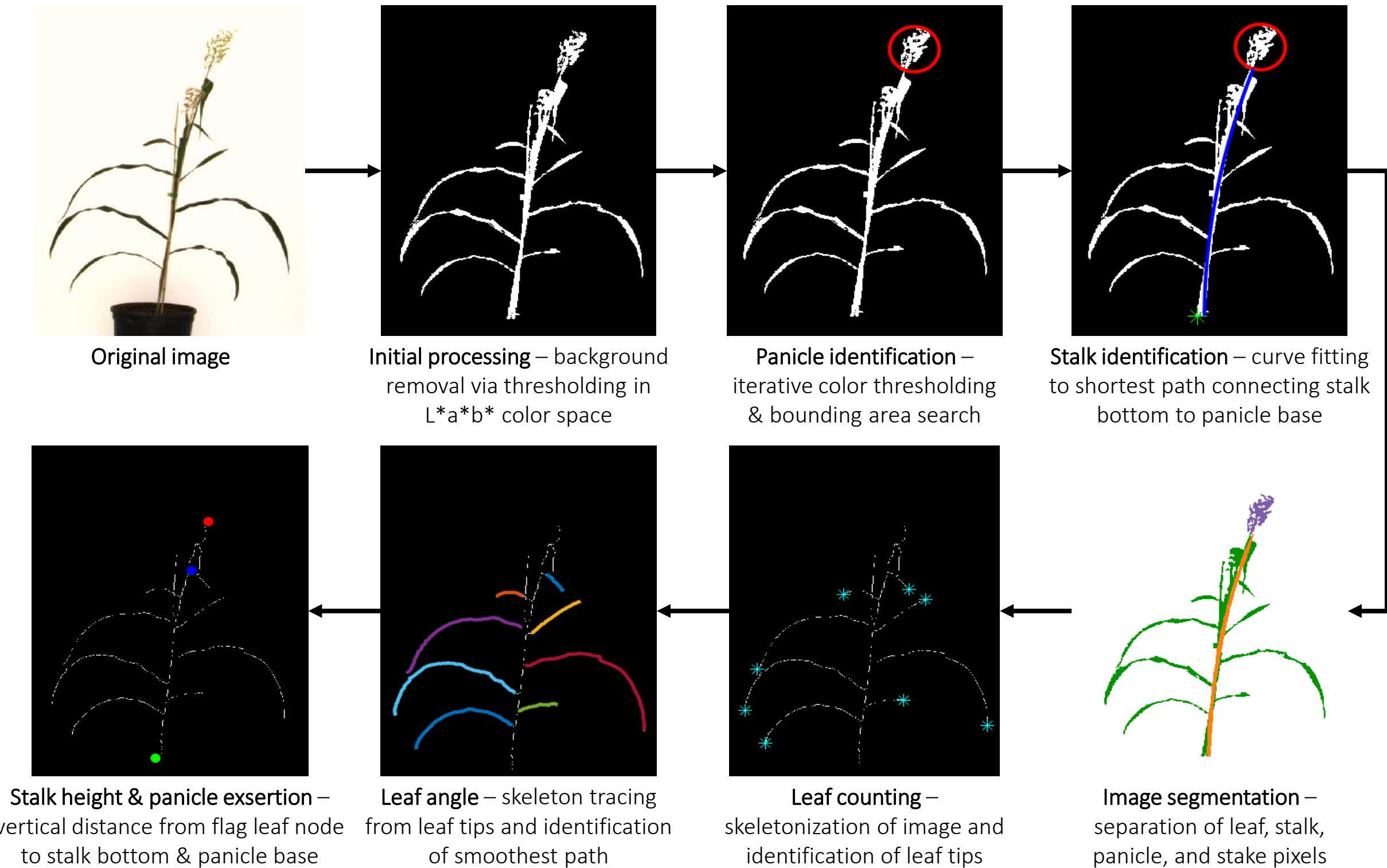


Figure 3. Flowchart of computer vision method with examples of images at various stages of processing

Genetic map construction and QTL mapping

- A genetic map was constructed using marker data at 616 SNP markers for 393 recombinant inbred lines (RILs) (Taylor & Butler, 2017)
 - Map accuracy was evaluated with Spearman's rank correlation coefficient
- A multiple QTL model was generated with a stepwise search (Broman et al, 2003)
 - LOD score penalties for each trait were calculated with a two-QTL scan
 - QTL effects on phenotypic traits were estimated

All code can be found in the repository: https://github.com/bryceaskey/sorghum_QTL

Results

Constructed genetic map

- Contains 10 linkage groups spanning a total length of 2639.5 cM
- Marker order closely matches that of physical marker order
 - Spearman's rho for each chromosome is almost equal to 1

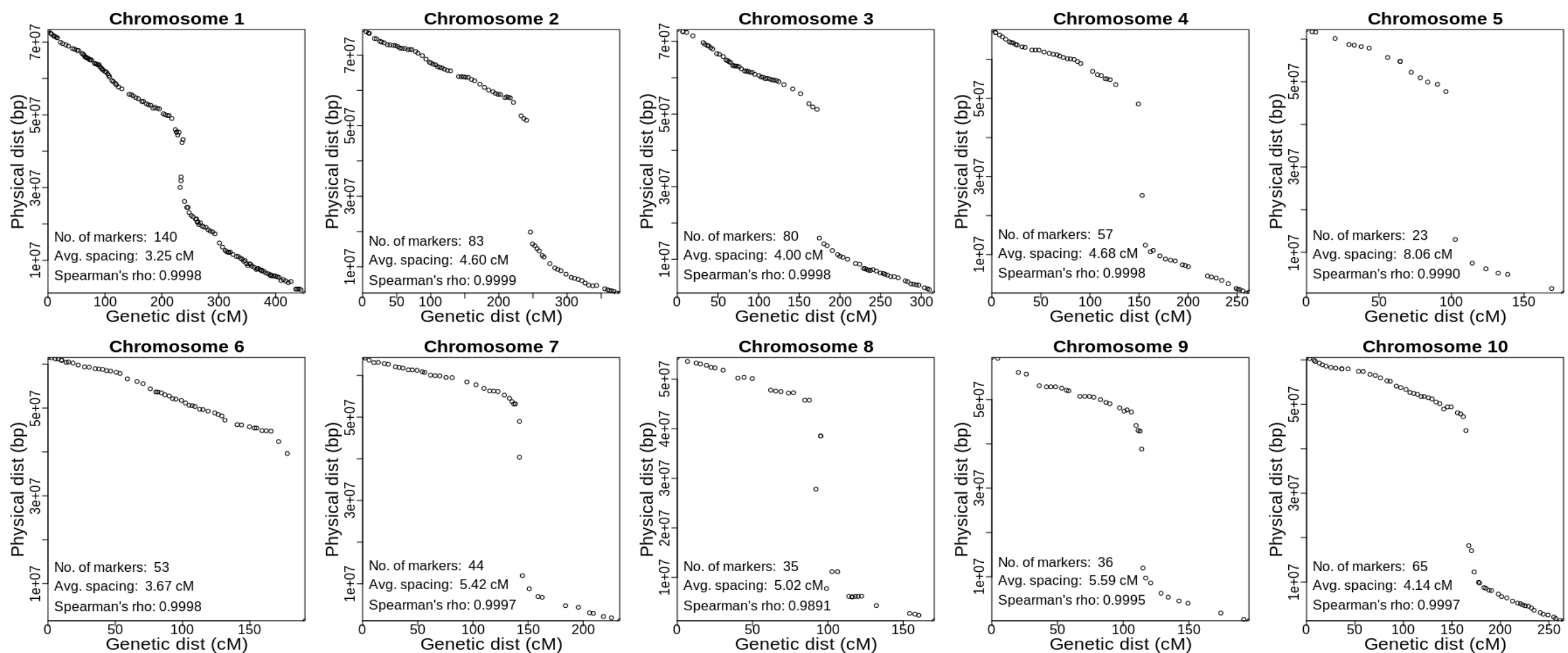


Figure 4. Genetic marker location plotted against physical marker location for each chromosome in constructed genetic map

Identified QTL and interactions

- 19 QTL and 3 interactions identified w/ many matches to previously identified QTL
 - e.g. loci on chromosome 7 @ 92.5 cM is likely dwarfing gene *dw3*
- Performed best for stalk height – 39.03% of phenotype variance explained by QTL

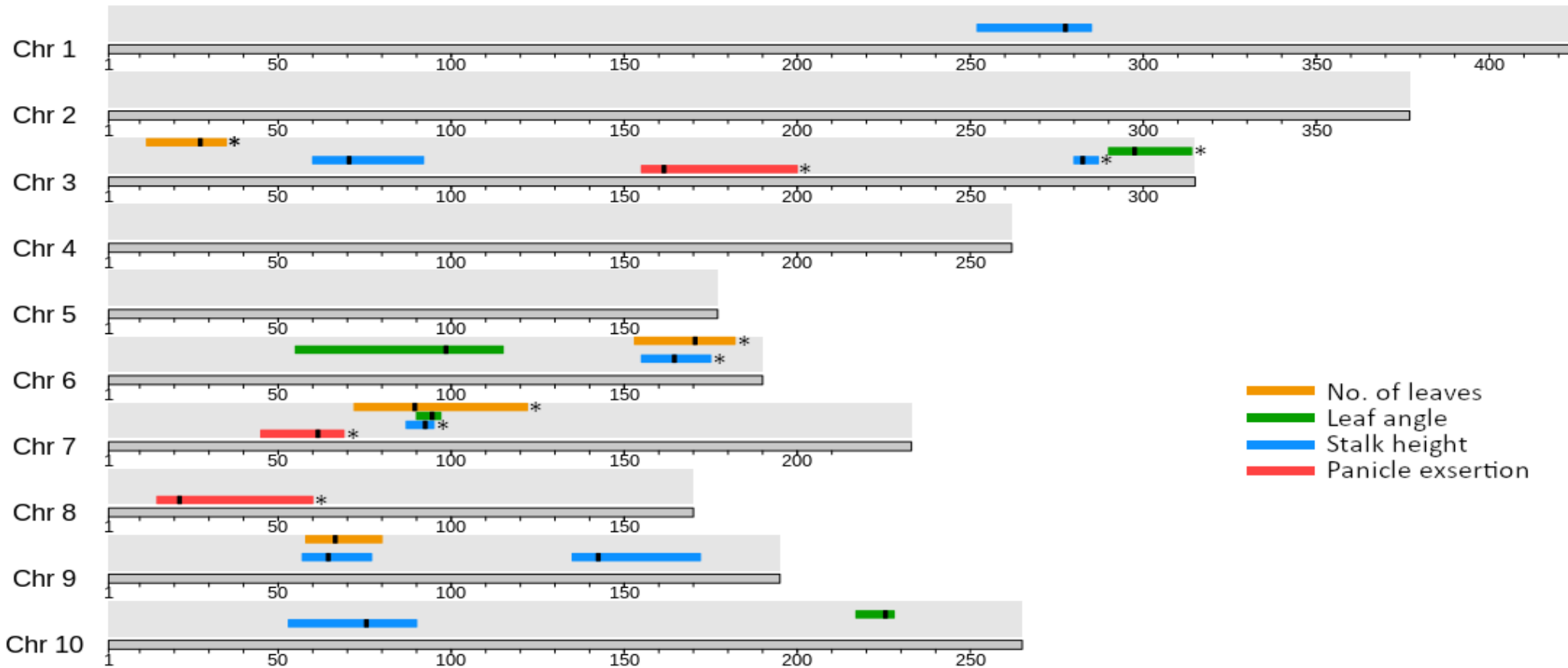


Figure 5. 1-LOD intervals in cM of identified QTL for leaf count (yellow), leaf angle (green), stalk height (blue), and panicle exertion (red) plotted on sorghum chromosomes. Asterisks (*) indicate that QTL location matches previous publications.

Estimated QTL effects

- Equal to 1/2 the difference between the measured trait averages for the two genotypes at each QTL/interaction
 - Useful for breeders when selecting lines for further development

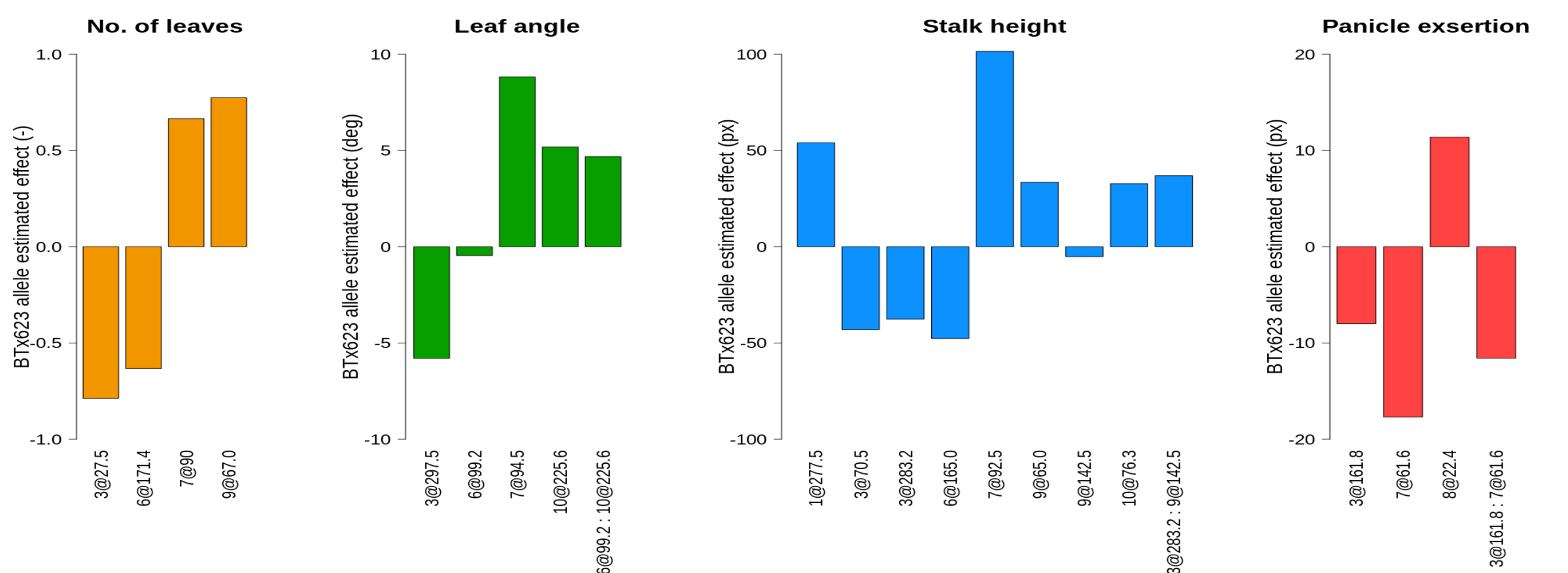


Figure 6. Estimated phenotypic effect of the BTx623 allele at each identified QTL and QTL interaction

Conclusions

In this study, we investigated the ability of a computer vision method to provide phenotypic measurements for a QTL mapping study. A total of 19 QTL and 3 interactions were identified with a multiple QTL mapping approach, and their effects on plant phenotype estimated. Because sensitivity and accuracy of QTL detection can often be improved by increasing the number individuals in the studied population, replacing labor-intensive manual measurements with an automated phenotyping system may allow researchers to increase the size of their studied population, and thus, increase their statistical power for identifying QTL. Future work will focus on direct comparisons of trait and QTL distributions generated by manual and automated phenotyping to further evaluate the efficacy of the developed method.

References & Acknowledgements

References

- Brink, C. (2012). *The food dries: red sorghum* [Electronic]. Retrieved from: <https://is.gd/PVfBhu>
- Broman, K. W., We, H., Sen, S., & Churchill, G. A. (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19, 889-890.
- Gel, B., Serra, E. (2017). karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics*, 33(19), 3088-3090.
- Kong, W., Kim, C., Zhang, D., Guo, H., Tan, X., Jin, H., Zhou, C., Shuang, L., et al. (2018). Genotyping by Sequencing of 393 *Sorghum bicolor* BTx623 × IS3620C Recombinant Inbred Lines Improves Sensitivity and Resolution of QTL Detection. *G3: Genes, Genomes, Genetics*, 8(8), 2563-2572.
- K-State Research and Extension. (2016). *KSU Beef Yard_019 copy* [Electronic]. Retrieved from: <https://is.gd/s5I3Ah>
- Ledbetter, K. (2016). *AgriLife Extension forage sorghum trials* [Electronic]. Retrieved from: <https://is.gd/7XRKMc>
- National Renewable Energy Lab. (2017). *9th Annual BioFuels Science and Sustainability Tour* [Electronic]. Retrieved from: <https://is.gd/WJqG15>
- Taylor, J., & Butler, D. (2017). R Package ASMap: Efficient Genetic Linkage Map Construction and Diagnosis. *Journal of Statistical Software*, 79(6), 1-29.

Acknowledgements

This work was supported by funding from the National Science Foundation, the Nebraska Food for Health Center at the University of Nebraska-Lincoln, and the Center for Root & Rhizobiome Innovation at the University of Nebraska-Lincoln.