SOYBEAN (*Glycine max*): "The Mighty Yellow Bean" – Analysis and Review

Alexander Saka

Writer's Comment: When I was presented with the prompt, to choose an essential crop from around the world and determine its origination, I knew I wanted to pick something with a rich history, and a deep cultural connection. I remembered reading the mythology of the soybean in Japan, and recognized that was a rabbit hole I had to follow. Being a biologist at heart, the first trail I followed was into genomic analysis and complex trait tracking for the soybean. By this point I knew that the research I wanted to write about the plant needed to be accessible, but uncompromising in scientific accuracy. This plant has had a complex intertwined history with humanity and appears in many different cultures' creation mythology. So I knew I had to do justice to that. Additionally, it reveals some amazing insights into the power humans have over their natural environment, how we have truly shaped the biosphere of our planet in amazing ways. Which belays the warning that we should care for our planet, our home.

Instructor's Comment: PLB 143 deals with the origins of plant agriculture and the domestication of crop plants. The transition from hunting-gathering to agriculture to procure food and other plant-based products some 10,000 years ago was a significant milestone in human evolution and a necessary precursor to the independent development of civilizations in several parts of the world. How, where, and when this transition took place and what the major changes were in the functionalities of crop plants is the subject of multidisciplinary research presented in this course. To demonstrate they have mastered the topic, students like Alexander Saka write a term paper that summarizes the *up-to-date knowledge about domestication and evolution of a crop of their choice. In his search of the origin of soybean, Alexander elegantly describes genetic, archaeological, ecological, and historical information, which lead him to conclude that soybean originated in what is now northern China. My personal thanks to Kimberly Gibson, Teaching Assistant, and Sheena Campbell, Student Services Librarian, for their support of student writing.*

—Paul Gepts, Department of Plant Sciences

Abstract:

The common soybean (*Glycine max)* is one of the world's most useful, widespread, culturally significant, and largest growing cultivations in the world. The primary thread of global soybean history traced a path out of East Asia before spreading across the globe and finding more modern agricultural homes in India, Brazil, Argentina, and the United States. However, despite its global fame, both the origin of the species, and the primary site of domestication in East Asia, are still hotly contested. As a result, an array of scientific, archaeobotanical, and anthropological methods have been used to assess the validity of the most suspected regions. A review of these significant findings and research may help elucidate the history of this mighty yellow bean. While the exact geographical origin of the soybean is currently a mystery, with further research and additional international collaboration, these locations may become less obfuscated.

Introduction:

Ōgetsu-hime, the Shinto goddess of food, was detailed to have been slain by the moon god Tsukiyomi, after he was displeased by her gifts. From her body sprouted the first soybean, and so many other culturally foundational plants of Japanese agriculture. [10] Careful to not let her death be in vain, Tsukiyomi's sister, the sun goddess Amaterasu, sewed the crops across the island for humans to enjoy. This tale from the Koji (古事記) or "Record of Ancient Matters," is just one of the many tales of the origin of the soybean that exist across East Asia. Other stories include the divine farmer Shennong retrieving it from the mythological Mount Weishi in China[22] as well as numerous tales in Korea. As shown by these mythic origins, Glycine max, the domesticated soybean, is a crop of singular significance.

Soybean plants stand, on average, three to five feet tall, they are typically verdant green, and each modern, genetically improved plant, on average branches three times[13]. From these branches you can expect around twenty nodes of growth. *G. max* has the observed maximum capacity of six hundred pods, but realistically only produces fifty to a hundred. Each pod normally contains three seeds.[20, 9]

Soybean has a large distribution, with their primary range being located between latitudes of 35–50° degrees, with a secondary range between latitudes of -20 to -40 degrees.[18] Soybean has a strict light requirement of about twelve to thirteen hours a day, of which at least eight hours need to be direct sunlight. Without this they will not flower, and often become prone to diseases like white mold.[12] Despite these seemingly rigorous requirements, the soybean has a distinct advantage over some plants, which allows them to enjoy a wide and global distribution. Like other legumes, soybean plants supply their own nitrogen via fixation of nitrogen in the atmosphere by a symbiotic soil bacterium nestled in their roots.[20] Additionally, when used in crop rotations, soybean can restore nitrogen to the surrounding soil, reducing the need for fertilizers when planting other crops.

Soybean is also an economically crucial crop. They provide 300 million metric tons of food a year, and are responsible for nearly 100 billion U.S. dollars worth of annual revenue.[4] Soybean is used as an

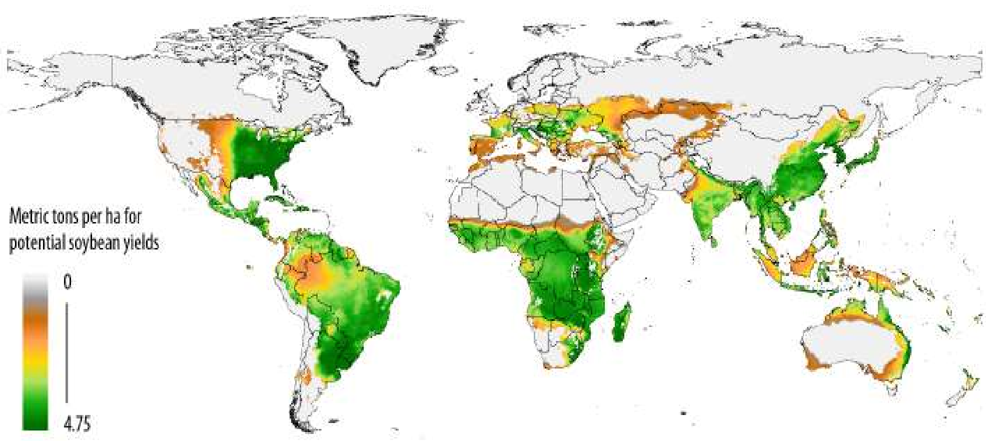

Figure 1. *Retrieved from the International Food Policy Research Institution. This image displays the global potential suitable land for Soy. Areas have been stained with blue to distinguish current ranges according to FOASTAT ranges. Most arable land supports soy production, but yield is restricted by above mentioned factors.*

ingredient in the formulation of a multitude of food, feed and industrial products.[21] There is a large and expanding range of different soyfoods, soy seed cooking oil, growing biodiesel applications, industrial farming uses as feed, as well as uses of vegetable protein substitutes for meat and dairy. Furthermore, soy food products have made more nutrients available to the economically disadvantaged in local markets by being "a primary source of high-value secondary co-products such as lecithin, vitamins, nutraceuticals and antioxidants."[3, 21] However, due to lack of genetic variation in the known modern cultivars, there has been somewhat of a plateau of global yield, even shrinking from 2018 to 2019.[26, 21, 4]

While the soybean has high current and future global distribution as seen in Figure 1, soybean origin and center of domestication research has been extensive and often conflicting. This is further complicated and compounded by the cultural importance of soy to China, South Korea, and Japan, as illustrated by the origin stories provided earlier. As a result, there is also a political interest in being able to say that one nation is the center of domestication as well as the birthplace for the crop.

Results and Discussion:

China, Korea, Japan; Understanding Trait Migration

In order to understand and narrow down the sites of origin, study of soybean phenotypes in these three countries can help trace the migration of soy. A key element of the domestication syndrome (the ensemble of traits that distinguish a crop from its wild ancestor) is the increased size of the yield from the plant,[6] particularly the seeds in plants in which the seed is the primary element of human consumption. These favorable traits are known as domestication related traits (DRTs).[14] In the case of the domestication of *Glycine max*, DRTs related to bean dimensions are a critical area for research. DRTs present in *Glycine max* are unique due to their loci positions being compressed. A study done by researchers from the Laboratory of Plant Genetics and Evolution in Japan used quantitative trait locus mapping to identify genotypic coalescence across loci.[15] They found that across twenty linkage groups, DRTs had a tendency to be clustered across just six. But since these regions were not clustered themselves, the introgression of useful genes was less hindered.[15]

Data collected from China, Korea, and Japan was analyzed in a study by dimensions (length and width), in order to track the progression of these DRTs, and hence their migration patterns.[14] This study of soy length and width was done across cultural eras in these three countries. Within this study, two main varieties were discovered: a more wild and weedy variety and the modern domesticate. Soybean had been collected from numerous other researchers, and data from other relevant research in question was used when direct observation of samples was not possible. They found that the earliest recovered *Glycine max* was present in the early Neolithic period in China (9000–8600 B.P.). Samples were recovered that spanned a large period of time. Figure 2 demonstrates their results when length, width, and time are considered. We can see a distinct increase over time in the dimensions of the seeds, which increases the confidence in the assumption that seed size is a DRT through integration. We also observe that the Chinese samples on a whole were smaller than the Korean samples, which were smaller in turn compared to the Japanese variety.[14] This suggests that as the larger crop varieties are preferred by humans, those would be the crops traded, effectively functioning as a genetic bottleneck by increasing the standard size threshold in the locations they arrive in.[17] This evidence points to the conclusion that China is likely the center of domestication, as seed size was likely much smaller when it was originally cultivated, as is similar to other legumes which have other DRT that are affected by domestication first.[5]

The Four Chinas Agricultural Origin Concept:

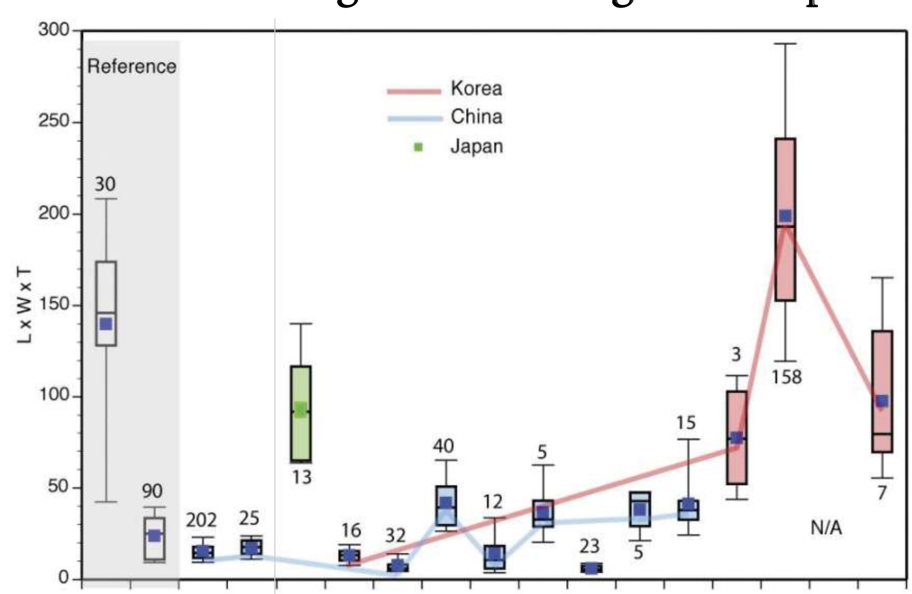

Figure 2: *Retrieved from Lee et al. (2011). This graph displays comparisons of length and width over time. China is indicated in blue, Korea in red, and Japan in green. Limited Japanese samples result in decreased confidence. This graph evidences the idea that earlier found samples also were the smallest, and all belonged to China, suggesting that is the region with earliest domestication.*

Key historical research done in 1975 placed the domesticated soybean's introduction to mainland China, and soon the rest of the world, to have taken place in the northeastern regions of China around 500 B.C.E.[4] In fact, it was speculated that soybean domestication was associated with the agricultural revolution in the Eastern Zhou Dynasty.[10] However, soybean is known to grow throughout China. With production being localized across four primary regions: the kaoliang-soybean region (Northeast China), the winter wheat-kaoliang region (Central China), the winter wheat-millet region (Eastern China), and the Szechwan rice region (South-Central China).[5] Additionally, the earliest appearance of soybean samples in the Kaoliang-soybean region are around 2000 B.C.E.,[24] which suggests that they were not domesticated there originally or had independent origin, as other samples from other regions were dated to be older. This is also somewhat supported by loci analysis performed in 2016, which discovered that soybean often has distinct regional traits.[26] Charred soybean specimens place the Huang-Huai basin (South of the winter wheat-Kaoliang) region as older than the Northeast regions.[19]

Additionally, phylogenetic and nucleotide diversity studies have shown that the Huang-Huai basin region has the highest observed diversity[7] when compared to the others.

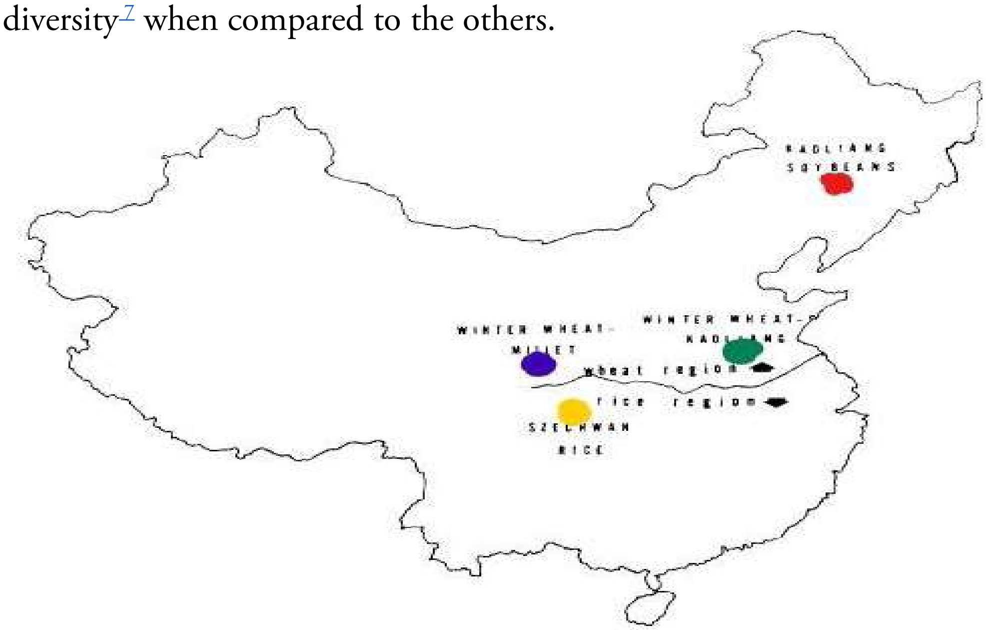

FIG. 3. The major soybean regions in mainland China.

Figure 3. *Retrieved from Hymowitz (1970). This map displays the prospective regions of the origin of soybean domestication. The kaoliang-soybean region is denoted in red, the winter wheat-kaoliang region in green, the winter wheat-millet region in blue, and the Szechwan rice region in yellow.*

This fits into Vavilov's method for finding the origin of a cultivated plant, as the most diverse area is likely the center of domestication. Furthermore, gene flow studies on *G. soja* and *G. gracilis* have shown that the flow of genetic material was passed from *G. soja* to *G. gracilis* before being passed to *G. max.*[8] This provided the insight that in the Huang-Huai basin, domesticated *G. max* had more genetic and phenotypic introgression from the wildtype than any other samples from other regions. This counters the claim of independent origin. The same loci analysis demonstrated that region specific traits also discovered with the use of principal component analysis (PCA) there were two main clades, in which the domesticated clade was shown to have lost 50 percent of its diversity.[26] This PCA suggests that a single domestication event likely occurred.

Figure 4. *"Shu"*

It is also of note that the archaic Chinese character for soybean, "Shu," has also had its origins traced separately from the plant. This pictograph has been broken down into 4 components, the middle line represents the plant, the upper and lower vertical lines represent the stem and roots respectively, with the three lines at the bottom signifying the root nodules, a very distinctive characteristic of soy plant[11] and other legumes because they harbor the nitrogen-fixing bacteria mentioned earlier. This symbol is particularly useful as its origin is found in the "Book of Odes," a text whose relevance spans 1200–900 B.P. and was discovered in the geological area of the winter wheat-kaoliang region.[11] Furthermore, the dating of the text also illuminates the usage of the word and character itself was developed in and around 1100 B.P.[2] Further morphological evidence has also been compiled, particularly with pod dehiscence. A study was conducted that tracked the concurrence of pod dehiscence with relative humidity of the environment in which they grew. *Pdh1*, the gene controlling pod dehiscence—an essential trait of soybean—only occurs in the Huang-Huai basin and originates in the wild progenitor species *G. Joja*.[23]

A Common Domestication, Uncommon Origin:

The suggested domestication origin of the common soybean seems most likely to be in the Huang-Huai basin. Specimens found at a Late Yangshao site near to the Huang-Huai basin were recovered from the floor of an ancient house; these are the earliest found in this region, however due to natural conditions of flooding and sun exposure, no earlier samples have been discovered.[14] While much evidence has been lost from this region, this seems to be the most likely origin of domestication in China. Which consequently would mean it is likely the origin in all of Asia and possibly the world. However, there are still complications to this theory. Namely, the common traits that track wild varieties of *G. soja* across Asia, which support the idea that at least partial domestication occurred in many sites across central and Northeast China. Which is supported further by the similar domestication track across at least two dozen other species of legumes.[1] This information likely makes the actual origin of the soybean impossible to discover with current technology and methods. However, with the current understanding and evidence we have, a common domestication of the soybean can be placed, likely just below the Yellow River, in the Huang-Huai basin between 7000 to 3000 years ago.

Recommended Lines of Future Research:

There were several problems present when analyzing the topic of the origin and domestication of the soybean, most of them had to do with the quality and quantity of evidence available. Numerous accounts and papers have revealed that the Huang-Huai basin is prone to both flooding and mudslides. This has two main negative effects: the destruction of evidence, as well as the movement of evidence away from its site of origin.[24] Further digs are most likely needed, preferably accounting for the potential movement of the specimens; however, this would require significant analysis and effort.

International collaboration is needed in order to better understand the evidence we do currently have. Many nations have evidence, and collaboration between more teams of researchers would benefit the study of the plant. In numerous studies, it is mentioned that a limitation is the researcher's lack of access to the actual specimens for two main reasons: the inability to facilitate collaboration with the previous team or the destruction of the pieces due to the distance in time from the original study. These are major hurdles that can be rectified by a singular, larger, and more inclusive study. Further studies should seek temporary access to the specimens collected by other institutions to perform experiments with a larger and more diverse sample size. Additionally, biases may be observed when it comes to the nation of origin of the studies, as each nation has a somewhat-vested interest in being able to claim the origination/domestication of *G. max*.

Another option that would not require access to the specimens, however, would be further research of the occurrence of vocabulary to do with soybean in the different regions of China. This could help place our marker for where and, more specifically, when the domestication may have occurred.

Finally, the use of more modern techniques by new studies would also benefit the discovery of the topic. Many published pieces about the origin and domestication of the soybean are not recent, some of which even occurred nearly a century ago, before the advent of mass spectrometry and genetic sequencing. Many of the samples have not been dated beyond their discovery in a site from a particular age. Further techniques like C14 dating could be employed to attempt to find a more accurate assessment of the age of certain samples. Overall, more collaboration and up-to-date methods need to be applied going forward to help clear up the crop origination of this mighty yellow bean.

References

1. Cannon, S. B., & Shoemaker, R. C. (2012). Evolutionary and comparative analyses of the soybean genome. *Breeding Science*, *61*(5), 437–444. doi.org/10.1270/jsbbs.61.437
2. Dobson, W. A. C. H. (1996). *The Language of the Book of Songs*. University of Toronto Press.
3. Du Bois, C. M. (2018). *The Story of Soy*. London: Reaktion Books.
4. Food and Agriculture Organization of the United Nations. (2021, March 8). *Value of Agricultural Production*. FAOSTAT. www.fao.org/faostat/en/#data/QV/visualize
5. Fuller, D. Q. (2007). Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World. *Annals of Botany, 100*(5), 903–924. doi.org/10.1093/aob/mcm048
6. Gepts, Paul. (2004). Crop Domestication as a Long-Term Selection Experiment. In Janick, J. (Eds.), *Plant Breeding Reviews, Part 2: Long-Term Selection: Crops, Animals, and Bacteria*, *(24)*2, pp. 1–44. John Wiley and Sons, Inc. psfaculty.plantsciences.ucdavis.edu/gepts/LTS.pdf
7. Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H., Wang, Y. (2010). A single origin and moderate bottleneck during domestication of soybean (*Glycine max*): implications from microsatellites and nucleotide sequences. *Annals of Botany, 106*(3), 505–514. doi.org/10.1093/aob/mcq125
8. Han, Y., Zhao, X., Liu, D., Li, Y., Lightfoot, D. A., Yang, Z., Zhao, L., Zhou, G., Wang, Z., Huang, L., Zhang, Z., Qiu, L., Zheng, H., & Li, W. (2015). Domestication footprints anchor genomic regions of agronomic importance in soybeans. *New Phytologist, 209*(2), 871–884. doi.org/10.1111/nph.13626
9. North Carolina Soybean Production Association. (Year, month day). *History of Soybeans*. North Carolina Soybeans. ncsoy.org/media-resources/history-of-soybeans/
10. Ho, P-t. (1975). *The cradle of the East: An inquiry into the indigenous origins of techniques and ideas of Neolithic and early historic China,*

5000-1000 B.C. University of Chicago Press.

11. Hymowitz, T. (1970). On the Domestication of the Soybean. *Economic Botany, 24*(4), 408–421. doi.org/10.1007/BF02860745
12. Jardine, D. J. (2020, July 1). *White Mold - Soybean Disease*. Soybean Research & Information Network. soybeanresearchinfo.com/soybean-disease/white-mold/
13. Kuswantoro, H., Indriani, F. C., Patriawaty, N. R., Sulistyo, A., Han, W. Y., Lee, P. Y., Cho, Y. H., & Baek, I. Y. (2014). Performance of Acid-Adaptive Soybean Expected Lines in South Lampung, Indonesia. *AGRIVITA Journal of Agricultural Science, 36*(2), 160–166. agrivita.ub.ac.id/index.php/agrivita/article/view/400/503
14. Lee, G-A., Crawford, G.W., Liu, L., Sasaki, Y., & Chen, X. (2011). Archaeological Soybean (Glycine Max) in East Asia: Does Size Matter? *PLoS ONE*, *6*(11), Article e26720. doi.org/10.1371/journal.pone.0026720
15. Liu, B., Fujita, T., Yan, Z. H., Sakamoto, S., Xu, D., & Abe, J. (2007). QTL Mapping of Domestication-related Traits in Soybean (*Glycine max*). *Annals of Botany*, *100*(5), 1027–1038. doi.org/10.1093/aob/mcm149
16. Philippi, D. L. (1969). Chapter 18. In *Kojiki*. (Trans.). University of Tokyo Press. www1.udel.edu/History-old/figal/Hist138/Text/er/kojiki.pdf
17. Piper, C. V., & Morse, W. J. (1943). *The Soybean.* Peter Smith Publishers.
18. Scurek, O. (2006, April 15). *Glycine Max: Distribution and Climate.* Glycine Max-Soybeans, bioweb.uwlax.edu/bio203/s2009/scurek_oliv/Habitat.htm
19. Sedivy, E. J., Wu, F., & Hanzawa, Y. (2017). Soybean Domestication: the Origin, Genetic Architecture and Molecular Bases. *New Phytologist, 214*(2), 539–553. doi.org/10.1111/nph.14418
20. Stowe, K. D., & Dunphy, J. (2017). The Soybean Plant. In *North Carolina Soybean Production Guide* (pp. 3–10). NC State Extension Publications. content.ces.ncsu.edu/north-carolina-soybean-production-guide/the-soybean-plant

21. Wilson, R. F. (2008). Soybean: Market Driven Research Needs. In *Genetics and Genomics of Soybean* (pp. 3–15). Springer. doi.org/10.1007/978-0-387-72299-3_1
22. Yang, L., An, D., & Turner, J. A. (2008). *Handbook of Chinese Mythology*. Oxford University Press. p. 197. img.fireden.net/tg/image/1560/73/1560731269210.pdf
23. Zhang, J., & Singh, A. K. (2020) Genetic Control and Geo-Climate Adaptation of Pod Dehiscence Provide Novel Insights into Soybean Domestication. *G3: Genes, Genomes, Genetics, 110*(2), 545–554.
24. Zhao, H., Huang, C. C., Zheng, Z., Hu., Y., Zhang, Y,. Guoa, Y., & Zhou, Q. (2017). New Evidence for the Catastrophic Demise of a Prehistoric Settlement (the Lajia Ruins) in the Guanting Basin, Upper Yellow River, NW China. *Journal of Asian Earth Sciences, 146*, 134–141. doi.org/10.1016/j.jseaes.2017.05.019
25. Zhao, Z. (2005). Discussion of the Xinglonggou site flotation results and the origin of dry farming in northern China. In *Antiquities of Eastern Asia, A,* 188–199.
26. Zhou, Z., Jiang, Y., Wang, Z., Gou, Z., Lyu, J., Li, W., Yu, Y,. Shu, L., Zhao, Y., Ma, Y., Fang, C., Shen, Y., Liu, T., Li, C., Li, Q., Wu, M., Wang, M., Wu, Y., Dong, Y. … Zhixi T. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nature Biotechnology, 33*(4) 408–414. doi.org/10.1038/nbt.3096