Enhancing Seed Quality and Application of Soybean through High Test Weight and Low Trypsin Inhibitors

Zachary Alexander Shea

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Crop and Soil Environmental Sciences

Bo Zhang
Bingyu Zhao
Michael Schwarz
Reza Ovissipour
MA Saghai-Maroof

12/7/22
Blacksburg, Virginia

Keywords: Crop Improvement, Seed Composition, Soybean, Test Weight, Trypsin Inhibitor

Enhancing Seed Quality and Application of Soybean through High Test Weight and Low Trypsin Inhibitors

Zachary Alexander Shea

ABSTRACT

Soybean is an integral oilseed crop that is used all over the world. Despite this, soybean has been experiencing a decrease in seed quality seen by the decreasing trend in test weight, and is limited in its uses due to the antinutritional factor, trypsin inhibitor (TI). This project first worked to identify single nucleotide polymorphisms (SNPs) associated with high test weight through a genome-wide association study (GWAS) as well as determine the relationship between important seed composition traits and test weight. Additionally, it focused on identifying and knocking out KTI genes specifically expressed in soybean seed tissue through CRISPR/Cas9. Lastly, this project investigated further use of soybean in aquaculture by evaluating the performance of Rainbow trout fed diets with low-TI, 'VT Barrack' soy meal. Nine SNPs on chromosome 15 were found to be significantly associated with high test weight as and eight potential gene candidates were identified. Test weight was found to be significantly, and negatively related to seed oil content and had inconsistent correlations with protein and sugar content. The KTI genes KTI1, Glyma01g095000, and KTI3, Glyma08g341500, were found to be only expressed in seed tissues and multiple KTI knock-out soybean plants with decreased seed TI content and activity were developed. Lastly, we found that our low-TI soymeal could replace 30% of fish meal in aquafeeds with no negative effects on trout growth and health.

Enhancing Seed Quality and Application of Soybean through High Test Weight and Low Trypsin Inhibitors

Zachary Alexander Shea

GENERAL AUDIENCE ABSTRACT

Soybean is an essential and lucrative crop due to its countless uses. Despite its importance and flexibility, it contains antinutritional factors, such as trypsin inhibitors (TI), and has been experiencing a decline in test weight, the amount of soybean seed in kilograms per hectoliter, that can limit the profits that growers receive. This dissertation seeks to contribute to the development of improved soybean varieties in addition to evaluating current low-TI varieties' application in aquaculture as a substitute for fish meal. A genome-wide association study (GWAS) was performed to identify molecular markers associated with high test weight and the relationship between high test weight and important seed composition traits was determined. Gene editing was conducted to knock-out expression of Kunitz trypsin inhibitor (KTI) genes specifically expressed in seed tissue. A fish feeding trial was conducted to evaluate the performance of Rainbow trout-fed diets using our low-TI soybean cultivar, 'VT Barrack'. We were able to successfully identify multiple molecular markers associated with high test weight as well as identify multiple candidate genes. We were able to determine that high test weight has a strong, negative relationship with seed oil content. We identified two KTI genes for gene knockout and were able to successfully develop a low-TI soybean line that has limited TI in just seed tissue. Lastly, we found that our low-TI soymeal could replace 30% of fish meal in aquafeeds with no negative effects on trout growth and health.

Attributions

Chapter 1: Literature Review: Soybeans: Production, Versatility, and Improvement

Zachary Shea – writing and editing
William Singer – writing
Bo Zhang – editing

Chapter 2: **Manuscript 1: Development of molecular markers of the Kunitz trypsin inhibitor mutant alleles generated by CRISPR/Cas9-mediated mutagenesis in soybean**

Zhibo Wang – Running experiments, writing, editing
Zachary Shea - Running experiments, writing, editing
Luciana Rosso – Running experiments and editing
Chao Shang – Help with experiments
Jianyong Li – Help with experiments
Patrick Bewick – Help with experiments and editing
Qi Li – Help with experiments
Bingyu Zhao – Experimental design and editing
Bo Zhang – Experimental design and editing

Chapter 3: Manuscript 2: Determining Genetic Markers and Seed Characteristics Related to High Test Weight in *Glycine max*

Zachary Shea – Experimental design, running experiments and writing
William M. Singer – Assisting with experiments and editing
Luciana Rosso – Editing
Qijian Song - Editing
Bo Zhang – Experimental design and editing

Chapter 4: Manuscript 3: Enhancing Soybean Meal Demand and Market by Developing Soy Meal-Based Aquafeeds

Zachary Shea – Experimental design, running experiments, writing
Ethan McAlhaney – Assisting with feeding trial
Steve Uric – Assisting with feeding trial
Michael Schwarz – Experimental design and editing
Reza Ovissipour – Experimental design and editing
Bo Zhang – Experimental design and editing

Table of Contents

1. Introduction……………………………………………………………………….....1
2. Chapter 1 Abstract…………………………………………………………………...2
3. Chapter 1 Section 1: Introduction…………………………………………………....2
4. Chapter 1 Section 2: Agronomic Practices…………………………………….......3
5. Chapter 1 Section 3: Pest Management……………………………………………...7
6. Chapter 1 Section 4: Soybean Utilization…………………………………………..11
7. Chapter 1 Section 5: Breeding Efforts……………………………………………...14
8. Chapter 1 Section 6: Conclusion……………………………………………………16
9. Chapter 2 Abstract…………………………………………………………………18
10. Chapter 2 Introduction……………………………………….........................19
11. Chapter 2 Materials and Methods………………………………………........20
12. Chapter 2 Results…………………………………………………………………22
13. Chapter 2 Discussion………………………………………………………………...29
14. Chapter 3 Introduction……………………………………………………………32
15. Chapter 3 Materials and Methods…………………………………………........34
16. Chapter 3 Results…………………………………………………………………39
17. Chapter 3 Discussion………………………………………...........................56
18. Chapter 3 Conclusion ……………………………………………………….........59
19. Chapter 3 Supplemental Figures……………………………………………………60
20. Chapter 4 Abstract………………………………………………………………...63
21. Chapter 4 Introduction……………………………………………………………63
22. Chapter 4 Materials and Methods…………………………………………........64
23. Chapter 4 Results…………………………………………………………………68
24. Chapter 4 Discussion………………………………………………………………...69
25. Future Directions…………………………………………………………………71
25. Bibliography……………………………………………………………........72

Introduction

Soybean is an important crop that is grown across the globe. Soybean is so important due to its excellent nutrient profile and its diverse range of applications. While yield is always the most important trait for soybean production, there are many different aspects of soybean that growers and end-users expect to be improved. This dissertation focuses on two traits in particular, test weight and trypsin inhibitor (TI). Test weight corresponds to seed quality, so seeds with high test weight needs to be selected to maintain good seed quality. TI is an antinutritional factor and limits soybean's use in many applications in human and animal consumption. Therefore, this dissertation studied these two traits for increased seed quality and more industrial applications of soybean meal.

Additionally, plant breeding encompasses all aspects of the particular species that includes enhancing traits of current varieties as well as diversifying uses or improving its current uses. This dissertation combines both breeding techniques that are used for developing new varieties as well as examining how we can improve its usage. Specifically, it looks at how soybean seed quality can be improved by performing a genome-wide association study (GWAS) to identify genetic markers related to high test weight for marker-assisted selection in breeding. It also utilizes gene editing to develop soybean lines that have lower amounts of TI in seed, healthier for human and animal consumption. Lastly, it evaluates the performance of Rainbow trout fed diets using low-TI soymeal to see if cheaper and more sustainable aquafeeds can be developed using value-added soybean to increase soy's presence in aquaculture

Chapter 1: Literature Review: Book Chapter

Soybean: Production, Versatility, and Improvement

Zachary Shea, William M. Singer, Bo Zhang

Abstract

Soybean is one of the most widely planted and used crops in the world due to its high protein and oil content. The many significant agronomic practices that are utilized in soybean production are highlighted with emphasis on those used during the pre-growing season and growing season. The various pests of soybeans and the pest management strategies used to control them are described with special attention to insects, weeds, bacteria, fungi, and nematodes. The multitude of soybean uses for livestock and human consumption as well as its industrial uses are discussed in this chapter. Additionally, the conventional breeding and genetic engineering attempts to improve soybean protein, oil, and sucrose content as well as eliminate the antinutritional factors, trypsin inhibitors, raffinose, stachyose, and phytate are examined.

Keywords: Agronomic practices, pest management, soybean uses, breeding, genetic engineering

1. Introduction

Soybean (Glycine max) is one of the most valuable, versatile, and nutritionally important legumes globally. It can be grown in a multitude of environments, using a variety of management practices, and for diverse end-user purposes. In 2018, roughly 398 million tons of soybeans were produced worldwide which accounted for 61 percent of overall oilseed production and 6 percent of the world's arable land use [1–3]. The United States, Brazil, and Argentina constituted approximately 81 percent of international soybean production, producing 34, 32, and 15 percent respectively [4,5]. Soybean seed composition and its main components, meal and oil, are the driving forces behind crop production that has increased nearly three hundred and fifty percent since 1987 [5]. Soybean meal is intricately connected to the food supply through direct food consumption and

indirect consumption as a large source of livestock feed. Soy oil provides greater versatility with uses in food and beverage, wax, construction, cosmetics, plastics, and fuel.

Soybean originated in East Asia and has been cultivated in China for millenia. It is estimated that the domestication event from wild soybean (Glycine soja) occurred during the Shang Dynasty, 1700-1100 B.C. [6]. While no longer the largest producer, China and other Asian countries continue to incorporate large quantities of traditional and innovative soy foods into their diet. In 2018, China was the largest customer for the United States whole soybeans, importing over $3 billion worth [7]. The United States and western countries mainly utilize soybean indirectly in the food supply as livestock feed and food ingredients such as textured vegetable protein and protein isolates. However, as more consumers are looking for plant-based protein in their diet, soy foods will become a globally viable alternative to animal protein. As the soybean appetite has increased and transformed, scientific developments have also improved soybean production through agronomic, management, and genetic methods to meet demand.

2. Agronomic Practices

2.1 Pre-Growing Season Practices

Soybeans are one of the most flexible crops in terms of production methods, geographical growing regions, and end use versatility. Therefore, there are multiple agronomic practices to consider when preparing a field for soybean production. While tillage and fertilization practices are common among producers, technique specifications can vary greatly due to preferences, environmental conditions, and cost. Historically, mechanized and non-mechanized tillage was considered a vital practice to maximize crop yield and value [8]. While tillage is still a useful tool, contemporary research has corroborated the dangers of over-tilling and the potential benefits from soil conservation and no-till operations. No-till practices and conservation tillage for soybean are wide-spread in areas of highly-erodible soil, and some research has shown that soybean yields remain the same or increase with decreased tillage [9–12]. However, other research has shown that rotational tillage practices will provide higher crop value than no-till practices, specifically because of herbicide costs and equipment requirements [13–15]. Given the need for proper soil maintenance, conservation tillage (<30% crop residue left on the soil surface) is a popular

compromise, especially in herbicide tolerant soybean production [16,17]. Research has further elucidated the benefits of conservation tillage and no-till practices on soil health by showing positive correlations with rhizobial and nematode populations [18,19].

Pre-plant fertilization for a variety of macro and micronutrients is another common practice in soybean production. Soil fertility programs are designed to provide sufficient nutrients for a crop's needs which maximizes crop yield and farm efficiency while also minimizing environmental impact. To prepare a field for soybean planting, a farmer must start by determining what nutrients are already present in the soil; this can be accomplished by a variety of soil sampling and analysis methods [20]. The primary macronutrients, nitrogen (N), phosphorous (P), and potassium (K), should be examined first alongside critical secondary macronutrients and micronutrients such as sulfur (S), calcium (Ca), magnesium (Mg), zinc (Zn), manganese (Mn), boron (B), iron (Fe), and copper (Cu). General field nutrient requirement guidelines for soybean production are summarized in Table 1. While soybeans require a large amount of nitrogen, fertilization is usually unnecessary because of the symbiotic relationship with Bradyrhizobium japonicum, a bacteria that performs nitrogen fixation and provides plant nitrogen [9,21,22]. Depending on field conditions, 25 to 75% of nitrogen in mature soybeans can originate from symbiotic nitrogen fixation [23]. Excess nitrogen has been linked to negative plant physiological conditions and inhibited rhizobia activity [24]. Thus, the best solution to limited field nitrogen is sometimes soil or seed-applied bacterial inoculation [21,22,24]. Soil pH is also a vital component of field management. It is well-documented that all nutrients have varying availability to plants depending on pH [22,25,26]. Generally, soybeans prefer a slightly acidic soil ranging from 6 to 7 pH [9,22]. Liming a field is the optimum technique to raise pH, while the most common practice for lowering pH is elemental sulfur application.

While yield is the driving factor for fertilization, recent market changes have adjusted soybean valuation with increased focus on seed composition quality. Amino acid profiles as descriptors for protein quality in human food and livestock feed as well as high oleic acid soybeans for increased functionality and performance are just two examples of possible premiums producers can receive through soybean seed composition. Research has shown that agronomic practices coupled with location-dependent, environmental variables can directly impact those premiums [27–29]. Nitrogen fertilization plays a limited role in seed composition as it is rarely needed due to the bacterial nitrogen fixation. However, excess nitrogen has been shown to decrease the levels of

sulfur-containing amino acids and has an inconsistent effect on fatty acid concentrations [30,31]. Phosphorous applications can increase protein quantity without adjusting the amino acid profile but also has a positive correlation with higher phytic acid and isoflavone concentrations [32–34]. Additionally, phytic acid has been shown to increase alongside zinc concentrations [32]. Pre-plant potassium applications have limited return on investment in regards to yield and seed composition, however, potassium-deficient soybean plants are at a greater risk to insect pests, specifically aphids [35,36]. Limited yield response is observed with sulfur applications. Although, researchers have found the use of sulfur fertilizers to be economically viable, particularly on coarse soils [9,21,37]. Soil sulfur levels have also been shown to greatly impact the ratio between 11S and 7S seed storage proteins [38]. As markets continue to change and value differing soybean seed compositions, it will be critical for producers to fertilize with both yield and seed components in mind.

2.2 Growing Season Practices

Soybean producers make decisions throughout the year that impact the final yield, value, and profit from their annual crop. Many of the most critical decisions occur at the beginning of and throughout the growing season. From the moment a soybean seed is planted to harvest, producers choose (or decide against) a multitude of practices including crop rotation, row spacing, population density, irrigation, post-emergence fertilization, and pest management. Maximizing a potential soybean crop is directly connected to previous field usage. Crop rotation or the process of growing different crops in sequenced seasons within the same field is a common practice in soybean production. Corn (Zea mays) and soybean rotations are advantageous because of corn's high nitrogen demand which can be alleviated through Bradyrhizobium japonicum nitrogen fixation in soybean nodules. Corn and soybean rotations also exhibit beneficial energy balance and grain yield improvement [13,39]. Rotations including corn and soybeans as well as wheat (Triticum), oats (Avena sativa), barley (Hordeum vulgare), cotton (Gossypium), and forageable pasture have also shown potential for economic and environmental gains [22,40,41]. Although depending on crop sequences, new management practices may be needed. For example, alfalfa or clover following soybeans would require liming for maximum production as those crops prefer a slightly higher pH [42].

Protecting and revitalizing the soil through non-harvested crops planted between soybean growing seasons or cover cropping is also beneficial. Cover crops protect the soil that would otherwise be farrow and replenish nutrients assimilated into the soybean plant [9,22,42]. Furthermore, cover crops can beneficially reduce weed pressure, lessen soil compaction, and improve water conservation [43–45]. However, cover crops increase annual cost and have not been shown to increase soybean yield which can negatively impact certain producer's net profit [46]. Many farmers who receive enough growing degree units throughout the year also limit farrow fields by double cropping with soybean. Soybean and wheat double-crop systems have exhibited high economic returns for producers in both field and modeling research [47,48] Double cropped soybeans exhibit lower yield due to late planting and decreased leaf-area-index potential, but this can be mitigated with early maturing varieties [49,50]. Intercropping or growing at least two crops simultaneously is another, less-common option for soybean production. Corn and soybean intercropping can increase yields for both crops with the proper seeding rates [51,52]. Wheat and soybean intercropping also displays a positive yield response [53,54]. Sugarcane (Saccharum officinarum) and soybean interspecific relationships increase sugarcane yield and improve rhizospheric activity while reducing soybean yield [55]. While intercropping can enhance value for soybean producers, it is unsuitable for most large-scale production systems.

After choosing a cropping system, soybean producers must then determine the proper row spacing and population density for their environment. The appropriate balance between row space and plant density is critical for maximum soybean production and reliable economic returns. Narrow rows and high plant densities both correlate with quickened canopy closure and weed suppression [56–59]. Increased plants per field also increases cost; however, subsequent increased yield and profit overcomes the cost [60–62]. As soybeans emerge and grow, the next consideration for producers is irrigation. This localized decision can be based upon historical precipitation records, predicted forecasts, day-to-day weather events, or a combination of factors. In the absence of natural precipitation, irrigation is vital to soybean production as water deficiencies inhibit yield potential [63–66]. Irrigation can also be optimized spatially throughout a field with variable rate techniques and temporally across the growing season by targeting specific growth stages [67,68]. Fertigation applications can be used to combine applications of post-emergent fertilizer with irrigation. Other methods of post-emergent fertilization including foliar spray and direct-to-soil applications are more common solutions for growing season nutrient issues. Plant tissue sampling

and analysis can be coupled with soil samples to determine in-season soil deficiencies and to prescribe further applications [9,21,25]. As soybean increases nitrogen uptake during reproductive stages when bacterial fixation may be diminishing, soil or foliar nitrogen applications are typical yet usually ineffective. While limited yield increases can be seen from supplemental nitrogen applications or various nutrient combinations, the economic returns generally fail to cover the cost of application [69–72]. Foliar nutrient applications have shown minor impacts on seed protein and oil content, however these results are inconsistent amongst experiments [73,74]. A location-specific, comprehensive nutrient management plan that accounts for all other agronomic practices is the best method for maximizing yield and economic returns in soybean production.

3. Pest Management

3.1 Insect Pests

Insect pests of soybean vary greatly ranging from aphids to stinkbugs to loopers to beetles. Which insects are the major pests and the impact that these pests can have on soybean varies significantly from year to year and depending on the region where soybean crops are grown. Total damage by insects is a little ambiguous but yield loss of up to 80% has been reported [75]. Some prominent insect pests include soybean aphids (*Aphis glycines*), Japanese beetle (*Popilla japonica Newman*), Mexican bean beetle (*Epilachna varivestis Mulsant),* two-spotted spider mites (*Tetranychus urticae*), brown marmorated and red banded stinkbug (*Halymorpha hals* and *Piezodorous guildinii*), bean leaf beetle (*Cerotoma trifurcata*), and kudzu bug (*Megacopta cribraria Fabricius)* [75–80].

Insecticides constitute a large portion of insect management as they are used to control most insect pests and in some cases are the primary method of control [79]. Integrated pest management (IPM) is becoming more common among growers due to its ability to reduce pesticide use, non-pests affected, workers exposure to pesticides, and the likelihood of insects developing resistance to the insecticides [81,82]. Additionally, it has been found to be effective at reducing the damage done by pests [83]. Integrated pest management works similarly for all pests. It involves monitoring fields to determine which pests are present, then determining which pesticides can and

should be used, and incorporating cultural management practices [82]. For insects, trap cropping and sweep nets are used to monitor and determine which insect pests are present [84,85]. The cultural practices used in insect management were mentioned in table 2 and they include altering planting date and row spacing, using no-till fields, and using resistant soybean cultivars [75–80].

3.2 Weeds

Weeds are considered one of the most damaging, if not the most damaging pests, in soybean [86]. About 37% of global production of soybean is affected by soybean, while 23% of global production is affected by other pests [87]. In the United States alone it has caused losses of several million US dollars each year [86]. Weeds pose a problem for soybean crops since they compete for nutrients, space, and other resources [88]. Many different weed pests compete with soybean, some of which include Common Waterhemp (Amaranthus rudis), Canadian Horseweed (Conyza canadensis), Giant Ragweed (Ambrosia trifida), Ivy-leaf Morning Glory (Ipomea hederacea), Common Cocklebur (Xanthium strumarium), Johnsongrass (Sorghum halepense), and pigweed (Amaranthus spp) [89,90]. It is important to note that which weeds are found in a particular field depends largely on where the soybean crops are grown.

Management of weeds is largely done through integrated pest management. This involves using herbicides along with herbicide-resistant soybean varieties and cultural practices [86,91]. There are many different classes of herbicides that range from enzyme inhibitors, lipid synthesis inhibitors, photosystems diverters, nucleic acid inhibitors, and auxin inhibitors [92]. Herbicides have been a large part of weed management and will most likely continue to be a large part of it due to their effectiveness and the other practices alone not being enough [86]. Herbicides' effectiveness are only improved given the use of herbicide-resistant soybean, such as glyphosate-resistant Roundup Ready soybean. Since weeds can develop resistance to herbicides, it is important to incorporate other management practices [86]. One such method is when to spray herbicides. A common management practice involves pre- and postemergence herbicide applications. This involves spraying herbicides before and a few days after the soybean plants have emerged to reduce any damage to the soybean plants [93]. Additionally, cultural control practices are used. Examples of these include crop rotations, planting in narrow rows and proper fertilization to promote crop competition, and cultivation [91]. Crop rotations allow for different herbicides to

be used which in turn helps to prevent the development of herbicide-resistant weeds [91]. Promoting crop competition allows soybean plants to grow enough to create a canopy to maximize the shading of weeds [91]. Cultivation is an effective and economical way to control weeds to help minimize herbicide use [91]. All of the aforementioned management practices are parts of integrated weed management and will continue to play a significant role in the control of weeds.

3.3 Diseases

Similar to insect pests, there is a wide variety of diseases in soybean. Most of the diseases are caused by fungal and bacterial diseases. Fungal diseases have been known to reduce yield up to 50%, while bacterial diseases have been known to cause yield loss of anywhere between 15-60% [75]. Which disease is the most devastating depends on the region and the year, but the most prevalent diseases include *Heterodea glycines*, *Phytophthora sojae*, Colletotrichum truncatum, *Septoria glycines,* and *Phakospora pachyrhizi* [75]. Of these five diseases, *Heterodea glycines*, or soybean cyst nematode is the most economic damaging disease being found in all countries that grow soybean and causing up to 90% yield reduction in some areas [75]. Table 3 provides an overview of some of the main soybean bacterial, fungal, and Oomycetediseases.

Disease Name	Bacterial, Fungal, Nematode	Causal Agent	Mode of Transmission	Symptoms	Current Treatment Strategies
Charcoal rot	Fungal	*Macrophomina phaseolina*	Soil born	Wilting, necrosis, black/dusty microsclerotia on stem/pods/seeds, brown lesions on emerged seedlings	Fungicides, resistant cultivars, reduce tillage, crop rotations,
Soybean Cyst Nematode	Nematode	*Heterodera glycines*	Soil born	Stunted roots, can increase sensitivity to some fungal diseases, presence of cysts on roots	Resistant cultivars, crop rotations,
Phytophthora Root and Stem Rot	Oomycete	*Phytophthora sojae*	Can overwinter in soil, water	Reddish-brown/black lesions on stem, chlorotic leaves, soft rot on roots, seed rot, emergence damping off of seedlings	Seed treated fungicides, improving soil drainage, resistant cultivars
Soybean Bacterial Blight	Bacterial	*Pseudomonas syringae pv. glycinea*	Water, can overwinter in plant residue	Affects mid to upper leaves, yellowish-brown, angular	Copper fungicides, resistant

				lesions on leaves, discolored/shriveled seeds, water soaking	cultivars, increasing tillage, crop rotations
Soybean Anthracnose	Fungal	*Colletotrichum spp*	Seedborn, can overwinter in plant residue	Brown lesions with setae, pods with fewer seed, brown cankers, defoliation, damping off	Crop rotation, fungicide treated seed
Brown spot/ Septoria Brown Spot	Fungal	*Septoria glycines*	Can be transmitted through infected seed, can overwinter in plant residue	Small, brown lesions on leaves, yellowing leaves, lesions contain pycnidia	Foliar fungicides, crop rotation, increased tillage
Soybean Rust	Fungal	*Phakospora pachyrhizi*	Spores are spread by wind	Reddish/brown lesions with pustules on leaves, pods, and seeds	Fungicides

Table 3: Overview of 8 Prevalent Diseases in Soybean [94–101]

From Table 3, it is evident that chemical pesticides still play a large role in treatment strategies against all of the main diseases in soybean even though there has been rising interest to incorporate other methods to prevent and treat diseases in soybean due to the harmful environmental and health effects of pesticides. Some of the other methods to control soybean diseases are seen in cultural control practices, such as increasing or decreasing tillage and crop rotation, drainage, and using resistant cultivars [102]. While the treatments listed in the above table are usually at least somewhat effective, some research have been done to continue to find more ways to improve the control of plant diseases. One such example is the development of using hyperspectral bands for the early detection of charcoal rot in soybean [103]. These researchers developed a method that involves analyzing spectral and spatial information of infected and healthy soybean in order to find wavebands that signify a soybean plant that is infected with charcoal rot [103]. This process identified six wavebands that were specific to plants infected with charcoal rot and can potentially allow for the detection of charcoal rot in crops in three days [103]. By being able to identify disease earlier, growers can minimize the damage done by that disease by removing infected plants and incorporating treatment strategies, such as pesticides or cultural controls.

The research above shows that there is interest in developing early detection for soybean pathogens. One of the other major areas of research for soybean diseases, is identifying resistance genes to promote resistant cultivars. Given that soybean cyst nematode is one of the most devastating soybean diseases, if not the most devastating one, there has been a lot of research done

to identify genes involved with resistance to soybean cyst nematode. The main resistance gene in soybean to cyst nematode is the Rhg1 gene, which encodes an amino acid transporter [104,105]. This gene confers partial resistance and has been shown to reduce the reproduction of soybean cyst nematode and improve yield in fields that are infected with soybean cyst nematode [77]. However, identifying resistance genes and having resistant cultivars do not permanently stop diseases. For soybean cyst nematode, it is advised to utilize cultural practices, such as using multiple resistant cultivars and rotating with non-host crops that are resistant to cyst nematode, and other methods [106]. This just shows that integrated pest management involves continuously incorporating new methods to control diseases to prevent the disease from overcoming any pesticides and resistant cultivars.

4. Soybean Utilization and Products

4.1 Livestock Feed

Soybean is a valuable crop worldwide mainly because of soybean meal's nutritional efficacy as a food and feed ingredient. A high protein content, balanced essential amino acid profile, and the presence of other beneficial nutrients all contribute to its economic and nutritional value. Soybean meal constitutes 70 percent of seed value while only being roughly 35 percent of seed dry weight [107,108]. Furthermore, in the United States, 97 percent of soybean meal is used for livestock feed [108]. This overwhelming usage rate as a livestock protein source is mainly due to the presence of essential amino acids. While some livestock require other amino acids, most livestock need nine essential amino acids: histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine [109]. All nine of these amino acids are found in some quantity in soybean meal [110]. For this reason, soybean meal can maximize livestock production in cattle, swine, poultry, and aquaculture. Generally, soybean meal and other soy byproducts use are limited to a supplementary or finishing role for cattle due to feed ration complications from other seed components [109,111,112]. Soybean meal use is highly prevalent in monogastric livestock production such as swine and poultry and is increasing in popularity for aquaculture [113,114]. However, soybean as feed has two main obstacles: methionine deficiency and trypsin inhibitor proteins. Albeit present in soybean, methionine content is deficient for livestock needs, is

considered the first limiting amino acid for soybean meal, and requires producers to supplement with synthetic methionine [115–119]. This has a variety of negative economic and environmental impacts, including increased cost and poor nitrogen-use efficiency [120,121]. Trypsin inhibitor proteins are an antinutritional factor present in raw soybean that decreases feed efficiency and can harm livestock. There are a variety of industrial processing methods used to overcome trypsin inhibitors in soybean such as thermal and infrared treatment [122]. In the future, soybean methionine deficiencies and trypsin inhibitor levels may both be solved via breeding and transgenic efforts.

Besides being used in livestock feed, soy meal has also been used in aquaculture to feed a range of species. Soy meal has been used in aquafeeds to replace fish meal due to its low cost of production and excellent nutrient profile [123]. Unfortunately, soy meal has been unable to fully replace fish meal in aquafeeds due to the presence of antinutritional factors [124]. There have already been multiple studies looking at how to increase soy's usage in aquafeeds. All have looked at the effects of replacing fish meal with soy meal at increasing rates. They have found that up to roughly 30% inclusion rates, soy meal does not have negative effects on the growth of Rainbow trout and salmon, but going higher than 30% can result in decreased growth [125–129].

4.2 Human Food

Soybean as human food exists to two different extents that are derived from geography and cultural tradition. Eastern hemisphere populations incorporate whole soybeans and processed soy foods into their lives daily, whereas Western hemisphere populations generally utilize processed soybeans as food ingredients. Eastern soy foods are divided into two main categories: fermented and non-fermented. Non-fermented soy foods include whole seed options such as whole dry soybeans, soy nuts, and edamame, processed items such as soy flour and soy milk, and vegetative soy sprouts [114,130]. Soy milk in its simplest form is a water extract from soybean that when further processed can make tofu and tofu byproducts such as okara (soy pulp) and yuba (tofu skin). Fermented soy products include miso, soy sauce, tempeh, natto, and tofu, and each product has a specific bacterial species that enables proper fermentation. For example, natto is associated with Bacillus subtilis, and soy sauce is associated with Aspergillus sp. [114,130]. Western cultures have assimilated many soy food products, and they are becoming more popular as consumers seek plant-

based protein sources. However, the vast majority of soybeans in western diets consist of food ingredients made from soybean meal and soy oil. Soybean meal can be processed into ingredients such as soy flour, protein concentrates, and protein isolates that are used in bakery mixes, breakfast cereals, baby food, and exercise supplements [113]. Soybean oil is widely used in vegetable oil and margarine mixes for a variety of cooking purposes. The importance of traditional and innovative soy food uses has been perpetuated because of the potential health benefits from soy consumption. Soy foods have been shown to play a role in chronic human disease prevention for conditions such as heart disease, osteoporosis, and cancer [131–133]. However, isoflavones, one of the most common seed components linked to disease prevention, are also negatively linked to hormonal health as a phytoestrogen. While large population subsets are concerned about isoflavones negatively impacting fertility, summarized data has shown inconsistent results [134–136]. As consumers continue to seek plant-based protein, soybeans will be the premier source for historically and culturally significant recipes as well as healthy, novel animal meat alternatives.

4.3 Industrial Uses

Even though soybean is classified as an oilseed, soybean oil has historically been an afterthought for soybean producers and processors. When markets for soybean meal would falter, researchers and other stakeholders would turn to soybean oil for added value or seek alternative uses for meal components. Modern sustainability and industrial goals have stimulated soy-based product usage in a variety of fields, as summarized in Table 4. Soybean oil as biodiesel has experienced the largest growth with United States consumers using over 2 billion gallons in 2017 [137]. Current biodiesel production methods can create soy-based fuel that performs nearly equal or equivalent to standard diesel fuels and have the potential to become a truly renewable resource when coupled with sustainable farming practices [138–140]. Constantly improving processing methods will continue to augment soybean seed component versatility and create new opportunities for soy-based products.

Test weight is a trait that can indirectly affect soybean's industrial uses. Test weight is the amount of soybean seed in pounds that fits into one bushel [141]. It is important because the higher the test weight the better quality the seed, the stronger the seed coat, less prone to getting mold, and the longer the seed can last in storage [141]. Due to this, test weight can also affect the pricing

that growers receive for their soybean since growers will receive a pricing penalty if their test weight drops below 54 lbs/bu [142]. With this in mind, it is integral for soybean seed to have higher test weight because the higher the test weight the better quality the seed which can ultimately lead to better end product quality for industrial uses. It was previously mentioned that test weight can affect the pricing of soybean and if it drops below 54 lbs/bu, growers can receive a pricing penalty. In the past, soybean test weight was typically observed to be around 60 lbs/bu but since the 1950s, growers have noticed a decreasing trend in test weight [141]. If this trend continues growers will start to receive more pricing penalties. Despite this not much has been done to prevent this in terms of breeding yet. Therefore, it is expected that test weight will receive more attention in the future so that soy varieties that will have higher test weight will be developed to protect farmer's profitability.

5. Soybean Seed Composition Improvement

5.1 Breeding Efforts

Soybean seed has many beneficial traits, such as high protein, oil, and soluble sugar content [143]. While soybean seed has a lot of favorable qualities, there have been breeding attempts to improve them. With regards to protein content, breeders have worked with soybean to increase total protein content as well as the amount of sulfur-containing amino acids, methionine and cysteine. [143]. Methionine and cysteine are of interest since the seed protein is lacking these two amino acids and these two amino acids can improve the nutritional value of soybean meal [143,144]. The research indicating that total protein content is negatively correlated with other favorable seed qualities, including yield, oil content, and potentially methionine and cysteine content [145,146], provides a stronger push to increase methionine and cysteine content. So far most of the breeding efforts that have been done to improve protein quality have involved identifying quantitative trait loci (QTL) that are associated with the amino acid content [147]. QTLs are regions of DNA that are associated with a particular trait and they allow breeders to select for particular cultivars that have a trait of interest allowing them to breed for plants that will have the trait of interest [148]. The composition of soybean seed oil primarily includes linolenic, steric, palmitic, linoleic, and oleic acid [149]. There have been breeding attempts mainly to

increase oleic acid in soybean seed while keeping linolenic acid relatively low [149]. Breeders want to keep oleic acid high due to it being beneficial for human health [150]. There has been work done involving the other fatty acids but those will be talked about in the next section. Lastly, the valuable, soluble sugar content, specifically sucrose content, has also been an area of interest in soybean breeding [143]. While sucrose is the main sugar found in soybean, fructose and glucose are also present but in trace amounts [151]. Similar to protein content, multiple QTLs have been identified associated with high sucrose [152]. Sucrose is a desirable seed composition trait due to it improving flavor of soy food products for human consumption [152]. Overall, conventional breeding has been used to improve protein quality, oil content, and sucrose content in soybean seed.

While soybean does have numerous profitable seed traits, it also contains several unfavorable traits that include trypsin inhibitors/TIs, indigestible carbohydrates, and phytate [153,154]. There are two trypsin inhibitors found in soybean, which include the Kunitz and Bowman-Birk trypsin inhibitor, and they are antinutritional factors due to their ability to interfere with protein digestibility and reduce the health of animals that are fed soybean meal containing these proteins [155,156]. Currently, growers can heat the soybean meal in order to inactivate the trypsin inhibitors, but this step is costly [157]. Due to TIs negatively affecting animal health and the costly heating steps that are required to remove them, more and more breeding efforts are being made to develop low-TI soybean lines [158]. Besides TIs, the indigestible carbohydrates, raffinose and stachyose, that are found in soybean seed are also a target for soybean breeding since they can cause flatulence and diarrhea when consumed [152]. There has been progress made in lowering these carbohydrates, which includes identifying QTLs associated with raffinose and stachyose [152]. Lastly, while phytate is an antinutritional factor found in soybean, there is not a lot of work being done to breed low phytate soybean lines since phytase supplements are an effective, inexpensive way to reduce the phytate found in soybean meal [159].

5.2 Genetic Engineering Efforts

Genetic engineering involves the process of artificially and intentionally manipulating the DNA of an organism with the purpose of modifying that organism [160]. Some of the methods used to transform plants include Agrobacterium, electroporating plant protoplasts, and microparticle

bombardment [161]. One relatively new field within genetic engineering is gene editing which involves using clustered regularly interspaced short palindromic repeats or CRISPR/Cas9 system [162]. Genetic engineering works by introducing a gene from one organism into another organism so that it can now express that gene product or by causing frameshifts or deletions to knockout a particular gene in an organism [160]. CRISPR/Cas9 has been gaining a lot of attention due to its promising ability to efficiently and effectively improve agronomic traits in crops [163,164]. Genetic engineering in soybean was first successfully accomplished in the 1990s [165]. Since this time, genetic engineering has been used a lot in soybean with about 90.7 million hectares of genetically modified/GM soybean being planted in 2014 [165]. Most of this genetic engineering has been done to create Roundup Ready soybean that is resistant to glyphosate herbicides [166]. Roundup Ready soybean has been used a lot because it allows growers to spray herbicides to kill any weeds in the field while not killing the soybean [166]. Genetic engineering has been used to additionally improve the protein quality of soybean by expressing zien proteins from corn to increase sulfur-containing amino acids and by altering biosynthetic feedback pathways to increase lysine [114,167]. Besides these examples, genetic engineering has been used to manipulate soybean oil content by increasing oleic acid content and decreasing linolenic acid content and to delay flowering time in soybean [168,169]. Given the ability of genetic engineering, especially gene editing, to successfully improve the qualities of soybean, it will most likely begin to be used to improve soybean by removing and/or modifying the expression of antinutritional factors. This can be accomplished through genetic engineering by knocking out particular genes responsible for the antinutritional factors preventing them from being expressed.

6. Conclusion(s)

Soybean is an essential crop that is grown globally due to its various and diverse uses. Given its importance, there are many pre-growing practices to prepare the field for the growing season, including tillage, pre-plant fertilization, and monitoring soil pH. Many agronomic aspects must be considered during growing season to ensure successful soybean growth including crop rotations, double cropping, cover crops, irrigation, row spacing, plant density, post-emergence fertilization. Additionally, integrated pest management involving the use of pesticides, resistant soybean cultivars, and cultural practices are vital to control the numerous pests of soybean. While soybean

is highly used in livestock due to its high protein content, it’s methionine deficiency and presence of antinutritional factors still present problems that need to be solved. Soybean versatility is highlighted in the many ways that it can be used for human consumption, biofuels, and other industrial uses. Due to the previously mentioned negative qualities of soybean, conventional breeding has been working to increase protein and oil content, while eliminating antinutritional factors. Genetic engineering and gene editing show promise to help improve soybean by introducing genes to improve protein and oil quality and knocking out genes to remove antinutritional factors.

Chapter 2:

Determining Genetic Markers and Seed Characteristics Related to High Test Weight in *Glycine max*

Zachary Shea, William M Singer, Luciana Rosso, Qijian Song, Bo Zhang

Abstract

Test weight is one of the primary indicators of soybean seed quality and is measured as the amount of soybean seeds in kilograms that can fit into one hectoliter. The price that growers receive for their soybean is dependent on test weight. This study examined the Southern Core Collection (SCC) grown in Blacksburg and Warsaw Virginia from 2019-2021 to identify molecular markers and seed composition traits associated with high test weight. A 2500 UGMA was used to measure test weight, and a DA7250 Near-Infrared Reflectance Spectroscopy (NIRS) instrument was used to measure protein and oil content. In addition to oil and protein content, the content of three sugars, sucrose, raffinose, and stachyose, was measured using HPLC. The TASSEL program and R were used to conduct a GWAS analysis to identify molecular markers related to high test weight. The Pearson's correlation coefficient was calculated to determine the relationship between test weight and the seed composition traits. Test weight values ranged from 62-77 kg/hL over the three years. Multiple single nucleotide polymorphisms (SNPs) significantly related to high test weight were found on chromosome 15 and eight candidate genes were found near these SNPs. Test weight was found to have a significant, negative correlation with oil content, inconsistent correlation with protein content, and negative correlation with all three sugars in both Blacksburg and Warsaw. Our GWAS results suggest there might be a QTL associated with high test weight and potential genes that control test weight located on chromosome 15. The correlation between test weight and oil content suggests that as we have been selecting for higher protein and oil in soybean seed, we may have been unintentionally selecting for lower test weight.

Introduction

Overview of test weight in soybean

Soybean (*Glycine max* (L.) Merr.) is the most important oilseed crop due to its low-cost production, and diverse uses in feed and food for its excellent nutrient profile and in paints and biofuel for its high oil content [170]. The United States has been one of the top two soybean producers for decades in the world [171]. Seed quality is one of the primary factors that affect the price that growers can receive for their soybean. Within this, multiple aspects can affect seed quality, including test weight, the amount of diseased seed, and damaged or disfigured seed [141]. Of these, test weight is the most important indicator of soybean seed quality because soybeans with higher test weight last longer in storage, having higher seed integrity, and being less prone to mold [141]. Additionally, every crop has a standard value for test weight, and if the crop's test weight is below the standard, the grower can receive a pricing penalty [142]. The standard for soybean is 60 lb./bu., and below 54 lb./bu can cause a pricing penalty [142]. This is important because farmers have noticed a slightly decreasing trend in test weight for soybean since the 1950s. While in the 1950s almost all soybean had test weight around 58-60 lb./bu., farmers are now seeing more test weight that is closer to 56 lb./bu. [141]. Even though this is still above the 54 lb./bu. limit, if this trend continues, farmers will start to see soybean test weight values at or below 54 lb./bu. and will be more likely to lose profit.

Breeding efforts to increase test weight

Although test weight is important in soybean, not much work has been done looking into the effect of genotypes on test weight in this crop. Some work has found that genotypes showed a significant effect on test weight in maize [172], wheat [173], and oat [174]. Recently, in soybean, genotypes were found to have a significant effect on test weight in most locations but not all depending on the environment [175]. With that said, all these findings indicate that it is possible to improve test weight through breeding. Despite this, little to no work has been done regarding how to assist breeding high-test weight soybean varieties. In order to address the concern of a decreasing trend of test weight in soybean over time, there is an urgent need to identify, validate and utilize genetic markers associated with high test weight in soybean breeding as a quick and effective approach. Thus, breeders will shorten the period of development and selection of soybean varieties for increased test weight to ensure the farmer's profitability.

Correlation with other traits

While test weight is important in most crops, other traits are also important in plant breeding and must be taken into consideration, such as protein, oil, and sugar content in wheat and soybean. Therefore, it is necessary to gain a better understanding of the relationship between test weight and seed compositions. Although some work has been done previously determining this relationship, reports have had varying degrees of significance. In wheat, some studies found that protein content had a significant, positive relationship with test weight [176,177], while others found a significant, negative correlation [178] or no significant relationship at all [179]. Oil content was found to have a significant positive correlation with test weight in sunflower [180] and no significant effect on test weight in oats [181]. For soybean, protein and sucrose content were found to have varying degrees of significant relationships with test weight, varying from positive to negative to not significant, while oil content and test weight had a significant, negative correlation in all experiments [175]. Test weight is known to be heavily impacted by the environment, but none of the research has been conducted in the Mid-Atlantic region, so it is important to determine the relationship between test weight and these traits in Virginia. Additionally, little to no work has been done to understand the relationship between test weight and raffinose family oligosaccharides (RFOs). While these two sugars are not as major as sucrose, reduced RFOs is a target when breeding new varieties for animal feed and human food.

Materials & Methods

Plant materials

All accessions were planted similarly to Singer et al. 2022 [182]. Briefly, a total of 390 soybean accessions from a total of 17 countries from the Southern Core Collection that has been maintained at Virginia Tech was grown in Blacksburg (BB) and Warsaw (W), VA for three years from 2019 to 2021. This panel contains soybean accessions that were grouped by maturity groups and all accessions belonged to either maturity groups IV and V. At both locations for all years, each sample was planted in two replications in 3 m two-row plots with 76 cm row spacing for each plot in BB and 3 m four-row spacing with 76 cm row spacing in W with cultivars Ellis and AG 4404 being used as commercial checks. All plots were checked for flower color and pubescent color and any plants that did not match the correct color for that plot were removed. All plots were planted in the beginning of May in both locations and all years. In 2019, all plots

were harvested around the beginning of October, but in 2020 and 2021 plots were not harvested until later October due to weather conditions.

Test weight

All samples were cleaned to remove split seeds, empty seed coats, pods, and sticks so that only intact seeds remained prior to measuring test weight. Sample seeds were also checked for seed coat and seed hilum color, which served as quality control to remove any contaminant seed. Blacksburg had 344, 311, and 241 accessions for 2019, 2020, and 2021, respectively. Warsaw had 313, 266, and 274 accessions for 2019, 2020, and 2021, respectively. Test weight was determined by using a 2500 AGRI model. For each sample, 414 mL was used to calculate the test weight. All samples were measured three times and the resulting test weight values were averaged to get the final test weight value for the sample. For 2019, due to the limited seed amount, the two replications had to be combined into a composite sample prior to measuring, but we were able to average two replications for a mean of test weight of each accession in 2020 and 2021. Test weight values were adjusted to a 13% moisture content according to Liu, et al., 2019 [175]. Lastly, test weight values were converted from lb/bu to kg/hL by multiplying the lb/bu values by a factor of 1.25.

Genome-wide association study (GWAS)

The genotypic data of all accessions screened by SoySNP50K iSelect Beadchip are publicly available at soybase.org (Song et al. 2013[183]) with a total of 35,570 SNPs. A GWAS was performed by first using the TASSEL 5.0 software to construct a kinship matrix, principal component analysis, and mixed linear model (MLM) [184]. The MLM was used to include a kinship matrix (K) with population structure (Q) to improve the statistical power using the Q+K approach [185]. A modified Šidák correction ($\alpha sid = 1-(1-\alpha)(1/m)$) for multiple testing was used to determine any significant markers, with the number of effective markers (Meff) being used instead of the total number of markers (m). Meff was determined to be 4,191 through the poolr package in R with the Li and Ji method [186]. A modified significant threshold at $\alpha = .05$ was constructed at $-\log10(P) > 4.91$ to determine single nucleotide polymorphisms (SNPs) significance. The qqman package was used to construct QQ and Manhattan plots [187]. In order to find candidate genes, genes located within 10kbp of significant SNPs were found through the Soybase database [188]. Many SNPs were found to be significant across the locations and years.

To limit the search, only SNPs that were found to be significant or close to being significant in multiple environments were used.

Protein, Oil, And Sugar Content

To measure protein and oil content, near-infrared reflectance spectroscopy (NIRS) was performed. Each sample and replication were run twice on the DA7250 NIR Analyzer from Perten, and the protein and oil content were averaged for the sample. For each sample, about 60 mL was used. Only samples that had a yellow or green seed coat could be used since the NIRS program was developed for seed for these seed coats and could not accurately measure protein and oil content for other seed coat colors. For both protein and oil content, 228 samples were used in both locations for 2019, 233 in both locations for 2020, 328 in BB2021, and 258 in W 2021.

For sucrose, raffinose and stachyose content, the protocol in Rosso, et al., 2018 was followed. Briefly, 180 seed samples were first ground using a water-cooled grinder until the seeds became fine powder. All samples were weighed out to 0.1 g and then 1.0 mL of HPLC-grade water was added to each sample. All samples were shaken for 15 minutes at 400 strokes per minute. Samples were centrifuged at 13.2 rpm for 15 minutes and 0.5 mL of the supernatant was transferred to a new 2.0 mL centrifuge tube. Then 0.7 mL of acetonitrile (ACN) was added and all tubes were mixed by inverting multiple times. Samples sat for 1 hour at room temperature and then centrifuged at 17000 g for 15 minutes. After centrifuging, 100 µL of the supernatant was mixed with 900 µL of 65% ACN and filtered through a 0.2 µm membrane into an HPLC-sample vial. HPLC was used to determine sucrose content according to Lord et al.2021 [189]. All samples were adjusted for moisture and technical replicates were averaged together. Sugar analysis was only done on 2019 samples due to there being no significance found between test weight and sugar. To determine the correlation between high test weight and the three seed composition traits, protein, oil, and sugar content, R was used to calculate the Pearson's correlation coefficient. An $\alpha<0.05$ was used to determine if the correlations were significant.

<u>Results</u>

Test weight

Test weight values showed a normal distribution across environments and years, ranging from 62 to 77 kg/hL with a grand mean of 70.2 kg/hL and an average standard deviation of 1.94 kg/hL.

Figure 1 shows the distribution of test weight for both locations (Fig. 1A), Blacksburg (Fig. 1B), and Warsaw (Fig. 1C). Blacksburg ranged from 64 to 77 kg/hL with an average of 70.3 kg/hL with a standard deviation of 1.72 kg/hL for all years, while Warsaw ranged from 62 to 77 kg/hL had an average of 70.1 kg/hL with a standard deviation of 2.14 kg/hL. Both locations had similar averages, but apparently, Warsaw had a higher spread. Throughout all years and both locations, there was one accession, PI87059, that consistently had high test weight.

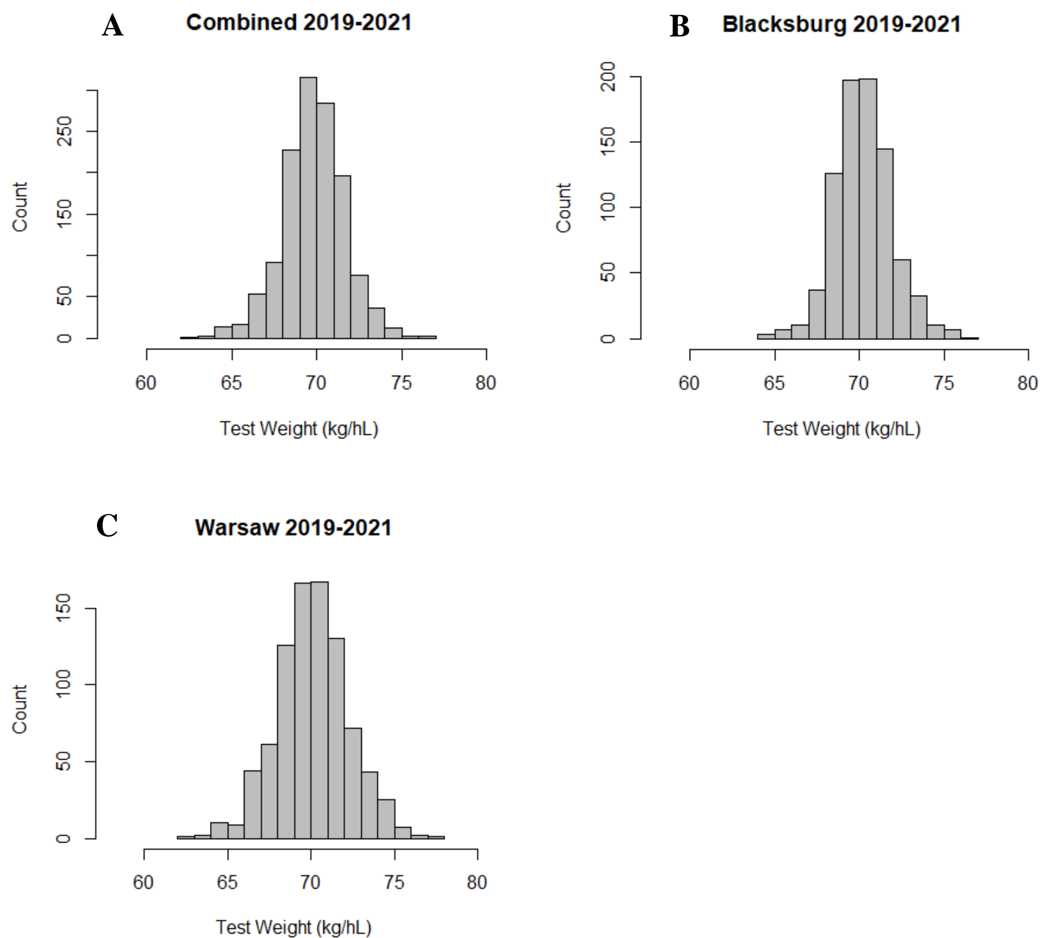
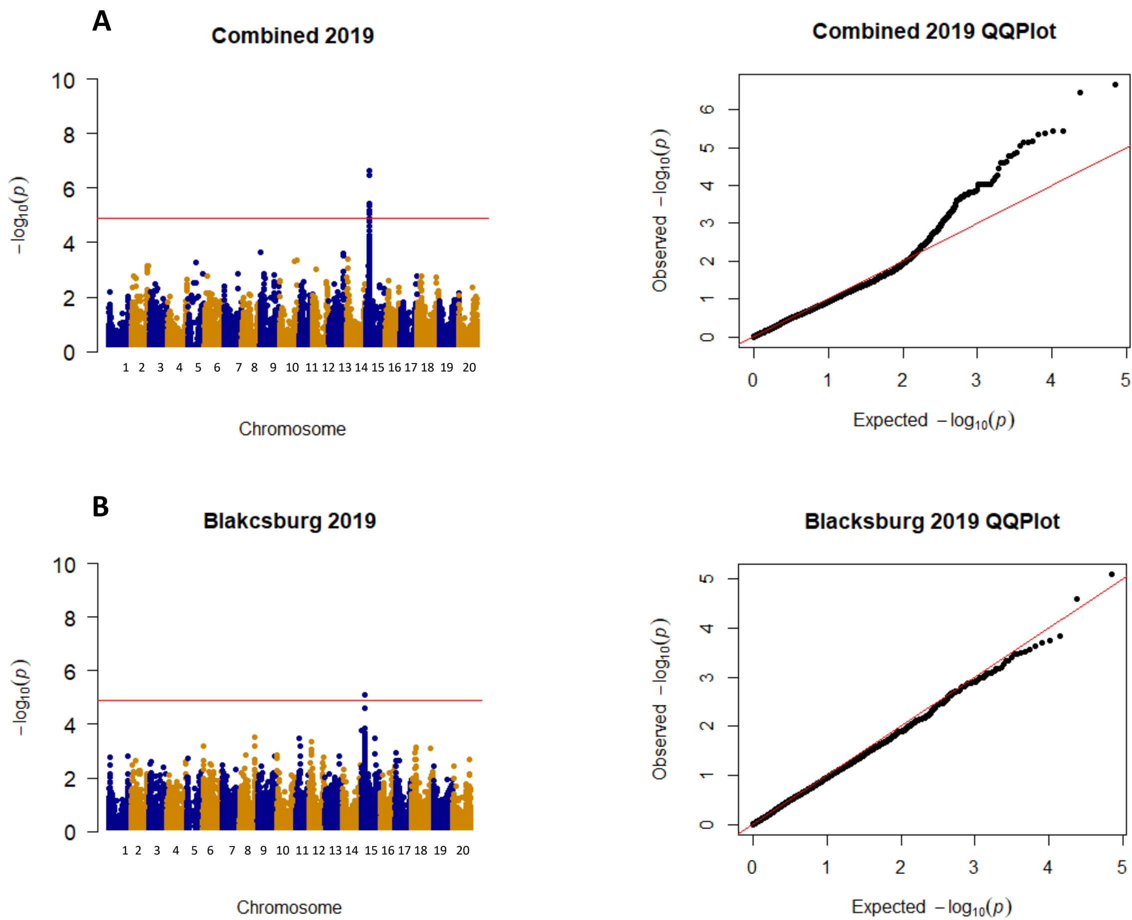


Figure 1. Frequencies of test weight values across both locations (A), Blacksburg (B), and Warsaw (C) for 2019-2021

Genome-wide association study

All significant SNPs were on Chr 15 with six SNPs being significant in more than one environment and three SNPs being just under the significance threshold (**α** = 4.91) in multiple environments (Table 1). There are two exceptions to this: SNP ss71562017 was included since it

had the highest LOD score of -log10(p) for W 2020 and no SNP from this environment was found to be significant. SNP ss715623162 was included since it was the only significant SNP in BB 2020. In 2019, all significant SNPs were found on chromosome 15 and significant SNPs were found on this chromosome in all locations (Fig. 2). In 2020 and 2021, there were many additional SNPs on other chromosomes found to be significant but were not found in any other environments and therefore were not included for further analysis (Fig. 3 and 4). Warsaw did not have any SNPs above LOD of 4.9 in 2020 and 2021 but had multiple SNPs that were close to the threshold (Fig. 3 and 4). QQ plots showed that the data were normally distributed (Fig. 2-4).



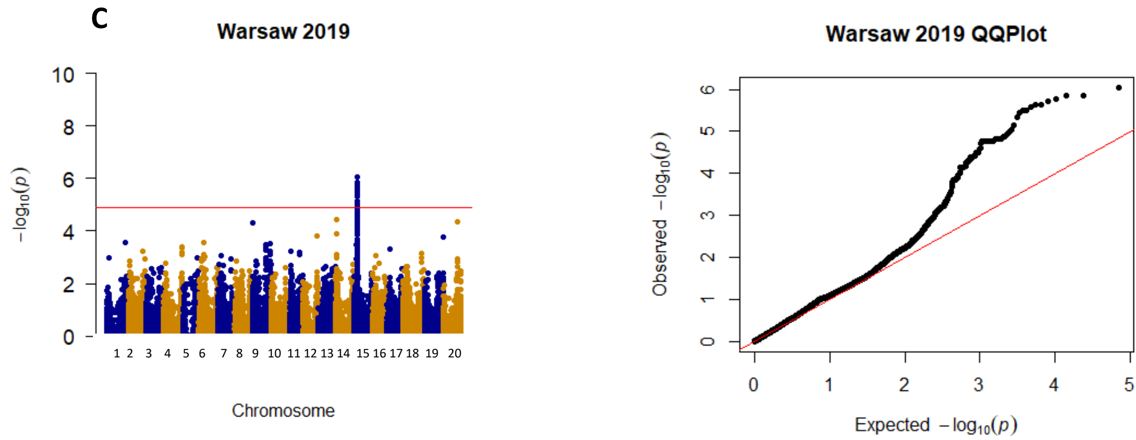
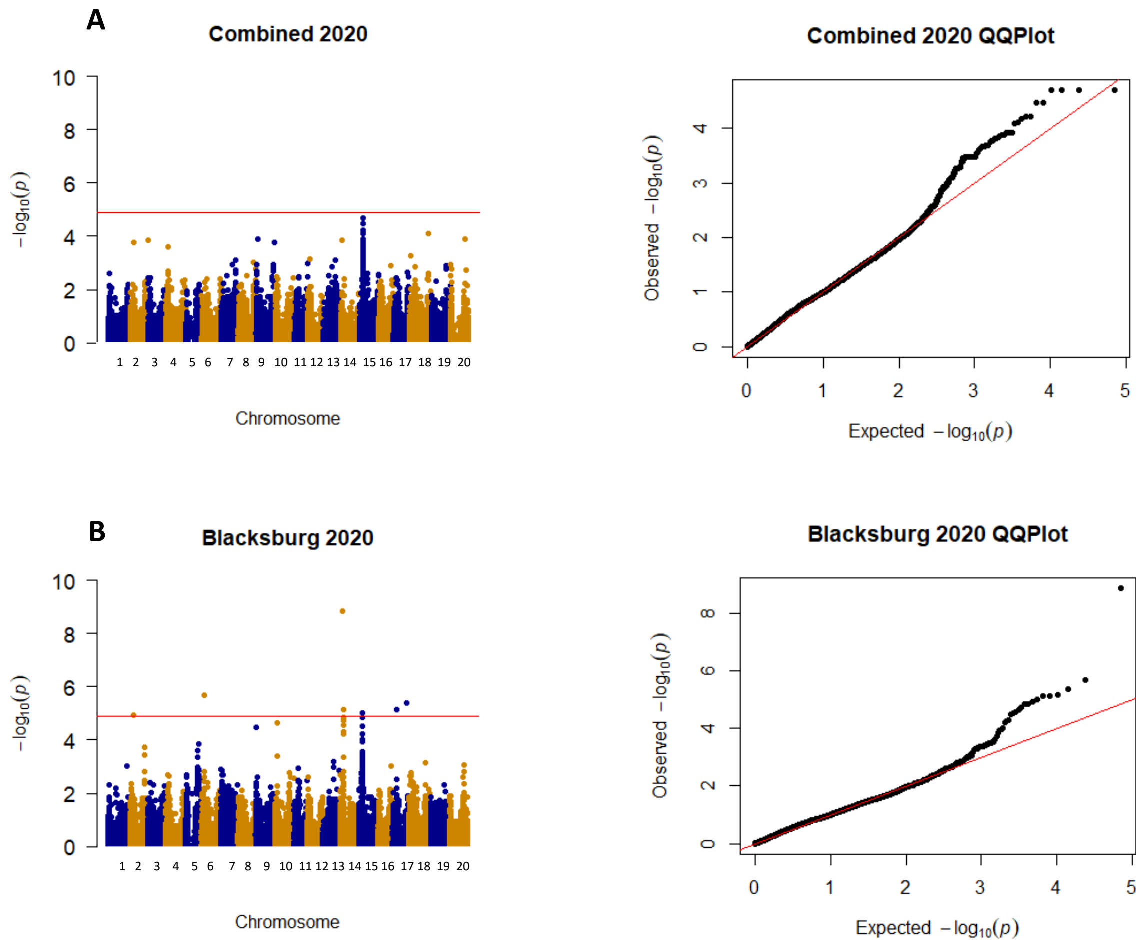


Figure 2: Manhattan and QQ plots for both locations (A), Blacksburg (B), and Warsaw (C) for 2019. Chromosomes in the Manhattan plots are shown in alternating colors and the significance threshold is represented by the red line at -log10(Pp of 4.9. QQ plots show observed -log10(p) plotted against expected -log10(p).



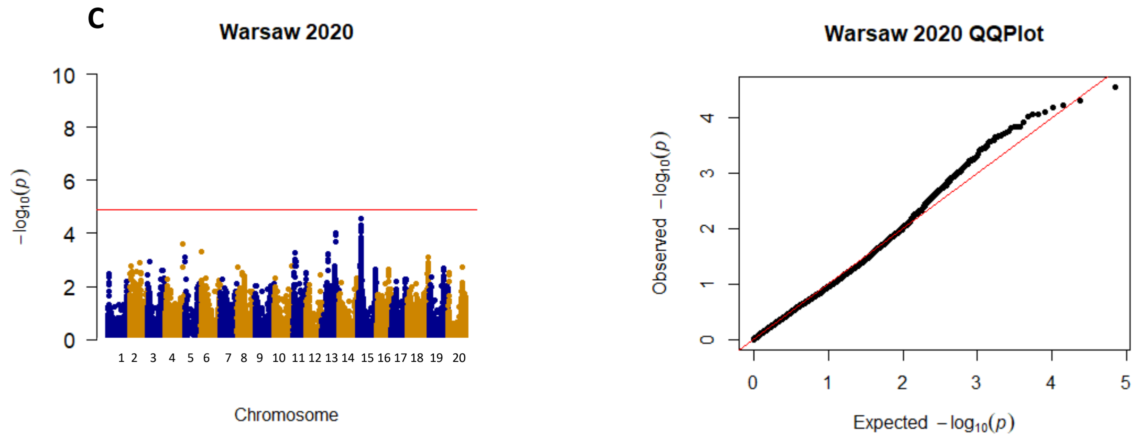
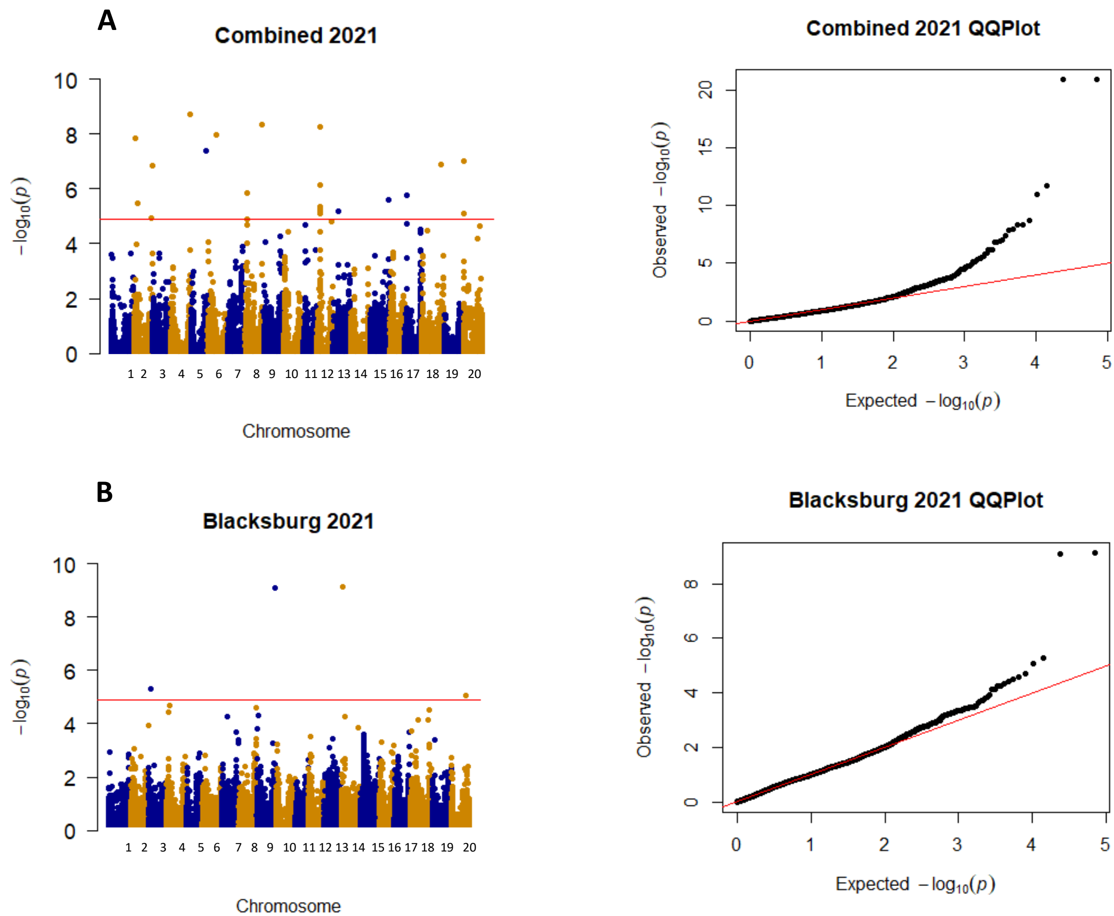


Figure 3: Manhattan and QQ plots for both locations (A), Blacksburg (B), and Warsaw (C) for 2020. Chromosomes in the Manhattan plots are shown in alternating colors and the significance threshold is represented by the red line at -log10(P) of 4.9. QQ plots show observed -log10(p) plotted against expected -log10(p).



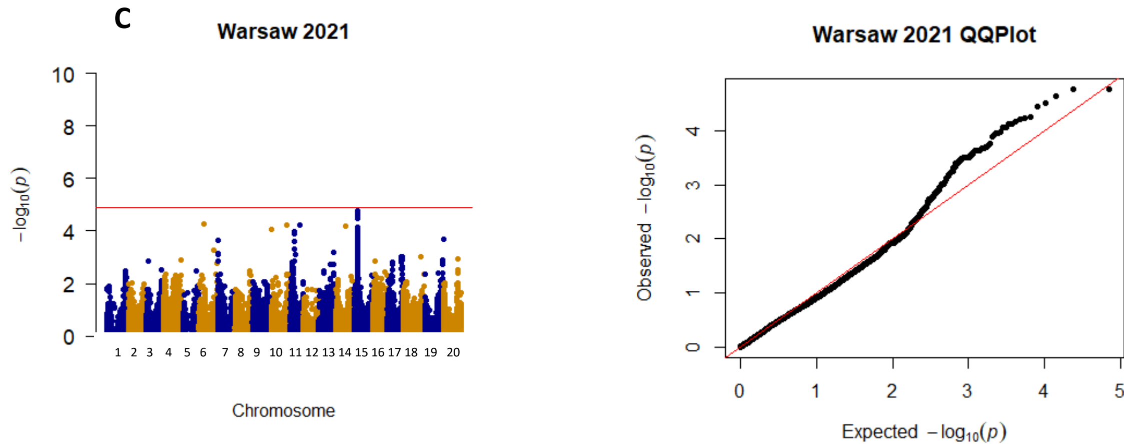


Figure 4: Manhattan and QQ plots for both locations (A), Blacksburg (B), and Warsaw (C) for 2021. Chromosomes in the Manhattan plots are shown in alternating colors and the significance threshold is represented by the red line at -log10(P) of 4.9. QQ plots show observed -log10(p) plotted against expected -log10(p).

Table 1: Significant SNPs associated with test weight in soybean across multiple environments.

Marker	Chr	Position	Environment (-log10(p))								
			BB 2019	BB 2020	BB 2021	W 2019	W 2020	W 2021	Combined[C] 2019	Combined 2020	Combined 2021
ss715623162	15	8758404	NS[b]	5.37	NS	NS	NS	NS	NS	NS	NS
ss715623211	15	9205168	NS	NS	NS	5.51	NS	NS	5.13	NS	NS
ss715623224	15	9279044	NS	NS	NS	5.85	NS	NS	5.44	NS	NS
ss715620221	15	9383632	NS	NS	NS	5.72	NS	NS	5.33	NS	NS
ss715623250	15	9557248	NS	NS	NS	4.99	NS	NS	5.39	NS	NS
ss715623269	15	9748128	4.60[a]	NS	NS	5.76	NS	NS	NS	NS	NS
ss715623270	15	9749617	5.09	NS	NS	6.04	NS	NS	6.66	NS	NS
ss715623292	15	9927090	NS	NS	NS	NS	NS	4.77[a]	NS	4.70[a]	NS
ss715620172	15	10176737	NS	NS	NS	NS	4.56[a]	NS	NS	NS	NS

[a]These values are just below the significant threshold [b]NS=Not Significant [C]Combined is Blacksburg and Warsaw together

Table showing SNPs that were above the threshold or just below it in more than one location with their corresponding name, chromosome location, position, and -log10(p-values). SNPs ss715623162 and ss715620172 were included because they were the markers with the highest -log10(p-value) in BB 2020 and W 2020 respectively and these locations did not have any SNPs that were found to be significant in any other environment.

Candidate Genes

A total of eight candidate genes were found on Chr 15 (Table 2). Only a few genes are located in within 10 kbp of the significant SNPs, but no genes are located within 10 kbp of ss715623250. The functions of the candidate genes mostly involve RNA and/or protein binding or regulation. Gene Glyma.15g119200 has a different function that involves seed storage and gene Glyma.15g127900 has an unknown function.

Table 2: Summary of candidate genes

Chromosome	SNP (position)	Gene	Function
15	ss715623162 (8758404)	Glyma.15g111700	Ribosomal Protein
15	ss715623211 (9205168)	Glyma.15g117100	Transcriptional regulation
15	ss715623224 (9279044)	Glyma.15g118300	Pentatricopeptide protein
15	ss715620221 (9383632)	Glyma.15g119200	Seed storage protein
15	ss715623269 (9748128)	Glyma.15g122800	RNA and protein binding
15	ss715623270 (9749617)	Glyma.15g122800	RNA and protein binding
15	ss715623292 (9927090)	Glyma.15g125000	Serves as a methyltransferase
15	ss715620172 (10176737)	Glyma.15g127900	Unknown

Table showing candidate genes near SNPs that were previously identified as significant or just below the threshold and their corresponding function.

Correlation of Test Weight with Seed Composition Traits

The Pearson's correlation between test weight and protein and oil content for 2019-2021 and between test weight and raffinose, sucrose, and stachyose content for 2019 varied depending on location and year (Table 3). The correlation between test weight and protein was not consistent due to a positive correlation in three environments and a negative correlation in other three environments. In addition, three environments showed significant correlation including BB and

W 2019 with a negative correlation and W 2020 with a positive correlation. The correlation between test weight and oil was found to be significantly negatively correlated in all environments except for BB 2021. The strongest correlation was found in W 2021 at -0.387. The correlation between test weight and all sugars in 2019 was negative, but only raffinose had a significant correlation coefficient of -0.175 in BB 2019.

Table 3: Correlation between test weight and five important seed compositions

Trait	BB 2019	W 2019	BB 2020	W 2020	BB 2021	W 2021
Protein	-0.115*	-0.136*	0.021	0.390*	-0.052	0.047
Oil	-0.174*	-0.297*	-0.203*	-0.265*	-0.024	-0.387*
Raffinose	-0.175*	-0.118	NA	NA	NA	NA
Sucrose	-0.086	-0.08	NA	NA	NA	NA
Stachyose	-0.003	-0.063	NA	NA	NA	NA

*These correlations were found to be significant at $\alpha = 0.05$.

Discussion

While test weight has not received as much attention as other traits such as protein and oil content in soybean, it is a crucial trait to threat farmers' profitability. This coupled with its power to affect the pricing of soybean and the decreasing trend that farmers have observed in test weight in past decades [141]. In this study, we were able to identify multiple genetic markers significantly associated with test weight through GWAS, potential candidate genes for test weight, and provide information regarding the relationship between test weight and important seed compositions.

Many studies have found a significant relationship between genotype and high test weight in wheat, soybean, and other crops [172–175]. However, no studies have found SNPs associated with test weight. In this study, we were able to find many significant SNPs, especially on chromosome 15. Environments including BB 2020, BB 2021, and Combined 2021, had multiple significant SNPs that were found to be associated with test weight but were not present in any other environment. SNPs on Chr 15 were either found to be significantly associated or close to being significantly associated with test weight across years and locations, which indicates that Chr 15 may have genes that control test weight. Furthermore, eight of these SNPs were found to be within 10 kbp of a gene. While most of these SNPS are located close to genes that encode proteins to bind and regulate RNA and other proteins, ss715620221 is near Glyma.15g119200,

which codes for a seed storage protein, so ss715620221 could be a promising genetic marker because test weight is related to seed durability. Additionally, SNP ss715620172 is located near a gene with an unknown function, which encourages future studies to determine if Glyma.15g127900 is related to test weight.

While Chr 15 was found to consistently have significant SNPs or close to the significance threshold, there were no SNPs that were found to be significant across all locations and years. This could be explained by the impact that the environment has on test weight. It is known that while genotypes affect test weight, the environment can also significantly influence test weight [175]. BB and W have different climates with BB being cooler and less humid than W. Three environments, Warsaw 2020, Combined 2020, and Warsaw 2021, had no SNPs that were above the threshold of 4.91. While they did have SNPs that were just below the threshold, it is interesting that no SNPs showed significant association in those environments despite all other environments having at least one significant SNP. Additionally, BB 2020, Combined 2021, and BB 2021 had multiple SNPs on other chromosomes that were significant but not consistently found in the other environments. This contrasts with 2019, where all environments in this year only had significant SNPs on chromosome 15. These two findings could largely be explained by the time of harvest. It is important to note that delayed harvest has been found to negatively affect test weight in wheat and corn [190,191]. Harvest occurred normally in 2019 but was delayed by a couple of weeks due to the weather in 2020 and 2021. Because delaying harvest can impact test weight [237, 238], these differences between years most likely resulted from delayed harvest in 2020 and 2021.

Our correlation study found that the correlation between test weight and protein content was inconsistent. In BB 2019 and W 2019, protein was found to be significantly, negatively correlated with test weight, but W 2020 showed significantly, positively correlation. No significant correlation was found between high test weight and protein content in the other environments. Additionally, sucrose content was found to have no significant correlation with test weight. Other studies have found inconsistent correlation results between either protein or sucrose and test weight [224, 226–228]. However, one study did find that sucrose could be significantly, positively related to test weight [175], but this study was conducted in Georgia, a different environment from Virginia, which might be the main reason to cause the different correlation results. No studies have looked at the relationship between raffinose and stachyose

with test weight. While we did find a significant negative correlation between test weight and raffinose in BB 2019, it was found not to be significant in W 2019, and stachyose was found to not be significant in either location. Based on the inconsistent correlation between sucrose and test weight, it is not surprising that we did not find consistent significant correlations for these traits since these compounds are both sugars and similar to sucrose. On the other hand, oil was found to be significantly, negatively correlated with test weight in all environments except for BB 2021. This is similar to the study conducted by Liu et al. 2019 that also found a significant, negative relationship between oil and test weight. This consistent negative relationship between oil content and high test weight could partly explain the decreasing trend in test weight that farmers have observed. One of the main traits that soybean breeders select for is higher oil content. Therefore, breeders may need to modify the selection objectives to main oil content with increased test weight.

In summary, multiple significant SNPs associated with test weight on chromosome 15 were identified, which could be used by breeders to quickly select high test weight progenies derived from high test weight parents in order to increase the overall test weight of the breeding programs' germplasm. Additionally, a consistent negative relationship between high test weight and oil content was found. This information could be useful for breeders while they make their selections to ensure high oil content, they should also pay attention to test weight to help offset its decreasing trend. Ultimately, incorporating parents that will have high test weight into breeding schemes and taking into consideration the relationship between test weight and oil content will enable breeders to develop varieties with increased test weight to avoid farmer's profitability in the long run.

Chapter 3:
Development of molecular markers for the Kunitz trypsin inhibitor mutant alleles generated by CRISPR/Cas9-mediated mutagenesis in soybean

Zhibo Wang, Zachary Shea, Luciana Rosso, Chao Shang, Jianyong Li, Patrick Bewick, Qi Li, Bingyu Zhao, Bo Zhang

Introduction

Soybean meal provides an excellent source of protein in animal feed since it is rich in amino acids with a high nutritional profile [192]. For instance, soy makes up 26% and 50% of swine and poultry feed, respectively [193]. However, reduced feed efficiency has been observed due to anti-nutritional and biologically active factors in raw soybean seeds [194]. Among these factors, trypsin inhibitor (TI) accounts for a substantial amount of this effect that cannot be ignored [158]. TI restrains the activity of trypsin in monogastric animals. Because this enzyme is essential for optimal protein digestion, its restriction can lead to animal growth inhibition of 30-50% due to pancreatic hypertrophy/hyperplasia when raw soybeans are used in feed [158,195,196]. In soybean meal processing facilities, TI in soybean meal is deactivated via a heating process at 90.5°C-100°C with the presence of 1% NaOH [197]. This process not only reduces the nutritional value of soybean meal due to thermal destruction of amino acids, but also increases the energy cost of meal production by 25% [198].

With the recent increase in feed price and shipping cost, livestock farmers and grain operations are reconsidering soybean varieties with low-TI or TI-free traits as a way to reduce farm expenses by using raw soybeans as feed. Raising low-TI or TI-free soybeans on farms creates a niche market for integrated crop and livestock farmers, increasing their farm's profitability. The use of soybean lines with genetically reduced levels of TI has proven as an effective strategy for improving animal growth. For instance, chicks fed soy-based diets with raw, unprocessed, low-KTI soymeal had higher feed efficiency ratios than chicks fed diets containing raw, unprocessed, conventional soybean meal [199]. Thus, soybean cultivars with low TI content in the seeds are a long-term breeding goal for higher protein digestibility, better economic benefits, reduced environmental pollution caused by phosphorus, and the pursuit of sustainability for humanity and nature.

Plants have evolved a group of TI genes encoding proteins that can suppress the enzyme activities of proteases found in plants, herbivores, animals and human beings [200,201]. The TIs in soybean can be classified into two families: the 21 kDa Kunitz trypsin inhibitor protein family (KTI) and the 7-8 kDa Bowman-Birk inhibitor protein family (BBTI) [193,202,203]. KTI proteins are thought to be largely specific for trypsin inhibition, while the major isoform of BBTI contains domains that interact with and inhibit both trypsin and chymotrypsin [193,204]. Currently, only the KTI genes are targeted for the selection of low-TI soybeans because KTI serves as the major contributor to trypsin inhibitor activity in soybeans. By far, the most significant success in reducing TI activity in soybean was the identification of a soybean accession (PI 157740) with dramatically reduced (~40%) TI activity [193]. A frameshift mutation in *KTI3* (Gm08g341500) gene was identified in PI 157740, which is responsible for the low TI phenotype [205]. PI 157740 has been used in feeding trials, and it was found that raw extruded protein meal with lower KTI3 protein is superior for animal weight gain when compared to raw soybean meal harboring functional KTI3 [195,206]. However, weight gain for young animals fed with non-heat-treated soybean materials including nonfunctional *KTI3* soybean materials is still inferior to those fed with heat-treated soybeans [195,206]. Another soybean germplasm accession (PI 68679) was identified to carry a nonfunctional mutation on *KTI1* (Gm01g095000) gene [193]. *KTI1* and *KTI3* genes were determined to be synergistically controlling the TI content in soybean seeds [193]. Therefore, it is desirable to breed new soybean cultivars carrying both *kti1* and *kti3* mutant alleles. However, it is time-consuming to breed low TI soybean cultivars by selecting progenies derived from crosses between PI 157740, PI 68679, and elite varieties. Besides, the linkage drags associated with *KTI1* and *KTI3* may introduce undesirable agronomic traits, which could be difficult to remove by backcrossing. Although Kompetitive Allele Specific PCR (KASP) markers associated with *KTI3* and its mutant allele with 86% efficiency are available [207], attempts at developing molecular markers associated with *KTI1* have not been successful [193]. Therefore, it is highly desirable to develop new *KTI1* mutant alleles that can be tagged with convenient molecular markers.

CRISPR/Cas9 mediated genome editing employs a Cas9 endonuclease and an 18-22 bp small guide RNA (sgRNA) that have a region that is complementary to a target gene sequence. The sgRNA binds to Cas9 and recruits the complex to target a gene. The Cas9 endonuclease generates DNA breaks, leading to mis-repaired target genes that contain deletions or insertions that disrupt gene function. In addition, several sgRNAs can be co-expressed in a single cell with Cas9, which

allows the multiplex mutations of different genes simultaneously [208]. Because genome-edited plants without transgenes are not considered genetically modified organisms (GMO) [209], mutant plants can be either directly released for field tests or served as valuable resources for further breeding selection. Thus far, the CRISPR/Cas9 mediated genome editing technology has been widely used for targeted gene mutagenesis in diverse crop plant species, including soybean [210–212], rice [213], wheat [214], maize [215], tomato [216], cotton [217], citrus [218], apple [219], grape [219], potato [220], and banana [221] to improve their agronomic performances. For example, Jacobs et al. (2015) reported the first targeted mutagenesis in soybean using the CRISPR/Cas9 technology [211]. Haun et al. (2014) generated a high oleic acid content soybean variety without transgenic components and improved the quality of soybean [210]. A soybean mutant with a late flowering phenotype was created using CRISPR/Cas9 technology to knock out the *GmFT2a* gene [212]. Thus far, there is no research reporting an attempt to apply CRISPR/Cas9 to target anti-nutritional factors in soybean.

In this study, we aimed to (1) simultaneously knockout *KTI1* and *KTI3* genes in soybean cultivar Williams 82 via CRISPR/Cas9-mediated genome editing and (2) develop molecular markers associated with *kti1* and *kti3* mutant alleles that can be used for marker-assisted selection (MAS). We successfully recovered transgenic soybean plants that are carrying both *kti1* and *kti3* mutations. KTI content and trypsin inhibition activities (TIA) are dramatically decreased in the *kti1* and *kti3* mutant lines. In addition, we also developed molecular markers for co-selection of the new *kti1* and *kti3* mutant alleles. These *kti1* and *kti3* mutant lines and the newly developed selection markers have great potential for breeding the low TI trait into elite soybean varieties in the future.

Materials & Methods

Plant materials and growth conditions

Soybean plants were grown in 2.5-gallon pots using Miracle-Gro all-purpose potting soil mix in Keck Greenhouse at Virginia Tech (14h/10h light/dark cycle at 25 °C/20 °C) for the experiments described herein. The plants were watered by an automatic irrigation system. Soybean transformation was performed at the plant transformation facility at Iowa State University as previously described [222,223]. The plant growth indicators and maturity period days of *WM82* and progeny plants of the T1 generation derived from lines #2 and #5 were measured in the green house. The 4-week-old T1 soybean plants were used for genotyping. The seeds of T1 plants were used for seed weight analysis.

Constructing a soybean KTI gene map

The gene map showing locations of *KTI* genes on soybean chromosomes was made using MapInspect. Locations of all *KTI* genes were obtained from the Phytozome database and plotted on their respective chromosomes.

Bacterial growth

E. coli strains *DH5α* and *C41* (DE3) (Lucigen, Middleton, WI) were grown on Luria agar medium at 37 °C. *Agrobacterium tumefaciens* (*A. tumefaciens*) *EHA105* was grown on Luria agar medium at 28 °C [224]. *E. coli* antibiotic selections used in this study were as follows: 50 µg/ml kanamycin, 100 µg/ml carbenicillin, 100 µg/ml spectinomycin. *A. tumefaciens* antibiotic selection were 100 µg/ml rifampicin, and/or 100 µg/ml spectinomycin.

Cloning

The open reading frames (ORFs) of *KTI1* and *KTI3*, were amplified from the genomic DNA of *WM82*. The $KTI1_{\Delta 66bp}$, truncated ORF of *KTI1*, was amplified from the genomic DNA of mutant soybean plant #2-1. All PCR primers with annotations are listed in Table S1. The genes/fragments were then cloned into a pDonr207 plasmid (Thermo Fisher Scientific) for future use.

T1 plant genomic DNA (gDNA) was used as the templates to amplify KTI1 and/or its mutant allele, and KTI3 and/or its mutant allele by PCR. The purified PCR fragments were used for genotyping by Sanger sequencing at Virginia Tech Genomic Sequencing Center and cloned to the PCR8/GW/TOPO vector by TA cloning (Invitrogen) for molecular marker tests.

In order to apply the CRISPR/Cas9 system to gene editing in soybean, we modified our current CRISPR/Cas9 construct [225]. The cassette consists of a MAS promoter, the bialaphos resistant gene, and a MAS terminator that was amplified using plasmid DNA of pEarleyGate101 as the template. All PCR primers with annotations are listed in Table S1. The cassette was assembled to the backbone of CRISPR/Cas9 construct using Gibson Assembly® Cloning Kit (New England Biolabs Inc). The gRNAs targeting *KTI1* and *KTI3* were synthesized in one cassette at GenScript Biotech Corp. The backbone of the new CRISPR/Cas9 construct and the fragment of gRNAs were assembled using Gibson Assembly® Cloning Kit.

Expression analysis of KTI Genes in WM82

RNA sequencing data, in FPKM (fragments per kilobase of transcript per million fragments mapped), of 38 *KTI* genes in 31 different tissue types from Williams 82 were acquired through the Gene Networks in Seed Development database (http://seedgenenetwork.net/sequence). Construction of the heatmap to visualize expression data was done using the heatmap.2 function from the ggplot2 package in R. A green/blue color gradient was chosen to show expression with blue representing little to no expression and green representing high expression. The code for the heatmap is as follows:

heatmap.2(x=KTI Expression, main = "KTI Expression In Different Soybean Tissue", notecol="black", density.info="none", trace="none", margins =c(12,9), col=my_palette, breaks=col_breaks, dendrogram="row",Colv="NA", ylab= "Genes", xlab= "Tissue Type", cexCol=.9,cexRow = .8)

RNA isolation and real-time PCR

All RNA was extracted from V98-9005 (high-TI line) and V03-5903 (low-TI line) seeds using TRIzol reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. Any DNA residue was eliminated by treating it with UltraPure DNase I (Thermo Fisher Scientific). The integrity and quantity of total RNA were determined by electrophoresis in 1% agarose gel and a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). cDNA synthesis was performed using the SuperScript III First-Strand RT-PCR Kit (Thermo Fisher Scientific) with an oligo-dT primer based on the manufacturer's instructions. Real-time PCR was conducted with cDNA as the template using the Quantitect SYBR Green PCR kit (Qiagen) according to the manufacturer's protocol. Oligo primers are listed in Table S1. The soybean *ELF1B* gene was used as a reference gene, and data are presented as ΔCT [226].

Expression and purification of KTI1, $KTI1_{\Delta 22aa}$, and KTI3 proteins

The *KTI1*, $KTI1_{\Delta 66bp}$, and *KTI3* genes in pDonr207 were subcloned into a Gateway compatible pET28a destination vector via a LR® Gateway cloning kit (Thermo Fisher Scientific) [225,227]. The plasmids were transformed into *E. coli C41* cells (Lucigen). KTI1, $KTI1_{\Delta 22aa}$, and KTI3 proteins were expressed and purified following a procedure as previously described. Protein purity

was evaluated by SDS-PAGE. The protein concentration was determined by a protein assay kit (Bio-Rad) using bovine serum albumin as standard [228].

Standard bioassay to measure trypsin inhibitor activity

A TI activity bioassay was performed following American Association of Cereal Chemists Official Method 22-40 (AACC, 1999) with some modifications previously reported by Rosso et al. (2018) [229]. Briefly, 30 mg of finely ground soybean seed powder was mixed with 3 mL of 9 mM HCl (pH 2.0). The mixture was shaken for 1 h at room temperature. 2 mL of the extracts was centrifuged at 10,350 rpm for 20 min at room temperature, and the supernatant was diluted by 10 times with 9 mM HCl for measuring TI activity. A TI activity assay was performed in a 96-well plate format following the same steps described in Rosso et al. (2018) [229]. Each sample row was repeated three times. Portions of diluted HCl extracts (0, 20, 30, 40, and 60 µL) or 50 µg recombinant proteins of KTI1, KTI1$_{\Delta 22aa}$, and KTI3 were pipetted into the microplate wells, and the volume was adjusted to 60 µL with 9 mM HCl. 60 µL of extractant was used as a sample blank and 60 µL of water was used as a substrate blank. 60 µL of trypsin (from bovine pancreas, Sigma-Aldrich T8003) solution was added to each sample well, and the microplates were placed in an oven at 37°C for 15 min. After the incubation, 150 µL of BAPNA substrate pre-warmed at 37°C was added to all wells, and the plates were incubated for exactly 10 min at 37°C. The reaction was stopped by adding 30 µL of acetic acid solution to all wells. The absorbance of each well was read on a plate reader (FLOUstar Omega, BMG Labtech) at 410 nm for 30 s after shaking at 700 rpm.

HPLC method to quantify kunitz trypsin inhibitor

The HPLC method to quantify *KTI* was performed following the method developed by Rosso et al. (2018) [229]. Briefly, 10 mg of finely ground soybean seed powder was mixed with 1.5 mL of 0.1 M sodium acetate buffer (pH 4.5). Samples were vortexed and shaken for 1 h at room temperature. The sample was centrifuged at 12,000 rpm for 15 min. 1mL of the supernatant was filtered through a syringe with an IC Millex-LG 13-mm mounted 0.2-mm low protein binding hydrophilic millipore (polytetrafluoroethylene [PTFE]) membrane filter (Millipore Ireland). The *KTI* in solution was separated on an Agilent 1260 Infinity series (Agilent Technologies) equipped with a guard column (4.6 x 5 mm) packed with POROS R2 10-mm Self Pack Media and a Poros R2/H perfusion analytical column (2.1 x 100 mm, 10 um). The mobile Phase A consisted of 0.01%

(v/v) trifluoroacetic acid in Milli-Q water, and the mobile Phase B was 0.085% (v/v) trifluoroacetic acid in acetonitrile. The injection volume was 10 mL and the detection wavelength was 220 nm.

Development of molecular selection markers with a gel electrophoresis free method for high throughput screening

The transgene free and double homozygous mutant line, #5-26, was selected for the development of molecular selection markers. Based on the genotyping data of #5-26, two pairs of markers were designed: ZW1 with ZW2 or ZW3. ZW1 is the common reverse primer for both *KTI1* and *kti1,* while ZW2 and ZW3 are two reverse primers matched with unique sequences in *KTI1* and *kti1*, respectively. Similarly, two pairs of molecular markers, ZW4 with ZW5 or ZW6, were designed. ZW4 is the common reverse primer for both *KTI3* and *kti3,* while ZW5 and ZW6 are two reverse primers matched with unique sequences in *KTI3* and *kti3*, respectively. The gDNA of *WM82*, #5-26 (homozygous mutants of both *kti1* and *kti3*), #5-9 (homozygous mutant of *kti1* while heterozygous mutant of *kti3*) and #2-30 (homozygous mutant of *kti3* while heterozygous mutant of *kti1*) were used as templates to test the efficiency and reliability of these markers in PCR.

PCR amplifications were performed in a total volume of 20 µl containing 50 ng of gDNA, 0.5 µM each of forward and reverse primers (Table S1), 10 µl 2X BioMix Red (Bioline) and ddH_2O. The PCR program was set to be 95 °C for 5 min for pre-denature, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, extension at 72 °C for 30 s, followed by a final extension at 72 °C for 5 min.

In order to screen the large-scale progenies derived from crosses between soybean elite cultivars carrying wild type *KTI1/3* and the newly developed mutant plant carrying *kti1/3*, a simple gel electrophoresis-free method was designed. 1X sybrgreen dye (Thermo Fisher) was added to complete PCR reactions, and the solution was incubated for 10 mins at 75°C before placing in the gel doc (Biorad) for imaging the fluorescent signals. Only the positive PCR products with the dye will display fluorescent signals while the failed PCR will not show signals.

Statistical data analysis

Analytical experiments were performed with at least three technical replicates. Statistical significance was based on one-way ANOVA test for multiple comparisons. Data was analyzed using JMP Pro14. Values of $P<0.05$ were considered significant

Results

KTI gene family consists of multiple members with distinct expression patterns

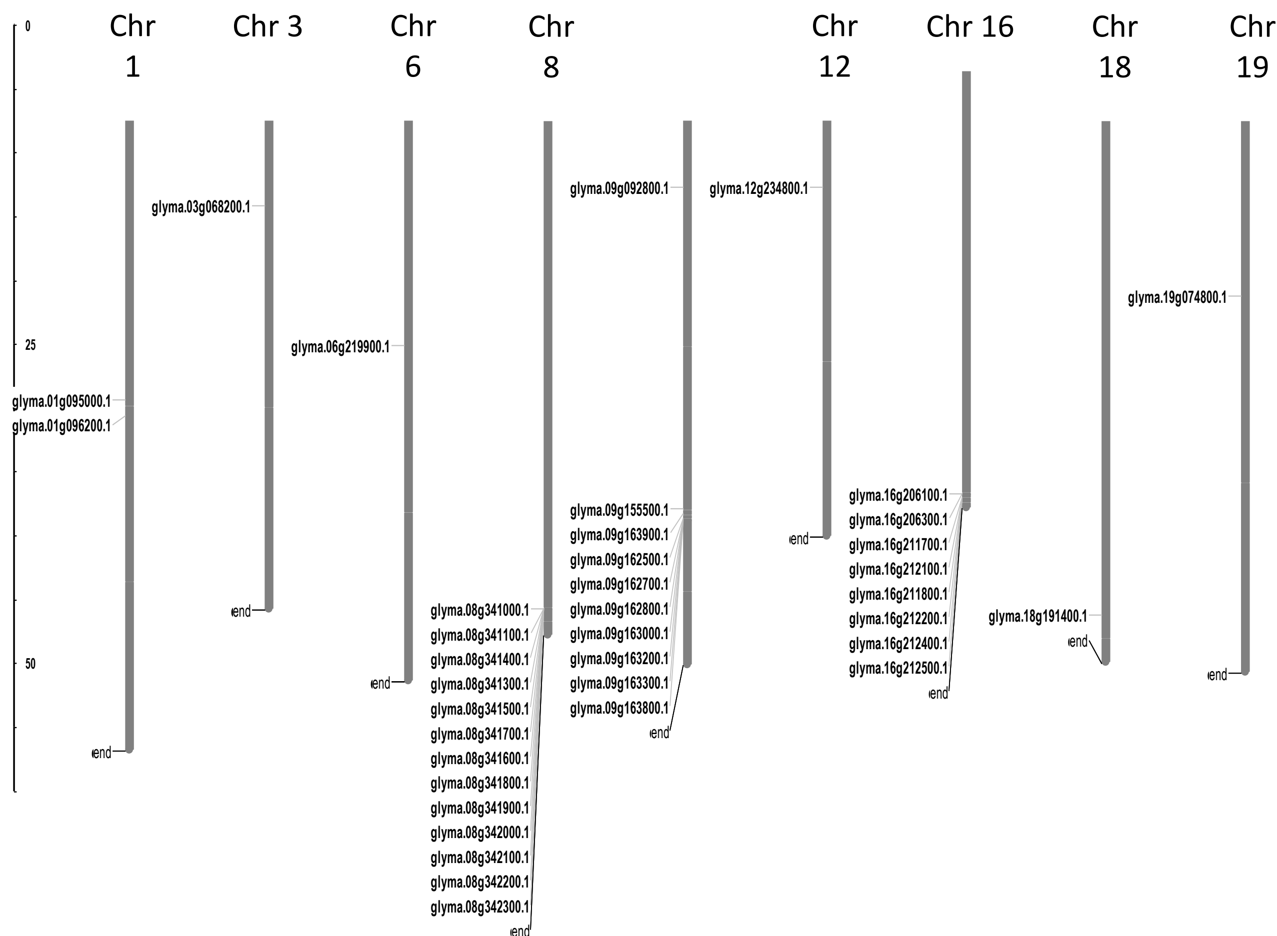


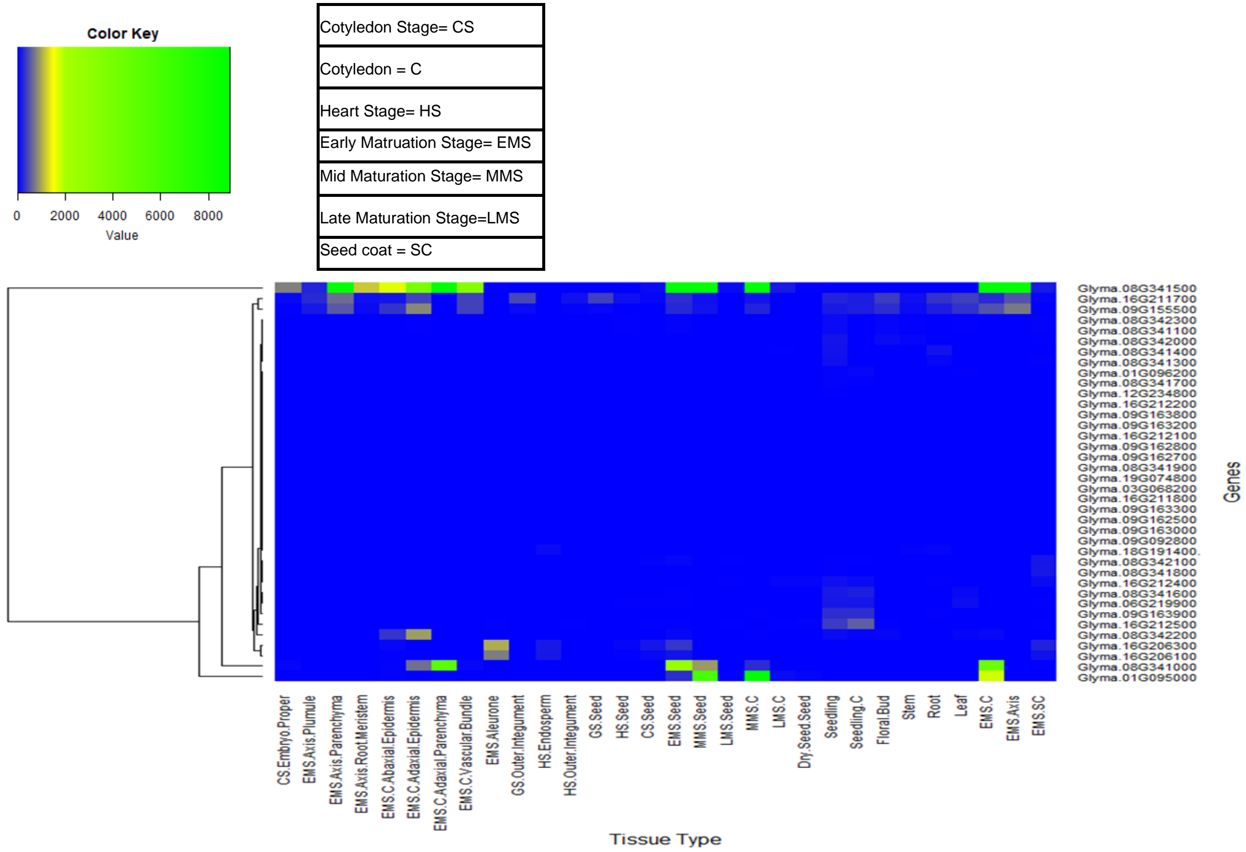
Figure 1. Physical mapping of 38 *GmKTI* genes on 20 soybean chromosomes.

The gene map showing the locations of *KTI* genes on soybean chromosomes was made using MapInspect. As displayed in the map, 38 *KTI* genes are located on 9 out of 20 chromosomes

Soybean *KTI* genes belong to a gene family with thirty-eight members (Soybase version Wm82.a4.v1) (Figure 1). As displayed in the genetic map, 38 *KTI* genes are located on 9 chromosomes (Chr) in the soybean genome. Specifically, Chr1 carries 2 *KTI* genes; Chr3 carries 1 *KTI* gene; Chr6 carries 1 *KTI* gene; Chr8 carries 13 *KTI* genes; Chr8 carries 10 *KTI* genes; Chr12

carries 1 *KTI* gene; Chr16 carries 8 *KTI* genes; Chr18 carries 1 *KTI* gene; and Chr19 carries 1 *KTI* gene (Figure 1).

A



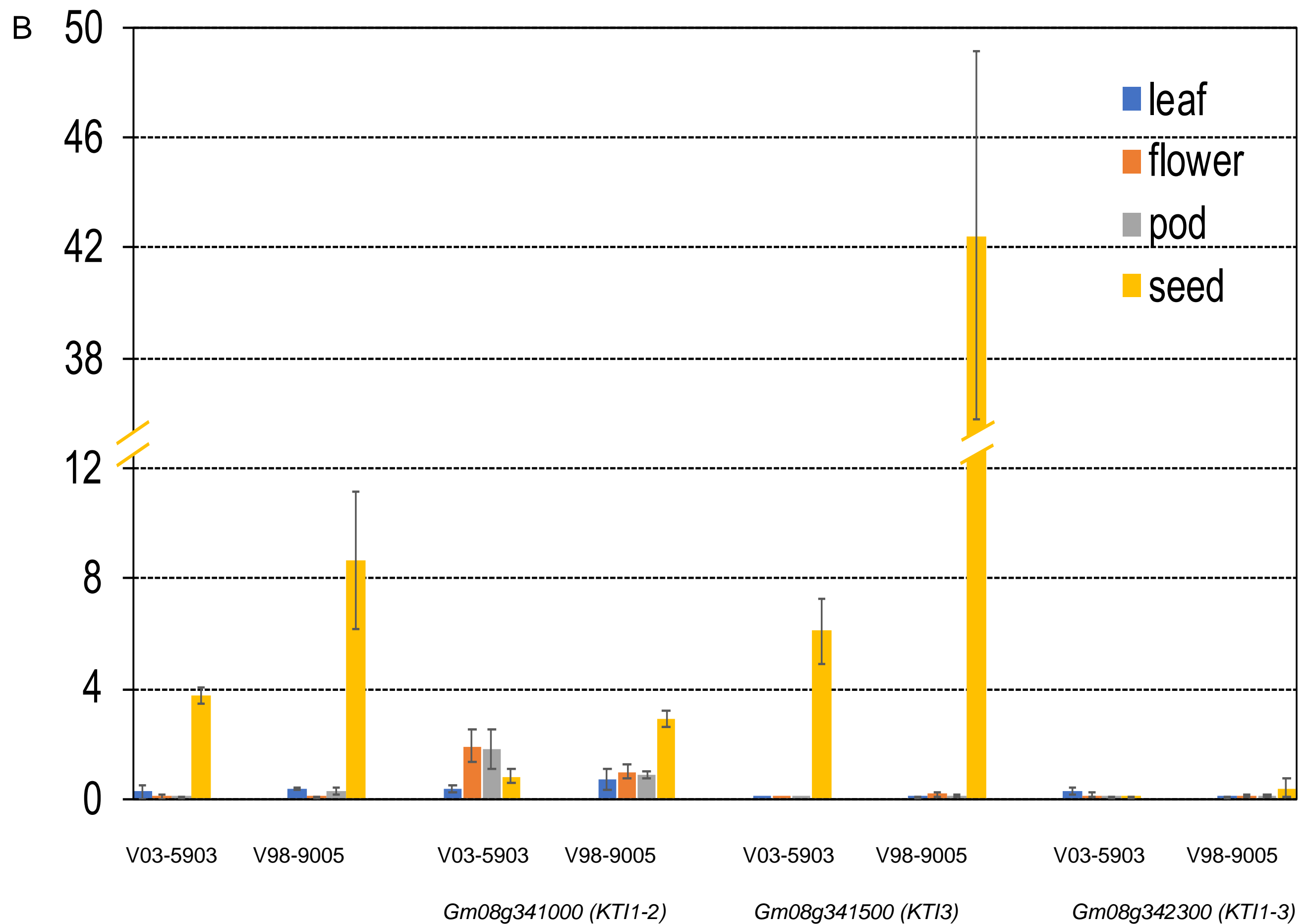
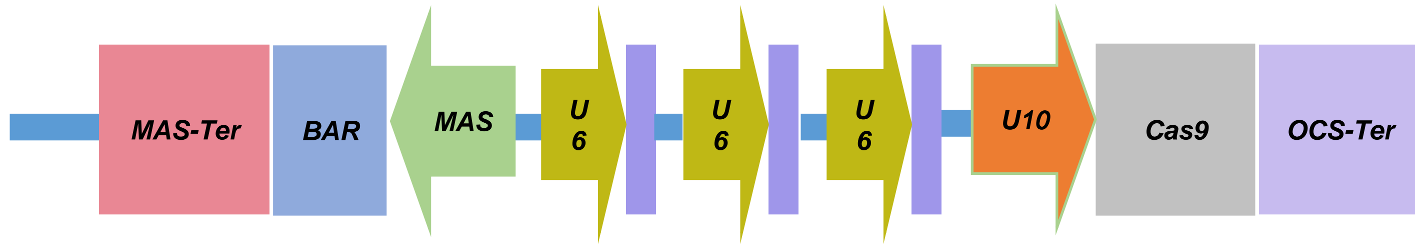


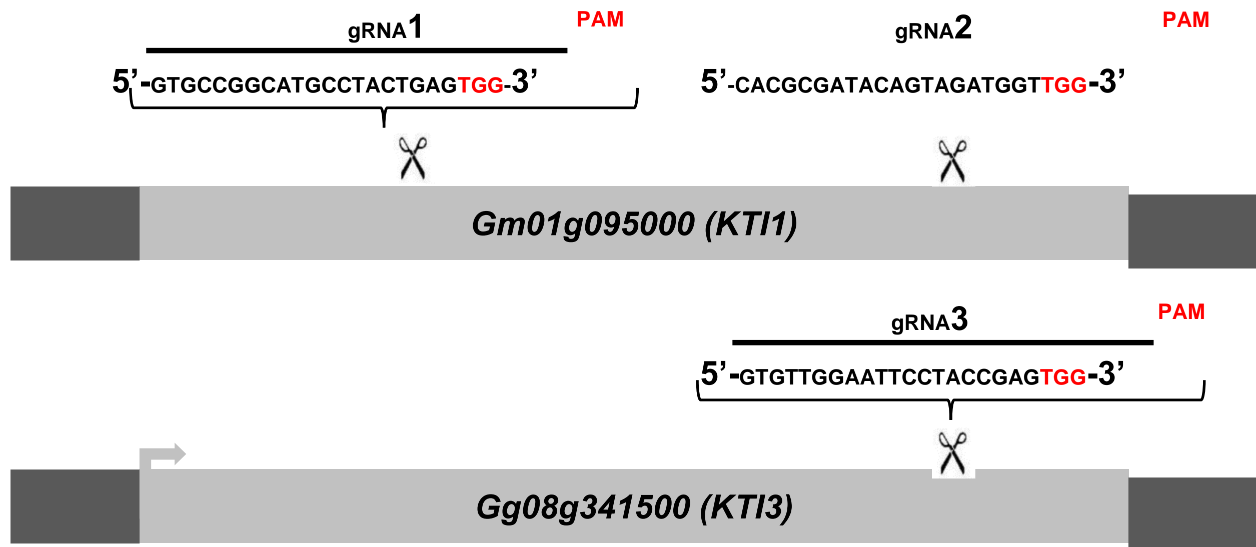
Figure 2. Expression levels of *KTI* genes in *WM82*.

(A) RNA sequencing data of 38 *KTI* genes in 26 different tissue types of cv. *Williams 82* acquired from Phytozome soybean database was used to construct the heatmap to visualize their expression patterns. (B) The expressions of four soybean *KTI* genes were monitored by real-time PCR. Samples of leaf, flower, pod, and seed tissues from 2 breeding lines, V98-9005 (normal-TI line) and V03-5903 (low-TI line), were collected for RNA extraction. After reverse transcription, real-time PCR was used to evaluate the expressions of 4 genes including *Gm01g095000*, *Gm08g341000*, *Gm08g342300*, and *Gm08g341500* in different tissues with the ELF1B as the reference gene. The expression data were normalized as ΔCT and shown as mean ± s.e. Experiments were repeated three times and obtained similar results.

A

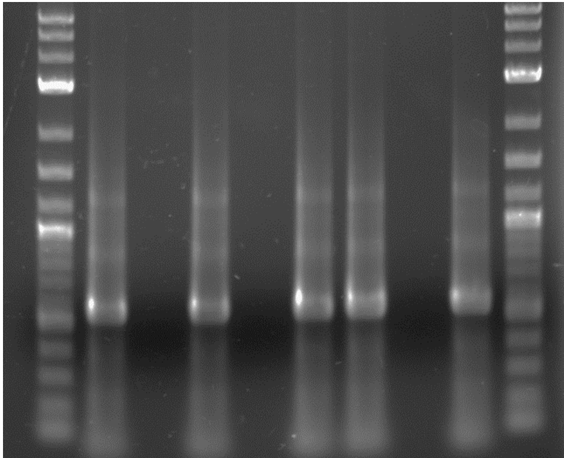


B



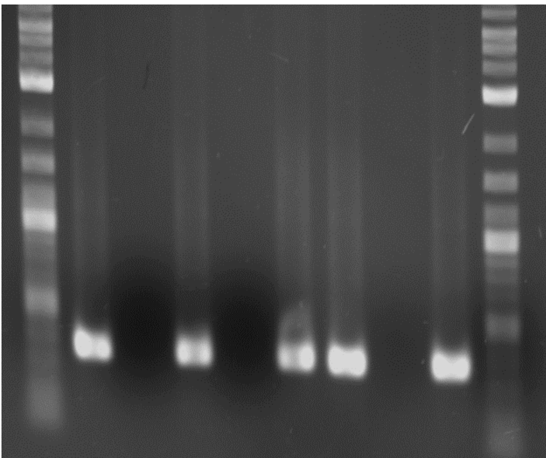
C

line # 2 4 5 7 11 17 WM82 pBAR-Cas9-*kti13*



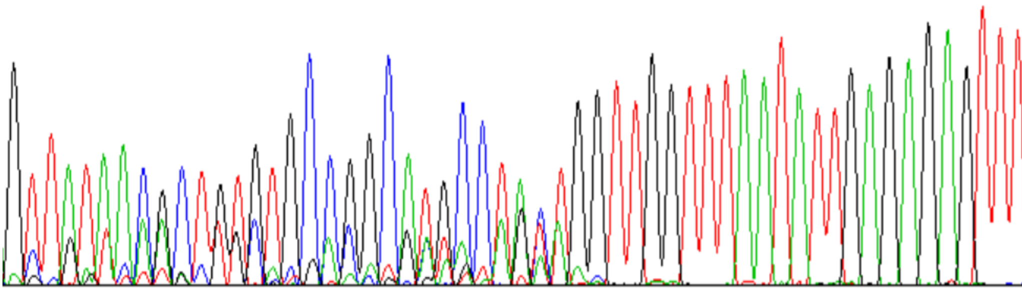
541 bp,
bar

D line # 2 4 5 7 11 17 WM82 pBAR-Cas9-kti13



371 bp,
Cas9 fragment

E



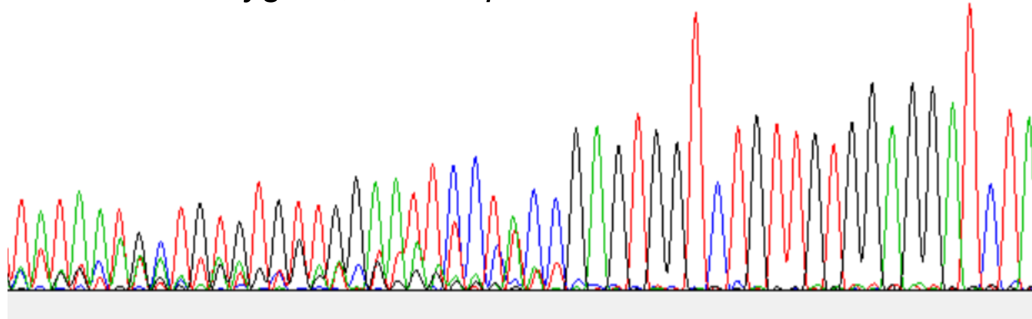


Figure 3. The scheme of binary vector used for CRISPR/Cas9 mediated gene editing on *KTI1*/*KTI3*, and transgenes and gene editing have been detected in the leaves of four T0 soybean plants. (A) The CRISPR/Cas9 construct harbors three necessary elements exhibited as below: the selection cassette consists of MAS promoter, *Bar* gene (soybean transformation selection marker), and MAS terminator; the Cas9 cassette consists of U10 promoter, Cas9 gene, and OCS terminator; three guide RNA cassettes and each of them consists of a U6 promoter, and one sgRNA. (B) The sequences of three sgRNAs is shown here. Two sgRNAs were designed, synthesized, and assembled to the plasmid to target on *KTI1*, while one sgRNA was designed, synthesized, and assembled to the plasmid to target on *KTI3*. The fragments of two transgenes, (C) *Cas9* and (D) *Bar*, have both been detected in lines #2, #5, #11, and #17 by PCR, but not lines #4 and #7. The *WM82* gDNA serves as the template for the negative control, while the plasmid DNA serves as the template for positive control. (E) The gene editing on *KTI1* and *KTI3* has also been observed in the leaf tissues of plants at T0 generation. The double peak sequence around the sgRNA region indicates the gene editing was ongoing but not completed.

To identify the *KTI* genes expressed in the seed, we analyzed the expression patterns of all *KTI* genes in cv. *WM82* based on the expression data deposited in USDA ARS's Soybase (Figure 2A). According to the expression patterns of *KTI* genes in various soybean tissues as displayed in Figure 2A, the four *KTI* genes, Gm01g095000 (*KTI1-1*), Gm08g341000 (KTI1-2), Gm08g342300 (KTI1-3, and Gm08g341500 (*KTI3*), were identified as seed-specific *KTI* genes. Soybean breeding lines V98-9005 (normal TI) and V03-5903 (low TI), presenting significantly different amounts of KTI concentration in seeds, were used to validate the tissue-specific expressions of the four *KTI* genes by real-time PCR. Gm01g095000 (*KTI1*) and Gm08g341500 (*KTI3*) were predominately

expressed in seeds compared to other tissues (Figure 2B). Both Gm08g341000 and Gm08g342300 had a relatively lower expression level in seeds than *KTI1* and *KTI3* but had higher expression in other tissue types (Figure 2B). Interestingly, both *KTI1* and *KTI3* had a relatively low expression level in the seeds of V03-5903 (low TI line), but the higher expression in V98-9005 (normal TI line). Thus, we conclude that *KTI1* and *KTI3* are two major genes that may directly contribute to the TI contents in soybean seeds.

Development of CRISPR/Cas9-based binary vector for genome-editing in soybean

To knock out the *KTI1* and *KTI3* genes from cv. *WM82* genome and create a new soybean cultivar with low TI content in soybean seeds, we developed a CRISPR/Cas9 construct, pBAR-Cas9-*kti13*, where the nuclease gene *Cas9* is expressed by Arabidopsis ubiquitin 10 (U10) promoter. A *bar* gene driven by a MAS promoter was used for the selection of the putative transformants with Bialaphos or phosphinothricin (Figure 3A). A tandem array of two sgRNAs targeting *KTI1* and one sgRNA targeting *KTI3* was expressed by the U6 RNA promoter (Figure 3B).

KTI1 and KTI3 genes are knocked out by CRISPR/Cas9-mediated gene editing

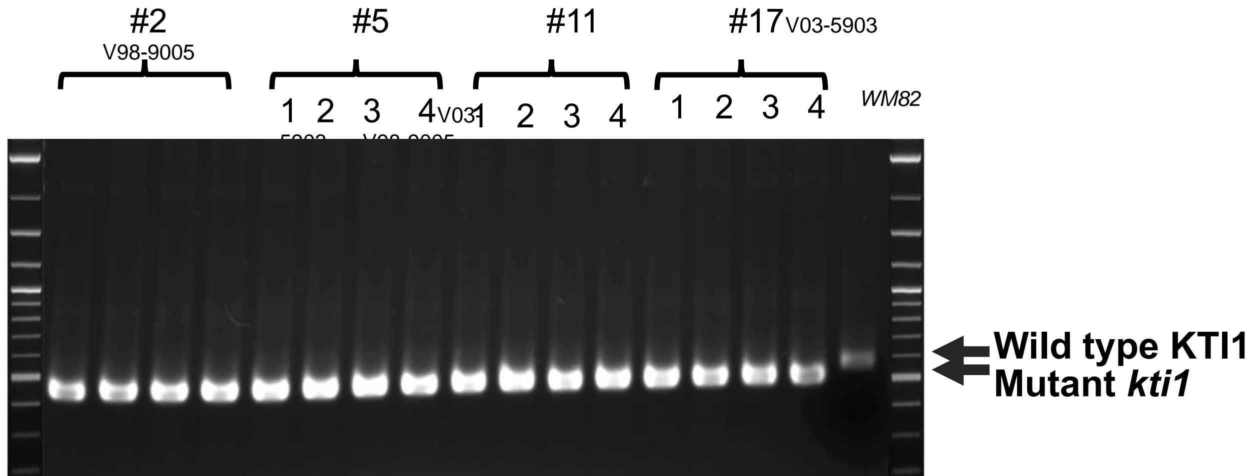
pBAR-Cas9-*kti1*/*kti3* was transformed into *WM82* via *Agrobacterium*-mediated transformation (Plant Transformation Facility at Iowa State University). Seventeen putative transgenic shoots were regenerated. Six shoots elongated and were transferred to rooting mediums. After further selection, they were transplanted into soil. Four lines, No. #2, #5, #11 and #17 were confirmed to be true transformants by positive amplification of the *bar* gene and a part of the *Cas9* gene (Figures 3C and D). The gene editing events in T0 plants were identified by amplification and sequencing of DNA fragments covering the sgRNA binding sites of *KTI1* and *KTI3*. The double peaks in the sequencing chromatograms suggest that both *KTI1* and *KTI3* genes were mutated and resulted in heterozygous alleles in the edited plant cells (Figure 3E). T0 seeds were harvested from T0 lines #2, #5, #11 and #17. Four T0 seeds of each line were randomly picked for DNA extraction and genotyping of the *KTI1* and *KTI3* genes via PCR amplification and DNA sequencing. The *KTI1* gene editing was completed and resulted in homozygous mutant alleles in all tested T0 seeds of the four lines. In addition, an identical gene editing pattern in *KTI1* was detected in all tested T0 seeds, in which a small DNA fragment (66bp) between two sgRNAs was lost after the gene editing (Figures 4A, B and C). Homozygous *KTI3* mutant alleles were only detected in T0 seeds from 2-

3, #5-4, #11-2, and #11-4 (Figure 4D, E, F and G). The gene editing patterns in *KTI3* included both small deletions and insertions that all resulted in frameshift mutations in KTI3.

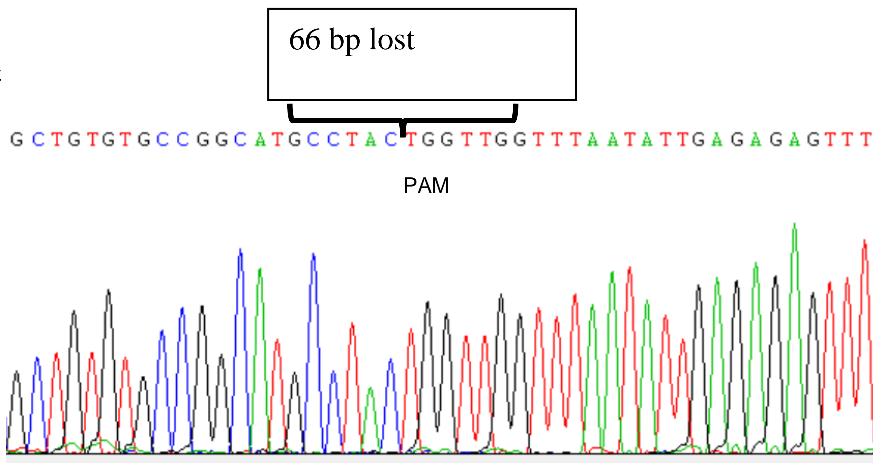
A

```
WM82      GGCATGCCTACTGAGTGGGCTATTGTGGAGAGAGAGGGTCTACAAGCTGTTAAACTTGCTGCACGCGATACAGTAGATGGTTGGTTT
Line2-1   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line2-2   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line2-3   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line2-4   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line5-1   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line5-2   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line5-3   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line5-4   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line5-26  GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line7-1   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line7-2   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line7-3   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line7-4   GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line11-1  GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line11-2  GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line11-3  GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
Line11-4  GGCATGCCTAC------------------------------------------------------------------TGGTTGGTTT
```

B



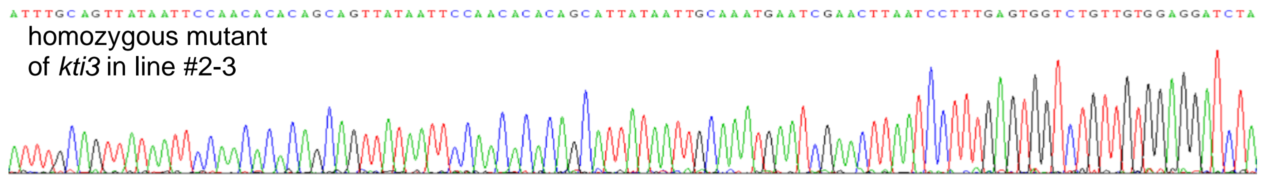
C



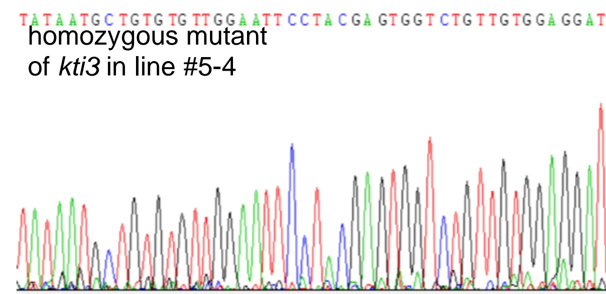
D Alignment of *kti3* mutants

```
WM82   TAAT---------------------------------------------------------------------GCTGTGTGTTGGAATTCCTACCGAGTGGTCTGTTGTGGAGGATCTACCGAGGATCTACCAGAAGGA
2-3    TAATTCCAACACACAGCAGTTATAATTCCAACACACAGCATTATAATTGCAAATGAATCGAACTTAATCCTTT----------------------GAGTGGTCTGTTGTGGAGGATCTACCGAGGATCTACCAGAAGGA
5-4    TAAT---------------------------------------------------------------------GCTGTGTGTTGGAATTCCTAC-GAGTGGTCTGTTGTGGAGGATCTACCGAGGATCTACCAGAAGGA
5-26   TAAT --------------------------------------------------------------------GCTGTGTGTTGGAATTCCTACCGA--------------------------------------AGGA
11-2   TAAT---------------------------------------------------------------------GCTGTGTGTTGGAAT------CGAGTGGTCTGTTGTGGAGGATCTACCGAGGATCTACCAGAAGGA
11-4   TAAT---------------------------------------------------------------------GCTGTGTGTTGGAAT------CGAGTGGTCTGTTGTGGAGGATCTACCGAGGATCTACCAGAAGGA
```

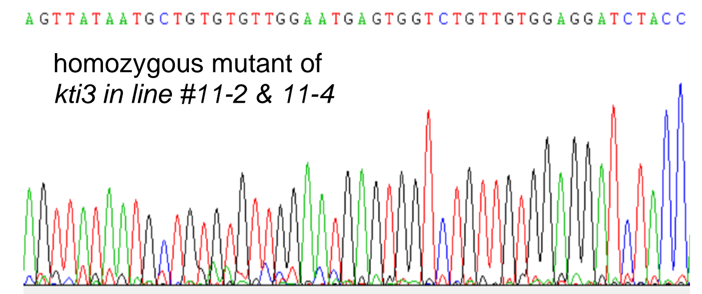
E



F



G



H

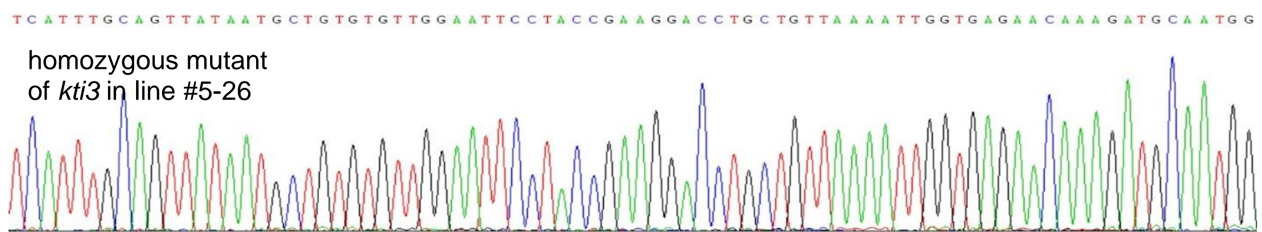


Figure 4. Gene editing on *KTI1* has been completed for all seeds of T0 generation while it has been completed on *KTI3* for some seeds of T0 generation. From each transgenic line (#2, #5, #11 and #17), four seeds of T0 generation were selected randomly for genotyping. (A) The alignment of mutant *kti1* in T0 seeds and T1 plant (#5-26) leaf, where the wild type *KTI1* in WM82 was the control. (B) Gel electrophoresis of *kti1* PCR products showed 16 seeds from line #2, #5,

#11, and #17 had the same mutant on *kti1*, in which 66 nucleotides are lost between two sgRNAs. (C) Sanger sequencing result displayed the identical mutant *kti1*. (D) The alignment of mutant *kti3* in T0 seeds (#2-3, #5-4, #5-26, #11-2 and #11-4) and T1 plant (#5-26) leaf, where the wild type *KTI1* in WM82 was the control. (E), (F), (G), (H) showed the sanger sequencing results of *kti3* mutant in #2-3, #5-4, #11-2, #11-4, and #5-26.

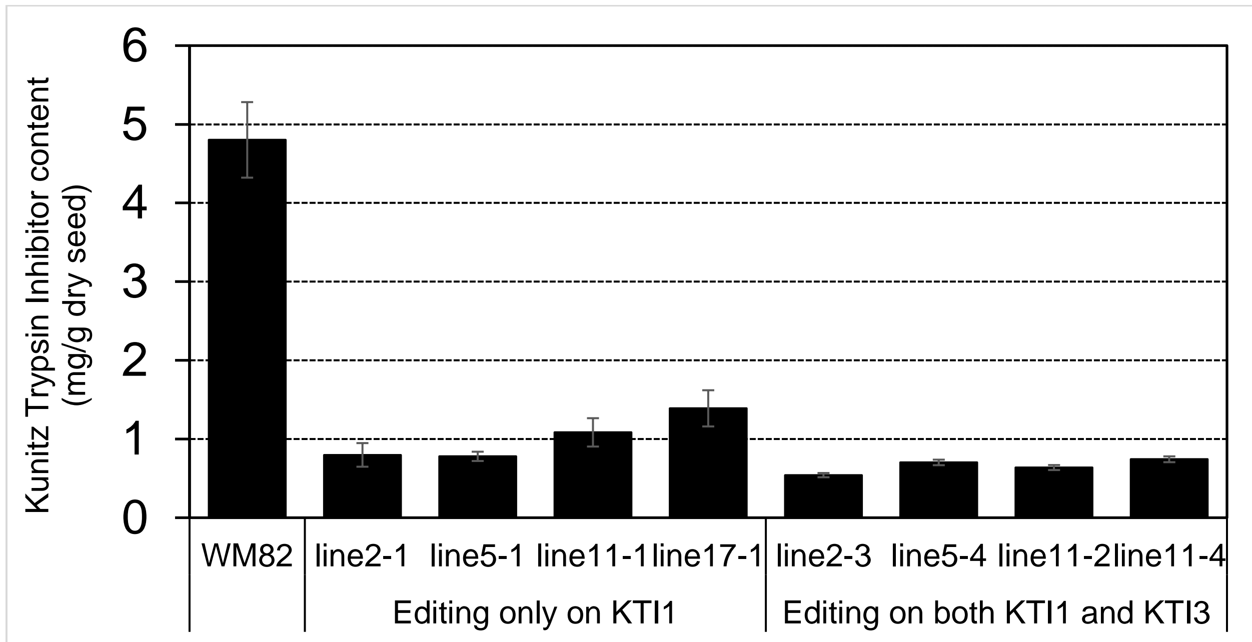


Figure 5. KTI content declined dramatically in gene-edited seeds. KTI content was measured in 4 double mutated seeds (#2-3, #5-4, #11-2, and #11-4) and 4 seeds with a single mutation on *KTI1* (#2-1, #5-1, #11-1, and #17-1), where the KTI content in *WM82* seed served as the control. Experiments were conducted with three technical replicates and showed comparable results, shown as mean ± s.e. Different letters indicate significant differences.

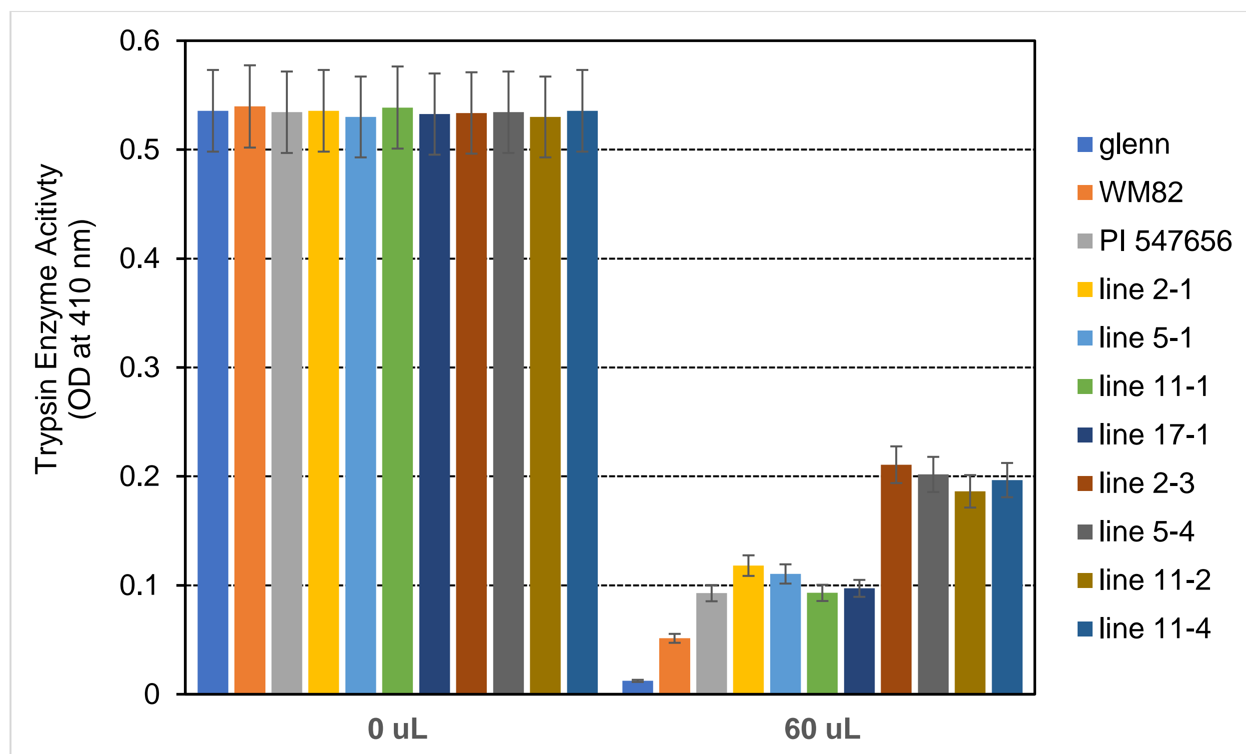


Figure 6. TIA declined dramatically in the gene-edited seeds.

Bovine trypsin enzyme activities were measured using crude extracts of 4 double mutated seeds (#2-3, #5-4, #11-2, and #11-4), 4 seeds with a single mutation on *KTI1* (#2-1, #5-1, #11-1, and #17-1), and *WM82*. Experiments were conducted with three technical replicates and showed comparable results, shown as mean ± s.e. Different letters indicate significant differences.

TI content and activity dramatically declined in the edited soybean seeds

T0 seeds were also used for quantification of the KTI content by using a HPLC-based approach (Rosso, Shang et al. 2018). The tested seeds of #2-3, #5-4, #11-2, and #11-4, which carried mutations on both *KTI1* and *KTI3* genes had the lowest KTI content (Figure 5). The tested seeds of #2-1, #5-1, #11-1, and #17-1, with only the *KTI1* mutation, also had lower KTI content than the wild-type *WM82* seeds (Figure 5). The KTI content in other genotyped seeds with editing only on *KTI1* was also lower than that in *WM82* seeds (data not shown). We further tested the trypsin inhibition activity (TIA) using crude protein extracts from the T0 seeds. As shown in Figure 6, the crude proteins of seeds with mutant *kti1* and *kti3* had the lowest TIA (Figure 6). The seeds with mutant *kti1* only also had reduced TIA (Figure 6) in comparison with *WM82* and Glenn (a commercial soybean cultivar as a control). The KTI content and TIA were ranked in order as:

kti1/3 double mutant < *kti1* single mutant ≤ PI 547656 (low TI accession) < *WM82* < Glenn. Taken together, we conclude that *KTI1* and *KTI3* are two major genes responsible for the KTI content and TIA in soybean seeds. Therefore, knockout of *KTI1* and *KTI3* reduced the KTI content and impaired the TIA in soybean seeds.

The edited *KTI1* gene lost 66 bp that may result in mutant proteins with deletion of 22 amino acids. Truncated *KTI1* may still possess some TIA. To rule out this possibility, we also tested the TIA of truncated $KTI1_{\Delta 22aa}$ protein *in vitro*. To this end, we cloned the open reading frames of $KTI1_{\Delta 66bp}$ and wild-type *KTI1* and *KTI3* into a protein expression vector, in which a 6xHis tag is fused to C-terminus of the expressed proteins. The purified proteins were subjected to a TIA assay which showed that while KTI1 and KTI3 both could inhibit trypsin activity, the truncated $KTI1_{\Delta 22aa}$ failed to suppress trypsin activity (Figure S1A and B). Therefore, the new *kti1* allele ($KTI1_{\Delta 66bp}$) encodes a truncated protein that loses its TI function.

Table 1. Knockout of *KTI1* and *KTI3* does not alter the plant growth indicator and maturity period days of cv. *WM82*.

	Plant growth indicator							Maturity period days					
Genotypes	Plant height	Main branches per plant	Pod bearing branches	No. of pods per plant	Leaf length	Leaf width	Petiole length	Plant to R1	R1 to R3	R3 to R5	R5 to R6	R6 to R8	Planting to R8
WM82	29.5 ± 2.2	2.2 ± 0.4	16.0 ± 1.6	30.0 ± 2.9	11.2 ± 1.1	6.8 ± 0.8	12.4 ± 1.5	43.8 ± 3.1	23.6 ± 2.1	18.6 ± 2.1	30.8 ± 2.6	19.6 ± 2.7	136.4 ± 4.7
Line 2 transgenic	29.2 ± 1.2	2.4 ± 0.5	16.4 ± 1.8	29.4 ± 3.8	11.6 ± 1.8	7.0 ± 1.0	11.4 ± 1.9	40.8 ± 3.3	22.4 ± 2.7	20.2 ± 3.3	32 ± 2.5	21.4 ± 2.3	136.8 ± 6.1
Line 2 non-transgenic	29.8 ± 1.2	2.5 ± 0.5	16.6 ± 1.7	30.7 ± 3.2	11.3 ± 1.5	7.1 ± 1.0	11.9 ± 2.0	42.1 ± 3.8	23.0 ± 3.0	19.4 ± 3.1	31.4 ± 2.2	20.1 ± 2.3	136.0 ± 5.5
Line 5 transgenic	29.1 ± 1.3	2.2 ± 0.4	17.2 ± 1.6	28.6 ± 3.4	11.4 ± 1.3	7.2 ± 0.8	11.2 ± 2.4	41.0 ± 4.1	20.6 ± 2.9	20.8 ± 2.8	31 ± 4.2	20.0 ± 2.1	133.4 ± 8.9
Line 5 non-transgenic	29.0 ± 1.5	2.4 ± 0.3	16.9 ± 1.8	29.1 ± 3.5	11.0 ± 1.4	6.9 ± 0.8	11.6 ± 1.8	41.7 ± 4.4	22.5 ± 2.8	20.5 ± 2.6	30.7 ± 3.2	19.9 ± 2.4	135.3 ± 6.7

R1: From planting to beginning bloom; R3: beginning bloom to beginning pod; R5: beginning pod to beginning seed; R6: beginning seed to full seed; R8: full seed to maturity. For each genotype, 5 plants were utilized for measuring plant growth indicators and maturity period days. An ANOVA test was employed here for statistical analysis.

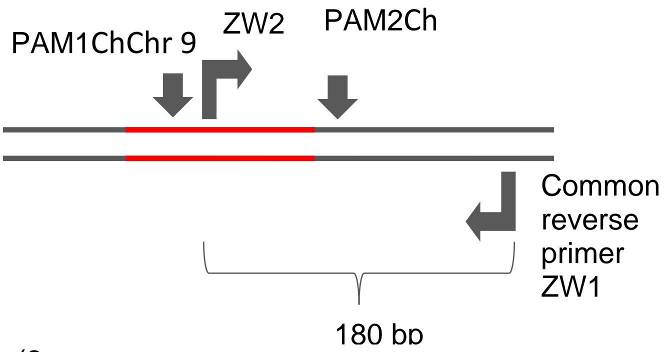
Knockout KTI1 and KTI3 did not affect plant growth and maturity period days of soybean

To examine whether the mutant *kti1/3* could significantly affect plant growth and maturity period days of soybean, we planted T0 seeds of line #2 and #5, and *WM82* in a greenhouse. By Bialaphos-mediated screening, we classified the T1 plants from line #2 and #5 as transgene-free plants or transgenic plants. We measured the agronomic traits of the transgenic plants including plant height, the number of main branches per plant, the number of pods bearing branches, the number of pods, leaf length, leaf width and petiole length. There was no significant difference in terms of all measured agronomic traits among the plants of *WM82*, Line 2 and Line 5 (Table 1). We also measured the maturity period days of the soybean plants by recording the dates from planting to beginning bloom (R1), to beginning pod (R3), to beginning seed (R5), to full seed (R6), to maturity (R8), and the total lifespan (from planting to maturity). There were no remarkable differences in terms of R1, R3, R5, R6, R8 and total life span among all tested plants (Table 1). Therefore, we conclude that knockout of *KTI1* and *KTI3* did not alter plant growth or the maturity period of soybean lines tested.

A

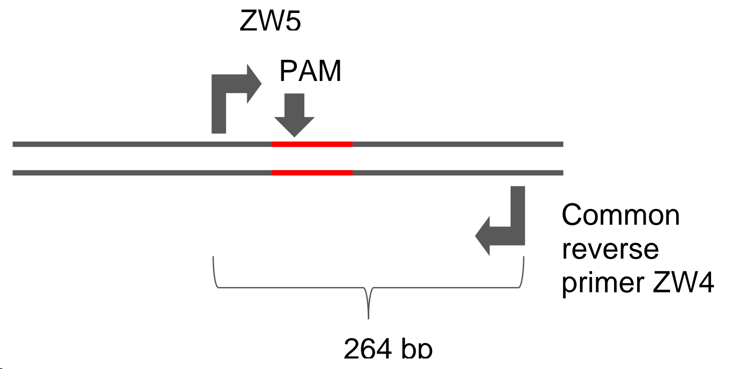
(1

Wild type



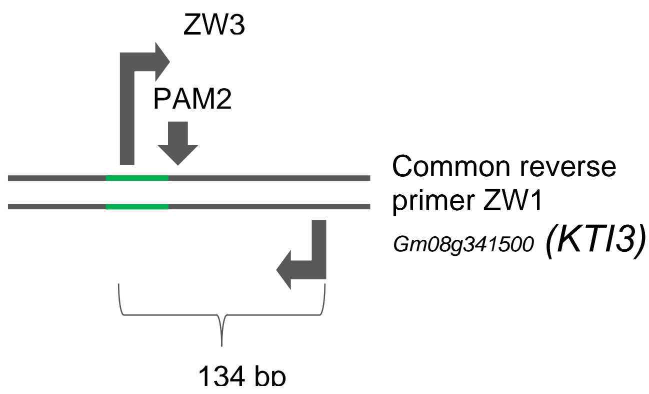
(3

Wild type



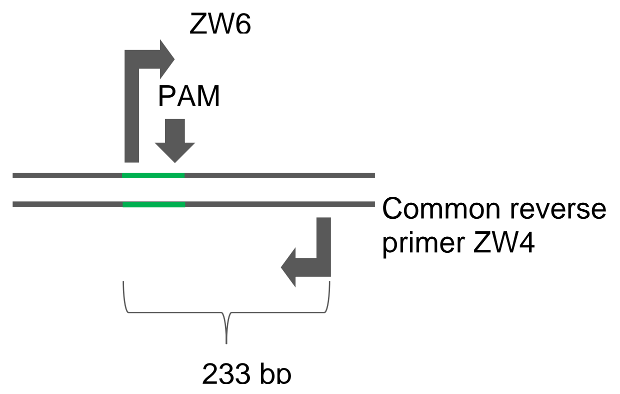
(2

Mutant



(4

Mutant *kti3*



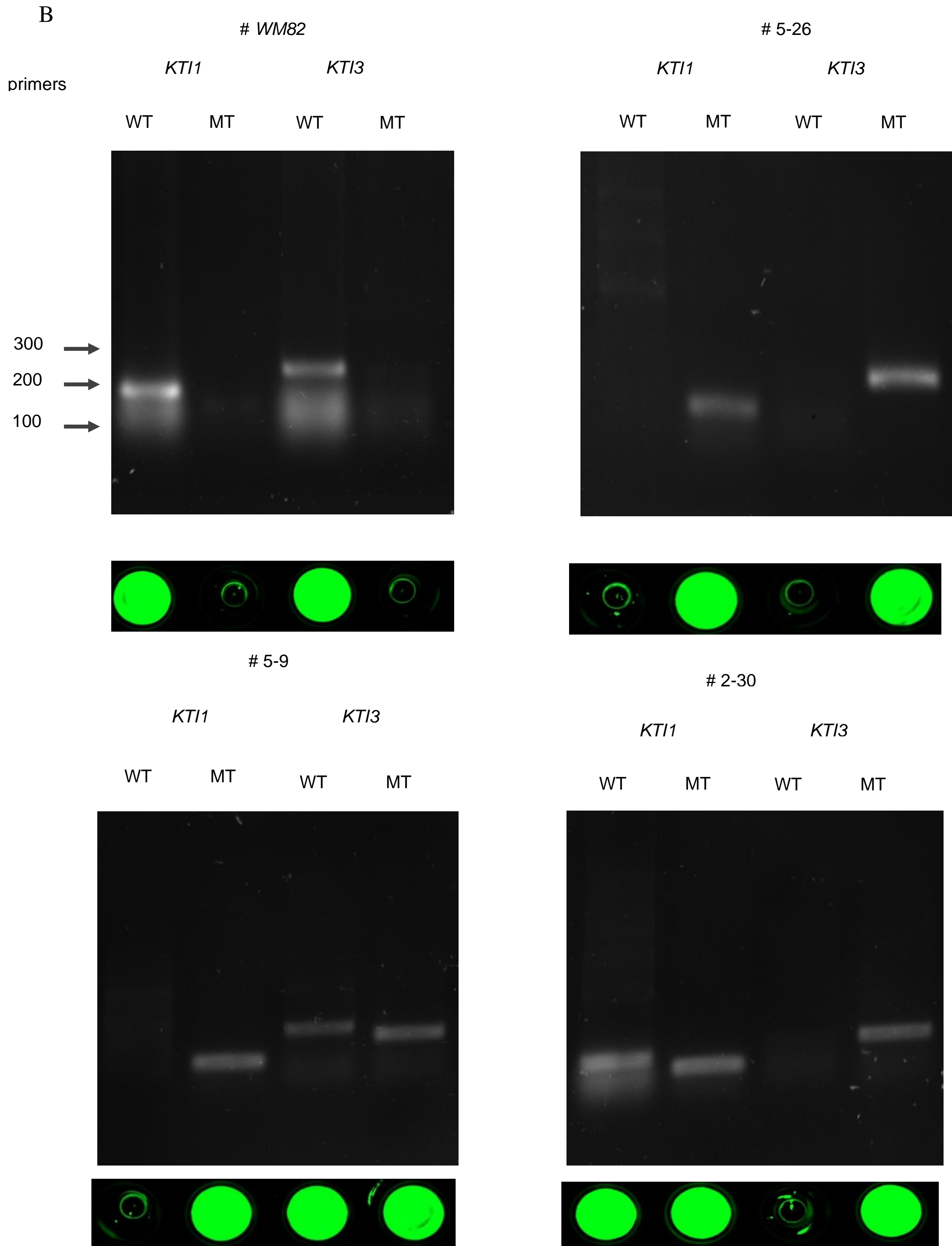

B
WM82
KTI1
KTI3
primers
WT
MT
WT
MT
300
200
100
5-26
KTI1
KTI3
WT
MT
WT
MT
5-9
KTI1
KTI3
WT
MT
WT
MT
2-30
KTI1
KTI3
WT
MT
WT
MT

Figure 7. The development of selection markers for breeding low KTI soybean varieties based on the *kti1* and *kti3* mutants generated by CRISPR/Cas9-mediated gene editing. (A) Schematic development of primers for amplification of wild type *KTI1* (1), mutant *KTI1* (2), wild type *KTI3* (3), and mutant *KTI3* (4). The red lines indicate the lost fragment in *KTI1* or *KTI3* during gene editing. The green lines indicate new DNA regions in *kti1* or *kti3* generated by splicing two fragments. (B) The 4 pairs of primers in (A) were utilized to amplify the alleles of *KTI1*, *kti1*, *KTI3*, and *kti3* with gDNA of four different soybean genotypes, including *WM82*, three transgenic lines #5-26, #5-9, and #2-30. Based on our genotyping data, #5-26 has homozygous mutations of *kti1* and *kti3*; #5-9 only has a homozygous mutation of *kti1* but carries the heterozygous mutation of *kti3*; #2-30 only has a homozygous mutation of *kti3* but carries the heterozygous mutation of *kti1*. Thus, it was clear that the pair of ZW1/ZW2 can amplify wild type *KTI1* from *WM82* and #2-30 gDNA in PCR tests, while the pair of ZW1/ZW3 can amplify mutant *kti1* from #5-9 and #5-26 gDNA. Also, the pair of ZW4/ZW5 can amplify wild type *KTI3* from *WM82* and #5-9 gDNA, while ZW4/ZW6 can amplify mutant *kti3* from #2-30 and #5-26 gDNA. As shown in the bottom panel, only the positive PCR products incubated with the dye of sybrgreen at 75 °C can display the fluorescent signals, suggesting the reliability of the developed gel-electrophoresis-free method for screening mutant alleles of *kti1* and *kti3*.

Development of molecular markers for selection of the kti1 and kti3 alleles

A double homozygous *kti1* and *kti3* mutant plant #5-26 that did not carry the Cas9 transgene was selected from T1 generation plants (Figure 4A, C, D and H). The ‘transgene-free’ soybean plants can be used to breed the low TI trait into other elite soybean cultivars. In order to co-select the *kti1* and *kti3* mutant alleles in the derived progenies, we attempted to develop co-dominant molecular markers that can distinguish between wild-type *KTI1*/*KTI3* and mutant *kti1*/*kti3* alleles.

In the genome of #5-26, the mutant allele of *kti1* had a 66 bp deletion. We designed three PCR primers, ZW1, ZW2 and ZW3 (Figure 7A and Table S1). ZW1 was a common reverse primer that can bind to the same region in both *KTI1* and *kti1* alleles, while ZW2 and ZW3 were both forward primers binding to unique sequences of *KTI1* and *kti1* alleles, respectively (Figure 7A). Soybean cultivars that carry the wild-type *KTI1* gene (*WM82*) amplified a 180 bp DNA fragment when hybridized with ZW1 and ZW2, but failed to amplify any fragments when hybridized with ZW1 and ZW3. On the contrary, the soybean lines carrying a homozygous *kti1* mutant allele (#5-26)

amplified a 134 bp DNA fragment with ZW1 and ZW3, but failed to amplify any fragments with ZW1 and ZW2 (Figure 7B).

In the genome of line #5-26, the mutant *kti3* allele had a 38 bp deletion, which allowed us to design PCR primers ZW4, ZW5 and ZW6 (Figure 7A and Table S1). ZW4 was a common reverse primer for both *KTI3* and *kti3*, while ZW5 and ZW6 were forward primers matched with unique sequences of *KTI3* and *kti3*, respectively (Figure 7A). A soybean cultivar carrying the wild-type *KTI3* gene (*MW82*) amplified a 264 bp DNA fragment with primers ZW4 and ZW5, but not with primers ZW4 and ZW6. In contrast, the soybean line carrying homozygous *kti3* allele (#5-26) can amplify a 233 bp DNA fragment with primers ZW4 and ZW6, but not ZW4 and ZW5 (Figure 7B).

We further tested these PCR primers by amplifying DNA fragments from two T1 plants that were genotyped by DNA sequencing. Line #5-9 had homozygous *kti1* alleles and heterozygous *KTI3/kti3* alleles, where the *kti3* allele was identical to the one in #5-26. Line #2-30 had homozygous *kti3* alleles and heterozygous *KTI1/kti1* alleles, where the *kti1* allele was identical to the one in #5-26. As shown in Figure 7B, PCR amplification with the different combinations of ZW1, ZW2, ZW3, ZW4, ZW5 and ZW6 can accurately identify the *KTI1/kti1* and *KTI3/kti3* genotypes of #5-9 and #2-30 (Figure 7B). Therefore, we successfully developed molecular markers to select the *kti1* and *kti3* mutant alleles generated by CRISPR/Cas9-mediated mutagenesis. These molecular markers can assist in the breeding selection of low TI soybean plants harboring *kti1/3*.

To simplify the procedure of marker-aided selection, we tested a gel-electrophoresis-free protocol that can be implemented for high throughput screening of progenies derived from a cross between a soybean cultivar carrying wild-type *KTI1/3* and one carrying the *kti1/3* mutant. In brief, all PCR products as described above were mixed with 1X SYBR Green and heated at 75°C for 10 mins and then visualized under UV light. As shown in Figure 7B, the fluorescent signals were the indications of positive amplifications in WM82 with primers ZW1/ZW2 and ZW4/ZW5, while in #5-26, the fluorescent signals can only be observed with primers ZW1/ZW3 and ZW4/ZW6 (Hirotsu, Murakami et al. 2010). Therefore, we identified the homozygous *kti1* and *kti3* alleles by directly staining the PCR products without the need of gel electrophoresis, which can significantly reduce the cost of labor and time.

Discussion

In this study, we optimized a CRISPR/cas9-vector for genome editing in soybean (Figure 3A). The modified vector allowed us to simultaneously knock out two seed-specific KTI genes (*KTI1* and *KTI3*). The *kti1/3* mutant plants grew normally in greenhouse conditions, and the seeds of *kti1/3* mutant had dramatically reduced KTI content and TI activities in comparison with wild type seed of *WM82*.

Soybean is one of the important sources of protein for animal and human consumption. However, in their evolution, soybeans have developed diverse defense components to protect seeds from being eaten by insects and animals including trypsin inhibitor, phytic acid, and raffinose family of oligosaccharides (RFOs). In agricultural practice, the anti-nutritional and biologically active factors are responsible for reduced feed efficiency when raw soybeans are fed to animals. Therefore, it is of great significance to increase feed efficiency, especially protein digestibility via assembling gene function exploration, application, and advanced genetic engineering together into the soybean industry. Proteinaceous plant trypsin inhibitors are a diverse family of (poly)peptides that play diverse roles in plant growth such as maintaining physiological homeostasis and serving the innate defense machinery [230–233]. Since TI proteins exert direct effects on pests and herbivores by interfering with their physiology, any food containing TI proteins will be avoided by these organisms. In alfalfa, the trypsin inhibitors Msti-94 and Msti-16 were demonstrated to act as a stomach poison, significantly reducing the survival and reproduction rates of aphids [233]. Mutant plants with reduced TI are usually more susceptible to pests. In wheat, α-amylase/trypsin inhibitors (ATIs) CM3 and 0.19 were identified as pest-resistance molecules, activating innate immune responses in monocytes, macrophages, and dendritic cells [231]. For example, the *Arabidopsis* lines containing silenced *atkti4* and *atkti5* were found to have a higher susceptibility to *T. urticae* (Spider mite) than wild-type plants [232]. RNAi silencing of the *AtKTI01* gene resulted in enhanced lesion development after infiltration of leaf tissue with the programmed cell death eliciting fungal toxin fumonisin B1 or the avirulent bacterial pathogen *Pseudomonas syringae* pv. *tomato DC3000* carrying *avrB* [230]. Although similar defense functions have not been reported on TI genes in the soybean genome, it is reasonably suspected that certain members of the KTI gene family have comparable protecting roles for soybean plants.

As previously discussed, the high concentration of TI proteins in soybean meal restricts the function of trypsin, which causes low digestibility and reduces its nutritional value. Thus, cultivars have been developed by the introgression of low TI traits into elite cultivars. We previously

developed a low TI line via conventional breeding: V12-4590. During field trials in 2017 and 2018, we observed that this low TI line is indeed more susceptible to multiple phytopathogens such as: all races of Soybean Cyst Nematode (SCN) (*Heterodera glycines*), Stem Canker (*Diaporthe aspalathi*), Cercospora leaf blight (*Cercospora kukuchii*), Soybean vein necrosis virus, and Downy Mildew (*Peronospora manshurica*) (Zhang, unpublished data). This suggests that the soybean *KTI* genes that are negatively selected do indeed have a role in plant immunity. It is also possible that some plant immunity related genes that genetically link with *KTI* genes are negatively selected during the breeding process. As shown in Figure 1, at least 13 KTI genes are clustered in a small region on chromosome 8. Interestingly, we also identified a putative TGACG-Binding (TGA) transcription factor (TF) that is tightly linked to the KTI gene cluster at chromosome 8. Arabidopsis TGA TFs play a positive role in systemic acquired resistance (SAR) which is crucial in plant immunity [234]. Therefore, breeding low TI soybean lines results in the loss or mutation of both of the KTI genes and the TGA TF gene, leading to an increased susceptibility of soybean challenged by phytopathogens.

The *kti1/3* mutant soybean line generated via CRISPR/Cas9-mediated mutagenesis is an isogenic line of wild type *WM82* (Table 1). Therefore, it will be an ideal test subject to see if *KTI1/3* has a direct role in plant immunity. Since *KTI1/3* were almost only expressed in seeds (Figure 2) [193], the knockout of these two genes may not interfere with plant immunity in non-seed tissue, which deserves to be further investigated in the future.

In this study, we identified a *kti1/3* double homozygous mutant along with *kti1* and *kti3* single homozygous mutants. The *kti1/3* double homozygous mutant has the lowest KTI content and trypsin inhibition activity (Figures 5 & 6). Therefore, it can be determined that (Hussain, Sheikh et al. 2018) *KTI1* and *KTI3* synergistically contribute to the KTI content and TI activity in soy proteins. The previous report suggests that the soybean line carrying natural mutations of *kti1* and *kti3* has increased Bowman-Birk (BBTI) content [193]. It is unclear if the increased BBTI content is caused by un-intentional selection during the breeding process or if the expression of *BBTI* genes is increased because of the mutations of two *KTI* genes. Therefore, it will be interesting to test the BBTI content and activity in the seed proteins of the *kti1/3* mutant generated in this study.

Despite the fact that the CRISPR/Cas9 technique has been successfully utilized to generate various soybean mutants [211], the current *agrobacterium*-mediated soybean transformation protocol is inefficient and genotype-dependent. This limits the wide implementation of CRISPR/Cas9

technique in soybean breeding programs [235]. The soybean transformation protocol employs Bialaphos as the selection agent [223]. The *Bar* gene is used as the selection marker gene and encodes a phosphinothricin acetyltransferase protein that can confer the transformants' resistance to bialaphos. It has been reported that the *Bar* gene expression must be fine-tuned in order to successfully select true transgenic plants [236]. The original CRISPR/Cas9 vector, pCut, has used a MAS (mannopine synthase) promoter to express the *Bar* gene (Peterson, Haak et al. 2016). However, for unknown reasons, the vector does not work well, even in *Arabidopsis thaliana* (Liu and Zhao, unpublished data).

With the intention of improving the transformation system, we modified the Bialaphos selection vector in the pMU3T [225]. Specifically, we replaced the Kanamycin selection marker gene with the *Bar* gene, whose expression was driven by a MAS promoter (Figure 3A). The MAS promoter is known to be most active in the roots of emerging seedlings and very active in the cotyledons and lower leaves [237]. Despite the MAS promoter having a lower level of expression than p35S, populations of transformants created with this promoter show normally distributed expression levels [238]. Thus, the MAS promoter can be used for functional screening of positive transformants in both our shoot re-generation and rooting medium supplemented with bialaphos.

It is noteworthy that, before the initiation of stable transformation, we evaluated the effectiveness of gRNAs by using a convenient *Agrobacterium*-mediated transient assay method (Wang et al, manuscript submitted for review). As soybean plants have a long-life cycle (4-6 months), the estimation of the gRNAs' effectiveness helps to avoid the waste of time and enhance the possibility of obtaining authentic gene-edited plants.

Although the soybean cultivars with natural variations on either *KTI1* (PI 68679) or *KTI3* (PI 542044) have been discovered, conventional breeding to develop new cultivars stacking with two mutant alleles via crossing will take a long time. In addition, linkage drag might lead to interference with the functions of genes located at the flanking sequences of mutant *kti1* or *kti3*. The limited genetic background of natural *kti1* and *kti3* mutants may also reduce the genetic diversity of soybean breeding lines with low TI trait, and it can be difficult to stack low TI trait with a bundle of various, desirable traits. In the present study, the *kti1/3* mutant line was created using cv. *WM82*, which has a genetic background distinct from accessions that harbor natural *kti1* and/or *kti3* mutations. Therefore, it offers a new recourse for breeding low TI traits in soybean practice.

Current soybean transformation protocol is genotype-dependent, and only a few cultivars (*WM82*, *Jack*, *Thorne*, etc.) can be efficiently transformed [235]. A mutant allele must be created in those transformable cultivars and bred into other elite cultivars via marker-assisted selection (MAS). Thus, creating a mutant allele tagged with convenient molecular markers is essential for MAS [239]. CRIPSR/Cas9-based genome editing can introduce small deletions/insertions to targeted genes, enabling us to develop molecular markers based on the sequences of the insertion and deletion mutation regions. In this study, we tested using single sgRNA and two sgRNAs for generation of mutagenesis on *KTI1* and *KTI3*, respectively (Figure 3B). Interestingly, we observed that all genotyped mutant lines carried an identical gene editing pattern of the *kti1* gene, where 66 nucleotides between the two gRNAs were deleted. The homozygous *kti1* allele can be identified in all tested seeds of the T0 generation, while homozygous *kti3* alleles were identified in some of those genotyped T0 seeds (Figure 4). Therefore, it is possible that two sgRNAs are more efficient for triggering the gene editing events in early generations of transgenic plants.

MAS has been widely implemented in plant breeding including soybean programs [239]. The selection marker of *kti3* has been developed based on its natural mutant allele, but the molecular marker for the natural mutation of *kti1* in PI 68679 is still not available [193]. Therefore, it is challenging to breed the natural *kti1*/*3* mutant alleles into a new cultivar via MAS. In this study, we created co-dominant markers that can distinguish between the wild and mutant alleles of *KTI1*/*KTI3* and *kti1*/*kti3* based on small deletions created by CRISPR/Cas9 machinery (Figure 7). In addition, a simple gel-electrophoresis-free method can be used to identify plants carrying mutant *kti1* and *kti3* alleles [240]. Thus, MAS makes it possible to effectively breed the new mutant *kti1*/*3* alleles into other elite cultivars. Taken together, the whole experimental design may serve as a practical example of how to create and select mutant alleles in crop plants in the future.

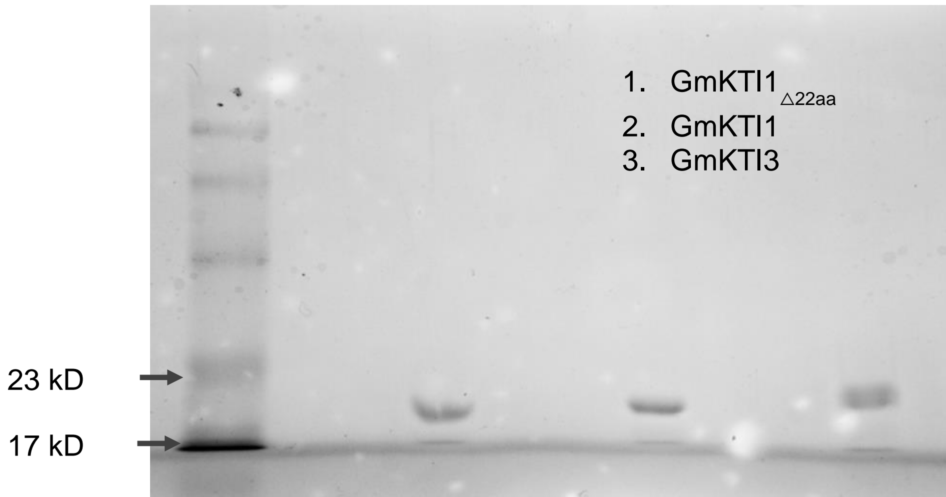
Conclusions

The present study developed non-transgenic, low TI soybean mutants in cv. William 82. The mutant gene alleles are tagged with convenient molecular markers that are suitable for high throughput marker-aided selection. We expect the low-TI soybean mutant will be widely used to

breed low-TI or TI-free soybean cultivars for commercial production in value-added meal industry and for stacking with other valuable agronomic traits in the future.

Supplemental Figures & Tables

A



B

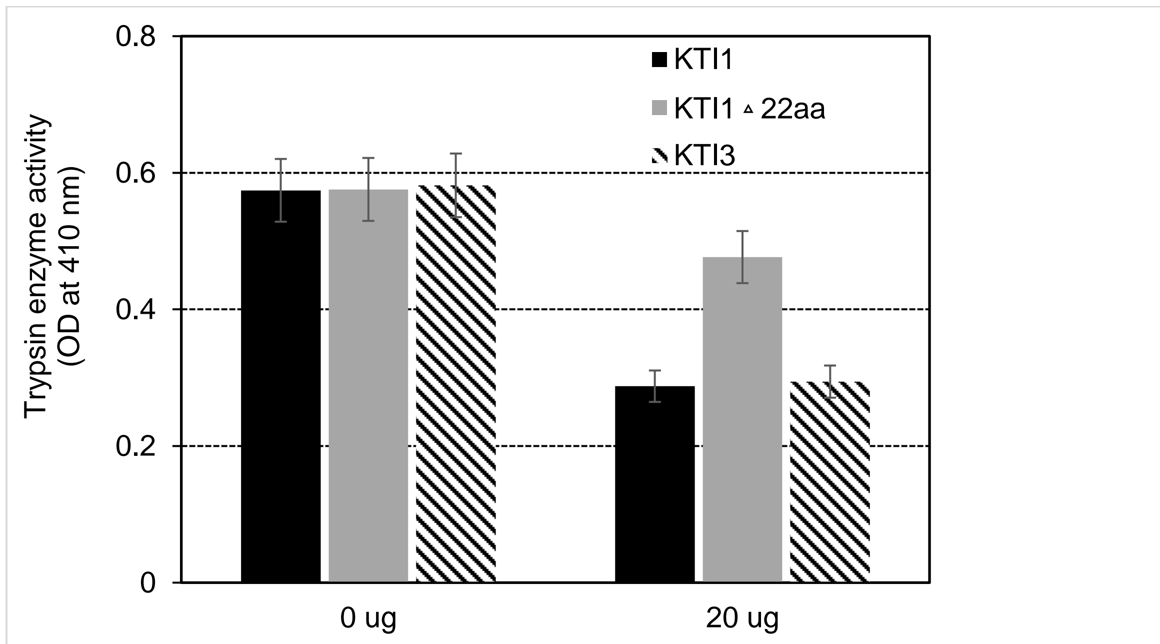


Figure S1 SDS-PAGE of purified recombinant proteins and in-frame mutated protein of KTI1$_{\Delta 22aa}$ nearly lost the TIA. (A) SDS-PAGE was used to assess the purity of three recombinant proteins, KTI1, KTI3, and KTI1$_{\Delta 22aa}$. (B) Purified proteins of KTI1 and KTI3, but not KTI1$_{\Delta 22aa}$

were able to inhibit the trypsin activity *in vivo*. Experiments were conducted with three technical replicates and obtained similar results.

Table S1. Oligo primers used in this study.

Primer	Sequence	Annotation
KTI1 For	GGGGACAAGTTTGTACAAAAAAGCAGGCTtgATGAAGAGTACTATCTTCTTTG	Forward primer for amplifying *KTI1*-1 ORF for cloning
KTI1 Rev	GGGGACCACTTTGTAcaAGAAAGCTGGGTaTGCAGTTGATGATCTAAATTTC	Reverse primer for amplifying *KTI1*-1 ORF for cloning
KTI3 For	GGGGACAAGTTTGTACAAAAAAGCAGGCTtgATGAAGAGCACCATCTTCTT	Forward primer for amplifying *KTI1*-1 ORF for cloning
KTI3 Rev	GGGGACCACTTTGTAcaAGAAAGCTGGGTaCTCACTGCGAGAAAGGCCATG	Reverse primer for amplifying *KTI1*-1 ORF for cloning
MAS pro For	ATCCGTAGCATACTAGCATCTATCAGCTAGCgtttaaacCGGCTACCGATCGACTGACTAGCATGATGATaaacTTTTCAAATCAGTGCGCAAGACG	Forward primer for amplifying the cassette for soybean transformation selection. Consists of MAS promoter, *Bar* gene, and MAS terminator.
MAS ter Rev	cgatctagtaacatagatgacaccgcgcgcgGATAATTTATTTGAAAATTCATAAGA	Reverse primer for amplifying the cassette for soybean transformation selection. Consists of MAS promoter, *Bar* gene, and MAS terminator
KTI1 (Gm01G095000) real-time PCR For	TCCTCTTCAAAACGGTGGCA	Forward primer for amplifying Gm01G095000 in real-time PCR
KTI1 (Gm01G095000) real-time PCR Rev	GTATCGCGTGCAGCAAGTTT	Reverse primer for amplifying Gm01G095000 in real-time PCR
Gm08G342300 real-time PCR For	GGAGTCGCTTTCATCACCCA	Forward primer for amplifying Gm08G342300 in real-time PCR
Gm08G342300 real-time PCR Rev	AAGCGCCTGATTCCCTTACC	Reverse primer for amplifying Gm08G342300 in real-time PCR
Gm08G341000 real-time PCR For	ACCTGGTGTTTGGATGTCGG	Forward primer for amplifying Gm08G341000 in real-time PCR
Gm08G341000 real-time PCR Rev	ACGGTTCACCAGTAACAGCA	Reverse primer for amplifying Gm08G341000 in real-time PCR
KTI3 (Gm08G341500) real-time PCR For	ATGAAGGTAACCCTCTTGAAAATG	Forward primer for amplifying Gm08G341500 in real-time PCR
KTI3 (Gm08G341500) real-time PCR Rev	ACAACAGACCACTCGGTAGG	Reverse primer for amplifying Gm08G341500 in real-time PCR
ELF1B (housekeeping gene) real-time PCR For	GTTGAAAAGCCAGGGGACA	Forward primer for amplifying ELF1B in real-time PCR
ELF1B (housekeeping gene) real-time PCR Rev	TCTTACCCCTTGAGCGTGG	Reverse primer for amplifying ELF1B in real-time PCR
KTI1 selection forward mutant	GGCTATTGTGGAGAGAGAGGG	ZW2, forward primer for wild type *KTI1* allele
KTI1/*kti1* selection common rev	CCATCATCGTCGATCTGAATC	ZW1, common reverse primer

kti1 selection forward wild type	GTGCCGGCATGCCTTGGT	ZW3, forward primer for mutant *kti1* allele
KTI3 selection forward mutant	CTACCGAGTGGTCTGTTGTG	ZW5, forward primer for wild type *KTI3* allele
KTI3/*kti3* selection common rev	TGAAACTGAACCACTAACGGT	ZW4, common reverse primer
kti3 selection forward wild type	GAATTCCTACCGAAGGAC	ZW6, forward primer for mutant *kti3* allele

Chapter 4:
Enhancing Soybean Meal Demand and Market by Developing Soy Meal-Based Aquafeeds

Zachary Shea, Ethan McAlhaney, Steve Uric, Michael Schwarz, Reza Ovissipour, Bo Zhang

Abstract

As the world population, we need to improve our food production methods to meet the growing food demand, not just in agricultural production but also seafood production. Aquaculture is one way to meet the growing seafood demand due to its cheaper cost of production and increased sustainability compared to wild-caught fisheries. While aquaculture can help meet the growing food demand, it relies on fish meal to make aquafeeds to feed most species which is expensive and not sustainable. Due to this, there has been an incentive to find cheaper alternatives, such as soybean meal. However, soybean meal cannot fully replace fish meal as of now due to antinutritional factors, such as trypsin inhibitors. This project focused on determining if it is possible to increase soybean meal inclusion in Rainbow trout diets by using low-TI soymeal. To do this, five diets were developed that varied in soybean meal amount and TI content. The diets developed included a fish meal control, a 30% SBM, a 30% VSBM, a 60% SBM, and a 60% VSBM diet with the SBM diets have high-TI soymeal and the VSBM diets having low-TI soymeal. Each diet was fed to 3 tanks with 10 fish each twice a day for 7 weeks. Weights were collected after 0, 2, and 7 weeks and specific growth rate, feed conversion ratio, total feed eaten, percentage growth rate, hepatosomatic index, and viscerosomatic index were calculated to determine fish growth. The 30% VSBM diet had no negative effects on fish growth but the 60% VSBM diet had the lowest specific growth rate, total feed eaten, percentage growth rate, and highest feed conversion ratio. This is most likely due to this diet having the lowest protein, lysine, and methionine content of all diets.

Introduction

The UN predicts that the world population will reach almost 10 billion by 2050 [241]. We will need food production to increase to feed this growing population. While we are changing aspects in agriculture to meet this demand, we also need to improve seafood production to supply the increasing population. Seafood consumption has been continuously growing for over 50 years and its consumption per capita has doubled from 10 kg to 20 kg [242,243]. Seafood is acquired

primarily through wild-caught fisheries and aquaculture. Wild-caught fisheries are not sustainable and have been one of the main causes of the depletion of wild fish species [244]. This coupled with the fact that fish reared in an aquaculture setting are cheaper, has caused aquaculture production to significantly increase and outproduce wild-caught fisheries [245,246]. Seafood from aquaculture now represents much of the seafood we eat and is estimated to reach about US $180 billion in value by 2022 [245]. Aquaculture relies heavily on aquafeed to feed the various species. The main ingredient in aquafeed is fish meal due to its high protein content and quality [247]. As aquaculture grows, the demand for aquafeed and fish meal increases as well. While fish meal has excellent nutritional value, it is very expensive and relies heavily on wild-caught fisheries to supply the fish meal resulting in fish feed consisting of about 50-70% of aquaculture production [248]. Due to this, there has been increased pressure to find a cheaper, more sustainable alternative for fish meal with the most promising alternative being plant-based protein. Many different plant sources can be used, such as soy, rapeseed, wheat, corn, and more [123,124]. Of these, soy is the most common due to its cheap cost of production and its excellent nutrient profile [123]. Soy meal costs approximately $530/metric ton compared to a metric ton of fish meal costing about $1900 [248,249]. Since the 1990s, soy as well as other plant sources have become incorporated into fish feed at higher and higher rates [250]. While soy does have many benefits for aquafeed, it is limited in how much can be used due to the presence of antinutritional factors, such as trypsin inhibitors (TI), phytate, and insoluble carbohydrates [251,252]. TI is the most influential of the antinutritional factors due to its ability to inhibit many gut proteases, interfere with protein digestion, cause decreased growth, and in extreme cases cause pancreatic hyperplasia [253]. Additionally, the palatability of diets is another factor that can affect fish growth since diets that are more palatable will be eaten more and allow for more growth [254]. It has been shown that plant-based protein sources tend to have lower palatability than fish meal due to the presence of antinutritional factors, like TI, and the fact that most plants are limited in other nutrients, such as methionine and lysine [254–256]. Therefore, there is a cause to determine if using low-TI soy meal in aquafeeds will allow for the greater replacement of fish meal while maintaining fish growth and health.

Materials & Methods

Fish Population

For this study, 180 Rainbow trout were used. The specific strain of Rainbow trout was Steelhead trout. This strain was used due to it being bred for aquaculture use. All fish were supplied by Riverance (Boise, ID).

Diet Formulation

Five different diets have been chosen for this study. These diets will be referred to as fish meal (FM), 30% and 60% soybean meal (SBM), and 30% and 60% of Virginia soybean meal using VT Barrick (VSBM). The percentages refer to how much of the fish meal was substituted with soy meal in each diet. The FM diet will be used as a control to see how soy meal compares to fish meal. The SBM diets will be used to see if VSBM can serve as a better substitution for the fish meal than standard soy meal. Besides the composition of fish meal and soy meal, all other components in the diets are the same, and all diets were created by Dyets (Bethlehem, PA). It was determined to do a 30% replacement for two diets since the typical inclusion rate for soy in aquafeeds is about 30-35% [257]. The exact formula for each diet is listed in Table 1. Proximate analysis was performed for each diet to find the composition of the diets and was done through the Agricultural Experiment Station Chemical Laboratories of University of Missouri (Columbia, MO). Table 2 list the proximate composition for all diets.

Table 1 – Composition of the FM, SBM, and VSBM diets

Ingredient	FM (g/kg)	30%SBM (g/kg)	60% SBM (g/kg)	30% VSBM (g/kg)	60% VSBM (g/kg)
Fish Meal	610	427	244	427	244
Soybean Meal	0	183	366	0	0
Virginia Soybean Meal	0	0	0	183	366
Wheat Grain Meal	251.5	251.5	251.5	251.5	251.5
Cornstarch	45	45	45	45	45
Fish Oil (with 0.02% TBHQ)	30	30	30	30	30
Cellulose	30	30	30	30	30
Vitamin Mix #390040	15	15	15	15	15
Mineral Mix #290040	15	7.5	7.5	7.5	7.5
Mineral Mix #290041	0	7.5	7.5	7.5	7.5
Vitamin C	3	3	3	3	3
Choline Chloride	0.5	0.5	0.5	0.5	0.5
total	1000	1000	1000	1000	1000

Table 2: Proximate Composition of Diets

Nutrient	FM	30% SBM	60% SBM	30% VSBM	60% VSBM
Taurine 8	0.565	0.43	0.34	0.44	0.345
Hydroxyproline	0.8	0.55	0.36	0.56	0.36
Aspartic Acid	3.685	3.635	3.52	3.515	3.27
Threonine	1.685	78.285	1.385	1.5	1.275
Serine	1.595	79.795	1.46	1.309	1.37
Glutamic Acid	6.145	6.215	6.295	6.015	5.81
Proline	2.275	2.09	1.955	2.045	1.82
Lanthionine 8	0.05	0.065	0.05	0.055	0.06
Glycine	3.445	2.765	2.21	2.75	2.09
Alanine	2.705	2.285	1.905	2.255	1.79
Cysteine	0.415	0.43	0.46	0.42	0.43
Valine	1.9	1.745	1.69	1.765	1.56
Methionine	1.135	0.9	0.725	0.91	0.69
Isoleucine	581.795	1.495	1.525	1.55	1.395
Leucine	2.855	2.765	2.6	2.635	2.365
Tyrosine	1.165	1.22	1.105	1.07	1.03
Phenylalanine	1.59	1.625	1.6	1.545	1.48
Hydroxylysine	0.13	0.095	0.07	0.085	0.07
Ornithine §	0.055	0.045	0.04	0.045	0.035
Lysine	3.085	2.775	2.42	2.715	2.25
Histidine	1.105	1.005	0.935	0.985	0.865
Arginine	2.525	2.43	2.305	2.36	2.23
Tryptophan	0.425	0.425	0.43	0.345	0.315
Total	40.94	38.12	35.435	37.025	32.905
Crude protein	43.145	39.72	36.845	38.845	33.8
Moisture	7.16	7.975	7.33	7.58	7.69
Crude Fat	7.46	6.33	5.525	9.4	11.79
Crude Fiber	2.265	3.075	3.64	3.16	3.905
Ash	14.29	11.21	8.71	11.185	8.11
Glucose	0.735	1.195	1	1.255	0.93
Fructose	0.085	0.295	0.22	0.275	0.24
Sucrose	0.185	3.15	1.425	2.63	1.69
Lactose	0.305	0.385	0.275	0.295	0.345
Maltose	0.275	0.38	0.365	0.41	0.26

Feeding Trial

Each diet was fed to three tanks each holding 10 fish giving a total sample size of n=30. Fish were fed twice each day, once at 9 am and again at 4 pm and were fed until satiety. All tanks will be

monitored for any debris, dead fish, or temperature fluctuations. All fish were fed the FM control diet for 7 days before starting the trial to acclimate them. After seven days, the actual feeding trial started. Due to resource availability, the trial was broken into two phases. The first phase consisted of the FM, 30% SBM, and 30% VSBM diets and the second phase consisted of the FM, 60% SBM, and 60% VSBM diets. Before the acclimation for the second phase started, all trout from the first phase were euthanized using tricaine methanesulfonate.

Determining Trout Growth and Health

All fish were weighed at the end of the acclimation week and weeks 2 and 7. The weights of each tank were averaged for each diet to monitor the growth of the fish. At the end of the seven-week treatment period, one fish was selected from each tank and weighed. The amount of feed that a tank consumed each day was recorded and the total feed consumed was calculated for all tanks. The total feed consumed was averaged for each diet. The fish weights were used to calculate specific growth rate (SGR) and weight gain, while fish and feed consumed weights were used to calculate feed conversion ratio (FCR). Additionally, one fish from each tank was selected to be euthanized, weighed, and had its liver and viscera removed and weighed to determine hepatosomatic index (HSI) and viscerosomatic index (VSI). The equations used to calculate weight gain, SGR<, FCR, HSI, and VSI are listed in Table 2.

Table 3: Equations Used to Calculate Various Measurements

Measurement	**Equation**
Weight Gain (%)	(Final Weight - Initial Weight)/ Initial Weight * 100
Specific Growth Rate	(ln(Final Weight)-ln(Initial Weight))/days of experiment
Feed Conversion Ratio	Weight of Total Food Eaten/(Final Weight-Initial Weight)*100
Hepatosomatic Index	Weight of Liver/Individual Fish Weight * 100
Viscerosomatic Index	Weight of Viscera/Individual Fish Weight *100

Analyzing Gut Microbiome

Fecal matter from the gut was collected from the fish that were used to determine HSI and VSI. To collect fecal samples, a scalpel was used to open the gut and the interior of the intestines was scraped from the anterior end to the posterior end. For this experiment, 3 fish from each diet were

selected. DNA was extracted from each sample using the QIAamp Fast DNA Stool Kit (Qiagen, Valencia, CA) according to Qiagen's protocol. Briefly, all samples were checked for purity and concentration using nanodrop. To determine microbial diversity among samples, all samples will have the 16s rRNA gene sequenced using Illumina gene sequencing at the Fralin Life Sciences Institute at Virginia Tech. All sequence data will be analyzed using QIIME [258].

Water Quality

To maintain healthy environments for all tanks and ensure that all tanks were in the same environment, water quality was checked regularly. Dissolved oxygen, temperature, salinity, pH, and water alkalinity was monitored daily, and total ammonia nitrogen, nitrates, and nitrites were monitored weekly. The range for each parameter is listed in Table 3.

Table 4: Water Quality Parameters for Both Phases

Water Quality	Condition
Temperature	15-16.7 °C
Salinity	0.3-0.6 ppt
Dissolved Oxygen	9-10 ppm
Alkalinity	80-100 mg/L CaC03
TAN	<0.5 ppm
Nitrates	<25 ppm
Nitrites	<0.1 ppm

Statistical Analysis

A one-way ANOVA was performed to determine if there is a significant difference between the weight gains, SGRs, FCRs, HSI values, and VSI values among all diets. When the ANOVA found a significant difference, the Fisher's Least Significant Difference (LSD) was conducted for several comparisons.

Results

Only the HSI values for the first phase were found to be significantly different ($P < 0.05$) with fish being fed the FM diet being significantly higher than the other two diets (Table 6). With that being said, the FM diet had the highest SGR, weight gain, and total feed eaten at 2.08, 177.56%, and 1133.78 g respectively with the 30% VSBM diet having the lowest for these values at 1.89, 152.52%, and 1026.87 g respectively (Table 6). The FM diet had the lowest FCR value at 1.09 and the 30% SBM diet had the highest FCR at 1.23 for this phase (Table 6).

In the second phase, the FM diet had a significantly higher SGR and weight gain than both 60% SBM and VSBM at 1.70 and 111.06% respectively and 60% VSBM had a significantly lower SGR and weight gain than the other two diets at 0.79 and 41.96% respectively ($P < 0.05$) (Table 7). In addition to having the best SGR, the FM diet also had the best FCR value at 1.49. The FM and 60% SBM diets' FCR values were significantly lower ($P < 0.05$) than 60% VSBM which had a FCR of 2.34 (Table 7). The 60% VSBM diet had a significantly lower ($P < 0.05$) total feed eaten at 800.85 g than the FM and 60% SBM diets (Table 7). The FM diet had the highest total feed eaten at 1133.78 g (Table 7). The HSI and VSI values were not significantly different between the diets for phase 2.

Table 6: Growth Parameters of Trout Fed Different Diets from Phase 1

Diet	SGR	FCR	Weight Gain (%)	HSI	VSI	Total Feed Eaten (g)
FM	2.08	1.09	177.56	2.25^{a}	13.37	1133.78
30% SBM	1.99	1.23	165.06	1.72^{b}	12.14	1167.18
30% VSBM	1.89	1.16	152.52	1.81^{b}	13.14	1026.87

*Numbers with different letters indicate significant differences between values. Numbers with no letters indicate that there was no significant difference from the ANOVA.

Table 7: Growth Parameters of Trout Fed Different Diets from Phase 2

Diet	SGR	FCR	Weight Gain (%)	HSI	VSI	Total Feed Eaten (g)
FM	1.70^{a}	1.49^{a}	111.06^{a}	1.31	11.79	1,358.37^{a}
60% SBM	1.38^{b}	1.84^{a}	83.85^{b}	1.26	12.44	1,269.50^{a}
60% VSBM	0.79^{c}	2.34^{b}	41.96^{c}	1.07	11.35	800.85^{b}

* Numbers with different letters indicate significant differences between values. Numbers with no letters indicate that there was no significant difference from the ANOVA.

Discussion

From our first phase, we found no significant difference between the growth of the trout fed the FM, 30% SBM, and 30% VSBM diets (Table 6). This was expected since it is already known that replacing 30% of fish meal with soy meal has not significantly affected fish growth when compared to a fish meal diet with no soy meal [257,259]. This does tell us that using a low-TI soy meal is usable at a 30% inclusion for aquafeeds. However, both the 60% SBM and 60% VSBM had much lower growth and weight gain than the FM diet indicating that replacing 60% of fish meal with soy meal will negatively affect fish growth (Table 7). Additionally, the 60% VSBM diet

had the lowest SGR, weight gain, and total feed eaten and the highest FCR of all diets (Table 7). This indicates that reducing the TI content in soy meal is not enough to increase fish meal replacement with soy meal and in order to increase the amount of soy meal in aquafeed, other traits need to be looked at as well. It is important to note that the 60% VSBM diet did have lower methionine and lysine content compared to other diets. This could partially explain the lower growth and higher FCR since methionine and lysine are both important for trout growth. Other studies have found that trout fed diets with deficient methionine and lysine content had lower growth [258, 259]. Furthermore, the 60% VSBM diet had significantly lower total feed eaten at 800.85 g. The 60% SBM diet almost had a 50% higher total feed eaten than 60% VSBM and the FM diet's total feed eaten was almost double the 60% VSBM's. This shows that the soy meal from VT Barrick has lower palatability than the other diets. Most likely this is the main reason why the 60% VSBM diet had much lower growth than the other diets. If a diet has lower palatability, it will result in less growth even if that diet has lower TI content. The two 60% soy meal diets having lower growth than fish meal diets is supported by other studies [254–256], but having low-TI is not enough to increase the palatability of the diet or allow for the greater replacement of fish meal in aquafeeds for Rainbow trout.

There was no significant difference in VSI values in either phase. This implies that while the 60% VSBM diet did result in lower growth of trout, it did not negatively affect their health. There was a slight difference in the health of the fish in this phase but there was no significant difference between HSI's in the second phase. This makes it more difficult to tell if there was a particular effect from the diets on HSI or not and requires more work to look into.

Conclusions/Future Directions

For chapter 2, the next steps would be to validate the significant SNPs that we found to see if they are associated with high test weight. Additionally, we can use the one accession, PI 87059, that had high test weight in all locations and all years in our breeding scheme to help breed for high test weight. Lastly, we can use the information we learned from the correlation study to make more informed breeding decisions. We can use it to ensure that we are balancing the need for high oil content with the need to prevent test weight from decreasing more. For chapter 3, the next step would be to determine how our edited lines perform in field trials. We would compare this to the unedited line as well as high and low-TI lines and see if it has less pest damage than the low-TI and unedited lines. We would also want to see if it has similar pest damage compared to the high-TI lines. Additionally, we would compare yield and other traits, like protein and oil content, to ensure that they were unaffected by gene editing. For chapter 4, we need to see if we can improve the palatability of the VSBM diets. One way could be to add a lysine and methionine supplement to it and repeat a feeding trial with Rainbow trout to determine if this increases palatability. Additionally, we could try using some aquafeed additives that help increase the palatability of plant-based diets. There has already been some work done in this field and it could potentially help our low-TI VSBM diet.

Bibliography

1. International: World Oilseed Production | [Internet]. [cited 2022 Nov 21]. Available from: http://soystats.com/international-world-oilseed-production/

2. Goldsmith PD. 5 - Economics of Soybean Production, Marketing, and Utilization. In: Johnson LA, White PJ, Galloway R, editors. Soybeans [Internet]. AOCS Press; 2008 [cited 2022 Nov 21]. p. 117–50. Available from: https://www.sciencedirect.com/science/article/pii/B9781893997646500081

3. Hartman GL, West ED, Herman TK. Crops that feed the World 2. Soybean-worldwide production, use, and constraints caused by pathogens and pests. Food Security. 2011 Mar;3(1):5–17.

4. International: World Soybean Production | [Internet]. [cited 2021 Oct 18]. Available from: http://soystats.com/international-world-soybean-production/

5. World Soybean Production [Internet]. Soy Meal Info Center. 2018 [cited 2021 Aug 30]. Available from: https://www.soymeal.org/soy-meal-articles/world-soybean-production/

6. Hymowitz T, Singh R. Taxonomy and Speciation. In: Soybeans: Improvement, Production, and Uses. 2nd ed. Madison, WI: ASA, CSSA, SSSA; 1987.

7. U.S. Exports: Soy Exports by Customer [Internet]. Available from: http://soystats.com/u-s-exports-soy-exports-by-customersda/

8. History of tillage and tillage research [Internet]. [cited 2022 Nov 21]. Available from: https://extension.umn.edu/soil-management-and-health/history-tillage-and-tillage-research

9. North Carolina Soybean Production Guide | NC State Extension Publications [Internet]. [cited 2022 Nov 21]. Available from: https://content.ces.ncsu.edu/north-carolina-soybean-production-guide

10. No-Till Research [Internet]. [cited 2022 Nov 21]. Available from: https://taes.tennessee.edu/recs/milan/research/notill.asp

11. Is tillage needed for your soybean crop? | Integrated Crop Management [Internet]. [cited 2022 Nov 21]. Available from: https://crops.extension.iastate.edu/encyclopedia/tillage-needed-your-soybean-crop

12. Yusuf RI, Siemens JC, Bullock DG. Growth Analysis of Soybean under No-Tillage and Conventional Tillage Systems. Agronomy Journal. 1999;91(6):928–33.

13. Rathke GW, Wienhold BJ, Wilhelm WW, Diepenbrock W. Tillage and rotation effect on corn–soybean energy balances in eastern Nebraska. Soil and Tillage Research. 2007 Nov;97(1):60–70.

14. Popp MP, Keisling TC, McNew RW, Oliver LR, Dillon CR, Wallace DM. Planting Date, Cultivar, and Tillage System Effects on Dryland Soybean Production. Agronomy Journal. 2002;94(1):81–8.

15. Barret J. Corn and soybean production up in 2020, USDA Reports [Internet]. National Agriculture Statistics Service. 2021 [cited 2021 Aug 25]. Available from: https://www.nass.usda.gov/Newsroom/2021/01-12-2021a.php

16. Corn and Soybean Response to Rotation Sequence, Row Spacing, and Tillage System | Agronomy Journal [Internet]. [cited 2022 Nov 21]. Available from: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2134/agronj2003.9650

17. Fernandez-Cornejo J, Hallahan C, Nehring R, Wechsler S, Grube A. Conservation Tillage, Herbicide Use, and Genetically Engineered Crops in the United States: The Case of Soybeans. :11.

18. Ferreira MC, Andrade D de S, Takemura M, Hungria M. Tillage method and crop rotation e€ects on the population sizes and diversity of bradyrhizobia nodulating soybean. Soil Biology. 2000;11.

19. Okada H, Harada H. Effects of tillage and fertilizer on nematode communities in a Japanese soybean field. Applied Soil Ecology. 2007 Mar 1;35:582–98.

20. Carter MR, Gregorich EG. Soil Sampling and Methods of Analysis [Internet]. second. 2007 [cited 2022 Nov 21]. Available from: https://www.taylorfrancis.com/books/mono/10.1201/9781420005271/soil-sampling-methods-analysis-gregorich-carter

21. Ferguson RB. EC06-155 Nutrient Management for Agronomic Crops in Nebraska. Historical Materials from University of Nebraska-Lincoln Extension [Internet]. 2006 Jan 1; Available from: https://digitalcommons.unl.edu/extensionhist/1711

22. Orf J h., Diers BW, Boerma HR. Managing Inputs for Peak Production. In: Soybean: Improvement, Production, and Uses. 3rd ed. Madison, WI: ASA,CSA,SSSA; p. 451–525.

23. Varco J. Nutrition and Fertility Requirements. In: Soybean Production in the Mid-South. Boca Raton, Fl: CRC Press; 1999. p. 53–70.

24. Zhou XJ, Liang Y, Chen H, Shen SH, Jing YX. Effects of rhizobia inoculation and nitrogen fertilization on photosynthetic physiology of soybean. Photosynthetica. 2006 Dec 1;44(4):530–5.

25. Brann D, Abaye A, Azenegashe O, Peterson P, Paul R, Chalmers D, et al. Agronomy Handbook. 2009.

26. Soils and soil fertility | Titel [Internet]. [cited 2022 Nov 21]. Available from: https://library.wur.nl/WebQuery/titel/1762424

27. Carrera CS, Reynoso CM, Funes GJ, Martínez MJ, Dardanelli J, Resnik SL. Amino acid composition of soybean seeds as affected by climatic variables. Pesq agropec bras. 2011 Dec;46(12):1579–87.

28. Goldflus F, Ceccantini M, Santos W. Amino acid content of soybean samples collected in different Brazilian states: harvest 2003/2004. Rev Bras Cienc Avic. 2006 Jun;8(2):105–11.

29. Wolf RB, Cavins JF, Kleiman R, Black LT. Effect of temperature on soybean seed constituents: Oil, protein, moisture, fatty acids, amino acids and sugars. J Am Oil Chem Soc. 1982 May 1;59(5):230–2.

30. Krishnan HB, Bennett JO, Kim WS, Krishnan AH, Mawhinney TP. Nitrogen lowers the sulfur amino acid content of soybean (Glycine max [L.] Merr.) by regulating the accumulation of Bowman-Birk protease inhibitor. J Agric Food Chem. 2005 Aug 10;53(16):6347–54.

31. Kaur G, Serson WR, Orlowski JM, McCoy JM, Golden BR, Bellaloui N. Nitrogen Sources and Rates Affect Soybean Seed Composition in Mississippi. Agronomy. 2017 Dec;7(4):77.

32. Raboy V, Dickinson DB. Effect of phosphorus and zinc nutrition on soybean seed phytic Acid and zinc. Plant Physiol. 1984 Aug;75(4):1094–8.

33. Changes in proteins and amino acids in developing soybean seed and effect of phosphorus nutrition - Kapoora - 1977 - Journal of the Science of Food and Agriculture - Wiley Online Library [Internet]. [cited 2022 Nov 21]. Available from: https://onlinelibrary-wiley-com.ezproxy.lib.vt.edu/doi/abs/10.1002/jsfa.2740280202?casa_token=PAfOzIKz5isAAAAA:7aG_V0wwRhaNAK1OsIGa7yZpA6EQuwVxjVYvcBQwaIHLbvpCF6zP_vS2RCPX8bZNboZ92ElAkn6L4A

34. Vyn TJ, Yin X, Bruulsema TW, Jackson CJC, Rajcan I, Brouder SM. Potassium fertilization effects on isoflavone concentrations in soybean [Glycine max (L.) Merr.]. J Agric Food Chem. 2002 Jun 5;50(12):3501–6.

35. Myers SW, Gratton C, Wolkowski RP, Hogg DB, Wedberg JL. Effect of soil potassium availability on soybean aphid (Hemiptera: Aphididae) population dynamics and soybean yield. J Econ Entomol. 2005 Feb;98(1):113–20.

36. Walter AJ, DiFonzo CD. Soil potassium deficiency affects soybean phloem nitrogen and soybean aphid populations. Environ Entomol. 2007 Feb;36(1):26–33.

37. Devi K, Nongdren L, Singh K, Singh S, Singh SB, Singh K. Influence of Sulphur and Boron Fertilization on Yield, Quality, Nutrient Uptake and Economics of Soybean (Glycine max) under Upland Conditions. Journal of Agricultural Science. 2012 Feb 2;4.

38. Gayler KR, Sykes GE. Effects of nutritional stress on the storage proteins of soybeans. Plant Physiol. 1985 Jul;78(3):582–5.

39. Wilhelm WW, Wortmann CS. Tillage and Rotation Interactions for Corn and Soybean Grain Yield as Affected by Precipitation and Air Temperature. Agronomy Journal. 2004;96(2):425–32.

40. Karlen DL, Hurley EG, Andrews SS, Cambardella CA, Meek DW, Duffy MD, et al. Crop Rotation Effects on Soil Quality at Three Northern Corn/Soybean Belt Locations. Agronomy Journal. 2006 May;98(3):484–95.

41. (PDF) Microbial biomass in soils under different tillage and crop rotation systems [Internet]. [cited 2022 Nov 21]. Available from: https://www.researchgate.net/publication/226009145_Microbial_biomass_in_soils_under_different_tillage_and_crop_rotation_systems

42. Hoeft RG, Aldrich SR, Nafziger ED, Johnson RR. Modern corn and soybean production. 1st ed. Champaign, Ill.: MCSP Publications; 2000. 353 p.

43. Mirsky SB, Curran WS, Mortenseny DM, Ryany MR, Shumway DL. Timing of Cover-Crop Management Effects on Weed Suppression in No-Till Planted Soybean using a Roller-Crimper. Weed Science. 2011 Sep;59(3):380–9.

44. Williams SM, Weil RR. Crop Cover Root Channels May Alleviate Soil Compaction Effects on Soybean Crop. Soil Science Society of America Journal. 2004;68(4):1403–9.

45. Davis AS. Cover-Crop Roller-Crimper Contributes to Weed Management in No-Till Soybean. Weed Science. 2010;58(3):300–9.

46. Reddy KN. Effects of Cereal and Legume Cover Crop Residues on Weeds, Yield, and Net Return in Soybean (Glycine max) [1]. Weed Technology. 2001 Oct;15(4):660–8.

47. Shapiro BI, Brorsen BW, Doster DH. Adoption of Double-Cropping Soybeans and Wheat. Journal of Agricultural and Applied Economics. 1992;24(2):33–40.

48. Kelley KW. Double-cropping Winter Wheat and Soybean Improves Net Returns in the Eastern Great Plains. Crop Management. 2003;2(1):1–7.

49. Jones BP, Holshouser DL, Alley MM, Roygard JKF, Anderson-Cook CM. Double-Crop Soybean Leaf Area and Yield Responses to Mid-Atlantic Soils and Cropping Systems. Agronomy Journal. 2003;95(2):436–45.

50. Potential of Early-Maturing Soybean Cultivars in Late Plantings | Agronomy Journal [Internet]. [cited 2022 Nov 21]. Available from: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2134/agronj2000.923532x

51. Hayder G, Mumtaz SS, Khan A, Khan S. Maize and Soybean Intercropping under Various Levels of Soybean Seed Rates. Asian Journal of Plant Sciences. 2(3):339–41.

52. Muoneke C, Ogwuche M, Kalu B. Effect of maize planting density on the performance of maize/soybean intercropping system in a guinea savannah agroecosystem. African Journal of Agricultural Research. 2008 Jan 1;2:667–77.

53. Li J, Sun J, Zhang F, Li X, Rengel Z, Yang S. Wheat/maize or wheat/soybean strip intercropping: II. Recovery or compensation of maize and soybean after wheat harvesting. Field Crops Research. 2001;71(3):173–81.

54. Zhang F, Li L. Using competitive and facilitative interactions in intercropping systems enhances crop productivity and nutrient-use efficiency. Plant and Soil. 2003 Jan 1;248(1):305–12.

55. Li X, Mu Y, Cheng Y, Liu X, Nian H. Effects of intercropping sugarcane and soybean on growth, rhizosphere soil microbes, nitrogen and phosphorus availability. Acta Physiology Plant. 2013;35(4):1113–9.

56. Hock SM, Knezevic SZ, Martin AR, Lindquist JL. Soybean Row Spacing and Weed Emergence Time Influence Weed Competitiveness and Competitive Indices. Weed Science. 2006;54(1):38–46.

57. Knezevic SZ, Evans S, Mainz M. Row Spacing Influences the Critical Timing for Weed Removal in Soybean (Glycine max)1. Weed Technology. 2009 Sep 2;17:666–73.

58. Arce GD, Pedersen P, Hartzler RG. Soybean Seeding Rate Effects on Weed Management. Weed Technology. 2009 Mar;23(1):17–22.

59. Holshouser DL, Whittaker JP. Plant Population and Row-Spacing Effects on Early Soybean Production Systems in the Mid-Atlantic USA. Agronomy Journal. 2002;94(3):603–11.

60. Cox WilliamJ, Cherney JH. Growth and Yield Responses of Soybean to Row Spacing and Seeding Rate. Agronomy Journal. 2011;103(1):123–8.

61. De Bruin JL, Pedersen P. Effect of Row Spacing and Seeding Rate on Soybean Yield. Agronomy Journal. 2008;100(3):704–10.

62. Lambert DM, Lowenberg-DeBoer J. Economic Analysis of Row Spacing for Corn and Soybean. Agronomy Journal. 2003;95(3):564–73.

63. Specht J e., Chase K, Macrander M, Graef G l., Chung J, Markwell J p., et al. Soybean Response to Water: A QTL Analysis of Drought Tolerance. Crop Science. 2001;41(2):493–509.

64. Soybean Irrigation. In: Soybean Production in the Mid-South. Boca Raton, Fl: CRC Press; 1999. p. 19–142.

65. Specht JE, Hume DJ, Kumudini SV. Soybean Yield Potential—A Genetic and Physiological Perspective. Crop Science. 1999;39(6):1560–70.

66. Arora VK, Singh CB, Sidhu AS, Thind SS. Irrigation, tillage and mulching effects on soybean yield and water productivity in relation to soil texture. Agricultural Water Management. 2011 Feb 1;98(4):563–8.

67. Karam F, Masaad R, Sfeir T, Mounzer O, Rouphael Y. Evapotranspiration and seed yield of field grown soybean under deficit irrigation conditions. Agricultural water management [Internet]. 2005 [cited 2022 Nov 21]; Available from: https://scholar.google.com/scholar_lookup?title=Evapotranspiration+and+seed+yield+of+field+grown+soybean+under+deficit+irrigation+conditions&author=Karam%2C+F.&publication_year=2005

68. Nijbroek R, Jones J. Optimizing irrigation management for a spatially variable soybean field. Agricultural Systems. 2003 Feb 1;76:359–77.

69. Mallarino AP, Haq MU, Wittry D, Bermudez M. Variation in Soybean Response to Early Season Foliar Fertilization among and within Fields. Agronomy Journal. 2001;93(6):1220–6.

70. Freeborn JR, Holshouser DL, Alley MM, Powell NL, Orcutt DM. Soybean Yield Response to Reproductive Stage Soil-Applied Nitrogen and Foliar-Applied Boron. Agronomy Journal. 2001;93(6):1200–9.

71. Salvagiotti F, Specht JE, Cassman KG, Walters DT, Weiss A, Dobermann A. Growth and Nitrogen Fixation in High-Yielding Soybean: Impact of Nitrogen Fertilization. Agronomy Journal. 2009;101(4):958–70.

72. Haq MU, Mallarino AP. Soybean Yield and Nutrient Composition as Affected by Early Season Foliar Fertilization. Agronomy Journal. 2000;92(1):16–24.

73. Haq MU, Mallarino AP. Response of Soybean Grain Oil and Protein Concentrations to Foliar and Soil Fertilization. Agronomy Journal. 2005;97(3):910–8.

74. Popovic V, Glamoclija D, Sikora V, Dekic V, Cervenski J, Simic D, et al. GENOTYPIC SPECIFICITY OF SOYBEAN [GLYCINE MAX. (L) MERR.] UNDER CONDITIONS OF FOLIAR FERTILIZATION. Romania Agircultural Research. 2013;(30).

75. Zhang X, Qiong L. Brief Introduction of Main Diseases and Insect Pests in Soybean Production in the Global Top Five Soybean Producing Countries. Plant Diseases & Pests. 2018;17–21.

76. Lahiri S, Reisig DD. Ecology and Management of Kudzu Bug (Hemiptera: Plataspidae) in Southeastern Soybeans. J Integr Pest Manag [Internet]. 2016 Jan 1 [cited 2020 Jan 9];7(1). Available from: https://academic.oup.com/jipm/article/7/1/14/2658149

77. Mexican bean beetle - Epilachna varivestis Mulsant [Internet]. [cited 2020 Jan 9]. Available from: http://entnemdept.ufl.edu/creatures/veg/bean/mexican_bean_beetle.htm

78. USDA APHIS | Japanese Beetle [Internet]. [cited 2020 Jan 9]. Available from: https://www.aphis.usda.gov/aphis/ourfocus/planthealth/plant-pest-and-disease-programs/pests-and-diseases/japanese-beetle/japanese-beetle

79. Managing spider mite on soybean [Internet]. [cited 2020 Jan 9]. Available from: https://extension.umn.edu/soybean-pest-management/managing-spider-mite-soybean

80. Leskey T, Hamilton G, Nielsen A, Polk D, Rodriquez-Saona C, Christopher J, et al. Pest Status of the Brown Marmorated Stink Bug, Halyomorpha halys in the USA. Outlooks in Pest Management. 2012;

81. Sustainable Solutions To Pest Problems [Internet]. Integrated Pest Management. 2019 [cited 2020 Jan 20]. Available from: https://www.sccgov.org/sites/ipm/about/Pages/benefits-of-ipm.aspx

82. Norris R. Integrated Pest Management. In: Encyclopedia of Biological Invasions [Internet]. 2004 [cited 2020 Jan 20]. Available from: https://books.google.com/books?hl=en&lr=&id=qgQmDQAAQBAJ&oi=fnd&pg=PA353&dq=integrated+pest+management&ots=T9pG1zlok0&sig=dl87ucNN0OuYWpZ9AhVwMaf6AHE#v=onepage&q=integrated%20pest%20management&f=false

83. Motaphale AA, Bhosle BB. Validation of Integrated Pest Management modules on defoliating insect pests of soybean. Agricultural Science Digest. 2016 Dec;36(4):303–6.

84. Hokkanen HMT. Trap Cropping in Pest Management. Annual Review of Entomology. 1991;(36):119–38.

85. Soybean Insect Management | Extending Agricultural Information [Internet]. Mississippi Crop Situation. 2017 [cited 2020 Jan 20]. Available from: https://www.mississippi-crops.com/insect-control-guide/soybean-insect-management/

86. Vivian R, Reis A, Kálnay PA, Vargas L, Ferreira ACC, Mariani F. Weed Management in Soybean — Issues and Practices. Soybean - Pest Resistance [Internet]. 2013 Feb 13 [cited 2020 Jan 6]; Available from: https://www.intechopen.com/books/soybean-pest-resistance/weed-management-in-soybean-issues-and-practices

87. Oerke EC, Dehne HW. Safeguarding production—losses in major crops and the role of crop protection. Crop Protection. 2004 Apr;23(4):275–85.

88. Hartzler B. Managing weeds to protect crop yields | Integrated Crop Management [Internet]. Iowa State University Extension and Outreach. [cited 2020 Jan 19]. Available from: https://crops.extension.iastate.edu/encyclopedia/managing-weeds-protect-crop-yields

89. Prins A. Identifying Troublesome Broadleaf Weeds in Soybeans [Internet]. IL Soy Advisor. 2018 [cited 2020 Jan 19]. Available from: https://www.ilsoyadvisor.com/on-farm/ilsoyadvisor/identifying-troublesome-broadleaf-weeds-soybeans

90. Identifying weeds in field crops [Internet]. Integrated Pest Management. [cited 2022 Nov 21]. Available from: https://www.canr.msu.edu/ipm/agriculture/field_crops/identifying_weeds_in_field_crops

91. Cultural Weed Management [Internet]. NC State Extension. 2013 [cited 2020 Jan 19]. Available from: https://soybeans.ces.ncsu.edu/cultural-weed-management/

92. Sprague C. Herbicide Classification Guide. Soy Checkoff; 2016.

93. Postemergence Herbicide Applications in Soybean [Internet]. [cited 2020 Jan 20]. Available from: https://www.roundupreadyplus.com/resourcecenter/postemergence_herbicide_applications_in_soybean/

94. Hemmati P, Zafari D, Mahmoodi SB, Hashemi M, Gholamhoseini M, Dolatabadian A, et al. Histopathology of charcoal rot disease (Macrophomina phaseolina) in resistant and susceptible cultivars of soybean. Rhizosphere. 2018 Sep 1;7:27–34.

95. Charcoal rot on soybean [Internet]. [cited 2022 Nov 21]. Available from: https://extension.umn.edu/pest-management/charcoal-rot-soybean

96. Giesler LJ, Broderick KC. Management of Phytophthora Root and Stem Rot of Soybean. University of Nebraska-Lincoln Extension. 2017;5.

97. Giesler LJ. Bacterial Diseases of Soybean. University of Nebraska-Lincoln Extension. 2011;2.

98. Malvick D. Soybean rust [Internet]. 2018 [cited 2020 Jan 16]. Available from: https://extension.umn.edu/pest-management/soybean-rust

99. Malvick D. Anthracnose on soybean [Internet]. 2018 [cited 2020 Jan 16]. Available from: https://extension.umn.edu/pest-management/anthracnose-soybean

100. Malvick D. Septoria brown spot [Internet]. 2018 [cited 2020 Jan 16]. Available from: https://extension.umn.edu/pest-management/septoria-brown-spot

101. Giesler LJ. Soybean Cyst Nematode. University of Nebraska-Lincoln Extension [Internet]. 2015 Sep 18 [cited 2020 Jan 18]; Available from: https://cropwatch.unl.edu/plantdisease/soybean/soybean-cyst-nematode

102. Athow KL. Soybean pest management. Journal of the American Oil Chemists' Society. 1981;58(3):130–5.

103. Nagasubramanian K, Jones S, Sarkar S, Singh AK, Singh A, Ganapathysubramanian B. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. Plant Methods. 2018 Dec;14(1):86.

104. Brucker E, Niblack T, Kopisch-Obuch F, Diers B. The Effect of rhg1 on Reproduction of Heterodera glycines in the Field and Greenhouse and Associated Effects on Agronomic Traits. Crop Science. 2005;45:1721–7.

105. Guo W, Zhang F, Bao A, You Q, Li Z, Chen J, et al. The soybean Rhg1 amino acid transporter gene alters glutamate homeostasis and jasmonic acid-induced resistance to soybean cyst nematode. Molecular Plant Pathology. 2019 Feb;20(2):270–86.

106. Niblack TL. Soybean Cyst Nematode Management Reconsidered. Plant Disease. 2005 Oct;89(10):1020–6.

107. Soybean Meal [Internet]. Feedapedia. [cited 2020 Jan 8]. Available from: https://www.feedipedia.org/node/674

108. Verbeck R. Soybean Meal: The One-Stop Solution for Nutrition [Internet]. United Soybean Board. 2020 [cited 2022 Nov 21]. Available from: https://www.unitedsoybean.org/hopper/soybean-meal-the-one-stop-solution-for-nutrition/

109. Buttery P, D'Mello JPF. Amino Acid Metabolism in Farm Animals: An Overview. In: Amino Acids in Farm Animal Nutrition. Wallingford, UK: CAB International; 1994. p. 1–10.

110. Kuiken KA, Lyman CM. Essential amino acid composition of soy bean meals prepared from 20 strains of soy beans. J Biol Chem. 1949 Jan;177(1):29–36.

111. Woclawek-Potocka I, Bah MM, Korzekwa A, Piskula MK, Wiczkowski W, Depta A, et al. Soybean-derived phytoestrogens regulate prostaglandin secretion in endometrium during cattle estrous cycle and early pregnancy. Exp Biol Med (Maywood). 2005 Mar;230(3):189–99.

112. Mcniven M, Charmley E, Mitchell A. Processing of soybean affects meat fatty acid composition and lipid peroxidation in beef cattle. Animal Feed Science and Technology - ANIM FEED SCI TECH. 2004 Oct 1;116:175–84.

113. Lusas EW. Soybean Processing and Utilization. In: Soybeans: Improvement, Production, and Uses [Internet]. John Wiley & Sons, Ltd; 2004 [cited 2022 Nov 21]. p. 949–1045. Available from: https://onlinelibrary.wiley.com/doi/abs/10.2134/agronmonogr16.3ed.c20

114. Liu K. Soybeans: Chemistry, Technology, and Utilization. New York City, New York: Chapman&Hall; 1997.

115. The limiting amino acids in soybean protein. [Internet]. [cited 2022 Nov 21]. Available from: https://www.cabdirect.org/cabdirect/abstract/19631403397

116. Boisen S, Hvelplund T, Weisbjerg M. Ideal Amino Acid Profiles as a Basis for Feed Protein Evaluation. Livestock Production Science. 2000 Jun 1;64:239–51.

117. Fernandez SR, Aoyagi S, Han Y, Parsons CM, Baker DH. Limiting order of amino acids in corn and soybean meal for growth of the chick. Poult Sci. 1994 Dec;73(12):1887–96.

118. Bunchasak C. Role of Dietary Methionine in Poultry Production. The Journal of Poultry Science. 2009;46(3):169–79.

119. Han Y, Suzuki H, Parsons CM, Baker DH. Amino acid fortification of a low-protein corn and soybean meal diet for chicks. Poult Sci. 1992 Jul;71(7):1168–78.

120. Managing Nutrient and Pathogens from Animal Agriculture. Natural Resource, Agriculture, and Enegineering Service. 2000.

121. Control Feed Costs with Amino Acids [Internet]. Drovers. [cited 2019 Aug 5]. Available from: https://www.drovers.com/article/control-feed-costs-amino-acids

122. Vagadia BH, Vanga SK, Raghavan V. Inactivation methods of soybean trypsin inhibitor – A review. Trends in Food Science & Technology. 2017;C(64):115–25.

123. Hardy RW. Utilization of plant proteins in fish diets: effects of global demand and supplies of fishmeal. Aquaculture Research. 2010;41(5):770–6.

124. Chakraborty P, Mallik A, Sarang N, Lingam S. A review on alternative plant protein sources available for future sustainable aqua feed production. 2019 Jan 1;1399–404.

125. Vielma J, Mäkinen T, Ekholm P, Koskela J. Influence of dietary soy and phytase levels on performance and body composition of large rainbow trout (Oncorhynchus mykiss) and algal availability of phosphorus load. Aquaculture. 2000 Mar 15;183(3):349–62.

126. Heikkinen J, Vielma J, Kemiläinen O, Tiirola M, Eskelinen P, Kiuru T, et al. Effects of soybean meal based diet on growth performance, gut histopathology and intestinal microbiota of juvenile rainbow trout (Oncorhynchus mykiss). Aquaculture. 2006 Nov 16;261(1):259–68.

127. Barnes ME, Brown ML, Neiger R. Comparative performance of two rainbow trout strains fed fermented soybean meal. Aquacult Int. 2015 Oct 1;23(5):1227–38.

128. Oliva-Teles A, Gouveia AJ, Gomes E, Rema P. The effect of different processing treatments on soybean meal utilization by rainbow trout, Oncorhynchus mykiss. Aquaculture. 1994 Jul 1;124(1):343–9.

129. Burr GS, Wolters WR, Barrows FT, Hardy RW. Replacing fishmeal with blends of alternative proteins on growth performance of rainbow trout (Oncorhynchus mykiss), and early or late stage juvenile Atlantic salmon (Salmo salar). Aquaculture. 2012 Mar 7;334–337:110–6.

130. Golbitz P. Traditional soyfoods: processing and products. J Nutr. 1995 Mar;125(3 Suppl):570S-572S.

131. Birt D, Hendrich S, Alekel D, Anthony M. Soybean and the Prevention of Chronic Human Disease. In: Soybeans: Improvement, Production, and Uses. 3rd ed. Madison, WI; 2004.

132. Messina MJ. Soyfoods: Their Role in Disease Prevention and Treatment. In: Soybeans [Internet]. Boston, MA: Springer US; 1997 [cited 2022 Nov 21]. p. 442–77. Available from: http://link.springer.com/10.1007/978-1-4615-1763-4_10

133. Messina M. Soy foods, isoflavones, and the health of postmenopausal women. Am J Clin Nutr. 2014 Jul;100 Suppl 1:423S-30S.

134. Messina M. A brief historical overview of the past two decades of soy and isoflavone research. J Nutr. 2010 Jul;140(7):1350S-4S.

135. Xiao Y, Zhang S, Tong H, Shi S. Comprehensive evaluation of the role of soy and isoflavone supplementation in humans and animals over the past two decades. Phytother Res. 2018 Mar;32(3):384–94.

136. Zhong XS, Ge J, Chen SW, Xiong YQ, Ma SJ, Chen Q. Association between Dietary Isoflavones in Soy and Legumes and Endometrial Cancer: A Systematic Review and Meta-Analysis. J Acad Nutr Diet. 2018 Apr;118(4):637–51.

137. Biodiesel [Internet]. United Soybean Board. [cited 2022 Nov 21]. Available from: https://www.unitedsoybean.org/issue-briefs/biodiesel/

138. Cavalett O, Ortega E. Integrated environmental assessment of biodiesel production from soybean in Brazil. Journal of Cleaner Production. 2010 Jan 1;18(1):55–70.

139. Özener O, Yüksek L, Ergenc A, Özkan M. Effects of soybean biodiesel on a DI diesel engine performance, emission and combustion characteristics. Fuel. 2014 Jan 31;115:875–83.

140. Pradhan A. Energy Life-Cycle Assessment of Soybean Biodiesel. :31.

141. Davidson D. Test Weight Matters in 2019 [Internet]. ilsoyadvisor. 2018 [cited 2021 Aug 23]. Available from: https://www.ilsoyadvisor.com/on-farm/ilsoyadvisor/test-weight-matters-2019

142. Staton M. Understanding soybean discount schedules [Internet]. MSU Extension - Soybeans. 2020 [cited 2021 Aug 23]. Available from: https://www.canr.msu.edu/news/understanding_soybean_discount_schedules

143. Sudarić A, Kočar MM, Duvnjak T, Zdunić Z, Kulundžić AM. Improving Seed Quality of Soybean Suitable for Growing in Europe. Soybean for Human Consumption and Animal Feed [Internet]. 2019 Oct 29 [cited 2020 Jan 18]; Available from: https://www.intechopen.com/online-first/improving-seed-quality-of-soybean-suitable-for-growing-in-europe

144. Zarkadas C, Voldeng H, Yu Z, Choi V. Assessment of the Protein Quality of Nine Northern Adapted Yellow and Brown Seed Coated Soybean Cultivars by Amino Acid Analysis. Journal of Agricultural and Food Chemistry. 1999;47(12).

145. Kurasch AK, Hahn V, Leiser WL, Starck N, Würschum T. Phenotypic Analysis of Major Agronomic Traits in 1008 RILs from a Diallel of Early European Soybean Varieties. Crop Science. 2017 04/01;57(2):726–38.

146. Wilcox JR, Shibles RM. Interrelationships among Seed Quality Attributes in Soybean. Crop Science. 2001 Jan 1;41(1):11–4.

147. Vaughn JN, Nelson RL, Song Q, Cregan PB, Li Z. The Genetic Architecture of Seed Composition in Soybean Is Refined by Genome-Wide Association Scans Across Multiple Populations. G3: Genes, Genomes, Genetics. 2014 Nov 1;4(11):2283–94.

148. Gupta PK, Kulwal PL, Mir RR. QTL Mapping: Methodology and Applications in Cereal Breeding. In: Cereal Genomics II [Internet]. New York City, New York: Springer; 2013 [cited 2020 Jan 20]. p. 275–318. Available from: https://link-springer-com.ezproxy.lib.vt.edu/chapter/10.1007/978-94-007-6401-9_11

149. Fehr WR. Breeding for Modified Fatty Acid Composition in Soybean. Crop Sci. 2007 Dec;47:S-72-S-87.

150. Lopez S, Bermudez B, Pacheco YM, Ortega A, Varela LM, Abia R, et al. Chapter 154 - Oleic Acid: The Main Component of Olive Oil on Postprandial Metabolic Processes. In: Preedy VR, Watson RR, editors. Olives and Olive Oil in Health and Disease Prevention [Internet]. San Diego: Academic Press; 2010 [cited 2020 Jan 18]. p. 1385–93. Available from: http://www.sciencedirect.com/science/article/pii/B9780123744203001546

151. Hou A, Chen P, Alloatti J, Li D, Mozzoni L, Zhang B, et al. Genetic Variability of Seed Sugar Content in Worldwide Soybean Germplasm Collections. Crop Science. 2009 May 1;49(3):903–12.

152. Yueqiang Wang, Pengyin Chen, Bo Zhang. Quantitative trait loci analysis of soluble sugar contents in soybean. Plant Breeding. 2014 Jan 1;133(4):493–8.

153. El-Shemy H, Abdel-Rahim E, Shaban O, Ragab A, Carnovale E, Fujita K. Comparison of nutritional and antinutritional factors in soybean and fababean seeds with or without cortex. Soil Science and Plant Nutrition. 2000 Jun 1;46(2):515–24.

154. Refstie S, Sahlström S, Bråthen E, Baeverfjord G, Krogedal P. Lactic acid fermentation eliminates indigestible carbohydrates and antinutritional factors in soybean meal for Atlantic salmon (Salmo salar). Aquaculture. 2005 May 18;246(1):331–45.

155. Barrows FT, Stone DAJ, Hardy RW. The effects of extrusion conditions on the nutritional value of soybean meal for rainbow trout (Oncorhynchus mykiss). Aquaculture. 2007 May 1;265(1):244–52.

156. Becker-Ritt AB, Mulinari F, Vasconcelos IM, Carlini CR. Antinutritional and/or toxic factors in soybean (Glycine max (L) Merril) seeds: comparison of different cultivars adapted to the southern region of Brazil. Journal of the science of food and agriculture [Internet]. 2004 [cited 2022 Nov 21]; Available from: https://scholar.google.com/scholar_lookup?title=Antinutritional+and%2For+toxic+factors+in+soybean+%28Glycine+max+%28L%29+Merril%29+seeds%3A+comparison+of+different+cultivars+adapted+to+the+southern+region+of+Brazil&author=Becker-Ritt%2C+A.B.&publication_year=2004

157. Snowdon M. Full-fat Soybeans for livestock - Agriculture [Internet]. Agriculture, Aquaculture, and Fisheries. 2012 [cited 2020 Jan 18]. Available from: https://www2.gnb.ca/content/gnb/en/departments/10/agriculture/content/crops/field_crops/Soybeans.html

158. Hymowitz T. Genetics and Breeding of Soybeans Lacking the Kunitz Trypsin Inhibitor. In: Friedman M, editor. Nutritional and Toxicological Significance of Enzyme Inhibitors in Foods [Internet]. Boston, MA: Springer US; 1986 [cited 2020 Jan 18]. p. 291–8. (Advances in Experimental Medicine and Biology). Available from: https://doi.org/10.1007/978-1-4757-0022-0_18

159. Lei XG, Porres JM. Phytase: An Enzyme to Improve Soybean Nutrition. Soybean and Nutrition [Internet]. 2011 Sep 12 [cited 2020 Jan 18]; Available from: https://www.intechopen.com/books/soybean-and-nutrition/phytase-an-enzyme-to-improve-soybean-nutrition

160. Augsutyn A. genetic engineering | Definition, Process, & Uses [Internet]. Encyclopedia Britannica. 2019 [cited 2020 Jan 19]. Available from: https://www.britannica.com/science/genetic-engineering

161. National Academies of Science. Methods and Mechanisms for Genetic Manipulation of Plants, Animals, and Microorganisms. In: Safety of Genetically Engineered Foods: Approachers to Assessing Unintended Health Effects [Internet]. National Academies Press; 2004 [cited 2020 Jan 19]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK215771/

162. What are genome editing and CRISPR-Cas9?: MedlinePlus Genetics [Internet]. [cited 2022 Nov 21]. Available from: https://medlineplus.gov/genetics/understanding/genomicresearch/genomeediting/

163. Arora L, Narula A. Gene Editing and Crop Improvement Using CRISPR-Cas9 System [Internet]. frontiersin. 2017 [cited 2020 Jan 19]. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2017.01932/full

164. Wang T, Zhang H, Zhu H. CRISPR technology is revolutionizing the improvement of tomato and other fruit crops | Horticulture Research. Horticulture Research [Internet]. 2019 [cited 2020 Jan 19]; Available from: https://www.nature.com/articles/s41438-019-0159-x

165. Biotech Crop Highlights in 2018 | ISAAA.org [Internet]. [cited 2020 Jan 19]. Available from: http://www.isaaa.org/resources/publications/pocketk/16/

166. Roundup Ready Soybeans - SourceWatch [Internet]. [cited 2020 Jan 19]. Available from: https://www.sourcewatch.org/index.php/Roundup_Ready_Soybeans

167. Falco SC, Locke M, Guida T, Sanders C, Ward RT, Webber P. Transgenic Canola and Soybean Seeds with Increased Lysine. Nature Biotechnology. 1995;(13):577–82.

168. Clemente T, Cahoon E. Soybean Oil: Genetic Approaches for Modification of Functionality and Total Content. Plant Physiology [Internet]. 2009 [cited 2020 Jan 19]; Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2773065/

169. Cai Y, Chen L, Guo C, Sun S, Wu C, Jiang B, et al. CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean - Cai - 2018 - Plant Biotechnology Journal - Wiley Online Library. Plant Biotechnology Journal. 2017;16(1):176–85.

170. Pagano MC, Miransari M. 1 - The importance of soybean production worldwide. In: Miransari M, editor. Abiotic and Biotic Stresses in Soybean Production [Internet]. San Diego: Academic Press; 2016 [cited 2021 Aug 23]. p. 1–26. Available from: https://www.sciencedirect.com/science/article/pii/B9780128015360000013

171. Shahbandeh. Global major soybean exporting countries share 2020 [Internet]. Statista. 2021 [cited 2021 Aug 23]. Available from: https://www.statista.com/statistics/961087/global-leading-exporters-of-soybeans-export-share/

172. Vázquez-Carrillo MG, Rojas-Martínez I, Santiago-Ramos D, Arellano-Vázquez JL, Espinosa-Calderón A, García-Pérez M, et al. Stability Analysis of Yield and Grain Quality Traits for the Nixtamalization Process of Maize Genotypes Cultivated in the Central High Valleys of Mexico. Crop Science. 2016;56(6):3090–9.

173. Jing Q, Jiang D, Dai T, Cao W. [Effects of genotype and environment on wheat grain quality and protein components]. Ying Yong Sheng Tai Xue Bao. 2003 Oct;14(10):1649–53.

174. Doehlert DC, McMullen MS, Hammond JJ. Genotypic and Environmental Effects on Grain Yield and Quality of Oat Grown in North Dakota. Crop Science. 2001;41(4):1066–72.

175. Liu L, Prenger E, Zhang J, Little B, Mian MAR, Li Z. Impact of genotype, seed composition, agronomic trait, and environment on soybean test weight. Journal of Crop Improvement. 2019 Nov 2;33(6):711–29.

176. Schuler SF, Bacon RK, Finney PL, Gbur EE. Relationship of Test Weight and Kernel Properties to Milling and Baking Quality in Soft Red Winter Wheat. Crop Science. 1995;35(4):cropsci1995.0011183X003500040001x.

177. Gürsoy S, Sessİz A, Malhi SS. Short-term effects of tillage and residue management following cotton on grain yield and quality of wheat. Field Crops Research. 2010;119(2/3):260–8.

178. Bonfil DJ, Abbo S, Svoray T. Sowing Date and Wheat Quality as Determined by Gluten Index. Crop Science. 2015;55(5):2294–306.

179. Rharrabti Y, Villegas D, Moral LFG del, Aparicio N, Elhani S, Royo C. Environmental and genetic determination of protein content and grain yield in durum wheat under Mediterranean conditions. Plant Breeding. 2001;120(5):381–8.

180. Miji A, Marjanovi A. QUANTITATIVE ANALYSIS OF OIL YIELD AND ITS COMPONENTS IN SUNFLOWER (HELIANTHUS ANNUUS L.). ROMANIAN AGRICULTURAL RESEARCH. 2009;(26):6.

181. Holland JB, Frey KJ, Hammond EG. Correlated responses of fatty acid composition, grain quality, and agronomic traits to nine cycles of recurrent selection for increased oil content in oat. 2000;12.

182. Singer WM, Shea Z, Yu D, Huang H, Mian MAR, Shang C, et al. Genome-Wide Association Study and Genomic Selection for Proteinogenic Methionine in Soybean Seeds. Front Plant Sci. 2022;13:859109.

183. Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, et al. Development and Evaluation of SoySNP50K, a High-Density Genotyping Array for Soybean. PLOS ONE. 2013 Jan 25;8(1):e54985.

184. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007 Oct 1;23(19):2633–5.

185. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 2006 Feb;38(2):203–8.

186. Li J, Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity. 2005 Sep;95(3):221–7.

187. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. Journal of Open Source Software. 2014;(3):1–731.

188. Grant D, Nelson RT, Cannon SB, Shoemaker RC. SoyBase, the USDA-ARS soybean genetics and genomics database. 2010.

189. Lord N, Shang C, Rosso L, Zhang B. Development of near-infrared reflectance spectroscopy calibration for sugar content in ground soybean seed using Perten DA7250 analyzer. Crop Science. 2021;61(2):966–75.

190. Cloninger FD, Horrocks RD, Zuber MS. Effects of Harvest Date, Plant Density, and Hybrid on Corn Grain Quality1. Agronomy Journal. 1975;67(5):693–5.

191. Lloyd BJ, Siebenmorgen TJ, Bacon RK, Vories E. HARVEST DATE AND CONDITIONED MOISTURE CONTENT EFFECTS ON TEST WEIGHT OF SOFT RED WINTER WHEAT. American Society of Agricultural and BIological Engineers. 1999;15(5):525–34.

192. Cromwell GL, Stahly TS, Monegue HJ. Amino acid supplementation of meat meal in lysine-fortified, corn-based diets for growing-finishing pigs. Journal of Animal Science. 1991 Dec 1;69(12):4898–906.

193. Gillman JD, Kim WS, Krishnan HB. Identification of a new soybean kunitz trypsin inhibitor mutation and its effect on bowman-birk protease inhibitor content in soybean seed. J Agric Food Chem. 2015 Feb 11;63(5):1352–9.

194. Liener IE. Effects of processing on antinutritional factors in legumes: the soybean case. Arch Latinoam Nutr. 1996 Dec;44(4 Suppl 1):48S-54S.

195. Cook DA, Jensen AH, Fraley JR, Hymowitz T. Utilization by Growing and Finishing Pigs of Raw Soybeans of Low Kunitz Trypsin Inhibitor Content. Journal of Animal Science. 1988 Jul 1;66(7):1686–91.

196. Liener I. Implications of antinutritional components in soybean foods. Critical Reviews in Food Science and Nutrition. 1994;34(1).

197. Chen Y, Xu Z, Zhang C, Kong X, Hua Y. Heat-induced inactivation mechanisms of Kunitz trypsin inhibitor and Bowman-Birk inhibitor in soymilk processing. Food Chem. 2014 Jul 1;154:108–16.

198. Chang CJ, Tanksley TD, Knabe DA, Zebrowska T. Effects of different heat treatments during processing on nutrient digestibility of soybean meal in growing swine. J Anim Sci. 1987 Nov;65(5):1273–82.

199. Batal AB, Parsons CM. Utilization of Various Carbohydrate Sources as Affected by Age in the Chick. Poultry Science. 2004 Jul 1;83(7):1140–7.

200. Jofuku KD, Goldberg RB. Kunitz trypsin inhibitor genes are differentially expressed during the soybean life cycle and in transformed tobacco plants. Plant Cell. 1989 Nov;1(11):1079–93.

201. Schuler TH, Poppy GM, Kerry BR, Denholm I. Potential side effects of insect-resistant transgenic plants on arthropod natural enemies. Trends Biotechnol. 1999 May;17(5):210–6.

202. Kunitz M. Crystallization of a Trypsin Inhibitor from Soybean. Science. 1945 Jun 29;101(2635):668–9.

203. Wei CH. Crystallization of two cubic forms of soybean trypsin inhibitor E-I, a member of the Bowman-Birk inhibitor family. Journal of Biological Chemistry. 1983 Aug;258(15):9357–9.

204. Hwang DL, Foard DE, Wei CH. A soybean trypsin inhibitor. Crystallization and x-ray crystallographic study. Journal of Biological Chemistry. 1977 Feb 10;252(3):1099–101.

205. Jofuku KD, Schipper RD, Goldberg RB. A frameshift mutation prevents Kunitz trypsin inhibitor mRNA accumulation in soybean embryos. Plant Cell. 1989 Apr;1(4):427–35.

206. Perez-Maldonado RA, Mannion PF, Farrell DJ. Effects of heat treatment on the nutritional value of raw soybean selected for low trypsin inhibitor activity. Br Poult Sci. 2003 May;44(2):299–308.

207. Rosso ML, Shang C, Song Q, Escamilla D, Gillenwater J, Zhang B. Development of Breeder-Friendly KASP Markers for Low Concentration of Kunitz Trypsin Inhibitor in Soybean Seeds. Int J Mol Sci. 2021 Mar 6;22(5):2675.

208. Liang G, Zhang H, Lou D, Yu D. Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. Sci Rep. 2016 Feb 19;6(1):21451.

209. Kim J, Kim JS. Bypassing GMO regulations with CRISPR gene editing. Nat Biotechnol. 2016 Oct 11;34(10):1014–5.

210. Haun W, Coffman A, Clasen BM, Demorest ZL, Lowy A, Ray E, et al. Improved soybean oil quality by targeted mutagenesis of the fatty acid desaturase 2 gene family. Plant Biotechnol J. 2014 Sep;12(7):934–40.

211. Jacobs TB, LaFayette PR, Schmitz RJ, Parrott WA. Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnology. 2015 Mar 12;15(1):16.

212. Cai Y, Chen L, Liu X, Guo C, Sun S, Wu C, et al. CRISPR/Cas9-mediated targeted mutagenesis of GmFT2a delays flowering time in soya bean. Plant Biotechnol J. 2018 Jan;16(1):176–85.

213. Xu RF, Li H, Qin RY, Li J, Qiu CH, Yang YC, et al. Generation of inheritable and “transgene clean” targeted genome-modified rice in later generations using the CRISPR/Cas9 system. Sci Rep. 2015 Jun 19;5:11491.

214. Upadhyay SK, Kumar J, Alok A, Tuli R. RNA-guided genome editing for target gene mutations in wheat. G3 (Bethesda). 2013 Dec 9;3(12):2233–8.

215. Chen R, Xu Q, Liu Y, Zhang J, Ren D, Wang G, et al. Generation of Transgene-Free Maize Male Sterile Lines Using the CRISPR/Cas9 System. Frontiers in Plant Science [Internet]. 2018 [cited 2022 Nov 22];9. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2018.01180

216. Vu TV, Sivankalyani V, Kim EJ, Doan DTH, Tran MT, Kim J, et al. Highly efficient homology-directed repair using CRISPR/Cpf1-geminiviral replicon in tomato. Plant Biotechnol J. 2020 Mar 16;18(10):2133–43.

217. Gao W, Long L, Tian X, Xu F, Liu J, Singh PK, et al. Genome Editing in Cotton with the CRISPR/Cas9 System. Frontiers in Plant Science [Internet]. 2017 [cited 2022 Nov 22];8. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2017.01364

218. Peng A, Chen S, Lei T, Xu L, He Y, Wu L, et al. Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene CsLOB1 promoter in citrus. Plant Biotechnol J. 2017 Dec;15(12):1509–19.

219. Osakabe Y, Liang Z, Ren C, Nishitani C, Osakabe K, Wada M, et al. CRISPR-Cas9-mediated genome editing in apple and grapevine. Nat Protoc. 2018 Dec;13(12):2844–63.

220. Nakayasu M, Akiyama R, Lee HJ, Osakabe K, Osakabe Y, Watanabe B, et al. Generation of α-solanine-free hairy roots of potato by CRISPR/Cas9 mediated genome editing of the St16DOX gene. Plant Physiol Biochem. 2018 Oct;131:70–7.

221. Shao X, Wu S, Dou T, Zhu H, Hu C, Huo H, et al. Using CRISPR/Cas9 genome editing system to create MaGA20ox2 gene-modified semi-dwarf banana. Plant Biotechnol J. 2020 Jan;18(1):17–9.

222. Paz MM, Martinez JC, Kalvig AB, Fonger TM, Wang K. Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-mediated soybean transformation. Plant Cell Rep. 2006 Mar;25(3):206–13.

223. Luth D, Warnberg K, Wang K. Soybean [Glycine max (L.) Merr]. Methods Mol Biol. 2015;1223:275–84.

224. Zhao B, Dahlbeck D, Krasileva KV, Fong RW, Staskawicz BJ. Computational and Biochemical Analysis of the Xanthomonas Effector AvrBs2 and Its Role in the Modulation of Xanthomonas Type Three Effector Delivery. PLOS Pathogens. 2011 Dec 1;7(12):e1002408.

225. Liu Y, Miao J, Traore S, Kong D, Liu Y, Zhang X, et al. SacB-SacR Gene Cassette As the Negative Selection Marker to Suppress Agrobacterium Overgrowth in Agrobacterium-Mediated Plant Transformation. Frontiers in Molecular Biosciences [Internet]. 2016 [cited 2022 Nov 22];3. Available from: https://www.frontiersin.org/articles/10.3389/fmolb.2016.00070

226. Jian B, Liu B, Bi Y, Hou W, Wu C, Han T. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Molecular Biology [Internet]. 2008 Jun [cited 2021 Sep 7]; Available from: https://bmcmolbiol.biomedcentral.com/articles/10.1186/1471-2199-9-59

227. Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, et al. Gateway-compatible vectors for plant functional genomics and proteomics. Plant J. 2006 Feb;45(4):616–29.

228. Han Q, Zhou C, Wu S, Liu Y, Triplett L, Miao J, et al. Crystal Structure of Xanthomonas AvrRxo1-ORF1, a Type III Effector with a Polynucleotide Kinase Domain, and Its Interactor AvrRxo1-ORF2. Structure. 2015 Oct 6;23(10):1900–9.

229. Rosso ML, Shang C, Correa E, Zhang B. An Efficient HPLC Approach to Quantify Kunitz Trypsin Inhibitor in Soybean Seeds. Crop Science. 2018;58(4):1616–23.

230. Li J, Brader G, Palva ET. Kunitz trypsin inhibitor: an antagonist of cell death triggered by phytopathogens and fumonisin b1 in Arabidopsis. Mol Plant. 2008 May;1(3):482–95.

231. Junker Y, Zeissig S, Kim SJ, Barisani D, Wieser H, Leffler DA, et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. J Exp Med. 2012 Dec 17;209(13):2395–408.

232. Arnaiz A, Talavera-Mateo L, Gonzalez-Melendi P, Martinez M, Diaz I, Santamaria ME. Arabidopsis Kunitz Trypsin Inhibitors in Defense Against Spider Mites. Front Plant Sci. 2018;9:986.

233. Zhao H, Ullah H, McNeill MR, Du G, Hao K, Tu X, et al. Inhibitory Effects of Plant Trypsin Inhibitors Msti-94 and Msti-16 on Therioaphis trifolii (Monell) (Homoptera: Aphididae) in Alfalfa. Insects. 2019 May 30;10(6):154.

234. Hussain RMF, Sheikh AH, Haider I, Quareshy M, Linthorst HJM. Arabidopsis WRKY50 and TGA Transcription Factors Synergistically Activate Expression of PR1. Frontiers in Plant Science [Internet]. 2018 [cited 2022 Nov 22];9. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2018.00930

235. Yamada T, Takagi K, Ishimoto M. Recent advances in soybean transformation and their application to molecular breeding and genomic analysis. Breed Sci. 2012 Jan;61(5):480–94.

236. Testroet A, Lee K, Luth D, Wang K. Comparison of transformation frequency using the bar gene regulated by the CaMV 35S or NOS promoter in Agrobacterium-mediated soybean (Glycine max L.) transformation. In Vitro Cellular & Developmental Biology - Plant. 2017 Apr 3;53.

237. Langridge WH, Fitzgerald KJ, Koncz C, Schell J, Szalay AA. Dual promoter of Agrobacterium tumefaciens mannopine synthase genes is regulated by plant growth hormones. Proc Natl Acad Sci U S A. 1989 May;86(9):3219–23.

238. Pérez-González A, Caro E. Benefits of using genomic insulators flanking transgenes to increase expression and avoid positional effects. Sci Rep. 2019 Jun 11;9(1):8474.

239. Hasan N, Choudhary S, Naaz N, Sharma N, Laskar RA. Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. J Genet Eng Biotechnol. 2021 Aug 27;19(1):128.

240. Hirotsu N, Murakami N, Kashiwagi T, Ujiie K, Ishimaru K. Protocol: a simple gel-free method for SNP genotyping using allele-specific primers in rice and other plant species. Plant Methods. 2010 Apr 21;6(1):12.

241. Nations U. World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100 [Internet]. United Nations. United Nations; [cited 2022 Apr 14]. Available from:

https://www.un.org/en/desa/world-population-projected-reach-98-billion-2050-and-112-billion-2100

242. The state of food security and nutrition in the world 2020 | UNICEF [Internet]. [cited 2022 Apr 6]. Available from: https://www.unicef.org/reports/state-of-food-security-and-nutrition-2020

243. Kawarazuka N, Béné C. Linking small-scale fisheries and aquaculture to household nutritional security: an overview. Food Sec. 2010 Dec 1;2(4):343–57.

244. Beddington JR, Kirkwood GP, Caddy JF, Seijo JC. This is more difficult than we thought! The responsibility of scientists, managers and stakeholders to mitigate the unsustainability of marine fisheries. Philosophical Transactions of the Royal Society B: Biological Sciences. 2005 Jan 29;360(1453):59–75.

245. Technavio report: Global aquaculture market's growth accelerating through 2022 | SeafoodSource [Internet]. [cited 2022 Apr 14]. Available from: https://www.seafoodsource.com/features/technavio-report-global-aquaculture-markets-growth-accelerating-through-2022

246. Thilsted SH, Thorne-Lyman A, Webb P, Bogard JR, Subasinghe R, Phillips MJ, et al. Sustaining healthy diets: The role of capture fisheries and aquaculture for improving nutrition in the post-2015 era. Food Policy. 2016 May 1;61:126–31.

247. Hardy R. 12 - Aquaculture feeds and ingredients: an overview. In: Burnell G, Allan G, editors. New Technologies in Aquaculture [Internet]. Woodhead Publishing; 2009 [cited 2022 Oct 31]. p. 370–86. (Woodhead Publishing Series in Food Science, Technology and Nutrition). Available from: https://www.sciencedirect.com/science/article/pii/B9781845693848500129

248. Alexandratos N, Bruinsma J. World Agriculture towards 2030/2050: the 2012 revision. 2012;154.

249. Soybean Meal Price (Any Origin) [Internet]. [cited 2022 Oct 31]. Available from: https://ycharts.com/indicators/soybean_meal_price_any_origin

250. Gatlin III DM, Barrows FT, Brown P, Dabrowski K, Gaylord TG, Hardy RW, et al. Expanding the utilization of sustainable plant products in aquafeeds: a review. Aquaculture Research. 2007;38(6):551–79.

251. Francis G, Makkar HPS, Becker K. Antinutritional factors present in plant-derived alternate fish feed ingredients and their effects in fish. Aquaculture. 2001 Aug 1;199(3):197–227.

252. Hajra A, Mazumder A, Ganguly D, Mohanty B, Sharma A. ANTINUTRITIONAL FACTORS IN PLANT ORIGIN FISH FEED INGREDIENTS: THE PROBLEMS AND PROBABLE REMEDIES. In 2013. p. 193–202.

253. Becker-Ritt AB, Mulinari F, Vasconcelos IM, Carlini CR. Antinutritional and/or toxic factors in soybean (Glycine max (L) Merril) seeds: comparison of different cultivars adapted to the southern region of Brazil. Journal of the Science of Food and Agriculture. 2004;84(3):263–70.

254. Glencross B d., Booth M, Allan G l. A feed is only as good as its ingredients – a review of ingredient evaluation strategies for aquaculture feeds. Aquaculture Nutrition. 2007;13(1):17–34.

255. Hardy RW. Worldwide Fish Meal Production Outlook and the Use of Alternative Protein Meals For Aquaculture. Avances en Nutricion Acuicola. 2006;

256. Refstie S, Korsøen ØJ, Storebakken T, Baeverfjord G, Lein I, Roem AJ. Differing nutritional responses to dietary soybean meal in rainbow trout (Oncorhynchus mykiss) and Atlantic salmon (Salmo salar). Aquaculture. 2000 Oct 25;190(1):49–63.

257. Using Soy to Feed Fish [Internet]. U.S. Soy. 2021 [cited 2021 Aug 30]. Available from: https://ussoy.org/about-aquaculture/using-soy-to-feed-fish/

258. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 2010 May;7(5):335–6.

259. Adelizi, Rosati, Warner, Wu, Muench, White, et al. Evaluation of fish-meal free diets for rainbow trout, Oncorhynchus mykiss. Aquaculture Nutrition. 1998;4(4):255–62.

260. Davies SJ, Morris PC. Influence of multiple amino acid supplementation on the performance of rainbow trout, Oncorhynchus mykiss (Walbaum), fed soya based diets. Aquaculture Research. 1997;28(1):65–74.

261. Cheng ZJ, Hardy RW, Usry JL. Effects of lysine supplementation in plant protein-based diets on the performance of rainbow trout (Oncorhynchus mykiss) and apparent digestibility coefficients of nutrients. Aquaculture. 2003 Jan 10;215(1):255–65.

Chairman: Dr. Bo Zhang VT Id #

Committee Member: Dr. Bingyu Zhao VT Id #

Committee Member: Dr. Michael Schwarz VT Id #

Committee Member: Dr. Mahmoudreza Ovissipour VT Id #

Committee Member: Dr. Saghai Maroof VT Id #