SOYBEAN SEED TRAIT IMPROVEMENTS AND THEIR EFFECTS ON HUMAN AND ANIMAL NUTRITION

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

MICHELLE FOLTA

Dr. Zhanyuan Zhang, Dissertation Supervisor

JULY 2016

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

SOYBEAN SEED TRAIT IMPROVEMENTS AND THEIR EFFECTS ON HUMAN AND ANIMAL NUTRITION

presented by Michelle Folta,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

_____

Dr. Zhanyuan Zhang

_____

Dr. Kristin Bilyeu

_____

Dr. James Schoelz

_____

Dr. Monty Kerley

Dedicated to my grandmothers,
one who is the wisest woman I know,
and the other who instilled a love of all things green.

ACKNOWLEDGEMENTS

I owe the most thanks to Dr. Zhang and all of the lab members of the Plant Transformation Core Facility. Dr. Zhang, thank you for being a great mentor; knowing when to push me a little further and when to step back and let me work out problems myself. I'm a much better scientist than I was five years ago, and working under your leadership is a big reason for my growth.

To all the transformation wizards, especially Liwen, Murug, Hyeyoung, Hua, and Hien, thank you for always being available to give the best advice. My project would not have been successful had I not been surrounded by the best plant transformation team in the world. Soyon, Kaixuan, Joann, Hanbing, and Phat, thank you for letting me bounce ideas off you daily, and always being willing to answer my millions of questions. I came into the lab with zero molecular biology experience, and because of all of you I feel much more confident in my skills. Neng, thank you for all of your help, creative solutions, and always having an encouraging word.

Special thank you to Dr. Kristin Bilyeu and her lab members Christi, Carrie, Hyun and Paul for helping with the soymilk project, HPLC, and again, answering my questions. Somehow, we got 200 college students to try unsweetened, unflavored soymilk, and I think that's a pretty big accomplishment in and of itself.

Dr. Ingolf Gruen, thank you for helping with the soymilk project – everything from experimental design, IRB approval, and data analysis to lending me your lab for the soymilk preparation. Your enthusiasm for science is contagious and I can't thank you enough for being so accommodating.

To Dr. Firman and all the animal science graduate students I've had the pleasure of working with the past few months: huge thank you! Thank you Dr. Firman for agreeing to help a plant science graduate student do an animal science study, you definitely didn't have to do anything, but you've been more than helpful. Gui, thank you for guiding me through pretty much the whole project and answering all of my daily questions. Adam, Corey, and Morgan, thank you for keeping the chickens alive and teaching me how to take care of them. It's a life skill, I'm sure. Also thank you to everyone who helped with the cecectomy surgeries, helped me find supplies, taught me how to use the instruments, and generally made the whole process fun.

Thank you to my parents for always being encouraging, even if you have no idea what I'm doing. Thank you for being a friendly voice on the phone and keeping me optimistic. More than most, I couldn't do this without you. Thanks Andy, for always keeping me laughing. Jeremy, thank you for being so supportive. Thank you for driving me to the lab in the middle of the night, humoring my need to show you all my plant "friends", and always being patient.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Literature Review

Origin and history of soybean cultivation

Modern cultivated soybean, *Glycine max*, is estimated to have been domesticated from its wild progenitor *Glycine soja* in China three to five thousand years ago, although the location of original domestication is yet unclear (Carter et al, 2004; Guo et al, 2010). *Glycine soja* grows wild in northeast Asia, and is an annual, black-seeded legume with a vining, creeping growth habit that exhibits a high degree of pod-shattering (Liu et al, 2007). *Glycine max* and *Glycine soja* are both diploids with 2n=40, and the two species are compatible for crossing and exhibit normal chromosome pairing during meiosis (Stupar, 2010). It has been shown that up to 50% of genetic diversity was lost during the domestication process (Hyten et al, 2006), and modern breeding efforts are underway to breed favorable alleles from *Glycine soja* lines back into the cultivated variety population. Further, of the 45,000 *Glycine max* Asian landraces identified, only 17 landraces contribute about 86% of the North American breeding population parentage, causing further erosion of genetic diversity (Carter et al, 2004). Successful breeding efforts to incorporate *Glycine soja* and *Glycine max* landraces into the North American breeding population have increased the values of traits such as aluminum tolerance, yield, soybean cyst nematode tolerance, drought tolerance, and other important traits (Bianchi-Hall et al, 2000; Smalley et al, 2004; Guo et al, 2005; Manavalan et al, 2009).

Glycine species are considered paleopolyploids, as they have experienced at least one ancient whole genome duplication around 13 million years ago, in addition to the early legume family whole genome duplication an estimated 59 million years ago (Schmutz et al, 2010). Some *Glycine* subgenera have also had novel genome duplications in the last

50,000 years (Shoemaker et al, 2006). RFLP-based studies have shown that an average of 2.55 and as many as six copies of each duplicated segment are present in the soybean genome, further evidence confirming multiple genome duplications during its evolution. Gene duplication can give rise to complex gene families which may accumulate mutations causing gene gain-of-function, loss-of-function, new function, and alteration of expression patterns of individual genes (Doyle et al, 2008). These complex networks of gene families can complicate genetic studies, as knockouts of several related gene family members may be necessary to observe a phenotype.

Domesticated soybean spread from northern China to Korea and southern China by the 1st century AD (Hymowitz, 1990), and then to other southeast Asian areas from the 1st to 16th centuries (Singh and Hymowitz, 1999). Landraces are available for most of these secondary centers of origin, contributing added available genetic diversity. Soybeans have been introduced to the Americas several times beginning in the 1700's, and sources cite Samuel Brown, Benjamin Franklin, and Professor Thomas Nuttall with bringing soybeans to the United States from the late 1700s to the mid-1800s (Singh and Hymowitz, 1999). Soybean cultivation in the United States began to prosper after the study conducted by Osborn and Mendel in 1917 showed that rats fed heated soybean meal gain more weight than those fed unheated soy, and positioned soybean as an important crop for animal feed. This study laid the foundation for one of the first commercial soybean processing plants in the United States in Decatur, Illinois in 1922, built by The A.E. Staley Manufacturing Company (Shurtleff and Aoyagi, 2008).

Today, soybean is the number one protein and oil crop in the world, with the United States producing about one third of the world's soy crop, followed by Brazil, Argentina, China, and Paraguay (Soystats, 2015 data). Soybean meal is used primarily for animal feed, with poultry and swine being the major consumers (Cromwell, 2012). Soybean oil is consumed mostly by humans, and is used in frying applications, salad dressing, baking shortening, and margarine. Other whole-bean applications of soy are the traditional Asian foods soymilk, tofu, tempeh, miso, natto, soy sauce and edamame. In western countries, modified versions of the traditional Asian soyfoods are becoming popular, along with soy-based meat and cheese alternatives, roasted soynuts, soy yoghurt, and other non-dairy frozen desserts.

Human soy foods

Soymilk

Soymilk is essentially a water extract of soybean, and the basic preparation involves first soaking the beans in an excess of water, wet grinding, filtering of solids and then heating to sterilize the product and reduce anti-nutritional factors. Traditional Chinese soymilk is cooked after the filtering step, while traditional Japanese soymilk is cooked before filtering (Shurtleff and Aoyagi, 2008). Asian soymilks can have a strong beany, astringent flavor and mouthfeel, and are often consumed warm for the breakfast meal (Liu, 2008). Modern soymilk sold in most western markets is prepared similarly to Asian milks, but with additional steps to reduce the beany flavors, add flavoring and sweetness, fortify with vitamins and minerals, and increase creaminess (Imram, 2003). North Americans primarily consume soymilk chilled, and use it as a dairy milk replacement.

Tofu

Soymilk has been produced in China for centuries, but was primarily produced as an intermediate of the more popularly consumed tofu. Tofu is a coagulation of soy protein, produced by the addition of salts or acids to soymilk. Common coagulants include calcium sulfate, magnesium chloride, and edible acids such as vinegar and citric acid (Shurtleff and Aoyagi, 2000; Guo and Ono, 2006). Tofu was first mentioned in Chinese literature in 965 CE, and mentioned in Japanese literature in 1183 CE. Today, tofu is produced on a commercial scale in Asian and western countries, but a large degree of small-scale, village-level production is still common.

Fermented traditional Asian soy foods

Fermented soy products such as natto, fermented tofu, fermented soymilk, and tempeh are common in Asian countries. Fermentation of soybean increases available carbohydrates as microbial activity digests raffinose family oligosaccharides, and also provides healthful probiotic activity and improved the flavor. Several strains of fermentation bacteria and fungi are common, including *R. oligosporus* for tempeh, *L. acidophilus* for soymilk, *Mucor hiemalis* for fermenting tofu, and *B. subtilis var. natto* for natto (Hosking, 1996; Han et al, 2004; Chang et al, 2009; Y.M. Chen et al, 2013).

Ethnicity-related preference of soy foods

Soyfoods are native to Asia, and their consumption in western markets has only existed in the last hundred or so years. The first mention of soymilk in the United States was in 1896 *American Journal of Pharmacy* article (Shurtleff and Aoyagi, 2013). The first US soymilk patent was issued to Yu Ying-Li in 1913, and the first commercial soymilk was produced in 1917 in New York (Shurtleff and Aoyagi, 2013). Because soymilk is a relatively new food to the United States, its acceptability has taken some time to become popular. Much of this can be explained by the term "food neophobia", which is the reluctance or avoidance of eating novel foods (Pliner and Hobden, 1992). Because traditional Asian soyfoods have a stronger beany flavor and astringent mouthfeel, Asian consumers prefer those flavors in soyfoods while western consumers tend to dislike those attributes. A study comparing the preferences of Vietnamese and French consumers of soy yoghurt found that Vietnamese who consume traditional soymilk regularly prefer soy yoghurt with bolder soy flavor than their French counterparts (Tu et al, 2010).

Soybean composition

Soybean is a nutritionally balanced plant source of protein, fat, and carbohydrate. Soybeans are crushed to create full-fat flakes, which can then be further processed via solvent extraction to produce soybean oil and defatted flakes. The oil is typically used for human consumption, and the defatted flakes are ground into soybean meal which is primarily used for animal feed. Soybean meal can also be further processed into soy protein isolate or textured vegetable protein for use in specialty foods (Medic et al, 2014).

Roughly 90% of the weight of soybean is the embryo, consisting of two cotyledons, the radicle, hypocotyl, and epicotyl. For normal soybean varieties, about 40% of the dry weight is protein, 20% oil, 35% carbohydrate, and 5% ash (Rotundo and Westgate, 2009). Oil is stored in small oil bodies of the cotyledon, and protein is stored in the larger protein bodies. The remaining 10% of the seed is comprised of the seed coat, which is predominately made of carbohydrate (Kawamara, 1967). The majority of seed protein are classified as storage proteins, predominantly glycinin and β-conglycinin, and their primary role is to serve as a nitrogen source for the germinating embryo (Murphy, 2008). Other less-abundant proteins include those needed for metabolic function, protease inhibitors, lectins, and lipoxygenase enzymes. There are at least 15 known soybean allergen proteins, including the common allergens P34, Gly-m-IA and Gly-m-IB, and the α-subunit of β-conglycinin (Gonzales et al, 1992; Ogawa et al, 1993; Herman, 2005). High-protein lines of soybean have been developed, but they often suffer a yield drag compared to control lines, indicating the negative relationship between protein content and yield (Chung et al, 2003). Soybean protein quality is limited by the relatively low content of sulfur-containing amino acids in the major storage proteins. Thus, animal diets using soybean as a protein source are supplemented with synthetic methionine to provide a balanced profile of amino acids (Krishnan, 2005).

Soybean oil is comprised of linoleic (55%), oleic (21%), palmitic (12%), linolenic (9%), and stearic (4%) fatty acids. High and low varieties of each fatty acid have been developed, but nutritionally, the unsaturated fatty acids oleic, linoleic, and linolenic are considered the most healthful for human consumption. However, the poly-unsaturated fatty acids linoleic

and linolenic are unstable at high temperatures and can be oxidized by lipoxygenase enzymes, which causes rancid off-flavors to develop. To combat this, food manufacturers hydrogenate the fatty acids, but this produces saturated fats which are implicated in a number of human health concerns (Kris-Etherton et al, 1999).

The final major soybean component is carbohydrate, of which about half are structural carbohydrates such as cellulose, hemicellulose, and pectin, and the other half are nonstructural carbohydrates such as sucrose and oligosaccharides. Structural carbohydrates are generally non-digestible to non-ruminant animals, and are used as a source of fiber in animal and human diets. Nonstructural carbohydrates, of which the most abundant are sucrose, raffinose, and stachyose, are soluble carbohydrates, and only sucrose has a positive metabolizable energy value in monogastric diets. Raffinose and stachyose are considered anti-nutritional factors, as their presence in animal diets can cause flatulence, digestive disturbance, and low weight gain (Coon et al, 1990). Breeding efforts to improve soybean carbohydrate content have been successful, and one particular mutation in the soybean *raffinose synthase 2* gene can increase sucrose content from about 50% to as much as 85%, and decrease raffinose plus stachyose content from about 47% to less than 6% of the total carbohydrate fraction (Hagely et al, 2013).

Soybean composition is highly affected by environmental factors, especially temperature and water stresses. High temperature during podfill is associated with decreased sucrose and increased oligosaccharide content (Kumar et al, 2010; Bilyeu and Wiebold, 2016). Increasing daily temperatures are also negatively correlated with oil content and positively

correlated with protein (Dornbos and Mullen, 1992; Gibson and Mullen, 1996). Further, high temperatures are associated with reduced contents of polyunsaturated fats such as linoleic and linolenic acid, and increased contents of oleic, palmitic, and stearic acid (Dornbos and Mullen, 1992; Ren et al, 2009).

Raffinose family oligosaccharides

The raffinose family of oligosaccharides, or RFOs, are a major anti-nutritional factor in mature soybean seed. Monogastric creatures such as poultry, swine, and humans are the primary consumers of soy, and they lack the α-1,6-galactosidase enzyme in the foregut needed to digest RFOs and obtain metabolizable energy from the feedstuff. Bacterial fermentation of RFOs in the hindgut of monogastric animals causes flatulence, bloating, and decreased weight gain. Thus, reduction or removal of RFOs from soybean seed is hypothesized to have a positive effect in animal and human nutrition.

Raffinose is a trisaccharide composed of galactose, fructose, and glucose, and is synthesized in soybean from sucrose and galactinol. Stachyose is a tetrasaccharide with two galactose units, one fructose, and one glucose, and is synthesized from raffinose and galactinol. In soybean seed, the predominant enzyme of raffinose biosynthesis is *raffinose synthase 2,* which is most highly expressed in developing seeds (*RS2*, Glyma.06g179200). The soybean *RS2* genomic structure has five exons, as does the *Arabidopsis* homolog, compared with four exons for grape (*Vitis* spp), three exons for pea (*Pisum sativum*), and two exons for the rice raffinose synthase (*Oriza Sativa*), (Dierking and Bilyeu, 2008). Mutations in *RS2* reduce total raffinose content in seed, improving the soybean

carbohydrate profile. One mutant *RS2* allele from plant introduction 200508 significantly reduces raffinose content across environment conditions compared to wildtype (Skoneczka et al, 2009; Bilyeu and Wiebold, 2016). The mutation is a three base pair deletion at position 991 in the DNA sequence, and results in the loss of a tryptophan residue at position 331 of the amino acid sequence. The tryptophan residue is in a highly conserved region of the gene, and loss of that tryptophan results in the low-raffinose phenotype (Dierking and Bilyeu, 2008). Another SNP mutation in *RS2* identified in a line called '397' causes a threonine to isoleucine conversion, and gives a weaker low-raffinose phenotype (Dierking and Bilyeu, 2009; Hagely et al, 2013). Further, mutations in the other putative raffinose synthases *RS3* and *RS4* combine with the W331- *RS2* mutation to give an ultra-low RFO phenotype (Schillinger et al, 2011). While raffinose synthase is considered the committed step in RFO biosynthesis, mutations have also been identified in the stachyose synthase gene to reduce stachyose content by up to 90% in mature seed (Qui et al, 2015).

Impact of RFOs on animal and human nutrition

Low raffinose content is desirable in soybean seed used for animal feed as it improves the digestibility of the carbohydrate fraction. Extensive research has been done on the impact of RFOs on poultry, swine, and human nutrition, and it has been well-established that RFOs have a negative nutritional impact. Because of the α-1,6 bond of raffinose family oligosaccharides, they are not digested well in the small intestine of monogastric animals. The RFOs are fermented by microorganisms in the gut of the animal, producing gas which causes flatulence, diarrhea, and bloating (Steggerda, 1968; Wagner et al, 1976; Leske and

Coon, 1999). Thus, the presence of RFOs in poultry diets is associated with decreased metabolizable energy and feed efficiency (Parsons et al, 1990; Chen et al, 2013).

Using 80% ethanol extraction to remove oligosaccharides from soybean meal has been shown to increase metabolizable energy by about 20%, dry matter digestibility by 13%, and slow the passage time of the diet by 44 minutes, allowing more time for nutrient absorption (Coon et al, 1990). Further, addition of raffinose and stachyose to low-oligosaccharide soy protein concentrate significantly lowers the true metabolizable energy of the feed, further confirming oligosaccharides' negative impact on animal nutrition (Leske et al, 1993).

In addition to altering the soybean meal processing procedure, research has been done using soybean meal derived from genetically improved low-raffinose soybean varieties. Precision-fed rooster assays using cecectomized and conventional birds showed a significant increase in true metabolizable energy for low-oligosaccharide soybean meals (Parsons, 2000 et al; Chen et al, 2013). The 2013 study also evaluated the use of low-raffinose soy on the growth of broiler chickens, and found that although there were no significant differences in broiler growth between low-raffinose and conventional soybean meals, the low-raffinose meal provided a significantly higher feed efficiency than conventional. These studies used soybean varieties produced by the Dupont Company and the United Soybean Board commodity organization, and found that reduction of raffinose in soybean seed has a positive effect on digestibility in poultry. Negative impacts

of RFOs could also be lessened by addition of α-galactosidase enzyme to the animal diet, but this has proven ineffective as of yet (Angel, 1988 et al; Waldroup et al, 2006)

In swine, which is also a monogastric animal, raffinose content in the diet has been shown to have no significant effect on or amino acid digestibility or fecal consistency (Smiricky et al, 2002), and RFOs have an ileal digestion rate of 77% (Smiricky-Tjardes et al, 2003). However, the later study did demonstrate that inclusion of RFOs in swine diets still significantly decreased the ileal digestibility of the feed and also increases gas and short-chain fatty acid production. Overall, removal or reduction of raffinose family oligosaccharides from soybean is known to have a positive impact on monogastric animal nutrition.

History of transgenic crops

The first genetically modified organisms were developed in 1971; they were *E. coli* infected with tumor-inducing viral DNA (Devos et al, 2008). The development of transgenic plants followed in 1983 with tobacco (Bevan et al, 1983; Fraley et al, 1983; Herrera-Estrella et al, 1983), and the first field trials of transgenic plants were planted in 1986. The first transgenic food product in the American market was the Flavr Savr tomato, which was a delayed-ripening variety approved for commercial sale in 1994. 1995 saw the first herbicide resistant cotton and canola, and in 1996, the first generation of insect resistant transgenic plants were sold commercially, carrying a portion of gene coding a protein derived from the bacterium *Bacillus thuringiensis* which confers resistance to certain species of insects (Vaeck et al, 1987; Duke, 2005). These *BT* crops have quickly become the most widely-used form of transgenic insect resistance in the world. Since those

initial introductions of transgenic crop plants, many other varieties have been produced, including other insect- and herbicide- resistant traits, traits to improve tolerance to biotic and abiotic stresses, transgenic plants that produce specific pharmaceuticals or industrial compounds, and plants with improved nutritional profiles (Herrera-Estrella et al, 2004).

Transgenic plants can be produced in several ways, the most common being *Agrobacterium*-mediated transformation and biolistic methods. *Agrobacterium tumefaciens* is a Gram-negative soil bacterium that is the causative agent of crown gall tumors in plants. In 1974, it was discovered that virulent strains of *Agrobacterium* capable of inducing tumors carried a megaplasmid that was not present in non-virulent strains. This plasmid was named the Ti plasmid, and it has since been disarmed and modified to carry desired genes for transformation rather than the tumor-inducing genes (Zaenen et al, 1974). These disarmed strains of Agrobacterium have been used to transform many species of plants, including both monocot and dicot species, and also some species of fungi (Bundock et al, 1995). Currently, *Agrobacterium*-mediated transformation is considered the most efficient and widely-used method of plant transformation because it offers the benefits of a relatively high transformation rate for some crop species, the high proportion of desirable single-copy events, and the ability to transfer large segments of DNA into the plant host genome.

Another common method of plant transformation is the biolistic, or "gene gun" method. This method involves the delivery of microprojectiles coated in DNA to directly deliver the transgene into the host genome. The microprojectiles are generally made of inert metals

such as tungsten or gold, and the "gun" is powered by high pressure carbon dioxide or helium (Klein et al, 1987; Slater et al, 2008). Biolistic transformation is generally used when the plant species to be transformed is recalcitrant to *Agrobacterium* infection. A major drawback of the biolistic method is that the integration of the transgene is often complex, with multiple copies at multiple insertion sites, and often the copies are arranged in tandem repeats or palindromic repeats that can complicate downstream analysis.

Other methods for plant transformation include the use of polyethylene glycol (PEG) to mediate the uptake of DNA into protoplasts for transient expression analysis (Kofer et al, 1998), and electroporation of protoplasts to also induce transient expression (Bates, 1999).

Transgenic methods of gene modulation

Overexpression

The first uses of transgenic technology were to add genes and their expression in the plant. Today, there are many techniques available to researchers to modulate the expression of one or many genes in transgenic plants. Overexpression of transgenes is usually accomplished using a strong promoter sequence to drive the transgene, such as the promoter from the 35S RNA of the Cauliflower mosaic virus (CaMV). The CaMV 35S promoter is a strong constitutive promoter that promotes gene expression in most tissue of dicots, but is not as effective in monocots (Odell et al, 1985). Monocot overexpression is driven chiefly by one of the ubiquitin promoters (Christensen and Quail, 1996).

Gene stacking

Driving the expression of multiple genes in parallel is another useful tool in transgenic technology. Perhaps the most simplistic ways to add several genes to a crop genome are to use iterative transformation steps, adding one gene at a time and then re-transforming the resulting transgenic progeny, or to cross two independently transgenic plants (Cao et al, 2002). However, these methods can be time-consuming, and re-transformation requires the use of separate selectable markers at each stage of transformation, and recombination or interaction between the transgenes can be problematic (Singla-Pareek et al, 2003). A similar approach is to use co-transformation of multiple constructs both within the same strain of *Agrobacterium* or in two separate bacterial strains, as was used in the development of the first generation of golden rice (Ye, 2000). This approach can produce both genetically linked and unlinked transgenes, but is limited by the efficiency of co-integration (Komari et al, 1996). Biolistic methods of co-integration are also available, but these are further limited by the unpredictable nature of transgene integration and copy number in biolistic transformation. More modern methods of gene stacking include the use of the Cre-*lox* and Bxb1-*att* systems to excise unneeded selectable markers from events for use in re-transformation (Hou et al, 2014), the use of polycistronic transcripts with multiple genes encoded by one polypeptide (Vemanna et al, 2013), and the use of "trait landing pads" which utilize zinc finger nuclease activity to precisely integrate multiple transgenes neighboring one another (Ainley et al, 2013). Further, maize minichromosomes, which can be derived from the maize supernumerary B chromosomes, could be used as artificial chromosome vectors, carrying much larger amounts of DNA than any known binary or super-binary vector (Kato et al, 2005).

Gene down-regulation

Although the addition and over-expression of genes has produced some of the most widely-used transgenic crops, for some applications, it is necessary to reduce endogenous gene expression. Most of the current methods of gene down-regulation involve the use of endogenous RNA pathways collectively called RNA interference, which likely evolved in the plant as a defense against invading viruses and transposable elements. These pathways are alike in that they involve small RNA species of 19 to 27 nucleotides and most are produced from intermediates by Dicer-like proteins (DCLs) and loaded into Argonaute protein complexes for action, but differ in their origin and biogenesis, degree of complementarity with the target gene, and the mechanism by which they reduce gene expression.

siRNA

Short interfering RNAs, or siRNAs can be endogenous or exogenously expressed, and serve as guide sequences for target and cleavage or repression of complementary mRNA. In 1998, Fire and Mello established that it was double-stranded RNA that reduced gene expression, not single-stranded complementary RNAs binding in a stoichiometric fashion and causing translational arrest (Fire et al, 1998). siRNAs are processed from double stranded RNA into short, 21 nucleotide fragments via the Dicer enzyme, which is a RNaseIII family ribonuclease (Hamilton and Baulcombe, 1999; Zamore et al, 2000; Bernstein et al, 2001). The double stranded RNA may originate from exogenous,

transgenic sources or endogenous DNA repeat sources. The two RNA strands are unwound, and the strand with the least thermodynamic stability at the 5' end is loaded into the RNA-induced silencing complex (RISC) as the guide strand (Siomi and Siomi, 2009). The RISC includes one of the Argonaute (AGO) protein family members, which contains a single-stranded RNA binding domain for association with the guide strand and a nuclease domain for cleavage of the target mRNA. Alternately, some AGO proteins have lost the cleavage ability but retain function in gene suppression by binding to targeted mRNA and causing translational arrest (Tang, 2005; Hutvagner and Simard, 2008).

miRNA

Micro RNAs, or miRNAs, begin as a primary miRNA transcribed by RNA polymerase II (Lee, 2002). The transcript folds back on itself via imperfect pairing to form stem-loop secondary structures. This longer primary miRNA is processed by *DCL1* into short, 20 to 24 nucleotide miRNAs (Park et al, 2002). Plant miRNAs are processed from imperfectly paired precursors, and regulate expression of perfectly or almost perfectly paired complementary mRNA sequences predominantly via activity of the the AGO1 protein (Broderson and Voinnet, 2006; Ha and Kim, 2014). miRNAs are capable of both translational arrest and mRNA cleavage, depending on the degree of complementary between miRNA and target mRNA and the associated AGO protein (Borges and Martienssen, 2015). miRNA gene silencing can be engineered by replacing the endogenous primary miRNA targets with designed targets complementary to the gene(s) of interest.

tasiRNA

Trans-acting small interfering RNAs (tasiRNAs) begin as an untranslated primary pri-tasiRNA transcript that is processed by miRNA-directed cleavage by an AGO protein. In Arabidopsis, TAS1 and TAS2 are cleaved by miR173, and AGO1 mediates cleavage of the target (Allen and Howell, 2010). TAS3 requires miR390 binding in two places on the transcript, with cleavage by AGO7 (Adenot et al, 2006). TAS4 is processed by miR828-guided, AGO1-mediated cleavage. For TAS1-4, the cleaved transcript is stabilized by SGS3, and then RNA-dependent RNA polymerase 6 copies the ssRNA into double-stranded RNA (Yoshikawa, 2005). The dsRNA is further processed into arrayed 21 nucleotide siRNAs by Dicer-like 4, beginning at the miRNA cleavage site. Like dsRNA and miRNA, tasiRNAs are loaded into the RISC complex and target mRNA for cleavage or translational arrest (Molnar et al, 2011).

Artificial trans-acting small interfering RNA (atasiRNA) are produced by engineering the TAS transcripts by replacing the small interfering RNA sequences downstream of the miRNA target site (de la Luz Gutiérrez-Nava et al, 2008). atasiRNAs have many of the same advantages as artificial miRNAs, as the short sequences allow for specific down-regulation of target genes. However, tasiRNAs have the added advantage of having multiple small interfering RNA sequences in one construct, allowing multiple genes to be targeted at once.

CRISPR/Cas9

Before 2013, the most popular tools for gene editing were TALENs, transcription activator-like effector nucleases, and zinc finger nucleases, which are both limited by the large size of DNA sequence coding for the protein and the difficulty in construct design. The complex of clustered regularly interspaced short palindromic repeats (CRISPR) and the CRISPR-associated protein 9 (Cas9) is a novel gene editing tool that causes targeted double-strand breaks in genomic DNA of animals, plants, and humans. The most commonly used versions of CRISPR/Cas9 are derived from the bacteria *Streptococcus pyogenes*, and they function as part of the adaptive immune system, warding off invading nucleic acids from viruses (Barrangou et al, 2007). CRISPR/Cas9 functions by cleaving those targeted invading DNA sequences, and that attribute has been harnessed by researchers to cleave specific genomic DNA sequences in other organisms.

The CRISPR/Cas9 pathway in bacteria begins with the incorporation of short DNA sequences acquired from invading sources being integrated into the CRISPR locus. The CRISPR arrays are transcribed into roughly 40 nucleotide interfering CRISPR RNAs, or crRNAs. These mature crRNAs, together with associated trans-acting RNAs, tracrRNAs, and the Cas9 protein are all that is required to target and cleave complementary DNA (Deltcheva et al, 2011). The targeted DNA must have the protospacer-adjacent motif (PAM), usually 5'-NGG-3' and less frequently 5'-NAG-3', downstream of the target cleavage site (Jinek et al, 2012; Hsu et al, 2013). crRNA specificity for the target is provided by the 'seed sequence' approximately 12 nucleotides upstream of the PAM, and the crRNA and target must have direct matches in that region.

Engineering of the CRISPR/Cas9 system for biotechnology use has involved switching the crRNA sequences for sequences complementary to target DNA, and combining the crRNA and tracrRNA into a single guide RNA. Together with the Cas9 protein, this guide RNA can cleave genomic DNA of plants (Feng et al, 2013), animals (Chang et al, 2013), fungi (Nødvig et al, 2015), and even humans (Liang et al, 2015). To achieve gene editing, the broken double stranded DNA is repaired by non-homologous end joining (NHEJ) or homologous recombination (HR). NHEJ is imperfect, resulting in insertions, deletions, mutations, and if the break leaves DNA overhangs, compatible DNA fragments can be incorporated into the break site (Cristea et al, 2013). Homologous recombination can be used to create specific mutations or DNA insertions, though the technique is less efficient than the NHEJ pathway (Puchta, 2004).

Catalytically inactive versions of the Cas9 protein (dCas9) lacking nuclease function have also been developed, so they lack the ability to cleave DNA. These chimeric proteins are a fusion of the C-terminus of dCas9 with transcriptional activators or repressors, and binding of the dCas9 complex to guide-RNA targeted genomic locations can modulate gene expression (Piatek et al, 2015). Also, fusion of the GFP gene to dCas9 can give live imaging of genetic loci and the chromatic structure in real time (B. Chen et al, 2013).

CRISPR/Cas9 design is limited by the availability of convenient PAM sequences in the gene of interest, design to reduce the probability of off-target effects, and work is still being done to determine the efficiency of Cas9 in different crop species. For each crop species, the ideal codon-optimization of the Cas9 protein sequence must be determined, as well as

the optimizations of vector, promoter, terminator, and selectable marker. The world regulatory situation of Cas9 is yet unclear, but in April 2016, the United States Department of Agriculture declared for the first time it would not regulate a Cas9-gene-edited food organism, a variety of button mushroom resistant to browning (Waltz, 2016). This decision could lower the barrier to entry in the biotech industry, as the high cost of deregulation is prohibitive for smaller, independent companies and traits without a large economic value.

Summary

In summary, while soybean is a superior source of plant protein for human and animal use, soybean composition can be greatly improved by breeding and biotechnology efforts. Specifically, the carbohydrate fraction of soybean can be improved by reduction of raffinose family oligosaccharides and increase of sucrose. These modifications should both increase the human acceptability of soyfood products in western countries and increase the nutritional value for animal feed applications. Because a single soybean gene, *RS2*, has a large effect on soybean raffinose and sucrose content, double-stranded RNA-mediated gene silencing could be employed to reduce *RS2* expression and improve the soybean carbohydrate profile.

References

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouché, N., Gasciolli, V., & Vaucheret, H. (2006). DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Current Biology: CB*, *16*(9), 927–932. http://doi.org/10.1016/j.cub.2006.03.035

Ainley, W. M., Sastry-Dent, L., Welter, M. E., Murray, M. G., Zeitler, B., Amora, R., … Petolino, J. F. (2013). Trait stacking via targeted genome editing. *Plant Biotechnology Journal*, *11*(9), 1126–1134. http://doi.org/10.1111/pbi.12107

Allen, E., & Howell, M. D. (2010). miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *Seminars in Cell & Developmental Biology*, *21*(8), 798–804. http://doi.org/10.1016/j.semcdb.2010.03.008

Angel, C. R., Sell, J. L., & Zimmerman, D. R. (1988). Autolysis of alpha-galactosides of defatted soy flakes: influence on nutritive value for chickens. *Journal of Agricultural and Food Chemistry*, *36*(3), 542–546. http://doi.org/10.1021/jf00081a034

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., … Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*(5819), 1709–1712. http://doi.org/10.1126/science.1138140

Bates, G. W. (1999). Plant Transformation via Protoplast Electroporation. In R. D. Hall, *Plant Cell Culture Protocols* (Vol. 111, pp. 359–366). New Jersey: Humana Press. Retrieved from http://link.springer.com/10.1385/1-59259-583-9:359

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, *409*(6818), 363–366. http://doi.org/10.1038/35053110

Bevan, M. W., Flavell, R. B., & Chilton, M.-D. (1983). A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature*, *304*(5922), 184–187. http://doi.org/10.1038/304184a0

Bianchi-Hall, C. M., Carter, T. E., Bailey, M. A., Mian, M. A. R., Rufty, T. W., Ashley, D. A., … Parrott, W. A. (2000). Aluminum Tolerance Associated with Quantitative Trait Loci Derived from Soybean PI 416937 in Hydroponics. *Crop Science*, *40*(2), 538. http://doi.org/10.2135/cropsci2000.402538x

Bilyeu, K. D., & Wiebold, W. J. (2016). Environmental Stability of Seed Carbohydrate Profiles in Soybeans Containing Different Alleles of the Raffinose Synthase 2 (RS2) Gene. *Journal of Agricultural and Food Chemistry*, *64*(5), 1071–1078. http://doi.org/10.1021/acs.jafc.5b04779

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nature Reviews Molecular Cell Biology*, *16*(12), 727–741. http://doi.org/10.1038/nrm4085

Brodersen, P., & Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. *Trends in Genetics: TIG*, *22*(5), 268–280. http://doi.org/10.1016/j.tig.2006.03.003

Bundock, P., den Dulk-Ras, A., Beijersbergen, A., & Hooykaas, P. J. (1995). Trans-kingdom T-DNA transfer from Agrobacterium tumefaciens to Saccharomyces cerevisiae. *The EMBO Journal*, *14*(13), 3206–3214.

Cao, J., Zhao, J.-Z., Tang, D., Shelton, M., & Earle, D. (2002). Broccoli plants with pyramided cry1Ac and cry1C Bt genes control diamondback moths resistant to Cry1A and Cry1C proteins. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *105*(2-3), 258–264. http://doi.org/10.1007/s00122-002-0942-0

Carter, TE, Nelson, R, Sneller, CH, & Cui, Z. (2004). Genetic diversity in soybean. In *Soybeans: Improvement, Production, and Uses* (Boerma HR , Specht JE V, pp. 303–416). Am Soc of Agronomy, Crop Sci Soc of Am, Soil Sci Soc of Am.

Chang, C.-T., Hsu, C.-K., Chou, S.-T., Chen, Y.-C., Huang, F.-S., & Chung, Y.-C. (2009). Effect of fermentation time on the antioxidant activities of tempeh prepared from fermented soybean using *Rhizopus oligosporus*. *International Journal of Food Science & Technology*, *44*(4), 799–806. http://doi.org/10.1111/j.1365-2621.2009.01907.x

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., … Xi, J. J. (2013). Genome editing with RNA-guided Cas9 nuclease in Zebrafish embryos. *Cell Research*, *23*(4), 465–472. http://doi.org/10.1038/cr.2013.45

Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G.-W., … Huang, B. (2013). Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*, *155*(7), 1479–1491. http://doi.org/10.1016/j.cell.2013.12.001

Chen, X., Parsons, C. M., & Bajjalieh, N. (2013). Nutritional evaluation of new reduced oligosaccharide soybean meal in poultry. *Poultry Science*, *92*(7), 1830–1836. http://doi.org/10.3382/ps.2012-02856

Chen, Y.-M., Shih, T.-W., Chiu, C. P., Pan, T.-M., & Tsai, T.-Y. (2013). Effects of lactic acid bacteria-fermented soy milk on melanogenesis in B16F0 melanocytes. *Journal of Functional Foods*, *5*(1), 395–405. http://doi.org/10.1016/j.jff.2012.11.012

Christensen, A. H., & Quail, P. H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research*, *5*(3), 213–218. http://doi.org/10.1007/BF01969712

Chung, J., Babka, H. L., Graef, G. L., Staswick, P. E., Lee, D. J., Cregan, P. B., … Specht, J. E. (2003). The Seed Protein, Oil, and Yield QTL on Soybean Linkage Group I. *Crop Science*, *43*(3), 1053. http://doi.org/10.2135/cropsci2003.1053

Coon, C. N., Leske, K. L., Akavanichan, O., & Cheng, T. K. (1990). Effect of Oligosaccharide-Free Soybean Meal on True Metabolizable Energy and Fiber Digestion in Adult Roosters. *Poultry Science*, *69*(5), 787–793. http://doi.org/10.3382/ps.0690787

Cristea, S., Freyvert, Y., Santiago, Y., Holmes, M. C., Urnov, F. D., Gregory, P. D., & Cost, G. J. (2013). In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnology and Bioengineering*, *110*(3), 871–880. http://doi.org/10.1002/bit.24733

Cromwell, Gary L. (2012). *Soybean Meal – An Exceptional Protein Source*. University of Kentucky. Retrieved from http://www.soymeal.org/ReviewPapers/SBMExceptionalProteinSource.pdf

de la Luz Gutiérrez-Nava, M., Aukerman, M. J., Sakai, H., Tingey, S. V., & Williams, R. W. (2008). Artificial trans-acting siRNAs confer consistent and effective gene silencing. *Plant Physiology*, *147*(2), 543–551. http://doi.org/10.1104/pp.108.118307

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., … Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, *471*(7340), 602–607. http://doi.org/10.1038/nature09886

Devos, Y., Maeseele, P., Reheul, D., Van Speybroeck, L., & De Waele, D. (2008). Ethics in the Societal Debate on Genetically Modified Organisms: A (Re)Quest for Sense and Sensibility. *Journal of Agricultural and Environmental Ethics*, *21*(1), 29–61. http://doi.org/10.1007/s10806-007-9057-6

Dierking, E. C., & Bilyeu, K. D. (2008). Association of a Soybean Raffinose Synthase Gene with Low Raffinose and Stachyose Seed Phenotype. *The Plant Genome Journal*, *1*(2), 135. http://doi.org/10.3835/plantgenome2008.06.0321

Dierking, E. C., & Bilyeu, K. D. (2009). New sources of soybean seed meal and oil composition traits identified through TILLING. *BMC Plant Biology*, *9*(1), 89. http://doi.org/10.1186/1471-2229-9-89

Dornbos, D. L., & Mullen, R. E. (1992). Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. *Journal of the American Oil Chemists Society*, *69*(3), 228–231. http://doi.org/10.1007/BF02635891

Doyle, J. J., Flagel, L. E., Paterson, A. H., Rapp, R. A., Soltis, D. E., Soltis, P. S., & Wendel, J. F. (2008). Evolutionary Genetics of Genome Merger and Doubling in Plants. *Annual Review of Genetics*, *42*(1), 443–461. http://doi.org/10.1146/annurev.genet.42.110807.091524

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., … Zhu, J.-K. (2013). Efficient genome editing in plants using a CRISPR/Cas system. *Cell Research*, *23*(10), 1229–1232. http://doi.org/10.1038/cr.2013.114

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), 806–811. http://doi.org/10.1038/35888

Fraley, R. T., Rogers, S. G., Horsch, R. B., Sanders, P. R., Flick, J. S., Adams, S. P., … Woo, S. C. (1983). Expression of bacterial genes in plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(15), 4803–4807.

Gibson, L. R., & Mullen, R. E. (1996). Soybean seed composition under high day and night growth temperatures. *Journal of the American Oil Chemists' Society*, *73*(6), 733–737. http://doi.org/10.1007/BF02517949

González, R., Polo, F., Zapatero, L., Caravaca, F., & Carreira, J. (1992). Purification and characterization of major inhalant allergens from soybean hulls. *Clinical and Experimental Allergy: Journal of the British Society for Allergy and Clinical Immunology*, *22*(8), 748–755.

Guo, B., Sleper, D. A., Arelli, P. R., Shannon, J. G., & Nguyen, H. T. (2005). Identification of QTLs associated with resistance to soybean cyst nematode races 2, 3 and 5 in soybean PI 90763. *Theoretical and Applied Genetics*, *111*(5), 965–971. http://doi.org/10.1007/s00122-005-0031-2

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H., & Wang, Y. (2010). A single origin and moderate bottleneck during domestication of soybean (Glycine max): implications from microsatellites and nucleotide sequences. *Annals of Botany*, *106*(3), 505–514. http://doi.org/10.1093/aob/mcq125

Guo, S.-T., & Ono, T. (2006). The Role of Composition and Content of Protein Particles in Soymilk on Tofu Curding by Glucono-δ-lactone or Calcium Sulfate. *Journal of Food Science*, *70*(4), C258–C262. http://doi.org/10.1111/j.1365-2621.2005.tb07170.x

Hagely, K. B., Palmquist, D., & Bilyeu, K. D. (2013). Classification of Distinct Seed Carbohydrate Profiles in Soybean. *Journal of Agricultural and Food Chemistry*, *61*(5), 1105–1111. http://doi.org/10.1021/jf303985q

Hamilton, A. J., & Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science (New York, N.Y.)*, *286*(5441), 950–952.

Ha, M., & Kim, V. N. (2014). Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology*, *15*(8), 509–524. http://doi.org/10.1038/nrm3838

Han, B.-Z., Kuijpers, A. F. A., Thanh, N. V., & Nout, M. J. R. (2004). Mucoraceous moulds involved in the commercial fermentation of Sufu Pehtze. *Antonie van Leeuwenhoek*, *85*(3), 253–257. http://doi.org/10.1023/B:ANTO.0000020157.72415.b9

Herman, E. (2005). Soybean Allergenicity and Suppression of the Immunodominant Allergen. *Crop Science*, *45*(2), 462. http://doi.org/10.2135/cropsci2005.0462

Herrera-Estrella, L., Depicker, A., Van Montagu, M., & Schell, J. (1983). Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature*, *303*(5914), 209–213. http://doi.org/10.1038/303209a0

Herrera-Estrella, L., Simpson, J., & Martínez-Trujillo, M. (2004). Transgenic Plants: An Historical Perspective. In L. Pe□a, *Transgenic Plants* (Vol. 286, pp. 003–032). New Jersey: Humana Press. Retrieved from http://link.springer.com/10.1385/1-59259-827-7:003

Hosking, R. (1996). *A dictionary of Japanese food: ingredients & culture* (1st ed). Rutland, Vt: Charles E. Tuttle Co.

Hou, L., Yau, Y.-Y., Wei, J., Han, Z., Dong, Z., & Ow, D. W. (2014). An Open-Source System for In Planta Gene Stacking by Bxb1 and Cre Recombinases. *Molecular Plant*, *7*(12), 1756–1765. http://doi.org/10.1093/mp/ssu107

Hsu, P. D., Scott, D. A., Weinstein, J. A., Ran, F. A., Konermann, S., Agarwala, V., … Zhang, F. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature Biotechnology*, *31*(9), 827–832. http://doi.org/10.1038/nbt.2647

Hutvagner, G., & Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology*, *9*(1), 22–32. http://doi.org/10.1038/nrm2321

Hymowitz, Theodore. (1990). *Soybeans: The success story* (Advances in new crops) (pp. 159–163).

Hyten, D. L., Song, Q., Zhu, Y., Choi, I.-Y., Nelson, R. L., Costa, J. M., … Cregan, P. B. (2006). Impacts of genetic bottlenecks on soybean genome diversity. *Proceedings of the National Academy of Sciences*, *103*(45), 16666–16671. http://doi.org/10.1073/pnas.0604379103

Imram, N. (2003). *Soya Handbook*. Singapore: Tetra Pak.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, *337*(6096), 816–821. http://doi.org/10.1126/science.1225829

Kawamura S. (n.d.). Quantitative paper chromatography of sugars of the cotyledon, hull, and hypocotyl of soybeans of selected varieties. *Kagawa Univ Fac Tech Bull*, (18), 117–131.

Klein, T. M., Wolf, E. D., Wu, R., & Sanford, J. C. (1987). High-velocity microprojectiles for delivering nucleic acids into living cells. *Nature*, *327*(6117), 70–73. http://doi.org/10.1038/327070a0

Kofer, W., Eibl, C., Steinmüller, K., & Koop, H.-U. (1998). PEG-mediated plastid transformation in higher plants. *In Vitro Cellular & Developmental Biology - Plant*, *34*(4), 303–309. http://doi.org/10.1007/BF02822739

Komari, T., Hiei, Y., Saito, Y., Murai, N., & Kumashiro, T. (1996). Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by Agrobacterium tumefaciens and segregation of transformants free from selection markers. *The Plant Journal*, *10*(1), 165–174. http://doi.org/10.1046/j.1365-313X.1996.10010165.x

Kris-Etherton, P. M., Pearson, T. A., Wan, Y., Hargrove, R. L., Moriarty, K., Fishell, V., & Etherton, T. D. (1999). High-monounsaturated fatty acid diets lower both plasma cholesterol and triacylglycerol concentrations. *The American Journal of Clinical Nutrition*, *70*(6), 1009–1015.

Krishnan, H. B. (2005). Engineering Soybean for Enhanced Sulfur Amino Acid Content. *Crop Science*, *45*(2), 454. http://doi.org/10.2135/cropsci2005.0454

Kumar, V., Rani, A., Goyal, L., Dixit, A. K., Manjaya, J. G., Dev, J., & Swamy, M. (2010). Sucrose and Raffinose Family Oligosaccharides (RFOs) in Soybean Seeds As Influenced by Genotype and Growing Location. *Journal of Agricultural and Food Chemistry*, *58*(8), 5081–5085. http://doi.org/10.1021/jf903141s

Lee, Y., Jeon, K., Lee, J.-T., Kim, S., & Kim, V. N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*, *21*(17), 4663–4670.

Leske, K. L., & Coon, C. N. (1999). Hydrogen gas production of broiler chicks in response to soybean meal and alpha-galactoside free, ethanol-extracted soybean meal. *Poultry Science*, *78*(9), 1313–1316.

Leske, K. L., Jevne, C. J., & Coon, C. N. (1993). Effect of Oligosaccharide Additions on Nitrogen-Corrected True Metabolizable Energy of Soy Protein Concentrate. *Poultry Science*, *72*(4), 664–668. http://doi.org/10.3382/ps.0720664

Liang, P., Xu, Y., Zhang, X., Ding, C., Huang, R., Zhang, Z., … Huang, J. (2015). CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein & Cell*, *6*(5), 363–372. http://doi.org/10.1007/s13238-015-0153-5

Liu, B., Fujita, T., Yan, Z.-H., Sakamoto, S., Xu, D., & Abe, J. (2007). QTL Mapping of Domestication-related Traits in Soybean (Glycine max). *Annals of Botany*, *100*(5), 1027–1038. http://doi.org/10.1093/aob/mcm149

Liu, KS. (2008). Food Use of Whole Soybeans. In *Soybeans: Chemistry, Production, Processing, and Utilization* (L.A. Johnson, P.J. White, and R. Galloway, pp. 441–481). Urbana, IL: AOCS Press.

Manavalan, L. P., Guttikonda, S. K., Phan Tran, L.-S., & Nguyen, H. T. (2009). Physiological and Molecular Approaches to Improve Drought Resistance in Soybean. *Plant and Cell Physiology*, *50*(7), 1260–1276. http://doi.org/10.1093/pcp/pcp082

Medic, J., Atkinson, C., & Hurburgh, C. R. (2014). Current Knowledge in Soybean Composition. *Journal of the American Oil Chemists' Society*, *91*(3), 363–384. http://doi.org/10.1007/s11746-013-2407-9

Molnar, A., Melnyk, C., & Baulcombe, D. C. (2011). Silencing signals in plants: a long journey for small RNAs. *Genome Biology*, *12*(1), 215. http://doi.org/10.1186/gb-2010-11-12-219

Murphy, P. (2008). Soybean proteins. In *Soybeans: Chemistry, Production, Processing, and Utilization* (Johnson L, White P, Galloway R, pp. 229–269). Urbana, IL: AOCS Press.

Nødvig, C. S., Nielsen, J. B., Kogle, M. E., & Mortensen, U. H. (2015). A CRISPR-Cas9 System for Genetic Engineering of Filamentous Fungi. *PLOS ONE*, *10*(7), e0133085. http://doi.org/10.1371/journal.pone.0133085

Odell, J. T., Nagy, F., & Chua, N.-H. (1985). Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature*, *313*(6005), 810–812. http://doi.org/10.1038/313810a0

Ogawa, T., Tsuji, H., Bando, N., Kitamura, K., Zhu, Y. L., Hirano, H., & Nishikawa, K. (1993). Identification of the soybean allergenic protein, Gly m Bd 30K, with the soybean seed 34-kDa oil-body-associated protein. *Bioscience, Biotechnology, and Biochemistry*, *57*(6), 1030–1033.

Osborne, TB, & Mendel, LB. (1917). The use of the soy bean as food. *J. Biol. Chem*, *32*, 369–376.

Park, W., Li, J., Song, R., Messing, J., & Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Current Biology: CB*, *12*(17), 1484–1495.

Parsons, C. M., Zhang, Y., & Araba, M. (2000). Nutritional Evaluation of Soybean Meals Varying in Oligosaccharide Content. *Poultry Science*, *79*(8), 1127–1131. http://doi.org/10.1093/ps/79.8.1127

Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., … Mahfouz, M. M. (2015). RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnology Journal*, *13*(4), 578–589. http://doi.org/10.1111/pbi.12284

Pliner, P., & Hobden, K. (1992). Development of a scale to measure the trait of food neophobia in humans. *Appetite*, *19*(2), 105–120. http://doi.org/10.1016/0195-6663(92)90014-W

Puchta, H. (2004). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *Journal of Experimental Botany*. http://doi.org/10.1093/jxb/eri025

Qiu, D., Vuong, T., Valliyodan, B., Shi, H., Guo, B., Shannon, J. G., & Nguyen, H. T. (2015). Identification and characterization of a stachyose synthase gene controlling reduced stachyose content in soybean. *Theoretical and Applied Genetics*, *128*(11), 2167–2176. http://doi.org/10.1007/s00122-015-2575-0

Ren, C., Bilyeu, K. D., & Beuselinck, P. R. (2009). Composition, Vigor, and Proteome of Mature Soybean Seeds Developed under High Temperature. *Crop Science*, *49*(3), 1010. http://doi.org/10.2135/cropsci2008.05.0247

Rotundo, J. L., & Westgate, M. E. (2009). Meta-analysis of environmental effects on soybean seed composition. *Field Crops Research*, *110*(2), 147–156. http://doi.org/10.1016/j.fcr.2008.07.012

Schillinger, John A, Dierking, Emily C, Bilyeu, Kristin D. (2013, June 25). Soybeans having high germination rates and ultra-low raffinose and stachyose content.

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., … Jackson, S. A. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, *463*(7278), 178–183. http://doi.org/10.1038/nature08670

Shoemaker, R. C., Schlueter, J., & Doyle, J. J. (2006). Paleopolyploidy and gene duplication in soybean and other legumes. *Current Opinion in Plant Biology*, *9*(2), 104–109. http://doi.org/10.1016/j.pbi.2006.01.007

Shurtleff, W., & Aoyagi, A. (2000). *Tofu & soymilk production: a craft and technical manual*. Lafayette, CA: Soyfoods Center.

Shurtleff, William, & Aoyagi, Akiko. (2008). *History of cooperative soybean processing in the United States*. Lafayette, CA: Soyinfo Center.

Singh, R. J., & Hymowitz, T. (1999). Soybean genetic resources and crop improvement. *Genome*, *42*(4), 605–616. http://doi.org/10.1139/g99-039

Singla-Pareek, S. L., Reddy, M. K., & Sopory, S. K. (2003). Genetic engineering of the glyoxalase pathway in tobacco leads to enhanced salinity tolerance. *Proceedings of the National Academy of Sciences*, *100*(25), 14672–14677. http://doi.org/10.1073/pnas.2034667100

Siomi, H., & Siomi, M. C. (2009). On the road to reading the RNA-interference code. *Nature*, *457*(7228), 396–404. http://doi.org/10.1038/nature07754

Skoneczka, J. A., Maroof, M. A. S., Shang, C., & Buss, G. R. (2009). Identification of Candidate Gene Mutation Associated With Low Stachyose Phenotype in Soybean Line PI200508. *Crop Science*, *49*(1), 247. http://doi.org/10.2135/cropsci2008.07.0403

Slater, A., Scott, N. W., & Fowler, M. R. (2008). *Plant biotechnology: the genetic manipulation of plants* (2nd ed). Oxford ; New York: Oxford University Press.

Smalley, M. D., Fehr, W. R., Cianzio, S. R., Han, F., Sebastian, S. A., & Streit, L. G. (2004). Quantitative Trait Loci for Soybean Seed Yield in Elite and Plant Introduction Germplasm. *Crop Science*, *44*(2), 436. http://doi.org/10.2135/cropsci2004.4360

Smiricky, M. R., Grieshop, C. M., Albin, D. M., Wubben, J. E., Gabert, V. M., & Fahey, G. C. (2002). The influence of soy oligosaccharides on apparent and true ileal amino acid digestibilities and fecal consistency in growing pigs. *Journal of Animal Science*, *80*(9), 2433–2441.

Smiricky-Tjardes, M. R., Grieshop, C. M., Flickinger, E. A., Bauer, L. L., & Fahey, G. C. (2003). Dietary galactooligosaccharides affect ileal and total-tract nutrient digestibility, ileal and fecal bacterial concentrations, and ileal fermentative characteristics of growing pigs. *Journal of Animal Science*, *81*(10), 2535–2545.

Soystats. (n.d.). *Soystats*. Retrieved from soystats.com

Steggerda, F. R. (1968). Gastrointestinal gas following food consumption. *Annals of the New York Academy of Sciences*, *150*(1), 57–66.

Stupar, R. M. (2010). Into the wild: The soybean genome meets its undomesticated relative. *Proceedings of the National Academy of Sciences*, *107*(51), 21947–21948. http://doi.org/10.1073/pnas.1016809108

Tang, G. (2005). siRNA and miRNA: an insight into RISCs. *Trends in Biochemical Sciences*, *30*(2), 106–114. http://doi.org/10.1016/j.tibs.2004.12.007

Tu, V. P., Valentin, D., Husson, F., & Dacremont, C. (2010). Cultural differences in food description and preference: Contrasting Vietnamese and French panellists on soy yogurts. *Food Quality and Preference*, *21*(6), 602–610. http://doi.org/10.1016/j.foodqual.2010.03.009

Vaeck, M., Reynaerts, A., Höfte, H., Jansens, S., De Beuckeleer, M., Dean, C., … Leemans, J. (1987). Transgenic plants protected from insect attack. *Nature*, *328*(6125), 33–37. http://doi.org/10.1038/328033a0

Vemanna, R. S., Chandrashekar, B. K., Hanumantha Rao, H. M., Sathyanarayanagupta, S. K., Sarangi, K. S., Nataraja, K. N., & Udayakumar, M. (2013). A Modified MultiSite Gateway Cloning Strategy for Consolidation of Genes in Plants. *Molecular Biotechnology*, *53*(2), 129–138. http://doi.org/10.1007/s12033-012-9499-6

Wagner, J. R., Becker, R., Gumbmann, M. R., & Olson, A. C. (1976). Hydrogen production in the rat following ingestion of raffinose, stachyose and oligosaccharide-free bean residue. *The Journal of Nutrition*, *106*(4), 466–470.

Waldroup, P. W., Keen, C. A., Yan, F., & Zhang, K. (2006). The Effect of Levels of -Galactosidase Enzyme on Performance of Broilers Fed Diets Based on Corn and Soybean Meal1. *The Journal of Applied Poultry Research*, *15*(1), 48–57. http://doi.org/10.1093/japr/15.1.48

Waltz, E. (2016). Gene-edited CRISPR mushroom escapes US regulation. *Nature*, *532*(7599), 293–293. http://doi.org/10.1038/nature.2016.19754

Ye, X. (2000). Engineering the Provitamin A (-Carotene) Biosynthetic Pathway into (Carotenoid-Free) Rice Endosperm. *Science*, *287*(5451), 303–305. http://doi.org/10.1126/science.287.5451.303

Yoshikawa, M. (2005). A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes & Development*, *19*(18), 2164–2175. http://doi.org/10.1101/gad.1352605

Yu Ying-Li. (1913, June 17). Method of manufacturing products from soja.

Zaenen, I., van Larebeke, N., Teuchy, H., van Montagu, M., & Schell, J. (1974). Supercoiled circular DNA in crown-gall inducing Agrobacterium strains. *Journal of Molecular Biology*, *86*(1), 109–127. http://doi.org/10.1016/S0022-2836(74)80011-2

Zamore, P. D., Tuschl, T., Sharp, P. A., & Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, *101*(1), 25–33. http://doi.org/10.1016/S0092-8674(00)80620-0

Chapter 2

Downregulation of soybean RS2 by RNA interference and characterization of transgenic events

Abstract

Soybean (*Glycine max*) is the number one oil and protein crop in the United States, but the seed contains several anti-nutritional factors that are undesirable for both human and animal nutrition. RNA interference technology has become an increasingly popular technique in gene silencing because it allows for both temporal and spatial targeting of specific genes. The objective of this research is to use RNA-mediated gene silencing to down-regulate the soybean gene raffinose synthase 2 (RS2), to reduce total raffinose content in mature seed. Raffinose is a trisaccharide that is indigestible to humans and monogastric animals, and as monogastric animals are the largest consumers of soy products, reducing raffinose would improve the nutritional quality of soybean. An RNAi construct targeting RS2 was designed, cloned, and transformed to the soybean genome via *Agrobacterium* infection. Resulting plants were analyzed for the presence and number of copies of the transgene by PCR and Southern blot. The efficiency of mRNA silencing was confirmed by real-time quantitative PCR. Total raffinose content was determined by HPLC analysis. Transgenic plant lines were recovered that exhibited dramatically reduced levels of raffinose in mature seed, and these lines were further analyzed for other phenotypes such as development and yield. This research serves as further confirmation of the contribution of the RS2 gene to the low-raffinose phenotype, and validates both the effectiveness and specificity of RNA interference, thus the technology could be applied to many other important genes in crop plants.

Introduction

Soybeans are the number one protein source for animal feed in the world, accounting for 69% of global protein consumption with poultry and swine – both monogastric animals – being the major consumers (Cromwell, 2012). However, soy's use in monogastric animal diets must be supplemented with alternate sources of carbohydrates such as corn or other cereal grains to compensate for the overall low level of digestible carbohydrates. One of the major limitations of soy carbohydrates are the presence of the indigestible raffinose family oligosaccharides (RFOs): raffinose, stachyose, and verbascose. These compounds are derived from sucrose, which has a positive effect on metabolizable energy, but due to their alpha-1,6 glycosidic bond, monogastric animals are unable to digest RFOs. The oligosaccharides pass undigested through the upper gut of the animal, and are then fermented by anaerobic microbes in the lower gut. The fermentation produces carbon dioxide, methane, and hydrogen, causing flatulence and digestive disturbance in the animal. Further, it has been shown that presence of RFOs in animal diets cause the feed to pass quicker through the digestive system, reducing the amount of other nutrients absorbed from the feed (Coon et al, 1990).

RFOs are removed from most human soy food products by ethanol extraction, which removes almost all soluble carbohydrates. The extraction produces soy protein concentrate, which has been shown to improve turkey growth compared to regular soybean meal (Jankowski et al, 2009), but the processing adds considerable cost to the animal diet. In addition, ethanol extraction removes desirable carbohydrates such as sucrose, lowering the overall metabolizable energy of the feed. Other methods for reducing the negative

impacts of RFOs include bacterial fermentation and supplementing the diet with alpha-galactosidase enzyme (Yoon and Hwang, 2008; Chen et al, 2013). Bacterial fermentation can reduce RFO content, improve amino acid digestibility, and reduce soy immunoreactivity. However, fermentation and enzyme degradation of RFOs are currently limited by pH, temperature, and time constraints, and need further research to test their application in animal agriculture (Frias et al, 2008; Song et al, 2008). The percent of soybean meal used in animal diets is restricted in large part due to soy's relatively low true metabolizable energy value (TME). Zero- to three-week-old broiler chickens require 3200 kcal/kg in the diet, of which 23% should be protein (National Research Council, 1994). On average, corn provides 3350 kcal/kg of true metabolizable energy and 8.5% protein, while soybean meal provides 2990 kcal/kg with 37% protein (Windows-based User Friendly Feed Formulation (WUFFDA), 2014). Although soybeans have a high protein level, the relatively low quantity of digestible soymeal carbohydrates and presence of RFOs requires chicken and pig diets to be supplemented with corn, barley, or wheat to add additional metabolizable energy.

In plants, raffinose and related compounds are believed to provide protection from various stresses such as drought tolerance (Wang et al, 2009), seed desiccation tolerance (Koster and Leopold, 1988), cold tolerance (Zuther et al, 2004), scavenging of reactive oxygen species (Nishizawa et al, 2008), and partitioning of carbohydrates during times of stress (ElSayed et al, 2014). In soybean, slow drying of immature seeds increases RFO accumulation, and a positive correlation between seed stachyose content and desiccation tolerance has been established (Blackman et al, 1992). RFOs may also be a readily

accessible energy source for germinating seeds, as inhibiting RFO metabolism drastically decreases germination of pea seeds (Blöchl, 2007). However, raffinose and stachyose metabolism is not required for soybean seed germination, as demonstrated in lines bred for low seed RFO content (Dierking and Bilyeu, 2009b). It is hypothesized that the primary functions of RFOs are storage and transport, and although high accumulations of RFOs during times of stress do provide protection, stress protection is not their exclusive role in plants (Sengupta et al, 2015).

Sucrose accumulation in soybean seeds is negatively impacted by shading (Egli, 2001), high temperatures in controlled (Wolf et al, 1982) and field (Kumar et al, 2010; Bilyeu and Wiebold, 2016) conditions, and has a negative correlation with elevated protein content (Hymowitz et al, 1972). Raffinose has been shown to be significantly affected by environmental conditions in chickpea (Gangola et al, 2013) quinoa (Miranda et al, 2013), and soybean (Bilyeu and Wiebold, 2016).

Raffinose biosynthesis in developing soybean seeds is catalyzed by raffinose synthase 2 (RS2), which is encoded by Glyma06g18890 (Dierking and Bilyeu, 2008). RS2 catalyzes the reaction: sucrose + galactinol → raffinose + myo-inositol. The subsequent conversion of raffinose to stachyose, and stachyose to verbascose produce the spectrum of raffinose family oligosaccharides, though verbascose content in soybean seeds is negligible (Kumar et al, 2010). RS2 is considered the committed step in RFO biosynthesis, and down-regulation of the gene should lead to increased sucrose, and decreased raffinose and stachyose. Successful soybean breeding efforts to reduce raffinose content in mature seeds

have lowered raffinose content from about 1 to 1.5% in wild-type to almost undetectable levels in the mutant lines. Two major mutations in the soybean gene RS2 have been identified and associated with the low-raffinose phenotype. The first is a deletion of the codon for a highly-conserved tryptophan at position 331 in the amino acid sequence identified in PI200508, with the allele named W331- (Kerr and Sebastian, 2000). This mutation results in a low-raffinose phenotype that is relatively stable across environments, but moderately variable depending on temperature at pod-fill (Bilyeu and Wiebold, 2016). Another mutant line, 397, has a missense SNP mutation that causes the incorporation of isoleucine instead of threonine at amino acid position 107 – allele T107I (Dierking and Bilyeu, 2009a). This mutant line has a weaker low-raffinose phenotype than W331-mutant and also shows significant environmental variability in sucrose and stachyose content. The absolute quantities of sucrose, raffinose, and stachyose are highly dependent on location and planting date, but overall, W331-mutants had a more severe low-raffinose phenotype than T107I mutants. Agronomic traits such as emergence, yield, maturity, seed protein, and oil content were not affected in W331-mutants (Neus et al, 2005).

Although two individual mutant lines with a low-raffinose phenotype have been identified, additional association of the RS2 gene with altered carbohydrate phenotype could provide separate confirmation of the gene-to-phenotype association. Transgenic knockdown of gene expression is a useful technique in reverse genetics to analyze the contribution of a gene to a particular phenotype. RNA interference has become an efficient tool for gene knockdown and regulation (Hannon, 2002). In plants, the RNAi pathway begins with double-stranded RNA, originating from viral, transgenic, or endogenous sources being

processed by the enzyme DICER into short, 20 to 25nt small interfering RNAs (Yin, 2005). These double-stranded siRNAs are unwound into the guide and passenger strands by argonaute, and the guide strand is loaded into the RNA induced silencing complex, or RISC. The siRNA-RISC complex targets complementary mRNA for degradation by RNase activity of argonaute, causing post-transcriptional gene silencing (Naqvi, 2009).

RNAi has been used to down-regulate gene expression in several major crop plants, including reduction of terpenoid compounds in cottonseed (Rathore et al, 2012), soybean fatty acid alteration (Flores, 2008), and reduction of gliadins in bread wheat to provide a low-allergen product for those with celiac disease (Gil-Humanes et al, 2010). RNAi is also used as a plant protection strategy to provide viral, fungal, or insect resistance. Transformation of papaya with a segment of the papaya ringspot virus coat protein provided resistance to subsequent viral infection (Fitch et al, 1992), and the mechanism of resistance was investigated by Waterhouse et al, 1998. Recently, transgenic plant RNAi constructs targeting insect mRNAs have become an effective tool for managing insect damage in crop plants by silencing key genes in insect physiology (Gordon and Waterhouse, 2007).

Benefits of RNAi technology include sequence specificity, the ability to modulate gene expression in a tissue- and time-dependent manner, and RNAi's usefulness in silencing multiple genes in parallel. RNAi also results in varying levels of partial knockdown, which is useful for determining functions of essential genes, where a full knockout mutant would be lethal. Micro RNAs in plants are more specific than those in animal systems, so off-

target effects can be greatly reduced with proper target design (Schwab et al, 2005). Using tissue- and developmental stage-specific promoters allows for targeted down-regulation of genes at key points in plant development. Further, inducible promoters have been shown effective in down-regulating genes that would be lethal if constitutively reduced (Wielopolska et al, 2005). Designing siRNA targets that complement several gene family members in parallel can greatly accelerate gene function annotation, as using traditional mutagenesis methods to knock out large gene families with redundant functions would take a massive effort. Further, RNAi phenotypes are dominant, so the desired phenotype can be observed in the first generation after transformation, rather than requiring the extra breeding effort to produce homozygous knockouts needed for mutagenesis (Small, 2007).

RNAi does have limitations. A major concern of RNA interference is the possibility of off-target effects, causing down-regulation of genes not related to the gene of interest. Off-target effects can be due to homology between the siRNA target and other genes in the genome, and potential off-target effects can be predicted by computer programs if the species' genomic sequence is available (Lu and Mathews, 2008; Srivastava et al, 2014). High expression of double-stranded RNA can cause off-target effects of even loosely-related genes if the concentration of siRNA is sufficient. Further, high expression of double stranded RNA could overload the endogenous siRNA pathway machinery and cause pleitropic effects by blocking endogenous RNA interference. When analyzing transgenic events, it is important to note that the level of gene down-regulation can vary widely between transgenic events transformed with the same construct. This variation can be harnessed to obtain an allelic series of phenotypes, but may also complicate analyses.

Further, RNAi efficiency can be reduced in subsequent generation by low heritability of transgenes or silencing of the transgene itself (Small, 2007). Because of the known limitations of RNA interference, it is important to carefully design the construct to reduce off-target effects, use as weak a promoter as possible to still observe a phenotype, and obtain many transformants to be able to select a specific, efficiently down-regulated event for future experiments.

To date, no successful RNA interference of the soybean RS2 gene has been reported. This research represents the first successful transgenic alteration of soybean raffinose content.

Materials and Methods

Vector design and construction

An RNA interference vector was constructed with two 300bp inverted repeats specific to the soybean raffinose synthase gene interrupted by an intron for optimal processing of the double-stranded hairpin DNA structure (Figure 2.1). The 300bp fragments were synthesized by Genscript and cloned into the vector pMU103 using the AvrII/AscI sites for the forward orientation repeat, and SpeI/SacI sites for the reverse orientation repeat (Figure 2.2a). pMU103 carries the glycinin promoter from soybean for seed-specific expression, an intron from the rice waxy-alpha gene, and the octopine synthase terminator from *Agrobacterium.* The entire fragment of promoter-repeat-intron-repeat-terminator was then cloned into pMU2T-bar via the NcoI/SpeI sites, producing the vector pMU2T-bar-RS2 (Figure 2.2b). pMU2T-bar carries the aadA gene for spectinomycin and streptomycin bacterial resistance and the *bar* gene for glufosinate resistance in transformed plants. The *bar* gene is driven by the 2XCaMV35S promoter for enhanced expression of the herbicide selection gene.

Important features of the pMU2T-bar vector are the two independent T-DNA cassettes. The two T-DNA nature of the vector allows for independent integration of the bar selectable marker and the RS2 transgene, and thus subsequent segregation of the marker gene and transgene (Xing et al, 2000). Such marker-free plants are desirable for commercial applications as they reduce the complexity of the transgene cassette, allow the use of alternative herbicide resistance in resulting crop plants, and permit the re-transformation of progeny plants using the same selectable marker.

Integrity of the vector was confirmed by restriction digestion (Figure 2.3) and sequencing, then transferred to *Agrobacterium tumefaciens* strain AGL1 via freeze-thaw transformation. AGL1 has native rifampicin resistance. After confirmation of the vector in *Agrobacterium*, glycerol stocks were maintained at -80C. An empty vector control, pMU2T-bar, was also transferred to *Agrobacterium*.

Soybean transformation

Agrobacterium-mediated soybean transformation followed our laboratory standard protocol, a modified version of Zhang et al, 1999. All media recipes are described in Table 2.1. Seeds were sterilized with chlorine gas for 16 hours and germinated on B5 media for five days. AGL1 *Agrobacterium* cultures were streaked on YEP plates supplemented with 50mg/L rifampicin, 100mg/L spectinomycin, and 100mg/L streptomycin for two days. 2mL cultures of *Agrobacterium* were shaken for eight hours, and 100uL of the starter culture was inoculated into 200mL of YEP+antibiotics for overnight shaking at 250rpm at 28C.

On the day of inoculation, *Agrobacterium* was grown until it reached an optical density at 650nm of 0.8 to 1.0. 50mL quantities of *Agrobacterium* culture were pelleted at 3500rpm for 10 minutes at room temperature. The bacterial pellet was resuspended in about 40mL of liquid cocultivation medium to a final OD650 of 0.6. Soybean cotyledons were excised from the hypocotyl, and the cotyledonary node was wounded with a scalpel dipped in the *Agrobacterium* resuspension. Explants were inoculated in the resuspension for 30 minutes,

and then placed on solid cocultivation media overlayed with sterile filter paper. The cocultivation plates were sealed with parafilm and explants were cocultivated with *Agrobacterium* for five days at 28 C in 16 hr day/8 hr night growth chamber conditions.

After five days of cocultivation, the explants were briefly washed with liquid shoot induction media and then stuck into solid shoot induction media, with the wounded cotyledonary node inserted into the media at a 45-degree angle. No herbicide selection was used at this stage. 14 days after initial shoot induction, the explants were transferred to fresh plate of solid shoot induction media, supplemented with 10mg/L of glufosinate. After 14 days of second shoot induction, the cotyledon was removed away from the regenerated shoot mass. The shoot mass was transferred to shoot elongation media amended with 3.5 or 4 mg/L glufosinate for up to 16 weeks, with subculture every two weeks. Resulting shoots over 4cm were excised and the cut edge was incubated in 1mg/mL IBA for 1 to 4 minutes, then inserted into rooting media. Shoots that developed sufficient roots were transferred to soil in a small Jiffy® pot, and acclimated in a growth chamber. After hardening in the growth chamber, putative transgenic plants were transferred to a greenhouse and allowed to mature and set seed.

Leaf-paint analysis of putative transgenic plants

After the soil-planted events had acclimated to growth chamber conditions for about one week, a solution of 100mg/L glufosinate (commercial name Liberty®) was painted on three individual leaves of different trifoliates using a cotton swab. Paintings were scored three days later for evidence of necrosis, and events were determined to be non-chimeric if each

of the three glufosinate leaf-paintings caused no necrosis of the leaf (Figure 2.4). Events which were resistant to glufosinate at least two of three paintings were transferred to the greenhouse after acclimation.

PCR analysis of putative transgenic plants

Primers were designed that flanked from the rice waxy-α intron to OCS terminator for detection of the gene of interest, and within the *bar* gene for detection of the selectable marker gene (table 2.2). Leaf hole-punches of putative transgenic plants were taken after events were transferred to greenhouse. DNA was extracted using an SDS/EDTA buffer, purified with chloroform/isoamyl-alcohol, and washed with isopropanol and ethanol. DNA was resuspended in 50uL of distilled deionized H_2O and 1uL of the resuspension was used for PCR. PCR was done using BullsEye Taq Polymerase and the following reaction conditions: 95°C for 30 seconds; followed by 35 cycles of denaturing 95°C for 10 seconds, annealing 55°C for 10 seconds, and extension 72°C for 1 minute; then 72°C for 10 minutes for final extension. PCR products were analyzed on agarose gel and events were considered transgenic if they displayed a 500bp band for the GOI and an 800bp band for the *bar* gene.

Progeny segregation of subsequent generations

To determine segregation of the gene of interest and selectable marker gene, at least 20 seeds from each T0 event were planted and screened using PCR and leaf-paint analysis. The leaf-paint and gene of interest PCR were carried out as in the T0 generation, but new PCR primers were designed for the *bar* gene that flanked from the CaMV35S promoter to the *bar* coding sequence. This new primer pair design was necessary to prevent false

positive amplification for *bar* gene from contaminant. T0 events whose progeny plants all showed PCR negative were determined to fail to pass the gene to the progeny. T2 progeny from the T1 generation were similarly analyzed to identify homozygous T1 lines. These lines were used for future studies.

RNA extraction of mid-mature seeds

To determine the relative expression of endogenous RS2 in wild-type versus transgenic developing seeds, qRT-PCR was employed. Plants were grown in the greenhouse, and seeds were harvested when green, but had reached maximum size (stage four of Figure 2.5). This time point roughly corresponds to the midway point of seed development and the maximum expression of the endogenous RS2 gene (Dierking and Bilyeu, 2008). For each experiment, nine seeds were analyzed from each plant. For all samples, the seeds were ground with liquid nitrogen and RNA was extracted with Trizol reagent. RNA was purified with phenol:chloroform:isoamyl-alcohol (24:1:1 v/v) and resuspended overnight in 50uL of ddH_2O at 4C. The DNA was then degraded by DNaseI treatment (Promega) and RNA was stored at -80C.

For T1 seed analysis, about 20 seeds from each plant were first seed chipped by removing about 1/6 of the seed and using that fragment for DNA extraction and PCR analysis as described above. For seed chips that were positive for the RS2 transgene, RNA was extracted from the remainder of the seed.

For the T3 greenhouse experiment, three mid-mature seeds were sampled from each of the top, middle, and bottom of the plant canopy. Seeds from each plant were pooled and flash frozen in liquid nitrogen, then stored at -80C until ready for RNA extraction. The pooled seeds were ground with liquid nitrogen and total RNA was extracted.

For T3 seeds grown in the field, nine mid-mature seeds were collected from each of ten field-grown plants of each wild-type Maverick and RS2 transgenic. To account for variability across the field plot, plants were sampled randomly across the plot space and the seeds for each genotype were pooled. A subsample of ten seeds was used for each of three RNA extractions for each genotype. Total RNA was extracted using Trizol as before, cDNA was reverse-transcribed, and three technical replicates were performed on a 96-well plate for qPCR.

Quantitative reverse-transcription PCR of mid-mature seeds

Random hexamers were annealed to DNase treated cDNA for reverse transcription by M-MLV RT (Promega) at 37C for one hour. After reverse transcription, cDNA was diluted 1:3 with ddH_2O and stored at -20C. 1uL of diluted cDNA was used for quantitative PCR in a BIO-RAD CFX96 Real Time System machine. qRT-PCR was performed with 5uL SsoAdvanced Universal SYBR Green Supermix (BioRad) in a 10uL total reaction volume. Forward and reverse primer sequences for endogenous RS2 are from Dierking and Bilyeu 2008. Elongation factor 1a was used and an internal control. Three technical replicates were used and three biological replicates for each study were run. The qPCR reaction conditions were: 95C for 3 minutes, followed by 40 cycles of: 95C for 10 seconds and 55C

for 30 seconds, then a melt curve was analyzed from 65C to 95C with measurements taken every 0.5 degrees. Each cDNA sample was triplicated on the 96-well plate to measure technical variability. A no-RT sample was included to confirm the lack of genomic DNA contamination.

Genomic DNA extraction and Southern blot

Genomic DNA (gDNA) was extracted from five small fresh leaves of each event. DNA was extracted with CTAB buffer and purified with 24:1 (v/v) chloroform:isoamyl alcohol. DNA was resuspended in TE buffer and treated with RNase to remove RNA contamination.

For Southern blot, 40ug of gDNA was digested with PstI (single cut reaction), and NcoI and PstI (double cut reaction) in 300uL total volume overnight at 37C. Digested DNA was ethanol precipitated and 30ug DNA was loaded onto 1% agarose gel for overnight separation at 20v. The next day, the gel was soaked in 0.2N HCl to depurinate DNA and improve blotting efficiency of long fragments. The gel was then denatured with 0.5N NaOH/1N NaCl and neutralized with 0.5M Tris-HCl/3M NaCl (ph7.4). Digested, separated, denatured DNA was blotted onto Amersham Hybond-XL membrane (GE LifeSciences) overnight.

The membrane was pre-hybridized for three hours at 65C in hybridization solution (6X SSC; 1% SDS; 0.04M $NaPO_4$; 1X Denhardt Solution; 1mg/mL Herring Sperm DNA). Radioactively labelled probe was prepared using Prime-It II Random Primer Labelling Kit

(Agilent Technologies) and P-32 labelled dATP. The probe used for both blots was the 300bp RS2 inverted repeat from pMU2T-bar-RS2. After radio-labelling, the probe was hybridized to the membrane for 16 hours at 65C. The next day, the membrane was washed and exposed onto x-ray film for five to six days. Film was developed in a dark room in the Bond Life Sciences Center.

Height experiment

Twenty T1 seeds from each T0 plant were germinated in flats of Promix soil in the greenhouse and watered as needed. Plant height was measured every three days for the first 30 days of development, and the height was recorded from the top node of each plant.

HPLC analysis of carbohydrate profiles

For the T1 generation, five seeds from each line were lyophilized and then pooled and ground with liquid nitrogen. 12.5mg of ground seed was mixed with 1mL of 50% ethanol and incubated at 70C for 30 minutes with occasional agitation. The samples were centrifuged to remove solids and 700uL of this extraction liquid was stored at 4C until HPLC analysis. For HPLC, 50uL of the extraction liquid was speed-vacuumed to remove ethanol and resuspended in 250uL ddH_2O. 10uL of the resuspension was used for injection in the HPLC machine. The column used was Dionex CarboPac PA10 4 X 250mm, and the guard column was Dionex CarboPac PA10 4 X 50mm. Results for each oligosaccharide were given as percent of total carbohydrate fraction by weight. Sucrose was evaluated using the Megazyme Sucrose/D-Glucose Assay Kit (Megazyme Co, Ireland) which uses a glucose oxidase and peroxidase method to detect sucrose and glucose content in foodstuffs.

Because soybeans have a very low glucose content, the assay can accurately measure seed sucrose content (Hou et al, 2009).

For the T3 generation, fifteen seeds from field-grown transgenic (MM3-2-15) and wild-type (Maverick) were pooled, ground, extracted, and HPLC analyzed as before. Each extraction sample was run on the HPLC machine twice. Two replicates of the Megazyme assay were also run to determine sucrose content.

Yield evaluation of T2 plants under greenhouse conditions

Thirty plants of each of three genotypes were planted on December 24, 2014 in the Sears Greenhouse. The genotypes planted were: MM3-2-15 (transgenic, RS2 RNAi), MM3-2-10 (transgenic, bar gene only), and Maverick (wild-type). Seeds were first sown in small five inch pots, then seedlings were transferred to large three gallon pots one week later. Large pots were arranged on a greenhouse bench in a randomized design (Figure 2.6). Plants were watered daily and fertilized per advisement of the greenhouse technician. Light conditions in the greenhouse were 16-hour day and 8-hour night for the duration of the growing period. When pods had dried, watering was withheld and plants were allowed to dry down. Seeds were harvested between August 8 and August 12, 2015. Total pods were harvested from each plant and pod and numbers and total seed weight were recorded.

Germination test under cold conditions

The germination test followed the protocol for the "Iowa State cold test", with minor adjustments (Iowa State University, 2015). 50 seeds of each genotype (transgenic RS2 and

wild-type) were sown in flats filled with saturated Promix that had been pre-chilled to the appropriate temperature. Seeds were incubated in continuous dark for seven days at 10° C, and then transferred to the greenhouse to analyze germination rate. Seeds were considered germinated if, after 5 days, the cotyledons had extended out of the soil. The experiment was replicated three times and results are given as an average of the mean germination rate.

Field increase of transgenic seed

USDA permit #15-154-102n was acknowledged to plant 240 seeds from each of transgenic RS2 (MM3-2-15) and wild-type at the University of Missouri Bradford Farm. Seeds were planted in June with 70 seeds in each of eight rows (30 inch spacing between rows and six inch spacing between seeds) (Figures 2.7a and 2.7b). Emergence rate was recorded 13 days after seed planting. All weeds were removed manually between rows and in the 20-foot border zone surrounding the plot. Plants were grown to physiological maturity and dry seeds were harvested manually in mid-October (Figure 2.7d). Seeds were triple-contained and transported back to campus for threshing, then stored in the Sears Greenhouse cold storage room.

Seed composition proximate analysis

To assess if alterations in carbohydrate content affected protein, oil, or fiber composition, ground seed was sent to the MU Agriculture Experiment Station Chemical Laboratory for analysis. Amino acid analysis was performed using AOAC Official Method 982.30, crude protein was measured by the Kjeldahl method, AOAC Official Method 984.13, crude fiber

was measured by AOAC Official Method 978.10, crude fat was measured by Ether Extraction, AOAC Official Method 920.39, ash was measured by AOAC Official Method 942.05, and moisture was measured by AOAC Official Method 934.01, 2006.

Results

Transformation

Using the soybean transformation protocol, 81 plantlets were recovered from various shoot elongation stages (Table 2.3). Of those 81, 63 plantlets were from the pMU2T-bar-RS2 plasmid and 18 shoots from pMU2T-bar empty vector control. Of the 63 shoots from pMU2T-bar-RS2, 35 survived through acclimatization stage and 26 of them were resistant to glufosinate, for an escape rate of 25.7%. Of the 15 plantlets for pMU2T-bar, 12 were resistant to glufosinate, for an escape rate of 20%. The overall transformation escape rate was 24%, which is consistent with typical soybean transformation escape rates (Zeng et al., 2004). All herbicide resistant plantlets were then transplanted to the greenhouse.

PCR of *bar* and *RS2* transgenes

Leaf-paint is useful for quickly determining presence of the *bar* selectable marker. However, because the two T-DNA regions within the vector could be separately integrated into the genome, it is possible for an event to carry *bar* gene alone and be resistant to glufosinate but not positive for the *RS2* transgene. Thus, PCR was done to confirm the presence of the *bar* and *RS2* transgenes, separately. Figure 2.8 is a representative gel from PCR screening. Of 26 glufosinate-resistant pMU2T-bar-RS2 plants sampled, 24 were positive for the *bar* transgene, indicating that two plants were unlikely real transgenic. Of 24 bar-positive *RS2*-transformed plants, only 9 also carried the *RS2* transgene, for a co-integration rate of 37.5%. Of 13 pMU2T-bar glufosinate resistant plants sampled, 12 were positive for the *bar* gene, indicating that one leaf-paint result was false-positive. No pMU2T-bar plants were positive for the *RS2* transgene, as expected.

Transgenic plant growth and development

Because transgenic plants can often have developmental delays and stunted phenotypes (Elmore et al, 2001), a growth assay of plant height for the first thirty days of development was conducted. The results are summarized in Figure 2.9. At the T1 generation, a stunting of plant growth for at least one measurement was observed for lines MM2-4, MM3-2, MM4-1, MM4-5, ZM1-1, ZM2-5, ZM2-6, and ZM4-1. Lines MM4-2 and ZM3-7 were not significantly different in height than wild-type at any point in early development. Visually, plants from the subsequent T2 and T3 generations had no obvious stunting (data not shown), so it is likely that the developmental delays exhibited in the T1 generation were due to seed quality from T0 plants.

Progeny segregation

Progeny segregation of T1 plants

T1 seeds collected off the original nine T0 plants that were positive for both the *bar* and *RS2* transgenes were planted in greenhouse flats and analyzed for resistance to glufosinate leaf-painting and PCR screen for presence of the *bar* and *RS2* transgenes. The results of the leaf-painting and PCR screen are summarized in Table 2.4. Of the nine lines sampled, only four were able to pass the *bar* and *RS2* transgenes to the progeny, for an inheritance rate of 44.4%. All T0 lines that carried the *bar* gene to the next generation also carried the *RS2* transgene and vice-versa; however, not all individual progeny plants carried both the *bar* and *RS2* transgenes. Lines MM2-4, MM4-5, ZM2-5, ZM2-6 and ZM3-4 did not pass the transgenes to the progeny.

Because the bar and RS2 transgenes should integrate separately into the genome, T1 progeny analysis could reveal genetic linkage of the transgenes. However, some T0 events had higher rate of transgene inheritance than others, complicating the analysis. A chi-square test was done to test for the likelihood of the model fitting a 9:3:3:1 [bar(+)/RS2(+);bar(+)/RS2(-);bar(-)/RS2(+):bar(-)/RS2(-)] distribution, indicating the bar and RS2 transgenes are inherited independently. Results are summarized in Table 2.5. For the MM3-2 line, a 13:5:0:1 ratio was observed and the chi-square value is 4.7, and the hypothesis of a 9:3:3:1 distribution is *not* rejected at p=0.05 with three degrees of freedom. For the line ZM3-7, a 3:5:6:3 ratio was observed, with a chi-square value of 11.55, and the hypothesis of a 9:3:3:1 distribution is rejected. For the line ZM4-1, the distribution of 6:1:1:5 has a chi-square value of 23.51, and the 9:3:3:1 hypothesis is rejected. For ZM1-1, the distribution of 2:3:0:15 has a chi-square value of 162, and the hypothesis of a 9:3:3:1 is rejected. The overall segregation ratio was: 24:14:7:24, with a chi-square value of 98.34, and the hypothesis of a normal 9:3:3:1 distribution for the bar and RS2 transgenes segregating independently is rejected at p=0.05 and 3 degrees of freedom.

Because chi-square values for ZM3-7, ZM4-1, and ZM1-1 do not fit the models for normal 9:3:3:1 distributions of bar and RS2 transgenes, and only seven bar(-)/RS2(+) plants were identified in all 185 plants of all lines sampled, it is assumed that the transgene segregation for bar and RS2 at the T1 generation does not follow Mendelian patterns. Progeny segregation to determine genetic linkage of the two T-DNAs was done on T2 seeds from

T1 plants that carried both the bar and RS2 transgenes, when it was assumed that transgenes were stably inherited.

Progeny segregation to estimate locus number

If the T0 event is hemizygous for the single transgene insertion, then T1 plants should have an expected genetic ratio for each transgene locus of 1:2:1 and phenotypic ratio of 3:1. Selfed T1 plants should produce an offspring pool with a 3:2:3 genotypic ratio (Figure 2.10) and 5:3 phenotypic ratio. So, if T2 progeny have a segregation ratio of about 5:3 for each transgene, it can be inferred that original T0 plant had only one integration event of each transgene, or that only one locus of multiple integration events is expressed. Significant deviations from the 5:3 ration indicate multiple inserts of the transgene were integrated.

However, for lines ZM1-1 and ZM3-7, because T0 to T1 inheritance of the transgenes was low, only progeny from T1 plants that expressed the bar gene and carried the RS2 transgene were used. Thus, homozygous-transgene-negative T2 progeny were omitted from the population by artificial selection (blue line on Figure 2.10), and a 5:1 phenotypic ratio of transgene(+):transgene (-) is expected for single-insertions at the T2 generation.

Additionally, for line ZM1-1, only two T1 plants inherited any transgene – and both lines inherited both the bar and RS2 transgenes. T2 segregation for these events showed neither event was homozygous positive for either transgene, so for ZM1-1, the progeny denoted by the yellow line on Figure 2.10 can also be omitted from the analysis. Thus, for ZM1-1,

a 3:1 ratio of transgene(+):transgene(-) is expected for single-locus insertions. Results are summarized in Table 2.7.

For ZM1-1, 31 T2 progeny were resistant to glufosinate and 6 were susceptible. Testing for a 3:1 segregation ratio, the ratio of 31:6 has a chi-square value of 1.523 and one degree of freedom, and the hypothesis of a 3:1 ratio is not rejected. For the RS2 transgene, 24 ZM1-1 progeny plants were positive and 13 were negative for the presence of the transgene. This 24:13 ratio has a chi-square value of 2.027 and the 3:1 hypothesis is not rejected. For ZM3-7, the 41:9 ratio of bar(+):bar(-) has a chi-square value of 0.064 and the hypothesis of a 5:1 ratio is not rejected. The 36:12 ratio of RS2(+):RS2(-) has a chi-square value of 2.4, and the hypothesis of a 5:1 ratio is also not rejected.

Testing for 5:3 segregation of line MM3-2, the ratio of bar(+):bar(-) is 138:73, which has a chi-square value of 0.759 and the 5:3 hypothesis is not rejected. For the RS2 transgene, 86 progeny plants were positive and 39 were negative. This 86:39 ratio has a chi-square value of 2.117 and the hypothesis of a 5:3 model is not rejected. For ZM4-1, the 21:15 ratio of bar(+):bar(-) has a chi-square value of 0.267, and the hypothesis of a 5:3 ratio is not rejected. The 11:5 ratio of RS2(+):RS2(-) also has a chi-square value of 0.267, and the 5:3 hypothesis is also not rejected.

For no transgene from any event was there a significant deviation from the expected phenotypic ratio at the T2 generation. Thus, for all lines, it appears that each event has a

single insertion locus of both the bar and RS2 transgenes, or that only one copy of the bar gene is expressed.

Progeny segregation of T2 plants to estimate genetic linkage of transgenes

From the same data set as the transgene copy number estimations, genetic linkage of the transgenes can be estimated. Because T1 inheritance of the transgenes did not appear stable, the T2 generation was used for linkage analysis. T2 seeds from original T0 lines ZM1-1, MM3-2, ZM3-7, and ZM4-1 were planted and analyzed as before with leaf-paint and PCR. The parental phenotype is bar(+)/RS2(+) or bar(-)/RS2(-), and the recombinant phenotype is bar(+)/RS2(-) or bar(-)/RS2(+). Results of the experiments are summarized in Table 2.6.

For line ZM1-1, only two T1 plants inherited any transgene – and both lines inherited both the bar and RS2 transgenes. Because of the low T0 to T1 inheritance of the transgenes, only these two T1 lines were used for further experiments. Twenty T2 seeds from each line, ZM1-1-1 and ZM1-1-6, were planted and analyzed with leaf-paint and PCR. Of the 37 surviving T2 plants from both lines combined, 9 had the recombinant genotype of bar(+)/RS2(-) or bar(-)/RS2(+) and 28 had the parental genotype, giving a 24.3% recombination frequency. Because T1 plants were artificially selected against the bar(-)/RS2(-) phenotype, and inheritance of the RS2 transgene is low, exact genetic linkage of the transgenes would be difficult to estimate. However, there is some degree of recombination, so it can be stated that the bar and RS2 transgenes are not immediately linked.

Line ZM3-7 was also artificially selected against the transgene-negative phenotype. Eight T1 lines of this event were planted out, of which all carried both transgenes. At the T2 generation, 15 out of 48 plants had the recombinant genotype, for a recombination frequency of 31.25%. Again, because of selection against transgene-null plants, direct genetic linkage cannot be established, but some degree of recombination indicates the bar and RS2 transgenes are not immediately linked.

For line MM3-2, seeds from 18 T1 plants were planted out. For this line, which appeared to show a normal 3:1 segregation ratio for both transgenes at the T1 generation, a randomly selected sample of all T1 lines were planted out. Thus, estimation of genetic linkage between the transgenes could be estimated from the data. To estimate genetic linkage of the original T0 events, the formula $x=2(r-r^2)$ was used, where r is the estimated T1 recombination frequency and x is the observed recombination frequency at T2 (Figure 2.11). Although the estimation may not be exact due to non-Mendelian inheritance of transgenes, it allows for an estimation of degree of recombination and genetic linkage. MM3-2 had an estimated T1 recombination frequency of 42.93%, indicating the genes are not linked.

Progenies from six randomly selected T1 plants of line ZM4-1 were analyzed, and the observed recombination frequency was 25%. Based on the estimation formula, the estimated frequency of recombination at T1 is 14.65%, showing the two transgenes are linked. For all events, there was some degree of recombination between the bar and RS2

transgenes, showing the two T-DNAs indeed integrated separately into the soybean genome.

The T2 progeny segregation also allowed for selection of a homozygous bar(-)/RS2(+) T1 line – carrying the RS2 transgene but lacking the selectable marker – which is commercially valuable. Of the 17 T2 progeny leaf-painted from the T1 line MM3-2-15, none were resistant, and 12 out of 12 PCR screened were positive for the RS2 transgene. The line MM3-2-10, which shared a T0 parent with MM3-2-15 was found to be resistant to glufosinate in 9/9 plants and RS2(-) for 4/4 plants. Based on these results and confirmed in T3 analysis, the line MM3-2-15 was chosen as the representative homozygous bar(-)/RS2(+) line and MM3-2-10 was chosen as the representative homozygous bar(+)/RS2(-) line for use in future experiments.

Southern blot

Two radiolabelled ^{32}P Southern blots were done on genomic DNA to confirm transgene integration into the soybean genome. Six T1 plants derived from individual T0 events were analyzed, five putative pMU2T-bar-RS2 transgenic events and one pMU2T-bar event. The samples were digested with PstI to cut the T-DNA once (single cut), and NcoI and PstI to flank the first T-DNA region (double cut). For the double cut, the expected fragment size was 3199bp. From the double cut blot, four of the five T0 events were confirmed as having the 3199bp transgenic RS2 construct (Figure 2.12a). Thus, from this Southern blot, it can be concluded that events ZM1-1, MM3-2, ZM3-7, and ZM4-1 are indeed transgenic and the gene is stably inherited to at least the T1 progeny. The event ZM3-4 was initially

identified as transgenic by PCR, but subsequent PCR screening and Southern blot analysis revealed the event was not transgenic. The event MM4-1 was transformed with pMU2T-bar, so it did not have the transgenic RS2 construct as expected.

In the double cut blot, the wild-type gDNA + 200pg plamid DNA lane had two additional bands that could not be attributed to the transgenic or endogenous RS2. The larger band at about 20kbp is likely due to incompletely digested plasmid or genomic DNA. The band at about 3.6kbp may be due to incomplete digestion of the PstI site at 8776bp on the pMU2T-bar-RS2 plasmid, causing the fragment to be lengthened to the second PstI site at 8385. This would give a band at 3590bp. Alternatively, the extra band may be due to supercoiled genomic DNA.

Because the probe used was homologous to the endogenous RS2 gene, all events, including the non-transgenic ZM3-4 and empty vector control MM4-1 had a band at about 6kb in the double cut blot. Based on the soybean reference genome (Phytozome v.11), the probe hybridizes to chromosome 6 at position 15223403 to 15223704 on the plus strand, which is the location of Glyma.06G179200, the RS2 gene. The enzyme NcoI cuts at positions Chr06:15223711..15223716 and Chr06:15217725..15217730 on the minus strand, giving a fragment of 5986bp. This roughly 6kb fragment from the endogenous RS2 gene is seen in all genomic DNA samples.

The results from the single cut Southern were not completely clear (Figure 2.12b). The expected size fragment for genomic region around the endogenous RS2 gene digested by

PstI is ~24kb, and a large band was observed in all samples. It appears that for ZM1-1, MM3-2, and ZM4-1 there is only one additional band, meaning there may be only one insertion site in those events. To more accurately estimate insertion site number, T2 progeny segregation results were used.

Real-time qPCR

T1 seeds from T0 plants

Mid-mature T1 seeds collected from six T0 plants were first confirmed to be transgenic by PCR screening, then mRNA levels of endogenous RS2 were evaluated by RNA extraction and qRT-PCR. Because the progeny segregation experiments showed that lines MM4-5, ZM2-5, ZM2-6 and ZM3-4 did not pass the transgenes to the progeny, but the seed-chipped DNA did possess transgenic DNA, it is likely that the PCR screening was contaminated by maternal DNA from the seed coat. However, all T0 lines had some degree of down-regulation of the endogenous RS2 gene. The results are summarized in Figure 2.13a. At the T1 generation, the level of mRNA down-regulation ranged from 34% (MM4-5) to 71% (MM3-2 and ZM3-4).

T3 seeds from T2 plants grown in the greenhouse

By the T2 generation, homozygous lines MM3-2-15 (transgenic RS2 homozygous) and MM3-2-10 (bar homozygous) were identified and grown for the yield test in the greenhouse. Mid-mature seeds were harvested from three plants and total RNA was extracted for RT-qPCR. The results of the experiment are summarized in Figure 2.13b. Seeds from MM3-2-10 plants had a non-significant 26% reduction in endogenous RS2

mRNA levels (p=0.084). Seeds from MM3-2-15 plants had a significant 80% reduction in RS2 mRNA levels, which is consistent with the level of downregulation (71%) for T1 seeds.

T3 seeds from T2 plants grown in the field

Nine mid-mature seeds were collected from each of ten field-grown plants of each wild-type Maverick and RS2 transgenic and pooled by genotype to account for variability within the plot space. Three RNA extractions from subsamples of 10 seeds for each genotype were performed. Figure 2.14 summarizes the data, with each bar on the graph representing a 10-seed subsample of the total genotype's sampled population. The pooling and subsampling method greatly reduced variability between plants and allowed for better quantification of RS2 down-regulation under field conditions.

Under field conditions, endogenous RS2 expression was 56% downregulated in transgenic compared to wild-type developing seeds. The reduction in downregulation from greenhouse to field conditions could be due to environmental impacts of carbohydrate synthesis in developing seeds (Thomas et al, 2003).

RT-qPCR of endogenous RS3 gene to evaluate RNAi transgene specificity

While the whole inverted repeat sequence has a low identity (58.6%) with RS3, it is possible that siRNAs generated from the inverted repeat might be partially complementary to RS3 transcripts. To evaluate the possible effect of RS3 downregulation, the cDNA samples from field-grown mid-mature seeds were analyzed (Figure 2.15). No significant

difference in RS3 mRNA levels were observed in wild-type versus transgenic plants at p=0.05. This result confirms the specificity of the RS2 inverted repeat, and also allows for observed carbohydrate phenotypes to be attributed to the contribution of RS2 and not RS3.

RT-qPCR of endogenous RS2 in different tissues

Although the glycinin promoter is known to be seed-specific (Nielsen, 1989), an experiment was conducted to evaluate RS2 down-regulation in mid-mature seed, leaf, root, and expanded cotyledon tissue for the T3 progeny of event MM3-2-15. Three samples from each tissue were obtained from: greenhouse-grown mid-mature seeds, young leaves from three-week old plants, total root mass from three-week old plants, and 7-day old expanded cotyledons. All samples were finely ground with liquid nitrogen and RNA was extracted using Trizol and the standard protocol. Resulting cDNA was pooled by genotype to obtain one sample per tissue used for RT-qPCR, and three technical replications were loaded on the plate. Developing seeds had significantly higher expression of RS2 than cotyledons, leaves, and roots, at 5.5-, 12.7-, and 20-fold, respectively. The experiment revealed a significant 71% reduction in RS2 expression for mid-mature seed as expected, and also a significant 61% reduction for expanded cotyledons (Figure 2.16). Leaves and roots did not show a significant difference in RS2 expression between wild-type and transgenic. Together these data show the specificity of the glycinin promoter for developing seed and also expanded cotyledon tissue.

HPLC analysis of carbohydrate profiles

T1 seeds from T0 plants grown in the greenhouse

Five individual seeds were analyzed from each of three lines, wild-type Maverick, empty vector control MM4-1, and RS2 transgenic MM3-2. HPLC was done on finely ground and ethanol extracted mature soybean seeds to determine the relative contents of galactinol, sucrose, raffinose, and stachyose. At this generation, the line MM3-2 was still segregating for the transgene, so one of the five seeds sampled had a carbohydrate profile similar to wild-type ("RS2-seg"). Values from each genotype were averaged and the results are presented in Figure 2.17.

No significant difference in galactinol content was observed for the three lines. Wild-type seeds had significantly more sucrose than transgenic RS2, with the average percent sucrose for wild-type, empty vector, and RS2 transgenic being 5.08, 4.55, and 3.66 respectively. Raffinose content in wild-type and empty vector was significantly higher than RS2 transgenic at 1.03, 0.92, and 0.05 percent respectively. Stachyose in wild-type and empty vector was also significantly higher than RS2 transgenic at 2.96, 2.43, and 1.37 percent respectively.

T3 seeds from T2 plants grown in the field

To more accurately measure carbohydrate content in mature soybean seeds, transgenic (MM3-2-15) and wild-type (Maverick) plants were grown side-by-side in a field plot at Bradford Research and Extension Center during the 2015 growing season. Fifteen mature seeds were collected from a bulk harvest of each genotype, ground, and analyzed with HPLC and Megazyme assay in two replications (Figure 2.18). Galactinol in transgenic seeds was significantly higher than wild-type at 0.31 and 0.03 percent, respectively

($p<0.001$). Sucrose was significantly increased in transgenic seeds compared to wild-type at 9.6 and 6.95 percent, respectively ($p=0.05$). Both raffinose and stachyose contents were significantly reduced in transgenic seeds as compared to wild-type (raffinose $p=0.0015$; stachyose $p<0.001$). Raffinose content in transgenic seeds was 0.11%, compared with 0.63% in wild-type. Stachyose in transgenic seeds was 1.21%, and 3.79% in wild-type.

Taken together with the T1 results, relative stachyose and raffinose contents in transgenic seeds remained reduced, while the sucrose was significantly increased in T3 seeds. Because the T3 experiment replicated real field conditions, and the wild-type plants had identical growing conditions as transgenic, the T3 results are likely more representative of the RS2 knock-down phenotype than T1.

Yield evaluation of T2 plants under greenhouse conditions

T3 seeds from three lines: MM3-2-15 (RS2 transgenic, bar null), MM3-2-10 (bar transgenic, RS2 null), and wild-type Maverick were sown in the greenhouse in late December 2014. Ten large pots for each genotype with three seeds each were germinated and PCR analysis confirmed that all plants had the expected genotype. After PCR analysis, pots were thinned to one plant each. Total pods were collected for each plant, and pod number, seed number, and seed weight were recorded.

Results are summarized in Figure 2.19. Seed number, total seed weight, and seeds per pod were all significantly reduced in transgenic RS2 compared to both wild-type and transgenic control plants. Pod number in transgenic RS2 was significantly also lower than transgenic

control plants. However, seeds from the RS2 transgenic plants were larger than both wild-type and transgenic control, but the result was not statistically significant (Figure 2.19b). Taken together, the greenhouse yield study showed an overall negative impact on yield for RS2 transgenic plants. The data could be complicated by the extremely long growing season in the greenhouse, with plants on the east side of the bench being closer to the maize greenhouse and receiving higher light conditions.

Germination test under cold conditions

It has been shown that inhibition of raffinose metabolism in pea leads to a decrease in germination frequency (Blöchl et al, 2007). To test if the RS2 transgene had a negative effect on seedling germination under cold-stress conditions, a modified "Iowa Cold Test" was performed (Iowa State University, Seed Testing Laboratory). After seven days of incubating seeds in moist soil at 4 degrees Celsius and then transferring the flats to a growth chamber, no significant difference in germination between wild-type and transgenic seeds was observed (Figure 2.20). Thus, the RS2 transgene and low-raffinose phenotype do not have a deleterious effect on seedling germination. This result is consistent with emergence rates under field conditions.

Field growth of transgenic seed

Transgenic (MM3-2-15) and wild-type (Maverick) plants were grown side-by-side in a field plot at Bradford Research and Extension Center during the 2015 growing season. Because of wet field conditions, planting was delayed until June 24. Emergence rate was counted on July 7, thirteen days after planting. Of the 280 seeds planted for each genotype,

167 Maverick seedlings emerged, and 159 transgenic RS2. Wet field conditions, a hard "crust" on the top layer of soil, and cool temperatures contributed to the relatively low emergence rates. Plant growth was observed throughout the season, and leaf samples were collected to confirm the absence of transgenes in wild-type and presence of RS2 transgene in MM3-2-15 plants. 10 plants for each wild-type and transgenic were leaf-painted for presence of the bar gene and, as expected, all leaves were susceptible. Mid-mature seeds, corresponding to roughly stage R5/R6 were collected on September 25 to be analyzed for RS2 and RS3 gene expression by RT-qPCR. No visible differences in plant growth or maturation were observed throughout the growing season (Figure 2.7). Seeds from mature, dry plants were harvested on October 15 and 16 and transported back to the MU campus for threshing and storage. Total seed yield for MM3-2-15 was 6.44kg and Maverick WT 5.77kg. This result contradicts the yield reduction observed in greenhouse-grown plants, however the field results are likely more representative of the true yield potential of transgenic, low-raffinose plants.

Composition analysis of transgenic and wild-type seed

To measure if the alterations in carbohydrate content affected seed fat, crude protein, or amino acid composition, samples of field-grown ground seed were sent to the MU Agriculture Experiment Station Chemical Laboratory for analysis. Results are summarized in Table 2.8. There were no major differences between the transgenic and wild-type lines in moisture content (6.13 and 6.14%, respectively), crude fat (19.18 and 19.79%, respectively), ash (5.21 and 5.27%, respectively), crude protein (37.26 and 36.14% respectively), or for any of the amino acids. Crude fiber appeared reduced in transgenic

(6.56%) compared to wild-type (7.07%), however it is unknown if the difference is significant. The reduced-oligosaccharide soybean meal used in experiments by Chen et al, 2013 also had reduced levels of fiber, so it is possible that the alterations in sucrose and oligosaccharide also affected total fiber content.

Discussion

In the current work, results for transgene co-integration, transgene inheritance, T1 seedling vigor, carbohydrate phenotype, and yield were somewhat unexpected and deserve additional examination. Of the 24 bar-positive events generated from transformation using the pMU2T-bar-RS2 plasmid, nine also had the RS2 gene of interest. This 37.5% co-integration rate is lower than the 47% rate reported for rice and tobacco in Komari et al, 1996, and the 70% rate for soybean described in Xing et al, 2000. Of the nine co-integration events, four passed both transgenes to the progeny, and marker-free progeny were observed in all four lines by at least the T3 generation. Thus, in this study, 4 out of 24 events, or 16.6% produced transgene-positive, marker-free progeny. Interestingly, Matthews et al, 2001 also reported 16% of transformed barley events producing transgene-positive, marker-free progeny using a "twin T-DNA" vector where the two T-DNA borders are immediately adjacent to one another. This information is useful to future researchers wishing to obtain marker-free transgenic events, as it shows that about 1 in 6 events produced will yield progeny with the desired genotype using the pMU2T-bar plasmid.

Transgene inheritance is known to sometimes not follow typical Mendelian segregation patterns, and several mechanisms have been proposed to explain why. Deletions, rearrangements, duplications, co-suppression, and epistatic interactions have all been observed for transgenes, and can explain their often low inheritance (Yin et al, 2004). Gamete viability could also reduce transgene inheritance if the insertion site interferes with an endogenous gene essential for pollen or ovule formation (Limanton-Grevet and Jullien, 2001). The tissue culture process itself has also been shown to cause epigenetic changes in

resulting plants, partially explaining the phenomenon of somaclonal variation (Stroud et al, 2013). Further, in several species of plants including soybean, a higher degree of recombination is noted for integration loci with more than one copy of the transgene, leading to loss of one or more transgene copies (Choffnes et al, 2001; Tizaoui and Kchouk, 2012). In this work, inheritance for the two transgenes was about 44%. This result is based on PCR and not observed phenotype, so epigenetic or expression variation cannot be a cause for the low inheritance. The current work's 44% rate of inheritance is comparable to the 10% to 50% rate of non-Mendelian inheritance for all transformed plants reviewed in Yin et al, 2004.

It has been well established that seed quality affects seedling vigor and growth (Abdul-Baki and Anderson, 1973; Parrish and Leopold, 1978), and factors such as temperature, genotype, humidity, and fungal contamination contribute to seed quality (TeKrony et al, 1984; Dornbos et al, 1989; Gibson and Mullen, 1996; Egli et al, 2005). Greenhouse conditions during summer months when the T0 plants were growing were hot, humid, and had high fungal contamination which likely contributed to the poor seed quality and seedling vigor observed for T1 seedlings. Though plant growth was stunted in T1 transgenic plants, no obvious stunting was noted in the T2 or T3 generations, thus seed quality and seedling vigor are assumed to be dependent on external factors or effects of tissue culture rather than the transgene.

For the T1 carbohydrate content results (Figure 2.17), the decrease of sucrose in transgenic RS2 seeds is unexpected given the known raffinose biosynthesis pathway. However, the

effect could be due to effects of tissue culture on the T0 plant or different growing conditions of transgenic and wild-type plants. Wild-type plants were sown in the greenhouse to be used as a control at about the same time as transgenic plants were transferred to the greenhouse, but transgenic plants often flowered sooner than their wild-type counterparts. The strong environmental influence on carbohydrate partitioning in developing seeds likely contributed much of the variability in sucrose content. To elucidate the effect of only the transgene on soybean carbohydrate content, samples were obtained from field-grown plants with wild-type controls planted at the same time in the same field. This comparison of transgenic and wild-type showed a relative increase in sucrose content (Figure 2.18). Thus, the field results are likely more indicative of the true low-raffinose, high-sucrose phenotype expected in transgenic plants, however future studies are warranted.

Overall, the result of this work is that a low-raffinose line of marker-free transgenic soybean has been developed. For commercial transgenic crops, yield parity is one of the most important traits, and though total seed yield was significantly reduced in the greenhouse yield study, yield was comparable between wild-type and transgenic plants under field conditions. Future experiments are needed to evaluate the inheritance of transgenes through several more generations, as well as the carbohydrate phenotype over generations and environmental conditions. Also, replicated, controlled experiments are needed to precisely measure agronomic traits such as emergence, stress tolerance, maturation, and yield for the transgenic plants. The line MM3-2-15 is certainly a strong candidate for use in improving soybean carbohydrate quality.

Figures

a

```
Glyma.06G179200.1       ATGTCACCACACCCTACCCAGCAAAACCAAACCATAGCAAACCTAAGCACCAAACCTCTT
RS2_IR                  ------------------------------------------------------------

Glyma.06G179200.1       TCTTTCAAGATCCTTGAATTCAGTCCCATGGCTCCAAGCATAAGCAAAACTGTGGAACTA
RS2_IR                  ------------------------------------GCATAAGCAAAACTGTGGAACTA
                                                            ***********************

Glyma.06G179200.1       AATTCATTTGGTCTTGTCAACGGTAATTTGCCTTTGTCCATAACCCTAGAAGGATCAAAT
RS2_IR                  AATTCATTTGGTCTTGTCAACGGTAATTTGCCTTTGTCCATAACCCTAGAAGGATCAAAT
                        ************************************************************

Glyma.06G179200.1       TTCCTCGCCAACGGCCACCCTTTTCTCACGGAAGTTCCCGAAAACATAATAGTCACCCCT
RS2_IR                  TTCCTCGCCAACGGCCACCCTTTTCTCACGGAAGTTCCCGAAAACATAATAGTCACCCCT
                        ************************************************************

Glyma.06G179200.1       TCACCCATCGACGCCAAGAGTAGTAAGAACAACGAGGACGACGACGTCGTAGGTTGCTTC
RS2_IR                  TCACCCATCGACGCCAAGAGTAGTAAGAACAACGAGGACGACGACGTCGTAGGTTGCTTC
                        ************************************************************

Glyma.06G179200.1       GTGGGCTTCCACGCGGACGAGCCCAGAAGCCGACACGTGGCTTCCCTGGGGAAGCTCAGA
RS2_IR                  GTGGGCTTCCACGCGGACGAGCCCAGAAGCCGACACGTGGCTTCCCTGGGGAAGCTCAGA
                        ************************************************************

Glyma.06G179200.1       GGAATAAAATTCATGAGCATATTCCGGTTTAAGGTGTGGTGGACCACTCACTGGGTCGGT
RS2_IR                  GGAATAAAATTCATGAGCATATTCCGGTTTAAGGTGT-----------------------
                        ***********************************
```

b

```
Glyma.05g003900.1       --ATGGGTCCAAGCTCGAAGAAAGCTTCACCTAAATCAGGTG-TGACAAAGCACATGAAG
RS2_IR                  GCATAAGCAAAACTGTGGAACTAAATTCATTTGGTCTTGTCAACGGTAATTTGCCTTT-G
                          **  *   **    * *   *  ****  *      *     *  **    * *   *

Glyma.05g003900.1       GGCTTCAGCCTCTGCAACTCAACCCTAAAAGTAAATGGGCAAGTCATCCTCTCCCAAGTC
RS2_IR                  TCCATAACCCTAGAAGGATCAAATTTCCTCGCCAACGGCCACCCTTTTCTCACGGAAGTT
                          * * * ***       ****   *    *  ** ** **     * *** *  ****

Glyma.05g003900.1       CCCAAGAACGTAACCCTCACCCCAT----------------------------GCACCTAC
RS2_IR                  CCCGAAAACATAATAGTCACCCCTTCACCCATCGACGCCAAGAGTAGTAAGAACAACGAG
                        *** * *** ***   ******* *                             ** * *

Glyma.05g003900.1       GACACTCACACCACCGGATGCTTCCTCGGTTTCCACGCCACCTCCCCAAAATCCCGCCAC
RS2_IR                  GACGACGACGTCGTAGGTTGCTTCGTGGGCTTCCACGCGGACGAGCCCAGAAGCCGACAC
                        ***    **  *   ** ****** * ** ********   *   ** * *  *** ***

Glyma.05g003900.1       GTGGCACCCTTAGGACAGCTTAAAAACATAAGCTTCACTTCCATCTTCCGGTTCAAGGTT
RS2_IR                  GTGGCTTCCCTGGGGAAGCTCAGAGGAATAAAATTCATGAGCATATTCCGGTTTAAGGTG
                        *****  ** * **  **** * *   ****  ****    *** ******** *****

Glyma.05g003900.1       TGGTGGACCACTCTCTGGACCGGCTCCAACGGCCGCGACCTGGAAACCGAAACCCAATTC
RS2_IR                  T-----------------------------------------------------------
                        *                                                           ?
?
```

c

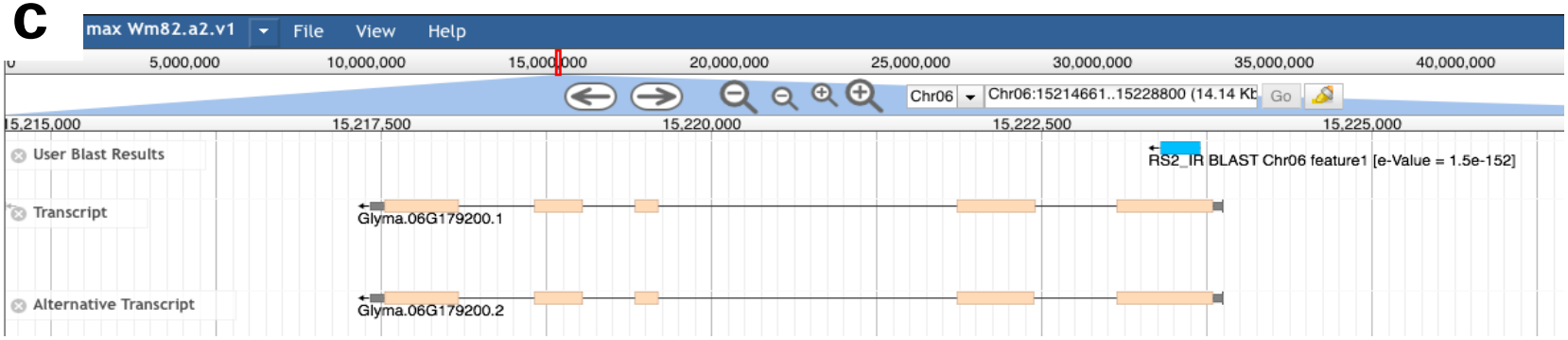


Figure 2.1
Sequence of 300 base-pair RS2 inverted repeat (RS2_IR) aligned with soybean genes (**a**) RS2 (Glyma.06G179200.1) and (**b**) RS3 (Glyma.05G003900.1) using ClustalW (http://www.genome.jp/tools/clustalw/). The inverted repeat sequence is an exact match for nucleotide 580 to 879 of RS2 genomic sequence, and nucleotide 97 to 396 of the coding sequence. Although the entire sequences of RS2 and RS3 have a high degree of identity (64%), the region of the inverted repeat sequence has low identity between RS2 and RS3 (58.6%) allowing for targeted down-regulation of only the RS2 endogenous gene. (**c**) Genomic structure of the soybean RS2 gene. Location of the 300bp RS2 inverted repeat is shown in blue in exon 1.

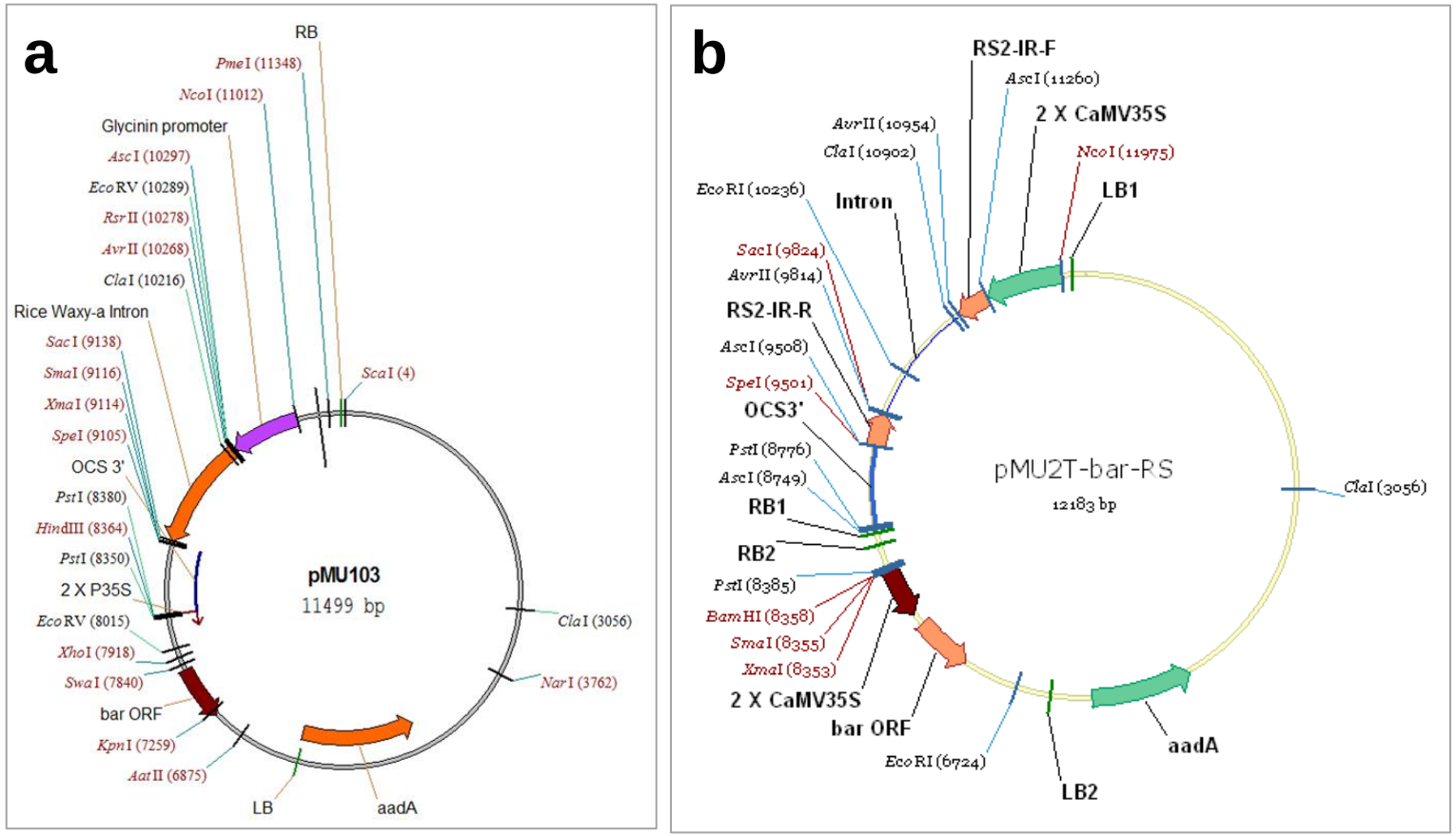


Figure 2.2
Plasmid maps used in construction and transformation.
a. pMU103 vector used for construction of the promoter-repeat-intron-repeat-terminator sequence. The synthesized RS2 inverted repeat was first inserted into the AscI/AvrII sites, and then into the SacI/SpeI sites.
b. Plasmid map of pMU2T-bar-RS2 used for soybean transformation.

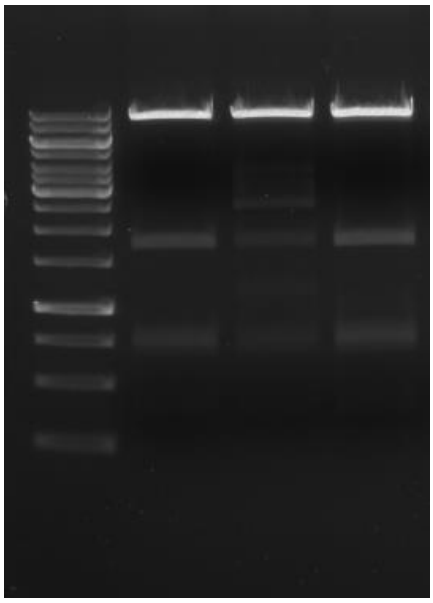


Figure 2.3
DNA digestion of pMU2T-bar-RS2 plasmid to confirm insert presence

1. 1kb ladder, Fermentas
2. pMU2T-bar-RS2 (sample A): digested with NcoI, BamHI, AscI - **correct**
3. pMU2T-bar-RS2 (sample B): digested with NcoI, BamHI, AscI - **incorrect**
4. pMU2T-bar-RS2 (sample C): digested with NcoI, BamHI, AscI - **correct**

Gel bands at 8566, 1752, 759, 715, and 391 confirmed the presence of RS2 inverted repeats, promoter, intron, and terminator.

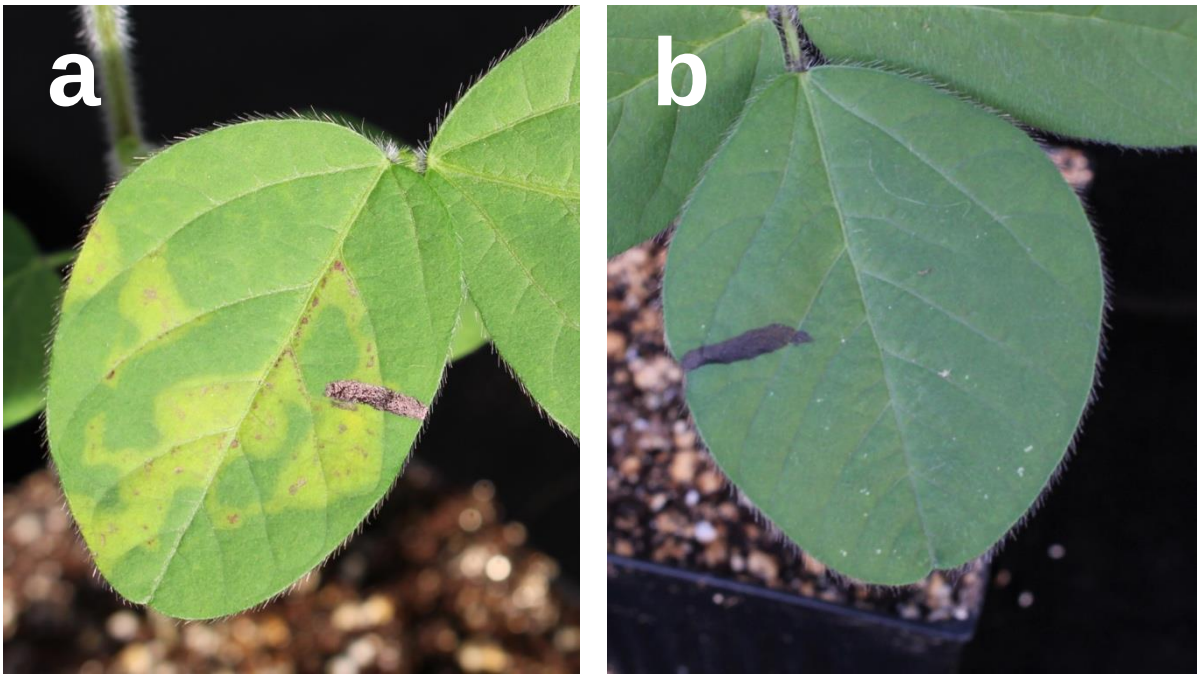


Figure 2.4
Representative result at three days post glufosinate leaf painting of (a) wild-type susceptible and (b) transgenic resistant. Susceptible leaves will eventually become necrotic and die.

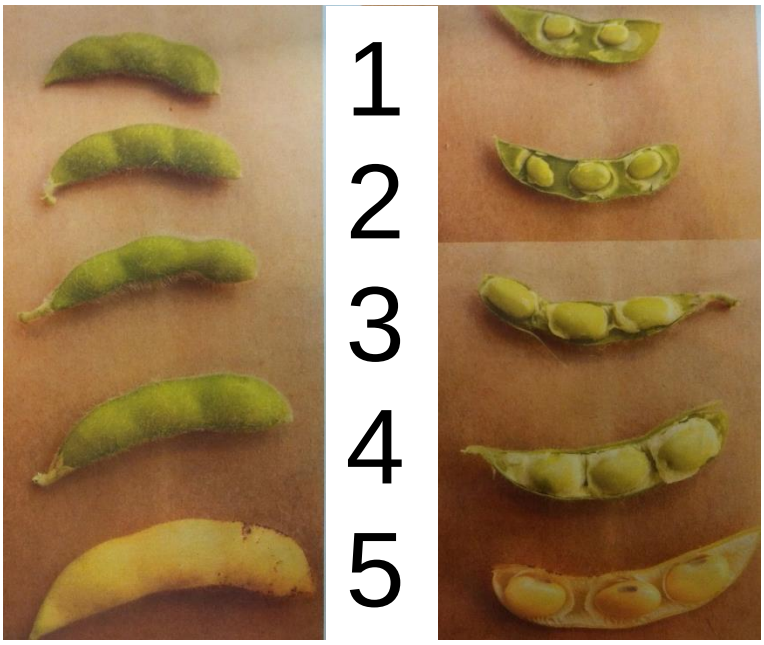


Figure 2.5
Stages of seed maturation. Mid-mature seeds were collected at stage 4, when the seed had reached maximum size but had not begun to turn a yellow color.

Corn greenhouse

MM3-2-10	Mav	MM3-2-15
Mav	MM3-2-15	MM3-2-10
MM3-2-15	Mav	MM3-2-10
MM3-2-15	MM3-2-10	Mav
MM3-2-10	MM3-2-15	Mav
MM3-2-15	Mav	MM3-2-10
Mav	MM3-2-10	MM3-2-15
MM3-2-10	MM3-2-15	Mav
MM3-2-15	Mav	MM3-2-10
MM3-2-10	Mav	MM3-2-15

Walkway

Figure 2.6
Layout of T3 greenhouse yield study. Plants were arranged in a randomized design on a single greenhouse bench.

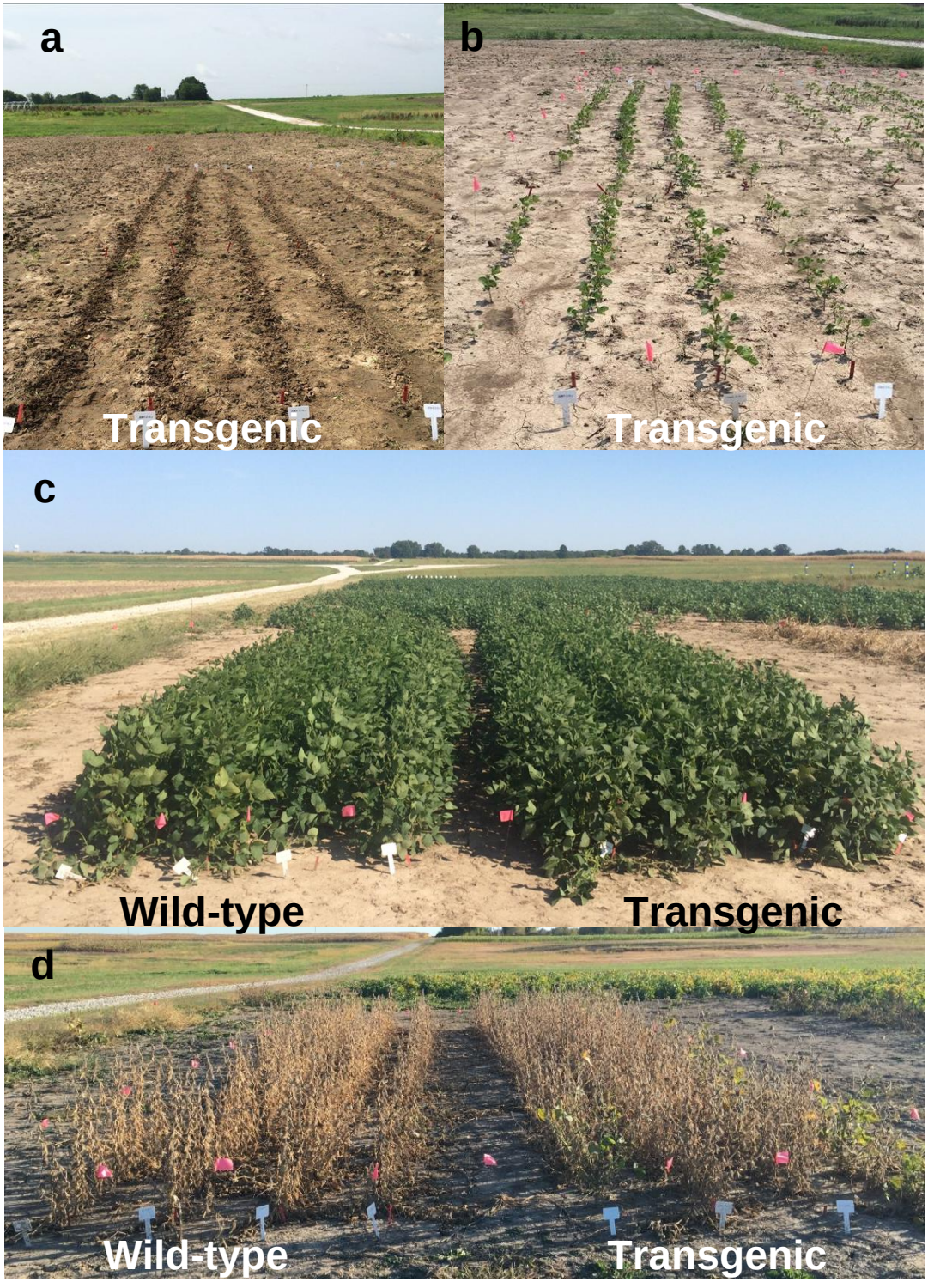


Figure 2.7
Field conditions on (**a**) June 24 (planting date), (**b**) July 23, (**c**) September 25, and (**d**) October 14, 2015. Four rows of 70 seeds for each transgenic and wild-type seeds were sown in the field on June 24.

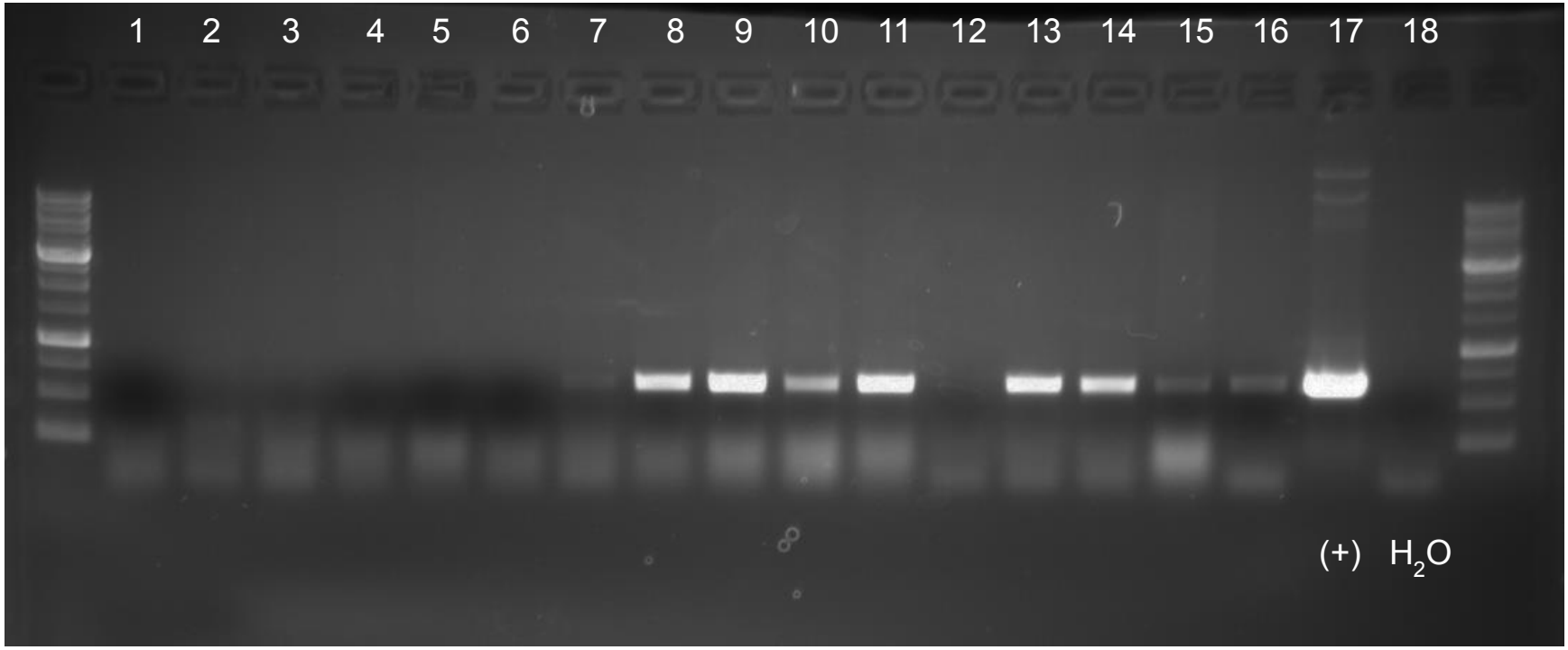
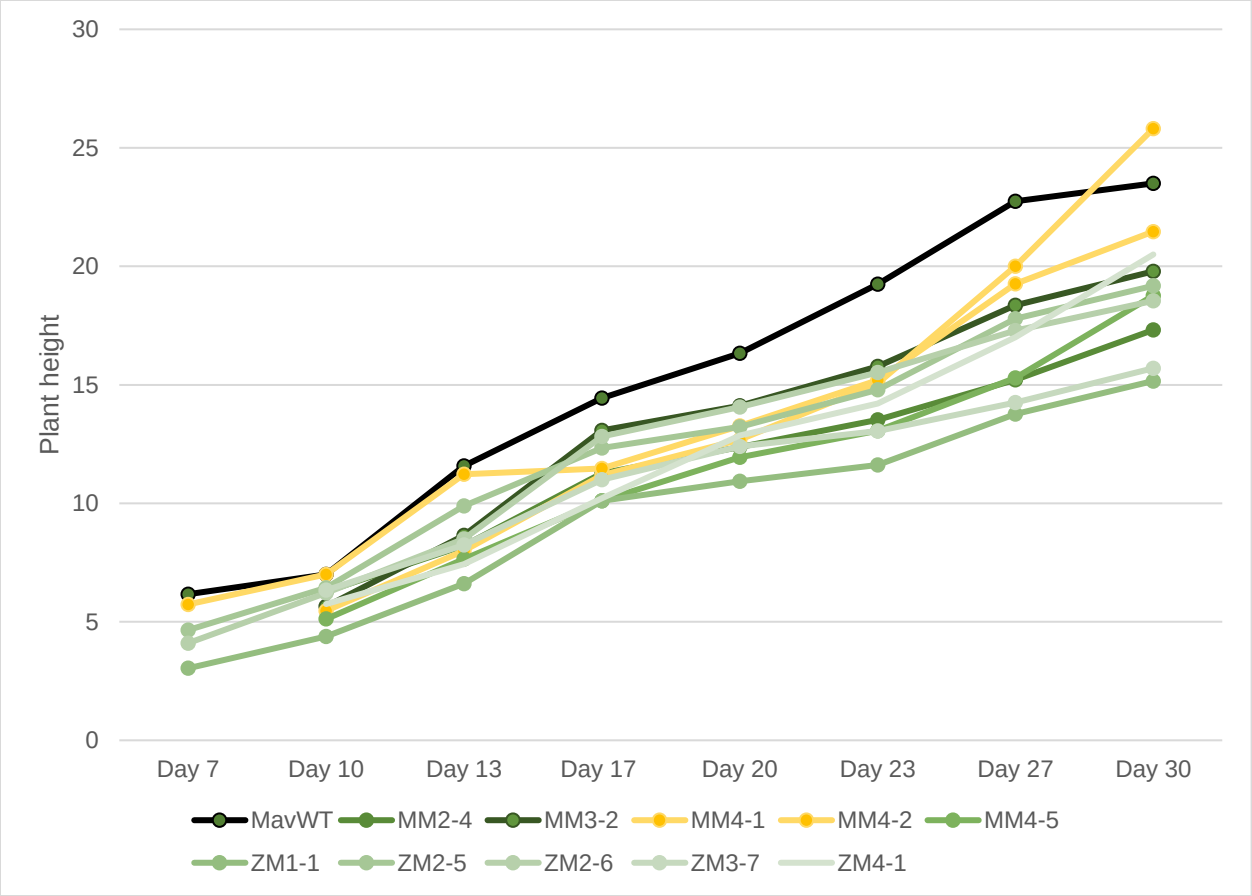


Figure 2.8
Representative gel result from PCR screen for presence of RS2 transgene.
Plasmid DNA was used as a positive control and ddH_2O for a negative control. Negative events (lanes 1-6) and positive events (lanes 8-11, 13, and 14) are identified by presence or absence of the ~600bp band. Ambiguous results, such as lanes 7, 15, and 16 were rescreened for confirmation.



	Day 7	Day 10	Day 13	Day 17	Day 20	Day 23	Day 27	Day 30
MavWT	6.17	7.00	11.58	14.44	16.33	19.25	22.75	23.50
MM2-4		6.28	8.21	11.26	12.37	13.53	15.21	17.32
MM3-2		5.65	8.65	13.08	14.11	15.78	18.36	19.79
MM4-1		5.47	8.00	11.19	12.69	15.06	20.00	25.81
MM4-2	5.74	7.00	11.22	11.47	13.27	15.20	19.27	21.47
MM4-5		5.12	7.65	10.11	11.94	13.06	15.29	18.76
ZM1-1	3.04	4.38	6.61	10.11	10.93	11.63	13.77	15.16
ZM2-5	4.65	6.43	9.89	12.34	13.23	14.79	17.79	19.18
ZM2-6	4.10	6.21	8.50	12.82	14.06	15.52	17.29	18.54
ZM3-7		6.33	8.25	11.00	12.40	13.05	14.25	15.70
ZM4-1		5.75	7.43	10.21	12.86	14.21	17.00	20.50

Figure 2.9
Plant heights of the T1 generation over the first 30 days of growth. Table lists average height for each sampling day, with data shaded in green being significantly different from wild-type. Lines MM4-1 and MM4-2 were pMU2T-bar empty vector transformed lines.

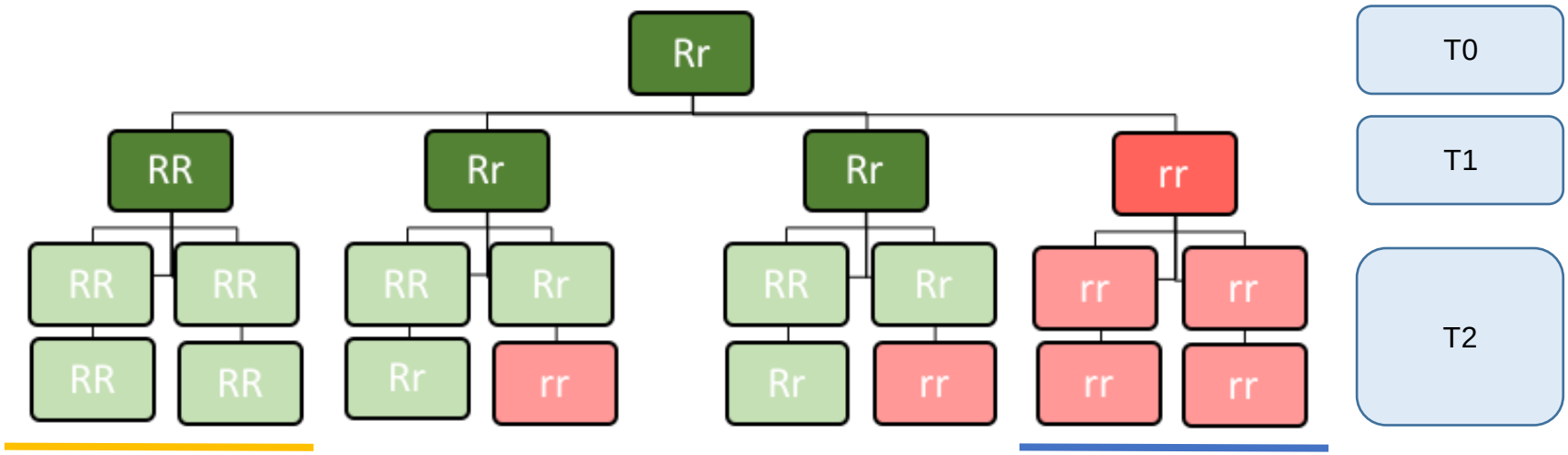


Figure 2.10
Expected genotypic and phenotypic inheritance for each transgene assuming normal Medelian segregation. At the T2 generation, the expected ratio of transgene(+):transgene(-) is 5:3. "R" denotes transgene positive and "r" denotes transgene negative.

Blue line indicates T1 plants that we not planted, and thus T2 progeny that were not included in the analysis for lines ZM1-1 and ZM3-7. Yellow line indicates T1 progeny that were known not to be present in the ZM1-1 population, and thus possible T2 progeny that were not included in the analysis.

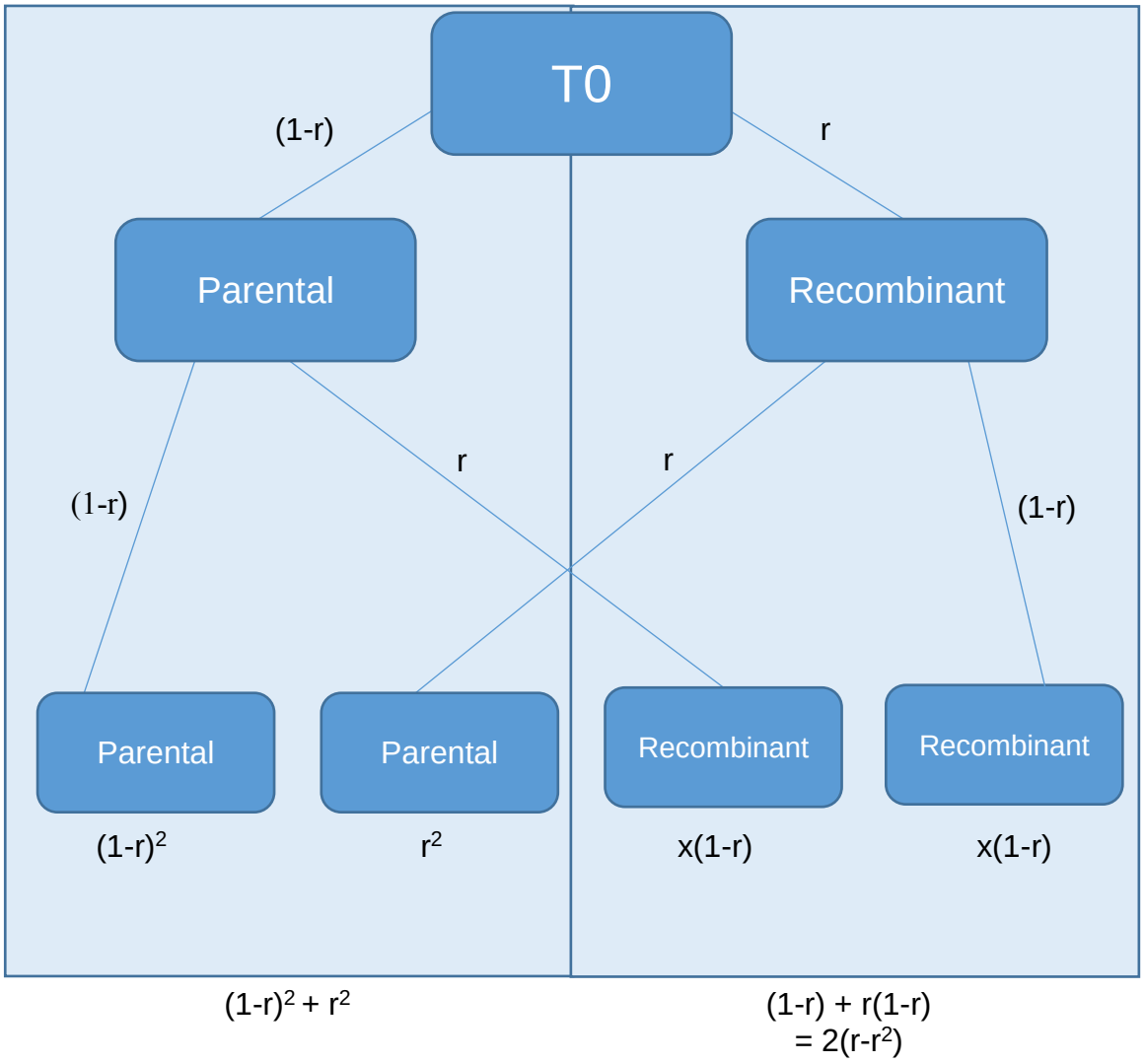


Figure 2.11
Calculations used to estimate genetic linkage of original T0 event, using recombination frequencies observed in T2 generation.

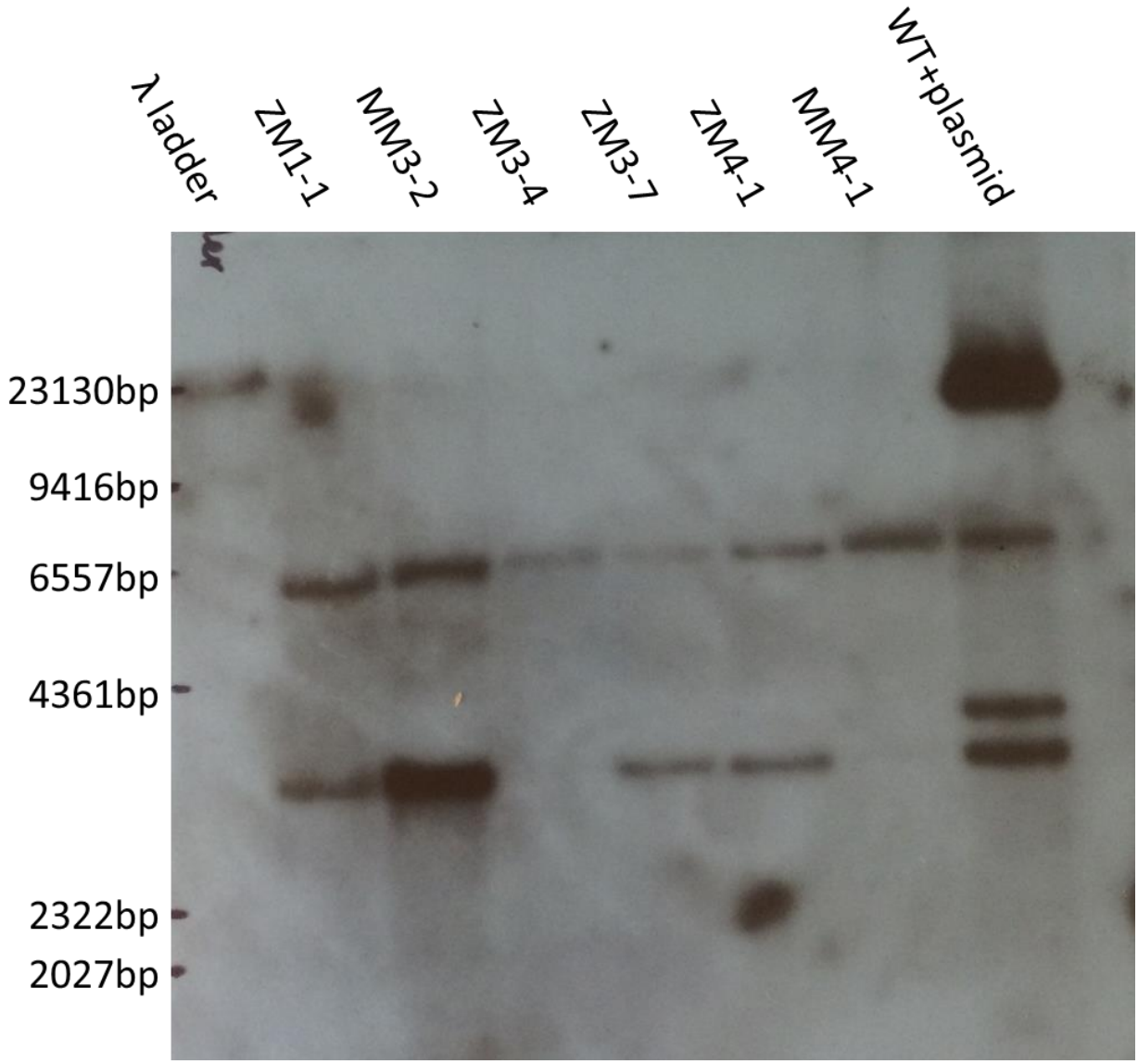


Figure 2.12a
Southern blot of soybean genomic DNA
Total soybean leaf genomic DNA was digested with NcoI and PstI. Lane 1 is the lambda DNA/HindIII ladder (NEB), lanes 2 through 7 are putative transgenic events, and lane 8 is wild-type genomic DNA spiked with 200pg of pMU2T-bar-RS2 plasmid.

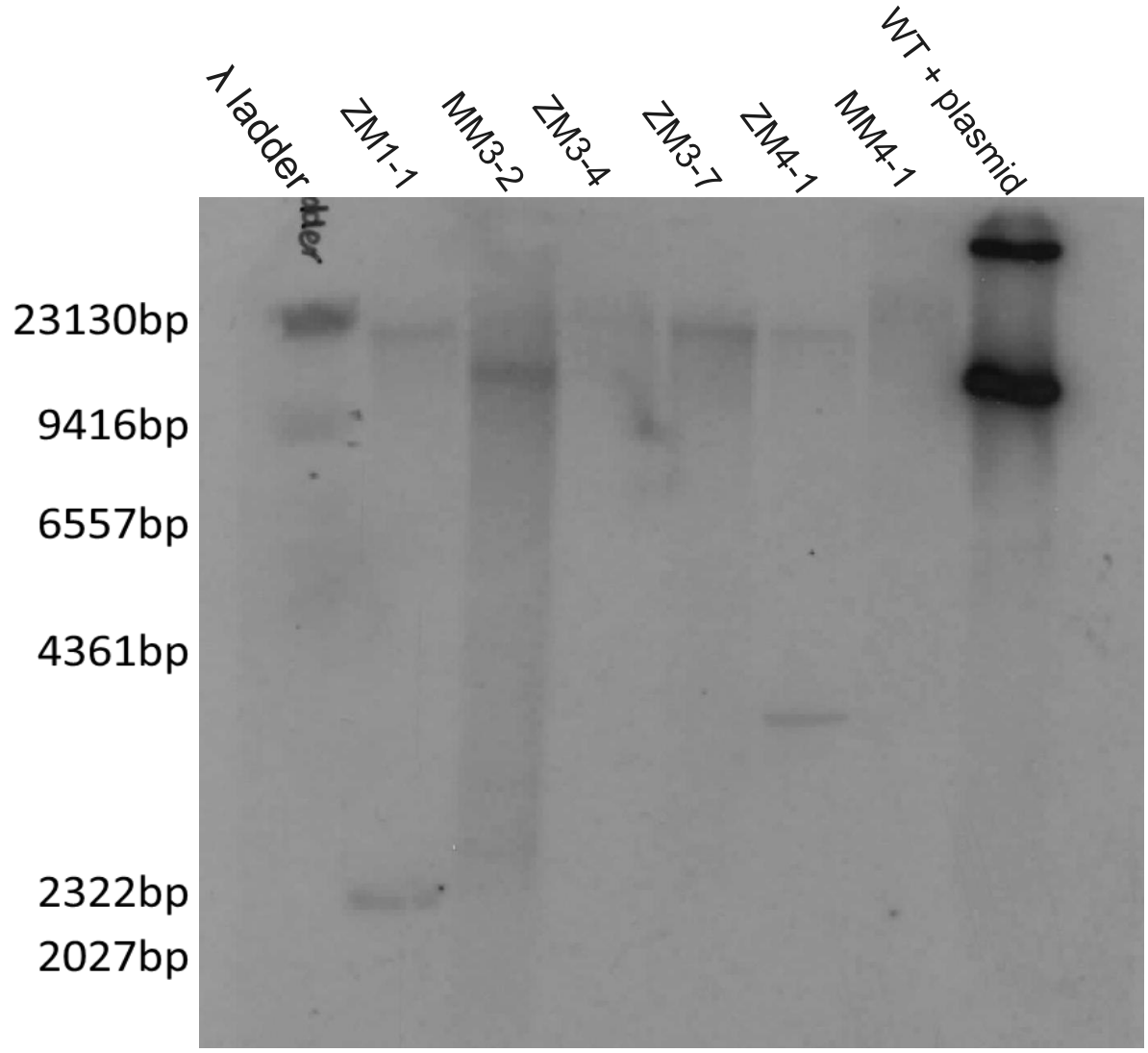


Figure 2.12b
Southern blot of soybean genomic DNA digested with PstI
Total soybean leaf genomic DNA was digested with PstI. Lane 1 is the lambda DNA/HindIII ladder (NEB), lanes 2 through 7 are putative transgenic events, and lane 8 is wild-type genomic DNA spiked with 200pg of pMU2T-bar-RS2 plasmid.

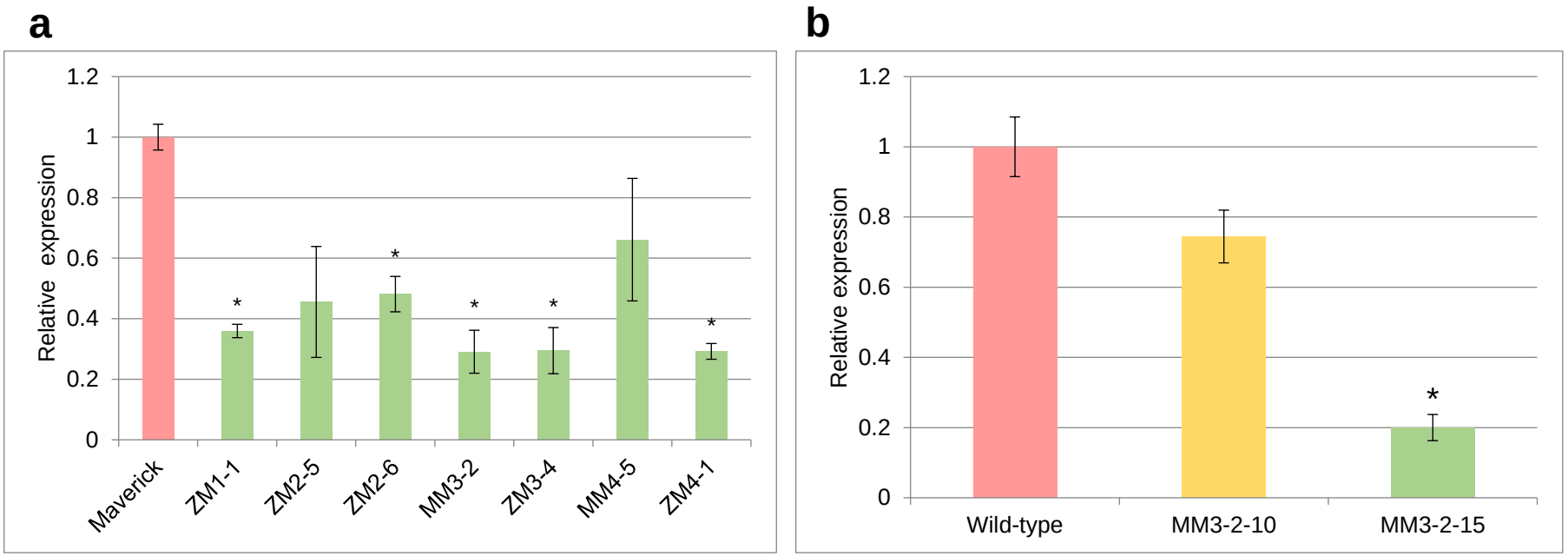


Figure 2.13
Real-time qPCR to quantify expression of endogenous soybean RS2 gene – greenhouse grown plants
Total RNA was extracted from mid-mature soybean seeds and reverse transcribed into cDNA. mRNA levels for RS2 from each of the T0 plants was evaluated by RT-qPCR using RS2 gene-specific primers and elongation factor 1alpha as an internal control.
(**a**) T1 seeds from T0 plants
(**b**) T3 seeds from T2 plants
Means with an (*) are significantly different from the wild-type value at p=0.05.

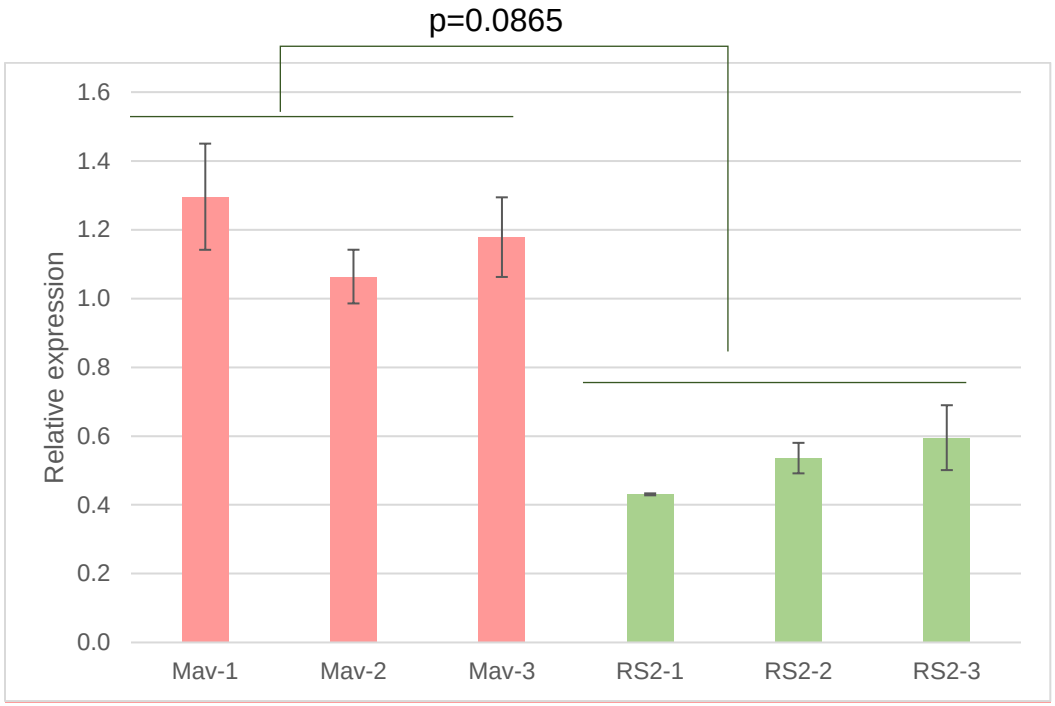


Figure 2.14
Real-time qPCR of soybean RS2 gene – T3 generation grown in the field
Each bar represents a 10-seed subsample of pooled mid-mature seeds. Level of down-regulation in RS2 transgenic seeds is, on average, 56% compared to wild-type.
Difference between average values for Maverick and RS2 transgene, p = 0.0865.

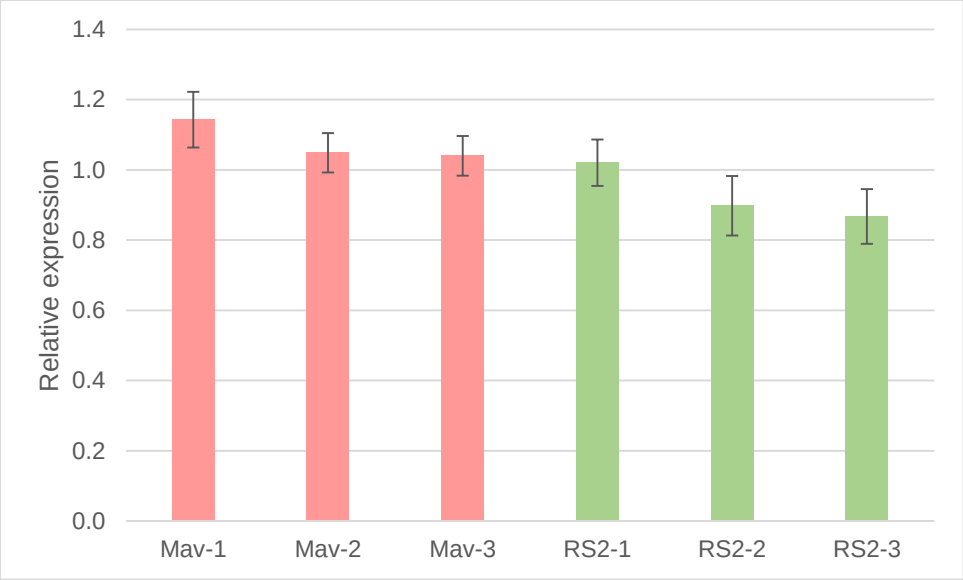


Figure 2.15
Real-time qPCR of soybean RS3 gene – T3 generation grown in the field
Each bar represents a 10-seed subsample of pooled mid-mature seeds. No significant difference at p=0.05 in expression of the RS3 gene was observed in wild-type versus transgenic plants.

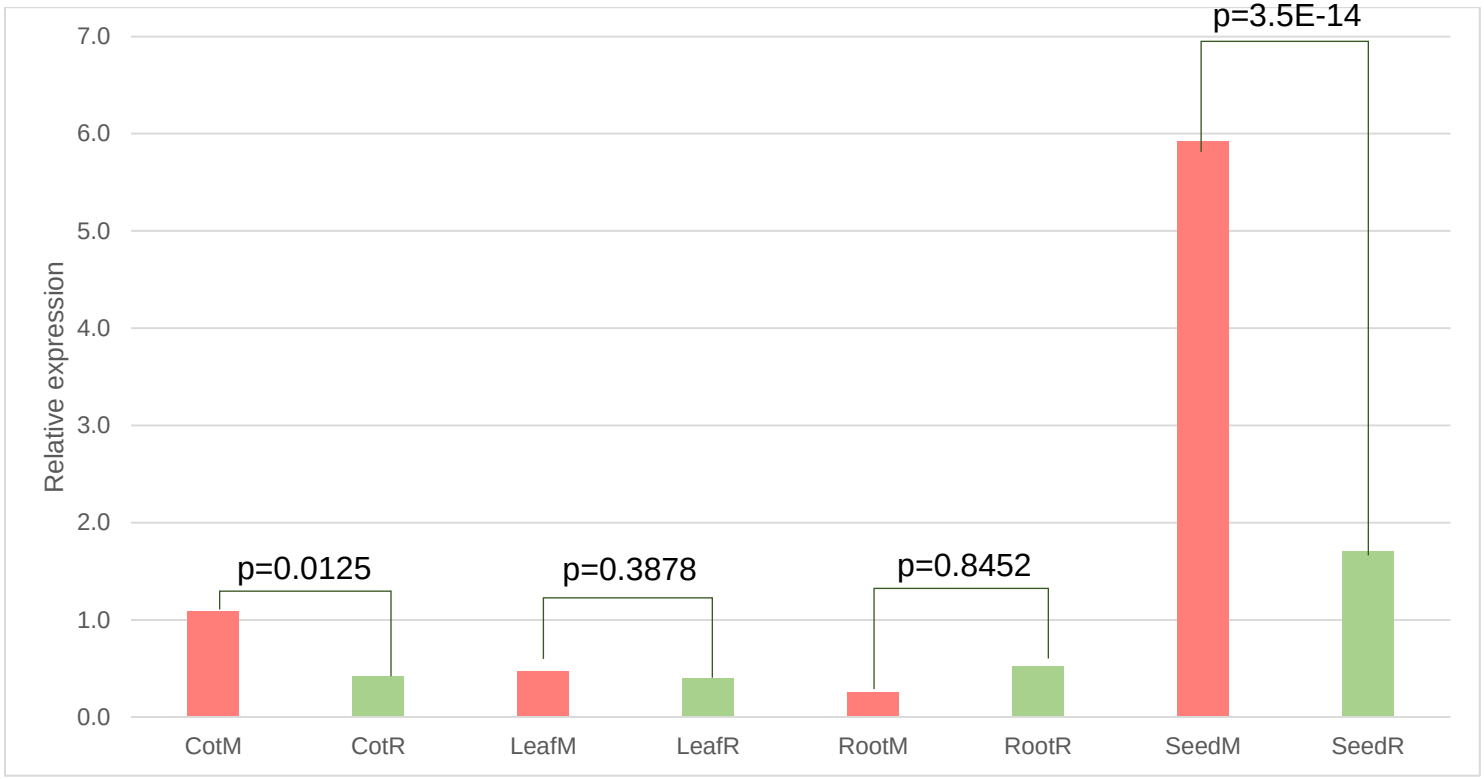


Figure 2.16
Real-time qPCR of endogenous soybean RS2 gene – cDNA extracted from different tissues. For each sample (cotyledon, leaf, root, and seed, samples were taken from Maverick wild-type plants (M) and RS2 transgenic plants (R).

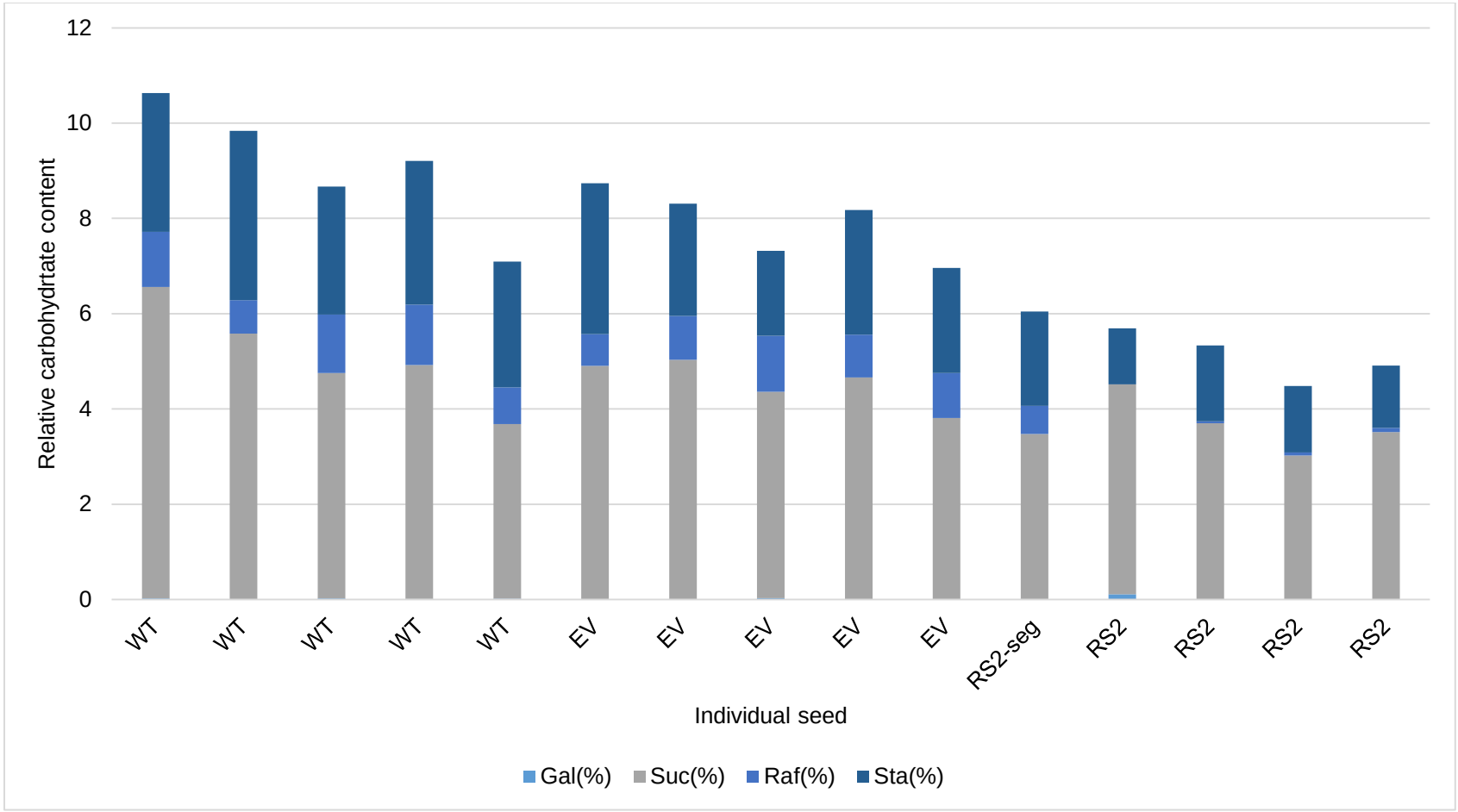


Figure 2.17
Carbohydrate content of individual T1 seeds grown in greenhouse.
WT = wild-type, EV = empty vector, RS2-seg = individual seed from a transgenic RS2 plant that did not have the RS2 transgene, RS2 = transgenic RS2.
Contents of sucrose, raffinose, and stachyose are significantly reduced in RS2 versus wild-type seeds.

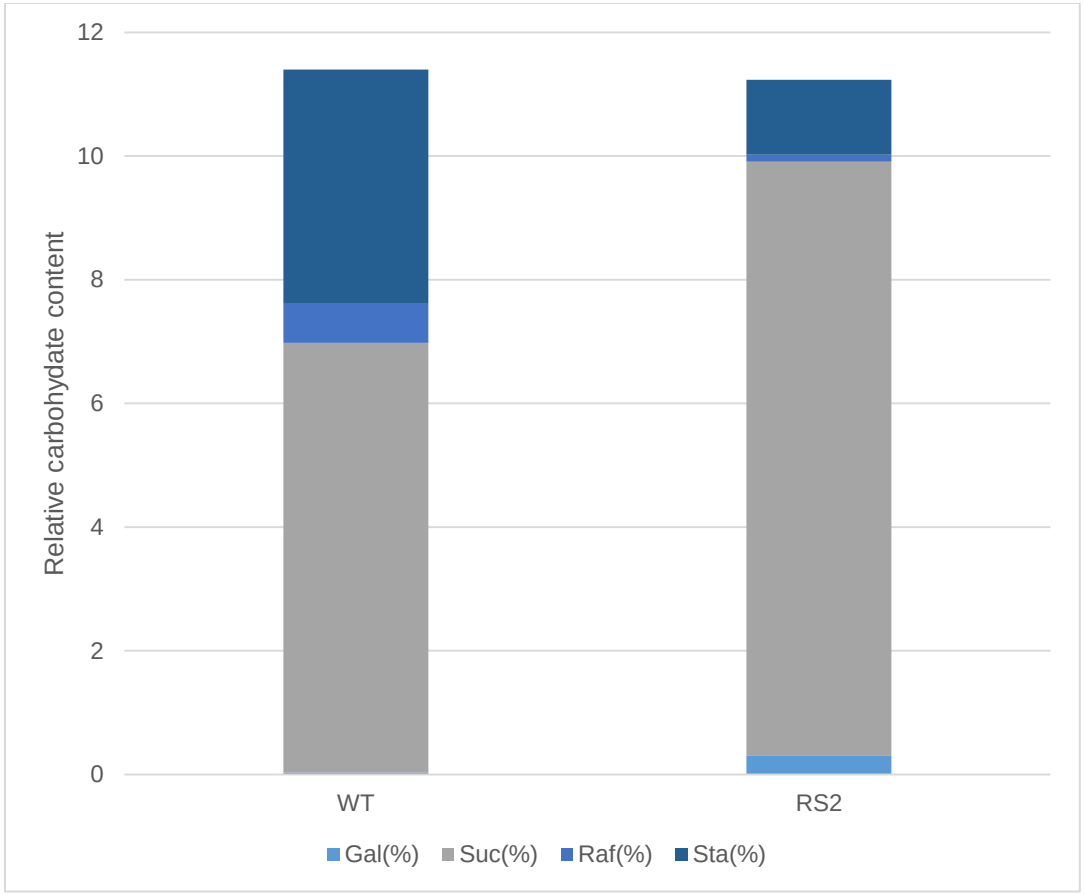


Figure 2.18
Carbohydrate content of field-grown, T3 seeds. WT = wild-type, RS2 = RS2 transgenic. Contents of galactinol and sucrose are significantly higher in transgenic than wild-type, and raffinose and stachyose are significantly reduced at p=0.05.

Figure 2.19
(a) Pod number, total seed weight in grams, and seed number for Maverick (wild-type), MM3-2-10 (bar transgenic, RS2-null) and MM3-2-15 (RS2 transgenic, bar-null). Calculate (b) 100 seed weight and (c) seeds per pod for the three genotypes. Means separated by different letters indicate significant differences at p=0.05.

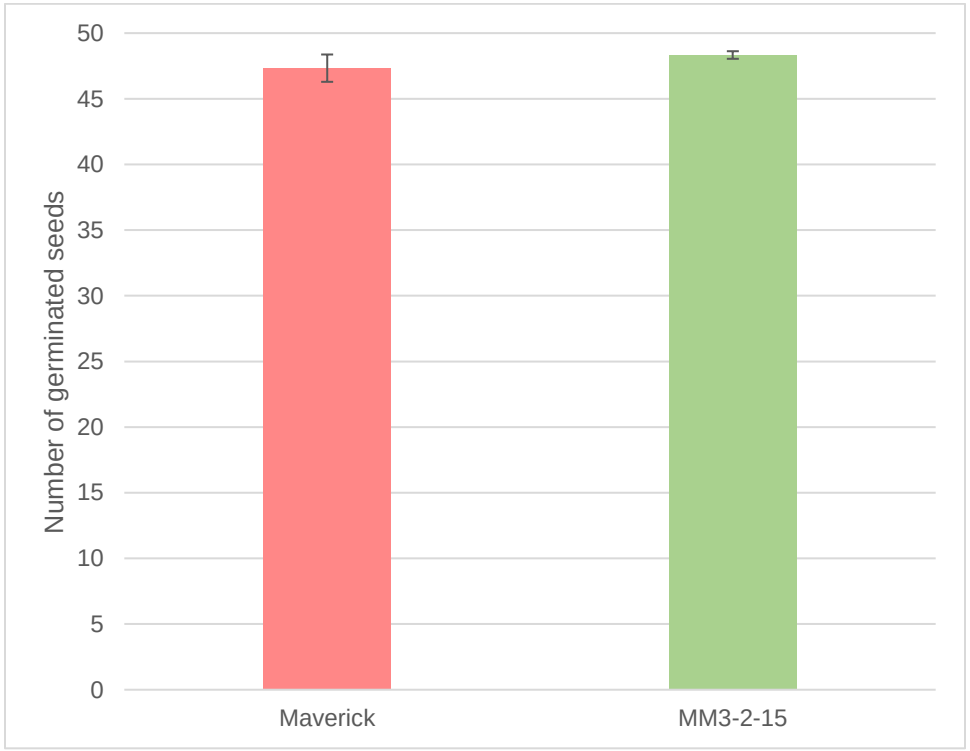


Figure 2.20
Germination rate of 50 wild-type and transgenic seeds under seven days of cold stress, then seven days of optimum temperature. The experiment was replicated three times. No significant difference was observed between wild-type and transgenic germination rates.

Tables

Table 2.1

	Germination	***Cocultivation***	***Shoot induction***	***Shoot elongation***	***Rooting***	***Unit***
B5 major salt (100x)	100	10 mL	100	100	100	mL
B5 minor salt (10x)	10	1 mL	10	10	10	mL
Iron-EDTA (10x)	10	1 mL	10	10	10	mL
Sucrose	20	30	30	30	20	g
pH	5.8	5.4	5.8	5.8	5.8	
MES	-	3.9	0.6	0.6	0.6	g
Gelling agent	Phytagel – 3	Washed agar – 8	Phytagel – 3	Phytagel – 3	Phytagel – 3	g
B5 vitamin mix (10x)	10	10	10	10	10	mL
Acetosyringone	-	0.04	-	-	-	g
BAP (1.67mg/mL)	-	1	1	-	-	mL
GA3 (1mg/mL)	-	0.25	-	1	-	mL
L-cysteine	-	0.4	-	-	-	g
Dithiothreitol	-	0.154	-	-	-	g
Na-thiosulfate	-	0.158	-	-	-	g
Zeatin-riboside (1mg/mL)	-	-	-	1	-	mL
Indole-3-acetic acid (1mg/mL)	-	-	-	0.1	-	mL
Aspartate/Glutamine (5mg/mL each)	-	-	-	10	10	mL
Timentin	-	-	0.05	0.05	0.05	g
Vancomycin	-	-	0.05	0.05	0.05	g
Cefotaxime	-	-	0.1	0.1	0.1	g

Soybean media recipes for 1L of each. Salt soutions, sucrose, MES, and gelling agent are added to ddH2O, then autoclaved at 250C and 20 PSI for twenty minutes. Remaining ingredients and appropriate concentration of glufosinate are filter-sterilized and added to media cooled in a water bath to 50°C.

Table 2.2
Primer sequences used in the project.

Primer name	***Primer sequence (5' to 3')***	***Description***
bar-ORF-sense	CACCATCGTCAACCACTACATCG	Detection of bar gene
bar-ORF-antisense	CAGCAGGTGGGTGTAGAGCGT	Detection of bar gene
Waxy-intron reverse	GTCATATCCCCTAGCCACCC	Detection of RS2 transgene
OCS-3' forward	TCATGCGATCATAGGCGTCT	Detection of RS2 transgene
CaMV-Right	GAAACCTCCTCGGATTCCAT	Contamination-free detection of bar gene
bar ORF reverse	CAGCAGGTGGGTGTAGAGCGT	Contamination-free detection of bar gene
1alpha-fwd	GACCTTCTTCGTTTCTCGCA	RT-qPCR for Elongation factor 1a control
1alpha-rvs	CGAACCTCTCAATCACACGC	RT-qPCR for Elongation factor 1a control
RS2 fwd	CTAGGGCCATCTCTGGTGGA	RT-qPCR for RS2 endogenous gene
RS2 rvs	CGTGTGGGGAGTGCATAGTG	RT-qPCR for RS2 endogenous gene
RS3 fwd	CACTGGAGTTCTTGGGGTGT	RT-qPCR for RS3 endogenous gene
RS3 rvs	GCTTGGCTGAAATACGAAGC	RT-qPCR for RS3 endogenous gene

Table 2.3
Summary of events derived from transformation experiments in 2013. Leaf painting results are denoted by an "R" for resistant, and an "S" for succeptible. PCR results are indicated by a "yes" for positive, "no" for negative, "?" for an ambiguous result, and "Control" for plants only transformed with the control vector, not the RS2 transgene. Empty boxes are plants that died in the rooting media, soil, or greenhouse. Lines shaded in green had both the bar gene and RS2 gene of interest.

Contributor	*Plasmid*	*Shoot Code*	*Rooting Date*	*Soil Date*	*Leaf Paint Results*	*Greenhouse Date*	*Bar Gene PCR*	*GOI PCR*
Murug	RS2	MM-1-1	4-Apr	24-Apr	SRR			
	RS2	MM-1-2	22-May					
	RS2	MM-1-3	22-May					
	RS2	MM-1-4	22-May	19-Jul				
	2T-bar	MM-2-1	6-May	22-May	SSS	19-Jun	No	Control
	RS2	MM-2-2	6-May	24-May				
	RS2	MM-2-3	6-May	24-May				
	RS2	MM-2-4	22-May	10-Jun	RRR	2-Jul	Yes	Yes
	RS2	MM-2-5	22-May					
	RS2	MM-2-6	22-May	21-Jun				
	RS2	MM-2-7	14-Jun	2-Jul				
	RS2	MM-2-8	21-Jun					
	RS2	MM-3-1	30-Apr	15-May	RRR	11-Jun	Yes	No
	RS2	MM-3-2	30-Apr	15-May	RRR	28-Jun	Yes	Yes
	RS2	MM-3-3	22-May	10-Jun	RRR	9-Jul	Yes	No
	RS2	MM-3-4	30-Apr					
	RS2	MM-3-5	21-Jun					
	2T-bar	MM-3-6	26-Jun	19-Jul				
	RS2	MM-3-7	9-Aug	22-Aug				
	RS2	MM-3-8	9-Aug	22-Aug				
	2T-bar	MM-4-1	7-May	22-May	RRR	11-Jun	Yes	Control
	2T-bar	MM-4-2	7-May	22-May	RRR	11-Jun	Yes	Control
	2T-bar	MM-4-3	7-May	3-Jun	RRR			
	RS2	MM-4-4	22-May	3-Jun	RRR	28-Jun	Yes	No

104

Murug	RS2	MM-4-5	22-May	3-Jun	RRR	28-Jun	Yes	Yes
	RS2	MM-4-6	22-May	3-Jun	RRR	28-Jun	Yes	No
	RS2	MM-4-7	7-May					
	RS2	MM-4-8	7-May					
	RS2	MM-4-9	7-May	5-Jul				
	2T-bar	MM-4-10	14-Jun	26-Jun	RRR	23-Jul	No	Control
	2T-bar	MM-4-11	14-Jun	26-Jun	RRR	23-Jul	Yes	Control
	RS2	MM-4-12	14-Jun	5-Jul	RRR			
	RS2	MM-4-13	21-Jun	5-Jul				
	RS2	MM-4-14	18-Jul					
	RS2	MM-5-1	22-May	3-Jun	RRR	28-Jun	Yes	No
	RS2	MM-5-2	22-May	10-Jun	RRR	9-Jul	No	No
	RS2	MM-5-3	22-May	10-Jun				
	RS2	MM-5-4	22-May	5-Jul				
	RS2	MM-5-5	18-Jul	22-Aug				
	RS2	MM-5-6	9-Aug	22-Aug				
	2T-bar	MM-6-1	22-May	3-Jun	RRR	25-Jun	Yes	Control
	2T-bar	MM-6-2	22-May	10-Jun	RRR	2-Jul		
	RS2	MM-6-3	22-May	10-Jun	RRR	2-Jul	Yes	No
	2T-bar	MM-6-4	14-Jun					
	RS2	MM-6-5	14-Jun	5-Jul	RRR	11-Oct	No	No
	RS2	MM-6-6	18-Jul	22-Aug				
	RS2	MM-6-7	18-Jul					
	2T-bar	MM-6-8	2-Aug	16-Aug				
	RS2	MM-6-9	2-Aug	6-Jan				
	2T-bar	MM-6-11	2-Apr	9-Sep				
Sandra	2T-bar	S-6-1	2-Jan	25-Jan	RRR	27-Feb	Yes	Control
	RS2	S-8-1	4-Feb	12-Feb	SSS	29-Mar	No	No
	RS2	S-8-2	4-Feb	20-Feb	RRR	29-Mar	Yes	No
	RS2	S-8-3	4-Feb	12-Feb	SSS	15-May	No	No
	RS2	Zhou-M-1-1	25-Jan	12-Feb	RRR	12-Mar	Yes	Yes
	RS2	Zhou-M-1-2	1-Mar	14-Mar	RRR	1-May	Yes	No

105

	2T-bar	Zhou-M-2-1	16-Jan	30-Jan	RRR	6-Mar	Yes	Control
	2T-bar	Zhou-M-2-2	29-Jan	12-Feb	RRR	12-Mar	Yes	Control
	RS2	Zhou-M-2-3	16-Jan	30-Jan	RRR	19-Mar	Yes	No
	2T-bar	Zhou-M-2-4	1-Feb	1-Mar	RRR	22-Apr	Yes	Control
	RS2	Zhou-M-2-5	1-Mar	14-Mar	RRR	1-May	Yes	Yes
	RS2	Zhou-M-2-6	1-Mar	14-Mar	RRR	1-May	Yes	Yes
	RS2	Zhou-M-2-7	1-Mar	14-Mar	SSR	15-May	Yes	No
	2T-bar	Zhou-M-3-1	2-May	20-May	RRR	25-Jun	Yes	Control
	2T-bar	Zhou-M-3-2	2-May	20-May	RSS	25-Jun	Yes	Control
Liwen	RS2	Zhou-M-3-3	23-May	10-Jun	SSS	28-Jun	No	No
	RS2	Zhou-M-3-4	23-May	10-Jun	RRR	28-Jun	Yes	Yes
	pZY101	Zhou-M-3-5	22-May	10-Jun	SSS	10-Jul	Yes	Control
	pZY101	Zhou-M-3-6	6-Jun	12-Jul	RRR	4-Sep	Yes	Control
	RS2	Zhou-M-3-7	17-Jul	2-Aug	RRR	4-Sep	Yes	Yes
	RS2	Zhou-M-4-1	29-May	10-Jun	RRR	16-Jul	Yes	Yes
	RS2	Zhou-M-4-2	2-May	20-May	RRR	16-Jul	Yes	No
	RS2	Zhou-M-4-3	25-Jun	12-Jul	SSS	9-Aug	?	No
	RS2	Zhou-M-4-4	25-Jun	12-Jul	SSS	13-Aug	No	No
	RS2	Zhou-M-4-5	11-Jun	2-Jul	SSS	23-Aug	Yes	No
	RS2	Zhou-M-4-6	12-Jul	26-Jul	RSS	11-Sep	Yes	?
	RS2	Zhou-M-5-1	4-Jun	24-Jun	SSS	30-Jul	?	No
	RS2	Zhou-M-5-2	23-May	10-Jun	RRR	9-Aug	Yes	No
	RS2	Zhou-M-5-3	25-Jun	12-Jul	RRR	23-Aug	Yes	No

106

Table 2.4
Summary of T1 genotype, RS2 expression phenotype, and progeny segregation. Shaded events indicate the transgene is inherited to the T1 progeny.
**Samples ZM3-7 and MM2-4 did not have sufficient T1 seeds to sample for qRT-PCR.*

T0 Event	*Leaf Painting Results*	*PCR bar*	*PCR RS2*	*qRT-PCR result (% downregulation)*	*No. Progeny*	*No. Progeny bar (+)*	*No. Progeny RS2 (+)*
ZM-1-1	RRR	Yes	Yes	64%	22	5	2
ZM-2-5	RRR	Yes	Yes	54%	30	0	0
ZM-2-6	RRR	Yes	Yes	52%	27	0	0
ZM-3-4	RRR	Yes	Yes	71%	17	0	0
ZM-3-7	RRR	Yes	Yes	*	20	8	9
ZM-4-1	RRR	Yes	Yes	71%	14	7	7
MM-2-4	RRR	Yes	Yes	*	19	0	0
MM-3-2	RRR	Yes	Yes	71%	19	18	13
MM-4-5	RRR	Yes	Yes	34%	17	0	0

Table 2.5
Summary of T1 segregation of bar and RS2 transgenes. Rejection of 9:3:3:1 model of independent segregation is at p=0.05 and 3 degrees of freedom.

T0 Event	*Bar (+)/RS2(+)*	*Bar (+)/RS2(-)*	*Bar(-)/RS2(+)*	*Bar(-)/RS2(-)*	*Chi-square value*	*Reject 9:3:3:1 model?*
ZM-1-1	2	3	0	15	162	Yes
ZM-3-7	3	5	6	3	11.55	Yes
ZM-4-1	6	1	1	5	23.51	Yes
MM-3-2	13	5	0	1	4.7	No
Overall	24	14	7	24	98.34	Yes

Table 2.6
Summary of T2 genotype segregation.

T0 Parent Plant	T2 progeny bar(+)/RS2(+)	T2 progeny bar(+)/RS2(-)	T2 progeny bar(-) /RS2(+)	T2 progeny bar(-) /RS2(-)	Recombination frequency at T2	Calculated recombination at T1
MM3-2	49	26	29	9	49%	42.93%
ZM1-1	23	8	1	5	24%	*
ZM3-7	30	9	6	3	31%	*
ZM4-1	10	3	1	2	25%	14.65%

Recombination frequency at T1 was calculated using the formula:

$$r = 2(x—x^2)$$

*where x = recombination frequency at T1, and r = observed recombination frequency at T2 (***see Figure 2.11***)*

*Lines ZM1-1 and ZM3-7 had a low inheritance of the transgenes, and only transgene(+) T1 plants were kept. Thus, T1 genetic linkage cannot be estimated using the above formula.

Not all plants were analyzed for both leaf-paint and RS2 PCR. Results listed are only for plants analyzed for presence of both transgenes.

109

Table 2.7
Summary of T2 genetic segregation of bar and RS2 transgenes. Chi-squared value was calculated and the hypothesis of 3:1, 5:1, or 5:3 models of segregation were rejected or not rejected at p=0.05 and three degrees of freedom.

	Expected ratio of bar(+):bar(-)	Observed ratio of bar(+):bar(-)	bar chi-square value	Reject expected model?	Expected ratio of RS2(+):RS2(-)	Observed ratio of RS2(+):RS2(-)	RS2 chi-square value	Reject expected model?
ZM1-1	3:1	31:6	1.523	No	3:1	24:13	2.027	No
MM3-2	5:3	138:73	0.759	No	5:3	86:39	2.117	No
ZM3-7	5:1	41:9	0.064	No	5:1	36:12	2.4	No
ZM4-1	5:3	21:15	0.267	No	5:3	11:5	0.267	No

110

Table 2.8
Composition proximate analysis of field-grown transgenic and wild-type soybean seeds

	Soybean sample	
Item	**Transgenic**	**Wild-type**
Moisture (W/W%)	6.13	6.14
Crude fat (W/W%)	19.18	19.79
Crude fiber (W/W%)	6.56	6.56
Ash (W/W%)	5.21	5.27
Amino acids		
Taurine	0.01	0.02
Hydroxyproline	0.01	0.01
Aspartic Acid	4.20	4.18
Threonine	1.47	1.47
Serine	1.70	1.72
Glutamic Acid	6.35	6.38
Proline	1.79	1.80
Lanthionine	0.00	0.00
Glycine	1.65	1.63
Alanine	1.62	1.61
Cysteine	0.61	0.63
Valine	1.90	1.87
Methionine	0.56	0.57
Isoleucine	1.78	1.79
Leucine	2.86	2.85
Tyrosine	1.30	1.30
Phenylalanine	1.82	1.81
Hydroxylysine	0.02	0.02
Ornithine	0.02	0.02
Lysine	2.51	2.51
Histidine	1.08	1.09
Arginine	2.59	2.53
Tryptophan	0.40	0.39
Total amino acids	36.25	36.20
Crude Protein	37.26	36.14

References

Abdul-Baki, A. A., & Anderson, J. D. (1973). Vigor Determination in Soybean Seed by Multiple Criteria1. *Crop Science*, *13*(6), 630. http://doi.org/10.2135/cropsci1973.0011183X001300060013x

Bilyeu, K. D., & Wiebold, W. J. (2016). Environmental Stability of Seed Carbohydrate Profiles in Soybeans Containing Different Alleles of the Raffinose Synthase 2 (RS2) Gene. *Journal of Agricultural and Food Chemistry*, *64*(5), 1071–1078. http://doi.org/10.1021/acs.jafc.5b04779

Blackman, S. A., Obendorf, R. L., & Leopold, A. C. (1992). Maturation proteins and sugars in desiccation tolerance of developing soybean seeds. *Plant Physiology*, *100*(1), 225–230.

Blöchl, A., Peterbauer, T., & Richter, A. (2007). Inhibition of raffinose oligosaccharide breakdown delays germination of pea seeds. *Journal of Plant Physiology*, *164*(8), 1093–1096. http://doi.org/10.1016/j.jplph.2006.10.010

Chen, L et al. (2013). *Value added products from soybean: Removal of anti-nutritional factors via bioprocessing.* (Vol. 90). Rijeka, Croatia: InTech.

Choffnes, D. S., Philip, R., & Vodkin, L. O. (2001). A transgenic locus in soybean exhibits a high level of recombination. *In Vitro Cellular & Developmental Biology - Plant*, *37*(6), 756–762. http://doi.org/10.1007/s11627-001-0125-8

Coon, C. N., Leske, K. L., Akavanichan, O., & Cheng, T. K. (1990). Effect of Oligosaccharide-Free Soybean Meal on True Metabolizable Energy and Fiber Digestion in Adult Roosters. *Poultry Science*, *69*(5), 787–793. http://doi.org/10.3382/ps.0690787

Cromwell, G.L. (2012). *Soybean Meal – An Exceptional Protein Source*. University of Kentucky. Retrieved from http://www.soymeal.org/ReviewPapers/SBMExceptionalProteinSource.pdf

Dierking, E. C., & Bilyeu, K. D. (2008). Association of a Soybean Raffinose Synthase Gene with Low Raffinose and Stachyose Seed Phenotype. *The Plant Genome Journal*, *1*(2), 135. http://doi.org/10.3835/plantgenome2008.06.0321

Dierking, E. C., & Bilyeu, K. D. (2009a). New sources of soybean seed meal and oil composition traits identified through TILLING. *BMC Plant Biology*, *9*(1), 89. http://doi.org/10.1186/1471-2229-9-89

Dierking, E. C., & Bilyeu, K. D. (2009b). Raffinose and stachyose metabolism are not required for efficient soybean seed germination. *Journal of Plant Physiology*, *166*(12), 1329–1335. http://doi.org/10.1016/j.jplph.2009.01.008

Dornbos, D. L., Mullen, R. E., & Shibles, R. E. (1989). Drought Stress Effects During Seed Fill on Soybean Seed Germination and Vigor. *Crop Science*, *29*(2), 476. http://doi.org/10.2135/cropsci1989.0011183X002900020047x

Egli, D. (2001). Source-sink Relationships, Seed Sucrose Levels and Seed Growth Rates in Soybean. *Annals of Botany*, *88*(2), 235–242. http://doi.org/10.1006/anbo.2001.1449

Egli, D. B., TeKrony, D. M., Heitholt, J. J., & Rupe, J. (2005). Air Temperature During Seed Filling and Soybean Seed Germination and Vigor. *Crop Science*, *45*(4), 1329. http://doi.org/10.2135/cropsci2004.0029

Elmore, R. W., Roeth, F. W., Nelson, L. A., Shapiro, C. A., Klein, R. N., Knezevic, S. Z., & Martin, A. (2001). Glyphosate-Resistant Soybean Cultivar Yields Compared with Sister Lines. *Agronomy Journal*, *93*(2), 408. http://doi.org/10.2134/agronj2001.932408x

ElSayed, A. I., Rafudeen, M. S., & Golldack, D. (2014). Physiological aspects of raffinose family oligosaccharides in plants: protection against abiotic stress. *Plant Biology (Stuttgart, Germany)*, *16*(1), 1–8. http://doi.org/10.1111/plb.12053

Fitch, M. M. M., Manshardt, R. M., Gonsalves, D., Slightom, J. L., & Sanford, J. C. (1992). Virus Resistant Papaya Plants Derived from Tissues Bombarded with the Coat Protein Gene of Papaya Ringspot Virus. *Bio/Technology*, *10*(11), 1466–1472. http://doi.org/10.1038/nbt1192-1466

Flores, T., Karpova, O., Su, X., Zeng, P., Bilyeu, K., Sleper, D. A., Zhang, Z. J. (2008). Silencing of GmFAD3 gene by siRNA leads to low α-linolenic acids (18:3) of fad3-mutant phenotype in soybean [Glycine max (Merr.)]. *Transgenic Research*, *17*(5), 839–850. http://doi.org/10.1007/s11248-008-9167-6

Frias, J., Song, Y. S., Martínez-Villaluenga, C., González de Mejia, E., & Vidal-Valverde, C. (2008). Immunoreactivity and amino acid content of fermented soybean products. *Journal of Agricultural and Food Chemistry*, *56*(1), 99–105. http://doi.org/10.1021/jf072177j

Gangola, M. P., Khedikar, Y. P., Gaur, P. M., Båga, M., & Chibbar, R. N. (2013). Genotype and Growing Environment Interaction Shows a Positive Correlation between Substrates of Raffinose Family Oligosaccharides (RFO) Biosynthesis and Their Accumulation in Chickpea (Cicer arietinum L.) Seeds. *Journal of Agricultural and Food Chemistry*, *61*(20), 4943–4952. http://doi.org/10.1021/jf3054033

Gibson, L. R., & Mullen, R. E. (1996). Soybean Seed Quality Reductions by High Day and Night Temperature. *Crop Science*, *36*(6), 1615. http://doi.org/10.2135/cropsci1996.0011183X003600060034x

Gil-Humanes, J., Piston, F., Tollefsen, S., Sollid, L. M., & Barro, F. (2010). Effective shutdown in the expression of celiac disease-related wheat gliadin T-cell epitopes by RNA interference. *Proceedings of the National Academy of Sciences*, *107*(39), 17023–17028. http://doi.org/10.1073/pnas.1007773107

Gordon, K. H. J., & Waterhouse, P. M. (2007). RNAi for insect-proof plants. *Nature Biotechnology*, *25*(11), 1231–1232. http://doi.org/10.1038/nbt1107-1231

Hannon, G. J. (2002). RNA interference. *Nature*, *418*(6894), 244–251. http://doi.org/10.1038/418244a

Hou, A., Chen, P., Shi, A., Zhang, B., & Wang, Y.-J. (2009). Sugar Variation in Soybean Seed Assessed with a Rapid Extraction and Quantification Method. *International Journal of Agronomy*, *2009*, 1–8. http://doi.org/10.1155/2009/484571

Hymowitz, T., Collins, F. I., Panczner, J., & Walker, W. M. (1972). Relationship Between the Content of Oil, Protein, and Sugar in Soybean Seed1. *Agronomy Journal*, *64*(5), 613. http://doi.org/10.2134/agronj1972.00021962006400050019x

Imai, K., & Lehmann, H. (1975). The oxygen affinity of haemoglobin Tak, a variant with an elongated beta chain. *Biochimica Et Biophysica Acta*, *412*(2), 288–294.

ISU. (2015). *Seed Health Testing*. Iowa State University. Retrieved from http://www.seeds.iastate.edu/seedtest/testing.html

Jankowski, J., Juskiewicz, J., Gulewicz, K., Lecewicz, A., Slominski, B. A., & Zdunczyk, Z. (2009). The effect of diets containing soybean meal, soybean protein concentrate, and soybean protein isolate of different oligosaccharide content on growth performance and gut function of young turkeys. *Poultry Science*, *88*(10), 2132–2140. http://doi.org/10.3382/ps.2009-00066

Kerr, P. S.; Sebastian, S. A. (2000). Soybean products with improved carbohydrate composition and soybean plants.

Komari, T., Hiei, Y., Saito, Y., Murai, N., & Kumashiro, T. (1996). Vectors carrying two separate T-DNAs for co-transformation of higher plants mediated by Agrobacterium tumefaciens and segregation of transformants free from selection markers. *The Plant Journal: For Cell and Molecular Biology*, *10*(1), 165–174.

Koster, Karen L., and A. Carl Leopold. (1988). Sugars and desiccation tolerance in seeds. *Plant Physiology*, *88*(3), 829–832.

Kumar, V., Rani, A., Goyal, L., Dixit, A. K., Manjaya, J. G., Dev, J., & Swamy, M. (2010). Sucrose and Raffinose Family Oligosaccharides (RFOs) in Soybean Seeds As Influenced by Genotype and Growing Location. *Journal of Agricultural and Food Chemistry*, *58*(8), 5081–5085. http://doi.org/10.1021/jf903141s

Limanton-Grevet, A, & Jullien, M. (2001). Agrobacterium-mediated transformation of Asparagus officinalis L.: molecular and genetic analysis of transgenic plants. *Molecular Breeding*, *7*(2), 141–150. http://doi.org/10.1023/A:1011310604393

Lu, Z. J., & Mathews, D. H. (2008). OligoWalk: an online siRNA design tool utilizing hybridization thermodynamics. *Nucleic Acids Research*, *36*(Web Server), W104–W108. http://doi.org/10.1093/nar/gkn250

Mansoor, S., Amin, I., Hussain, M., Zafar, Y., & Briddon, R. (2006). Engineering novel traits in plants through RNA interference. *Trends in Plant Science*, *11*(11), 559–565. http://doi.org/10.1016/j.tplants.2006.09.010

Matthews, Peter R., Wang, Ming-Bo, Waterhouse, Peter M., Thornton, Sarah, Fieg, Sarah J., Gubler, Frank, & Jacobsen, John V. (2001). Marker gene elimination from transgenic

barley, using co-transformation with adjacent `twin T-DNAs' on a standard Agrobacterium transformation vector. *Molecular Breeding*, *7*(3), 195–202.

Maupin, L. M., Rosso, M. L., & Rainey, K. M. (2011). Environmental Effects on Soybean with Modified Phosphorus and Sugar Composition. *Crop Science*, *51*(2), 642. http://doi.org/10.2135/cropsci2010.07.0396

Miranda, M., Vega-Gálvez, A., Martínez, E. A., López, J., Marín, R., Aranda, M., & Fuentes, F. (2013). Influence of contrasting environments on seed composition of two quinoa genotypes: nutritional and functional properties. *Chilean Journal of Agricultural Research*, *73*(2), 06–07. http://doi.org/10.4067/S0718-58392013000200004

Naqvi, A. R. (2009). The Fascinating World of RNA Interference. *International Journal of Biological Sciences*, 97–117. http://doi.org/10.7150/ijbs.5.97

National Research Council Nutrient Requirements of Poultry, Ninth Revised Edition, National Academies Press: Washington, DC, 1994, http://www nap edu/openbook php?isbn=0309048923. (n.d.).

Neus, J. D., Fehr, W. R., & Schnebly, S. R. (2005). Agronomic and Seed Characteristics of Soybean with Reduced Raffinose and Stachyose. *Crop Science*, *45*(2), 589. http://doi.org/10.2135/cropsci2005.0589

Nielsen, N. C. (1989). Characterization of the Glycinin Gene Family in Soybean. *THE PLANT CELL ONLINE*, *1*(3), 313–328. http://doi.org/10.1105/tpc.1.3.313

Nishizawa, A., Yabuta, Y., & Shigeoka, S. (2008). Galactinol and Raffinose Constitute a Novel Function to Protect Plants from Oxidative Damage. *PLANT PHYSIOLOGY*, *147*(3), 1251–1263. http://doi.org/10.1104/pp.108.122465

Parrish, David J., & Leopold, A. Carl. (n.d.). On the Mechanism of Aging in Soybean Seeds. *Plant Physiology*, *61*(3), 365–368.

Paz, M. M et al. (2004). Assessment of conditions affecting Agrobacterium-mediated soybean transformation using the cotyledonary node explant. *Euphytica*, *132*(2), 167–179.

Rathore, K. S., Sundaram, S., Sunilkumar, G., Campbell, L. M., Puckhaber, L., Marcel, S., … Wedegaertner, T. C. (2012). Ultra-low gossypol cottonseed: generational stability of the seed-specific, RNAi-mediated phenotype and resumption of terpenoid profile following seed germination: Ultra-low gossypol cottonseed. *Plant Biotechnology Journal*, *10*(2), 174–183. http://doi.org/10.1111/j.1467-7652.2011.00652.x

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., & Weigel, D. (2005). Specific Effects of MicroRNAs on the Plant Transcriptome. *Developmental Cell*, *8*(4), 517–527. http://doi.org/10.1016/j.devcel.2005.01.018

Sengupta, S., Mukherjee, S., Basak, P., & Majumder, A. L. (2015). Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Frontiers in Plant Science*, *6*. http://doi.org/10.3389/fpls.2015.00656

Small, I. (2007). RNAi for revealing and engineering plant gene functions. *Current Opinion in Biotechnology*, *18*(2), 148–153. http://doi.org/10.1016/j.copbio.2007.01.012

Song, Y.-S., Frias, J., Martinez-Villaluenga, C., Vidal-Valdeverde, C., & de Mejia, E. G. (2008). Immunoreactivity reduction of soybean meal by fermentation, effect on amino acid composition and antigenicity of commercial soy products. *Food Chemistry*, *108*(2), 571–581. http://doi.org/10.1016/j.foodchem.2007.11.013

Srivastava, P. K., Moturu, T., Pandey, P., Baldwin, I. T., & Pandey, S. P. (2014). A comparison of performance of plant miRNA target prediction tools and the

characterization of features for genome-wide target prediction. *BMC Genomics*, *15*(1), 348. http://doi.org/10.1186/1471-2164-15-348

Stroud, H., Ding, B., Simon, S. A., Feng, S., Bellizzi, M., Pellegrini, M., … Jacobsen, S. E. (2013). Plants regenerated from tissue culture contain stable epigenome changes in rice. *eLife*, *2*. http://doi.org/10.7554/eLife.00354

TeKrony, D. M., Egli, D. B., Balles, J., Tomes, L., & Stuckey, R. E. (1984). Effect of Date of Harvest Maturity on Soybean Seed Quality and Phomopsis sp. Seed Information1. *Crop Science*, *24*(1), 189. http://doi.org/10.2135/cropsci1984.0011183X002400010045x

Thomas, J. M. G., Boote, K. J., Allen, L. H., Gallo-Meagher, M., & Davis, J. M. (2003). Elevated Temperature and Carbon Dioxide Effects on Soybean Seed Composition and Transcript Abundance. *Crop Science*, *43*(4), 1548. http://doi.org/10.2135/cropsci2003.1548

Tizaoui, K., & Kchouk, M. E. (2012). Genetic approaches for studying transgene inheritance and genetic recombination in three successive generations of transformed tobacco. *Genetics and Molecular Biology*, *35*(3), 640–649. http://doi.org/10.1590/S1415-47572012000400015

Wang, Z., Zhu, Y., Wang, L., Liu, X., Liu, Y., Phillips, J., & Deng, X. (2009). A WRKY transcription factor participates in dehydration tolerance in Boea hygrometrica by binding to the W-box elements of the galactinol synthase (BhGolS1) promoter. *Planta*, *230*(6), 1155–1166. http://doi.org/10.1007/s00425-009-1014-3

Waterhouse, P. M., Graham, M. W., & Wang, M.-B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense

RNA. *Proceedings of the National Academy of Sciences*, *95*(23), 13959–13964. http://doi.org/10.1073/pnas.95.23.13959

Wielopolska, A., Townley, H., Moore, I., Waterhouse, P., & Helliwell, C. (2005). A high-throughput inducible RNAi vector for plants. *Plant Biotechnology Journal*, *3*(6), 583–590. http://doi.org/10.1111/j.1467-7652.2005.00149.x

Windows User-Friendly Feed Formulation (WUFFDA). (2010). College of Agricultural and Environmental Sciences, University of Georgia. Retrieved from: http://www.caes.uga.edu/publications/displayHTML.cfm?pk_id=7886.

Wolf, R. B., Cavins, J. F., Kleiman, R., & Black, L. T. (1982). Effect of temperature on soybean seed constituents: Oil, protein, moisture, fatty acids, amino acids and sugars. *Journal of the American Oil Chemists' Society*, *59*(5), 230–232. http://doi.org/10.1007/BF02582182

Xing, A., Zhang, Z., Sato, S., Staswick, P., & Clemente, T. (2000). The use of the two T-DNA binary system to derive marker-free transgenic soybeans. *In Vitro Cellular & Developmental Biology - Plant*, *36*(6), 456–463. http://doi.org/10.1007/s11627-000-0082-7

Yin, Y., Chory, J., & Baulcombe, D. (2005). RNAi in transgenic plants. *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, *Chapter 26*, Unit 26.6. http://doi.org/10.1002/0471142727.mb2606s72

Yin, Z., Plader, W., & Malepszy, S. (2004). Transgene inheritance in plants. *Journal of Applied Genetics*, *45*(2), 127–144.

Yoon, M., & Hwang, H. (2008). Reduction of soybean oligosaccharides and properties of α-d-galactosidase from Lactobacillus curvatus R08 and Leuconostoc mesenteriodes JK55. *Food Microbiology*, *25*(6), 815–823. http://doi.org/10.1016/j.fm.2008.04.008

Zeng, P., Vadnais, D. A., Zhang, Z., & Polacco, J. C. (2004). Refined glufosinate selection in Agrobacterium-mediated transformation of soybean [Glycine max (L.) Merrill]. *Plant Cell Reports*, *22*(7), 478–482. http://doi.org/10.1007/s00299-003-0712-8

Zhang ZY, Xing AQ, Staswick P, Clemente TE. (1999). The use of glufosinate as a selective agent in Agrobacterium-mediated transformation of soybean. *Plant Cell, Tissue and Organ Culture*, *56*(1), 37–46. http://doi.org/10.1023/A:1006298622969

Zuther, E., Büchel, K., Hundertmark, M., Stitt, M., Hincha, D. K., & Heyer, A. G. (2004). The role of raffinose in the cold acclimation response of Arabidopsis thaliana. *FEBS Letters*, *576*(1-2), 169–173. http://doi.org/10.1016/j.febslet.2004.09.006

Chapter 3

Consumer preference of soymilks made from soybeans with differing seed traits

Abstract

Soybean seed quality traits have been shown to significantly affect soymilk flavor, color, nutritional value, and overall acceptability. In this study, several soybean traits – high oleic, low linolenic, lipoxygenase-null, and low raffinose/high sucrose – and combinations thereof were evaluated for their impact on soymilk consumer acceptance. Six soybean varieties representing a control food-grade tofu line and five improved lines were processed into plain soymilk, assayed for nutritional qualities, and subjected to a consumer acceptability panel. Results show that soybean genotype information can accurately predict some aspects of soymilk composition. Consumers preferred soymilks with increased sucrose and low lipoxygenase, and disliked the high oleic trait. This research evaluates key soybean seed traits and their impact on soymilk acceptability in the United States, and provides a framework for future soybean seed trait improvements.

Introduction

Soybeans are an abundant, inexpensive, and nutritionally superior plant source of protein, and have been widely consumed in Asia for centuries. As a food source, soybeans are noted for their balanced profile of fat, protein, and carbohydrates, high fiber content, and possible health-promoting compounds such as saponins and isoflavones. The popularity of soy foods in the US market is increasing, in part due to the 1999 US Food and Drug Administration recognition of the cardiovascular health benefits of soy food consumption (CFR 21:101.82, USFDA 1999). US consumers are also turning to soy because of increased availability, purported health benefits for women, and desires to reduce animal product consumption (Schyver et al, 2005). A major limitation to soymilk acceptability in western countries is the comparison of soymilk to the more popular dairy milk, and thus soymilks with more dairy-like attributes such as sweetness and high viscosity are preferred (Villegas et al, 2009).

Soymilk and other soy foods are often described as having “beany”, “grassy”, and “rancid” flavors. In 2004, N'Kouka et al. developed a standardized lexicon for soymilk attributes, using 31 terms to describe soymilk flavor, mouthfeel, and aroma. The authors sought to discover the specific notes of more ambiguous flavor terms such as “beany”, and found that panelists identified more descriptive terms such as “green”, and “raw soy” to describe the beany flavor. A more recent study has determined that US consumers prefer soymilks with sweet and vanilla flavors, and a thicker consistency. Beany, green, grassy, and astringent attributes are less preferred (Lawrence et al, 2016).

While soy flavor is becoming more familiar to western palates, there are several barriers to soy food acceptance. Soybeans have relatively high levels of raffinose family oligosaccharides (RFOs), which are not digested by humans and can cause flatulence and digestive disturbance (Naczk, 1997). Raffinose synthesis is catalyzed by the RS2 gene, converting sucrose plus galactinol into raffinose, which contains the indigestible α-1,6 glycosidic bond. Soybeans also contain three genes (*SL-1, SL-2,* and *SL-3*) encoding lipoxygenase enzymes, which catalyze the formation of hydrogen peroxide intermediates in processed soybeans. These hydrogen peroxide intermediates create reactive oxygen singlets which oxidize polyunsaturated fatty acids containing a 1,4-cis,cis-pentadiene structure, namely the major soybean lipids linoleic and linolenic acid (Iassonova et al, 2009). The oxidation of these fatty acids produces medium chain length alcohols and aldehydes such as hexanal, which is primarily responsible for the characteristic beany, grassy flavor of soy foods (Lin and Wilkens, 1970).

Further, linoleic and linolenic acid are less suitable for heated applications such as cooking or frying, as their chemical structure degrades at high temperatures. To overcome this, food manufacturers hydrogenate soybean oil to stabilize the shelf-life. This hydrogenation produces trans-saturated fats, which are implicated in a number of human health problems such as heart disease and high cholesterol. Therefore, the US Food and Drug Administration has recommended limited consumption of trans fats. Conversely, monounsaturated fatty acids, such as oleic acid, do not need hydrogenation for stabilization, and have a positive impact on human health (Kris-Etherton et al, 1999). In

developing soybean seeds, oleic acid is converted to linoleic and linolenic acid via the FAD2-1A and FAD2-1B genes, and mutations of these genes increases oleic acid content (Pham et al, 2010).

Manufacturers currently use processing methods to reduce these anti-nutritional factors in commercial soymilk. High temperature cooking is used to lessen the amount of trypsin inhibitor, reduce beany flavor, and sterilize soymilk (Kwok et al, 2002; Yuan et al, 2008; Zhang and Chang, 2016). Prolonged heating, however, is shown to reduce the color and flavor acceptability of the product, and degrade nutritional compounds such as antioxidants (Xu et al, 2010). Sprouting of soybeans before processing has been shown to increase protein content, decrease trypsin inhibitor activity, and increase flavor acceptability (Murugkar, 2014). Further modifications to soymilk processing include the temperature at which soybeans are ground, the inclusion of a soaking step, high-pressure heat treatment, and α-galactosidase enzyme treatment (Zhang et al, 2012; Matsuura et al, 1989; Van der Ven and Van den Berg, 2005; Kulkarni et al, 2006), however additional processing can be cost-prohibitive.

To overcome barriers to soy food acceptance in the United States, soybean breeders have selected varieties with low anti-nutritional factors. Soybean lines lacking enzymes for lipoxygenase synthesis, and lines with altered carbohydrate and lipid profiles have been bred. A mutation in the soybean RS2 gene dramatically reduces raffinose and stachyose content in soybean seed (Dierking and Bilyeu, 2008). Because sucrose is required for raffinose synthesis, raffinose-free soybeans have been shown to have higher levels of the

digestible sugar sucrose, which positively affects metabolizable energy and taste. Soybean varieties lacking the three enzymes for lipoxygenase synthesis have been demonstrated to produce less hexanal and better tasting soymilk (King et al, 1998; Torres-Penaranda et al, 2006). Varieties have also been developed with low levels of polyunsaturated fatty acids and an increased level of monounsaturated oleic acid (Wilcox, 1984 et al; Pham, et al 2010).

Soybean lines bred at the University of Missouri have been developed with varying combinations of these desirable nutritional characteristics. The purpose of this research is to evaluate the contributions of high oleic, low linoleic, lipoxygenase-null, and low raffinose/high sucrose soybean traits on the consumer acceptance of soy milk.

Materials and Methods

Soybean varieties and composition analysis

Six soybean varieties were selected for the study, representing a variety of improved seed composition traits. The lines selected were: 534545, a tofu line with clear hilum; KB13-15, high oleic and low linolenic; KB13-7, high oleic and low raffinose; KB13-16 high oleic and lipoxygenase null; KB11-8, high oleic, low linolenic, low raffinose, and lipoxygenase null; and KB11-14, lipoxygenase null and low raffinose (Table 3.1). The soybean lines were developed at the University of Missouri, and grown in the field in 2015. Harvested seed was stored in a cold storage room to prevent seed quality degradation.

Soybean sucrose content was measured by enzymatic assay using the Megazyme® kit (Megazyme, Bray, Ireland). Briefly, total carbohydrate was extracted from 12.5mg of lyophilized, ground soybean with 1mL of 50% ethanol for 30 minutes at 70°C. After centrifugation to remove solids, 100uL of supernatant was diluted with 100uL of H_2O. 10uL of the dilution was mixed with 10uL of Megazyme Solution II in a new 96-well plate and incubated at 50°C for 20 minutes. 20uL of standards were prepared using the supplied 1mg/mL glucose standard, and 150uL of GOPOD reagent was added to all wells. After incubation at 50°C for 20 minutes, 125uL of the sample was analyzed on a spectrophotometer at 510nm. The assay measures glucose concentration, and percent sucrose was calculated assuming no endogenous glucose in the seed.

Soymilk sucrose content was also measured by Megazyme assay, but with modifications to the extraction procedure. Carrez's reagents I (3.6 g of potassium hexacyanoferrate(II) trihydrate in 100mL H_2O) and II (7.2g of zinc sulfate in 100uL H_2O) and 100mM NaOH were used to clarify the samples. 1mL of clarified soymilk was diluted in 50mL of ddH_2O and the dilution was mixed with Megazyme Solution II, then samples were analyzed with GOPOD as before.

Galactinol, raffinose, and stachyose content in seeds were determined by HPLC. 12.5mg of lyophilized, ground seed was extracted with 1mL of 50% ethanol at 70°C for 30 minutes with occasional agitation. 50uL of the extraction supernatant vacuum-dried to remove ethanol, and then resusupended in 250uL of ddH_2O. 10uL of the resuspension was used for injection for HPLC. A Dionex CarboPac PA10 4 X 250mm column and Dionex CarboPac PA10 4 X 50mm guard column were used to analyze samples. Soybean fatty acid composition was analyzed by FAME gas chromatography as in Beuselinck et al, 2006.

Soymilk preparation

Soymilk was prepared using a SoyaJoy® G3 soymilk maker, following manufacturer instructions. For each line, about 150g of dry seed was soaked overnight in an excess of room-temperature water. The imbibed soybeans were then rinsed with fresh water, and soymilk was prepared in the machine with 1.5L of fresh deionized water. The SoyaJoy® machine cycle is about 20 minutes, with a maximum temperature of 95ºC to reduce trypsin inhibitor activity. After the machine cycle, soymilk was strained twice through a fine mesh

sieve to remove solids. Milk was then transferred to a plastic food container and stored in a 4ºC refrigerator overnight.

Consumer panel

On the days of the study, volunteering participants were informed of the purpose of the study, potential risks associated with participating, and compensation for participation. After consent was obtained, participants were given a consumer ballot (Figures 3.1a and 3.1b), a tray with 6 randomly ordered samples of 15mL ice-cold prepared soymilk served in 30mL plastic food service containers, and a cup of water to drink between soymilk samples. Consumers were asked to indicate their gender, typical soymilk consumption, and then rate their liking of each soymilk on a 15cm scale from "extremely dislike" to "extremely like" on the ballot. Ballots were identified with only a number to maintain participant confidentiality. The study was conducted over three days at two locations on the MU campus.

Statistical analysis

SAS software (Version 9.4) was used for all data analysis. PROC GLM was used to separate means and calculate the effects of traits on rating and their interactions. PROC CORR was used to model the correlations between carbohydrate components, lipid components, and rating. Values were considered significant at $p=0.05$.

Results

Soybean composition

The sucrose content of soybean seeds used varied from 5.3 to 8.77%, with the food-grade tofu line, 534545 having the highest (Table 3.2). For other lines, the low-raffinose genetic trait was an accurate predictor of both low-raffinose and high sucrose content. While 534545 had increased sucrose and reduced raffinose content, stachyose levels in this line were similar with wild-type varieties. The genetic basis of the 534545 soybean variety's high sucrose, low raffinose phenotype is yet unknown. The sucrose content of soymilk was also analyzed, and the relative extraction of sucrose from the seed to the milk was examined (Table 3.3). The relative rate of extraction ranged from 69.3 to 90.6% of seed sucrose extracted into the soymilk, which is consistent with reported extraction rates for sucrose (Ku et al, 1976) and protein (Bourne, 1976).

Soybean seed lipid content of palmitic (16:0), stearic (18:0), oleic (18:1), linoleic (18:2), and linolenic (18:3) fatty acids were analyzed by gas chromatography (Table 3.4). Genotype is a strong predictor of fatty acid content; the lines with high oleic and low linolenic genetic traits had significantly increased oleic and decreased linolenic contents, respectively. Though line 534545 had altered carbohydrate composition, it did not appear to be significantly different from the fatty acid wild-type line KB11-14.

Soymilk rating

The rating of each soymilk on the 15 cm line was scored by measuring the distance from the "extremely dislike" anchor to the panelist's mark in millimeters. So, a score of 75 would be halfway between "extremely dislike" and "extremely like" on the 15 cm scale. To account for panelists' different interpretations of the ranking scale, several data processing methods are common. One option is to standardize data to zero mean and unit variance for all samples analyzed by each panelist. To do this, individual sample rating is subtracted by the average rating for that panelist, and then divided by the standard deviation of all rankings by that panelist. This correction is used to account for panelist differences in overall mean and range of rankings, and can be useful in reducing the effect of consuming samples in different orders (Powers, 1988; Rohm et al, 1994; Naes et al, 2010). However, centering the data around zero removes the attributes of the panelists from the analysis, and as panelist attributes such as gender and typical soymilk consumption were of interest to this study, data standardization was not performed. Instead, variability among panelists was incorporated into the statistical model as a random effect of panelist. The same ranking and significance levels of soymilks was obtained from both standardized data and non-standardized data incorporating panelist into the model.

195 people participated in the study; 111 females, 76 males, and 8 choosing not to respond. Of those responding, 78 reported never consuming soymilk, 49 consume soymilk a few times per year, 34 a few times monthly, 20 a few times weekly, and only 9 of participants reported consuming soymilk daily.

Soybean varieties 534545 and KB11-14 ranked the highest in the consumer acceptability study, followed by KB11-8, and then KB13-16, KB13-15, and KB13-7 (Figure 3.2). Interestingly, the only trait differences between the more preferred KB11-14 and less preferred KB11-8 are the traits for altered fatty acid content. High oleic and low linolenic acid soybeans are desirable for their improved nutritional qualities, but consumers did not prefer the taste of the enhanced variety.

Trait effects on rating

PROC GLM was used to test for the positive or negative effects of soybean traits on consumer preference, and results are summarized in Table 3.5. The genetic traits high oleic and low linolenic had significant negative effects on rating. The RS2 trait had no significant effect, and the lipoxygenase-null trait had a significant positive effect on rating.

The total polyunsaturated fat contents (18:1 + 18:2) for each sample were summed and the soybean varieties were divided into two categories: "high PUFA", with PUFA content above 50%, and "low PUFA", lower than 50% PUFA content. A significant interaction was observed between total polyunsaturated fat content and lipoxygenase-null trait presence (Figure 3.3a). A significant interaction was also observed between the low-linolenic and lipoxygenase-null traits (Figure 3.3b). As PUFA content increases, consumer rating also increases, and while the lipoxygenase-null trait has a significant positive effect on low PUFA samples, the trait has no significant effect for high PUFA samples. This response is unexpected given the known pathway of lipoxygenase-associated off-flavors.

Composition effects on rating

The carbohydrate and lipid profiles of soybean seeds were quantified, and correlation analysis revealed the correlation of those components with consumer preference, and their interactions with each other. All five tested lipids had a significant effect on rating (Table 3.6). Increasing concentrations of oleic acid had a significant negative effect, and concentrations of the other fatty acids had significant positive effects. As biosynthesis pathways of the different lipids are highly linked, all fatty acids were also highly significantly correlated with one another. So, with the given set of soybean lines, it is difficult to measure the specific effect of each fatty acid on rating.

For quantified carbohydrate content, seed sucrose had a significant positive effect on rating, and galactinol and raffinose had significant negative effects (Table 3.7). Stachyose did not have a significant effect on rating. Like fatty acids, soybean seed carbohydrates are significantly correlated with one another. Thus, to specifically measure each fatty acid and carbohydrate's effect on consumer preference, a different experimental approach would be needed.

Consumer attributes effects on rating

Gender did have a significant effect on rating, with males rating soymilks higher overall than females (Figure 3.4a). However, there was no significant interaction between gender and preference for individual soymilks.

To simplify analysis, the panelists were divided into two groups, “frequent soymilk consumers”, reporting consumption daily, weekly, or monthly (n=63), and “infrequent soymilk consumers”, reporting consumption yearly or never (n=127). Although there was a significant difference in average rating of all soymilks between the two groups (Figure 3.5), there were no significant differences in preference for particular soymilks between frequent and infrequent consumers.

Discussion

Based on results of this study, the ideal soybean variety for soymilk is: lipoxygenase-null, not high oleic or low linolenic, and a high in sucrose. The genetically improved line KB11-14 is preferred statistically similar to the food-grade tofu line 534545, which demonstrates that the modulation of specific soybean composition traits can drastically improve their human food use capability. The desirable composition traits can be quickly and efficiently bred into high-yielding, agronomically-desirable soybean varieties using molecular markers more easily than trying to bring complex agronomic traits such as yield into current lower-yielding food-grade lines (Rao et al, 2002).

The interaction of the lipoxygenase-null and low-linolenic traits, and of lipoxygenase-null with total PUFA content is surprising given the known mechanism of lipoxygenase-involved oxidation of PUFAs. The oxidation of *cis,cis*-1,4-pentadiene structures, such as linoleic and linolenic acids, produces hydroperoxide intermediates that are converted into compounds associated with objectionable flavor in soyfoods (Kitamura, 1984). Thus, it is expected that soybeans with low PUFA content and low lipoxygenase activity would produce the most acceptably flavored soymilk. In this study, lipoxygenase-null lines were preferred as expected, but high PUFA content was more preferred by panelists. Several studies have shown lipoxygenase-null soybeans produce more acceptable soymilk (Davies et al, 1987; Torres-Penaranda et al, 2006; Yang et al, 2016), however these have used soybeans with normal levels of linoleic and linolenic acids. In 2007, Yuan and Chang tested the contributions of lipoxygenase activity and fatty acid composition on odor-producing

volatile production, and found that linoleic acid content is highly positively correlated with hexanal production.

One explanation for the current study's result is that the observed consumer preference for high PUFA is actually a result of other fatty acids, as all fatty acids are highly correlated with one another. In the selected soybean varieties, high PUFA is significantly correlated with high palmitic and stearic acids, and so the contributions of these oils may be the significant factor. The major fatty acid in normal dairy milk is palmitic (16:0), accounting for about 30% of the total lipid content, followed by oleic (18:1), stearic (18:0), and myristic (14:0) acids at 24, 12, and 11%, respectively. (Månsson, 2008). Therefore, the high PUFA soymilks in this study had fatty acid profiles more similar to dairy milk than the low PUFA milks. Westerners tend to prefer dairy products over their soy counterparts (Wu et al, 2005; Palacios et al, 2009; Tu et al, 2010), and dairy milk with higher fat content is preferred over nonfat varieties (Pangborn et al, 1985; Phillips et al, 1995). In part, the preference for dairy products over their alternatives may be due to the improved creaminess and mouth-feel of dairy (Palacios et al, 2009). So, one strategy to improve the consumer acceptability of soymilk is to improve the texture and mouth-feel of the product. Another explanation for the consumer preference for low PUFA milks is that the soymilks were freshly prepared no more than 24 hours before serving and stored and served at cold temperatures. Oxygenation of fatty acids is accelerated by high temperatures, and studies using lipoxygenase-null soybean oil saw significant degradation of the oil at two weeks (King, et al 1998). Therefore, it is possible that the short storage time did not allow for full development of PUFA-associated off-flavors.

It is also possible that the significant negative correlations between soybean raffinose and galactinol content and rating are due to the positive effect of increased sucrose. The current study raises interesting questions for more research into the flavor and textural qualities of soy-based beverages.

Future studies are warranted to determine the contributions of each component of soybean seed composition on consumer acceptability. Because of the known pathways of soybean lipid and carbohydrate biosynthesis, breeding efforts to improve each constituent individually may not be successful. For example, breeding for high oleic content reduces linoleic content, as oleic acid is a precursor of linoleic acid biosynthesis (Schlueter et al, 2007). To more accurately measure the relative contributions of each component, the levels of each would need to be manually manipulated after preparation of the milk. Knowing the exact contribution of each component would give soybean breeders valuable information on what traits and interactions of traits are important for soymilk consumer acceptability.

Figures

Judge # _____ Session # _____ Date_____

Study of consumer acceptance of soymilks made from lines of soybeans with different seed traits

Introduction

This is a sensory test about unsweetened, unflavored soymilk. Soymilk is a good source of protein and is lactose free, but drinking soymilk is often an acquired taste. To improve the acceptance of soymilk, new varieties of soybeans are being developed to improve product flavor. Your participation in this sensory test will facilitate the development of new varieties of soybeans for soymilk application in the future.

Directions

1. Please answer the following questions:
   What is your gender?
   - ☐ Male
   - ☐ Female
   - ☐ Prefer not to respond

   How often do you consume soymilk?
   - ☐ Every day or every other day
   - ☐ A few times per week
   - ☐ A few times per month
   - ☐ A few times per year
   - ☐ Never
2. Write the number of the soymilk sample on line provided.
3. Please swallow **ALL** of the soymilk at one time for each sample, and rate the sample by making a mark on the line provided.
4. Remember, **do not** re-taste the samples during the test.
5. Thank you for participating in the study!

Figure 3.1a
Front of consumer ballot used in the study

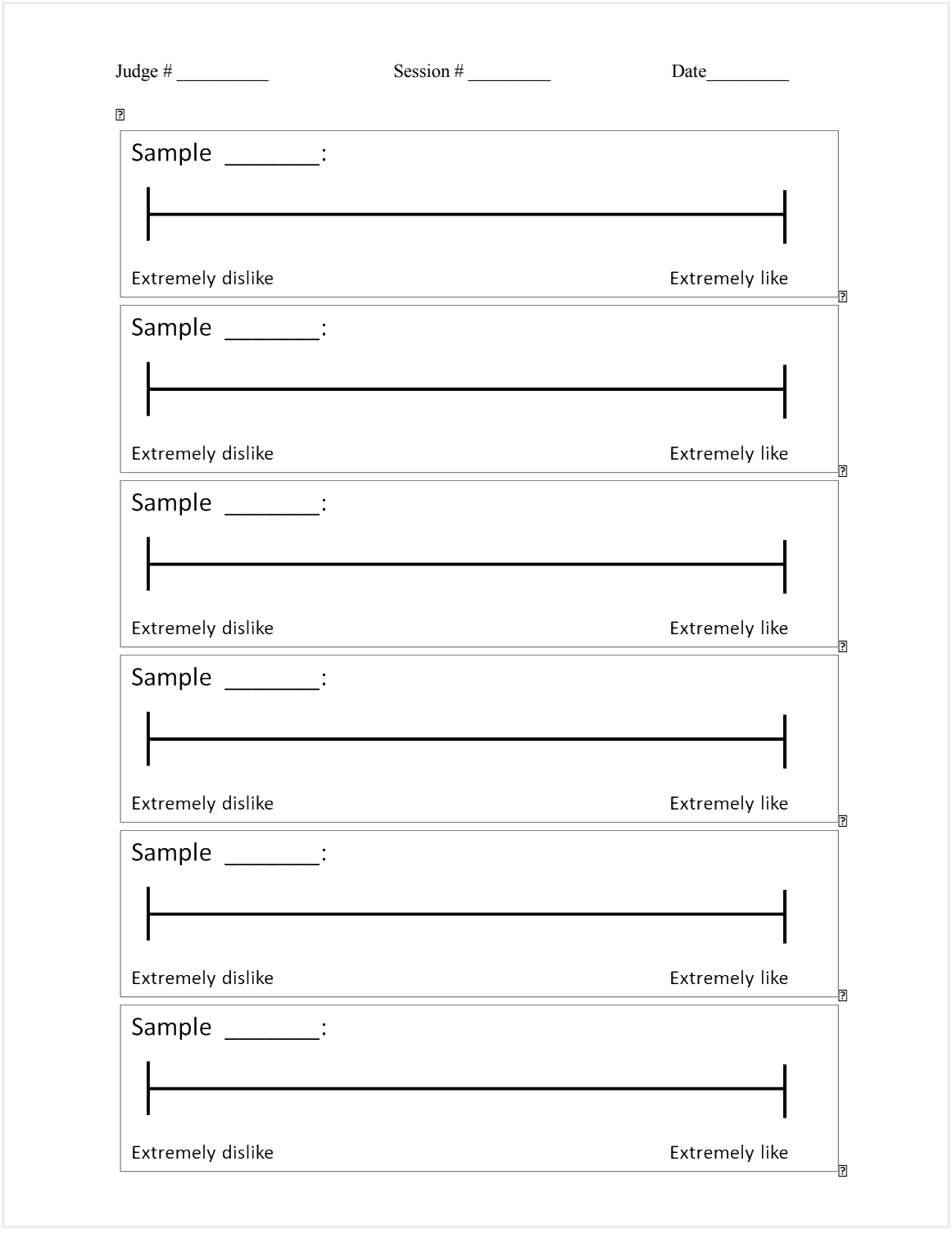


Figure 3.1b
Back of consumer ballot used in the study

Figure 3.2
Consumer preferences of soymilks made from soybeans with differing seed traits
Means with different letters are significantly different at p=0.05.

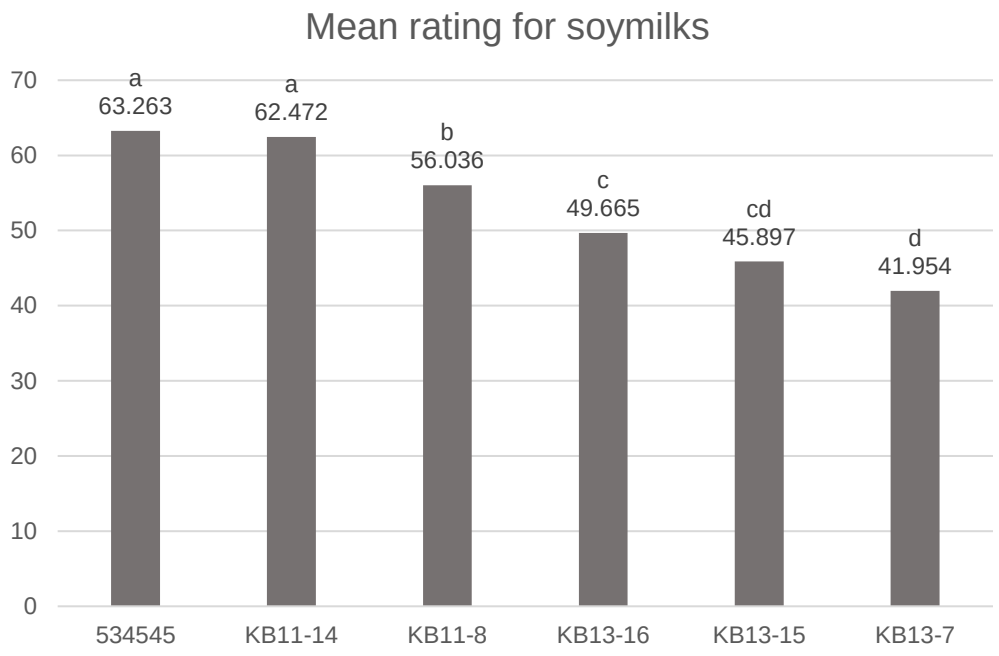


Figure 3.3
Interactions of (a) lipoxygenase-null and low-linolenic traits, and of (b) the lipoxygenase-null trait with total PUFA content

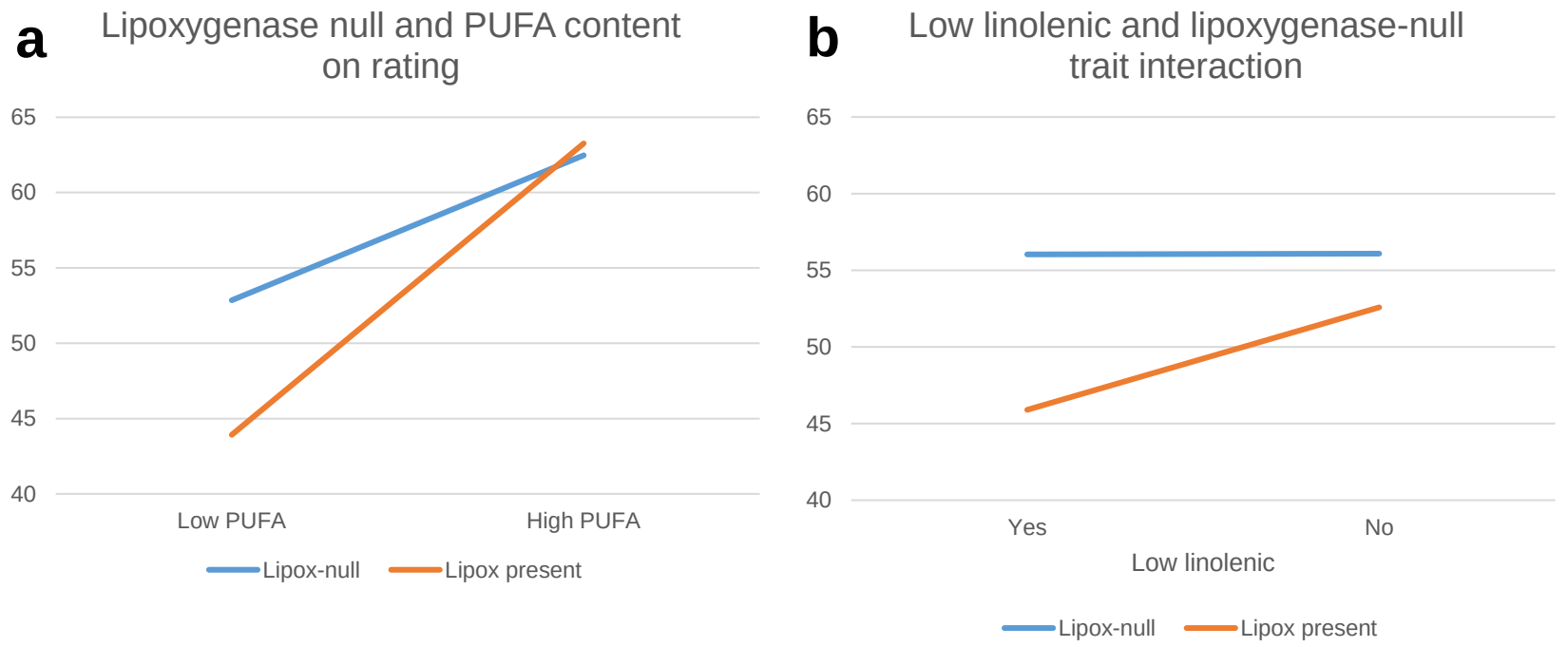


Figure 3.4
Differences in overall mean rating for all soymilks, by (a) gender and (b) typical soymilk consumption.
Means with different letters are significantly different at p=0.05.

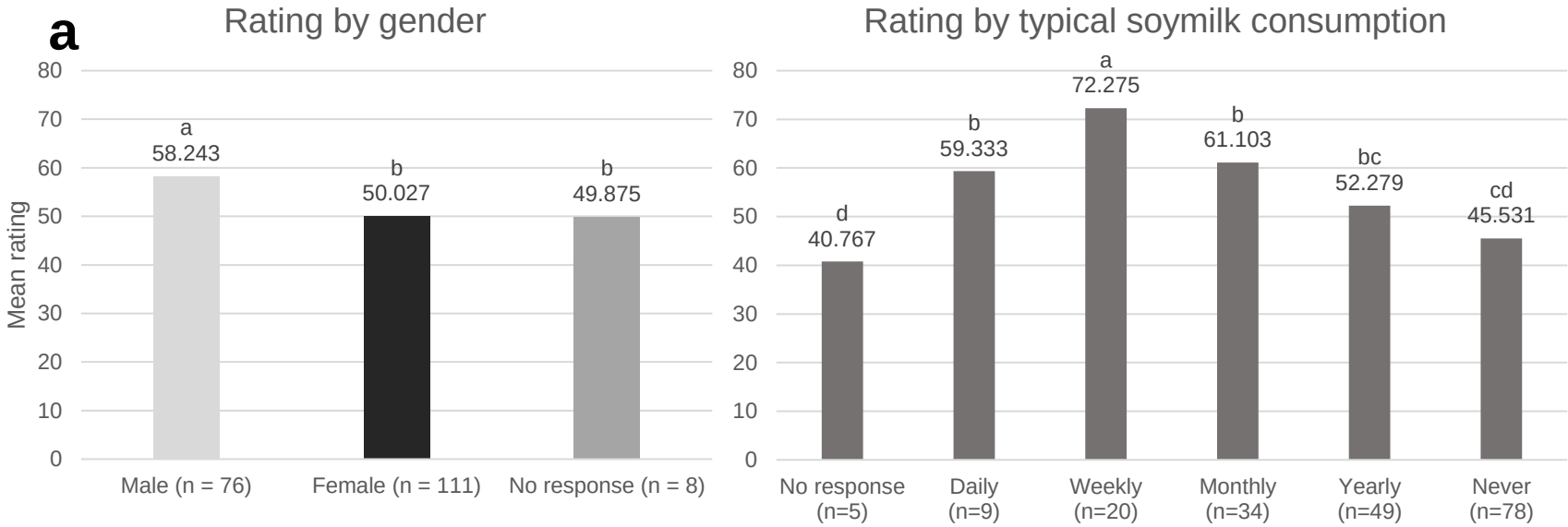
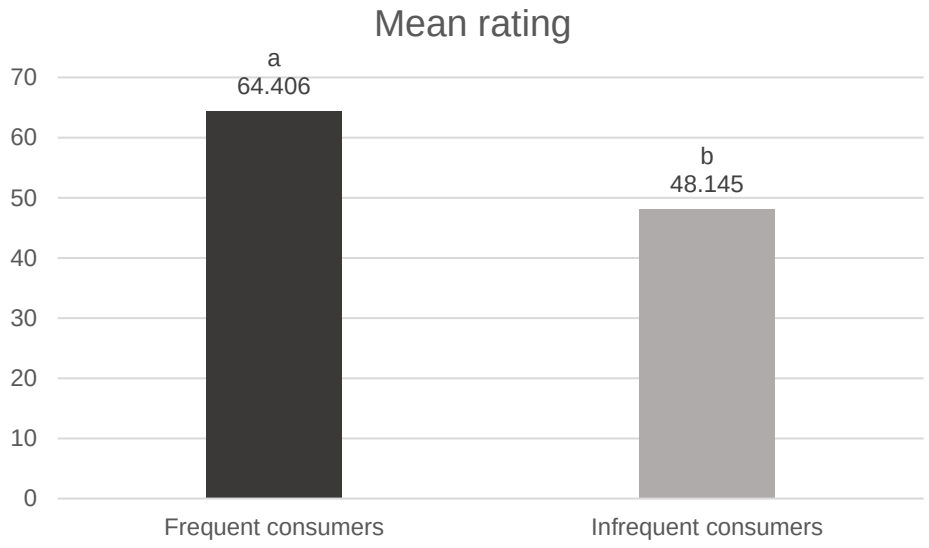


Figure 3.5
Means with different letters are significantly different at p=0.05.



Tables

Table 3.1
Summary of soybean genotypes used in the study

Soybean Line	*High Oleic*	*Low Linolenic*	*Lipoxygenase null*	*Low raffinose*
534545				
KB11-14			X	X
KB11-8	X	X	X	X
KB13-16	X		X	
KB13-15	X	X		
KB13-7	X			X

Table 3.2
Carbohydrate profile of soybean lines used
Values are percent of total seed composition

Soybean Line	*Galactinol*	*Sucrose*	*Raffinose*	*Stachyose*
534545	0.00	8.77	0.52	5.17
KB11-14	0.54	8.31	0.34	0.15
KB11-8	0.11	8.41	0.40	0.74
KB13-16	0.03	5.30	1.14	6.26
KB13-15	0.02	5.99	1.81	6.36
KB13-7	0.90	7.80	0.00	0.17

Table 3.3
Sucrose content of prepared soymilk
[*]*Values are percent of total soymilk or seed composition*

Soybean Line	*% Sucrose in milk*[*]	*% Sucrose in seed*[*]	*Grams sucrose in seed used*	*Grams sucrose in 1.5L milk*	*% Sucrose extraction from seed*
534545	0.624	8.77	12.8	9.4	73.1
KB11-14	0.566	8.31	12.2	8.5	69.3
KB11-8	0.720	8.41	12.1	10.8	89.2
KB13-16	0.478	5.30	7.9	7.2	90.6
KB13-15	0.484	5.99	8.9	7.3	81.5
KB13-7	0.566	7.80	11.5	8.5	73.7

Table 3.4
Lipid profile of soybean lines used
Values are percent of total lipid content

Soybean Line	*16:0*	*18:0*	*18:1*	*18:2*	*18:3*
534545	12.4	4.2	23.5	52.9	7.1
KB11-14	14.3	3.8	21.0	52.6	8.3
KB11-8	10.3	3.1	81.1	2.8	2.8
KB13-16	9.9	3.2	78.9	3.6	4.3
KB13-15	9.2	2.8	81.9	3.1	2.9
KB13-7	9.4	3.4	80.8	2.5	3.9

Table 3.5
Trait effects on rating

Trait	*Mean rating with trait*	*Mean rating without trait*	*p-value*	*Positive or negative effect on rating*
High oleic	48.380	62.866	<.0001	(-)
Low linolenic	50.966	54.333	0.0421	(-)
Low raffinose	53.483	52.942	0.4880	NS
Lipoxygenase null	56.069	50.357	0.0049	(+)

Table 3.6
Correlation of fatty acid components with each other and rating

Pearson Correlation Coefficients, N = 1166
Prob > |r| under H0: Rho=0

	Rating	Palmitic	Oleic	Stearic	Linoleic	Linolenic
Rating	1.00000					
Palmitic	0.19529*	1.00000				
Oleic	-0.18945*	-0.94479*	1.00000			
Stearic	0.15946*	0.77361*	-0.88926*	1.00000		
Linoleic	0.16753*	0.93803*	-0.96420*	0.87330*	1.00000	
Linolenic	0.18974*	0.93453*	-0.99920*	0.88894*	0.95426*	1.00000

*Denotes significance at p=0.05

Table 3.7

Correlation of carbohydrate components with each other and rating

Pearson Correlation Coefficients, N = 1166
Prob > |r| under H0: Rho=0

	Rating	Stachyose	Galactinol	Sucrose	Raffinose
Rating	1.00000				
Stachyose	-0.02487	1.00000			
Galactinol	-0.06939*	-0.76720*	1.00000		
Sucrose	0.13517*	-0.66108*	0.30973*	1.00000	
Raffinose	-0.05960*	0.83155*	-0.67558*	-0.77271*	1.00000

*Denotes significance at p=0.05

References

Beuselinck, P. R., Sleper, D. A., & Bilyeu, K. D. (2006). An Assessment of Phenotype Selection for Linolenic Acid Using Genetic Markers. *Crop Science*, *46*(2), 747. http://doi.org/10.2135/cropsci2005-04-0041

Bourne, M. C. (1976). SURVEY OF SUITABILITY OF THIRTY CULTIVARS OF soyBEANS FOR soyMILK MANUFACTURE. *Journal of Food Science*, *41*(5), 1204–1208. http://doi.org/10.1111/j.1365-2621.1976.tb14418.x

Davies, C. S., Nielsen, S. S., & Nielsen, N. C. (1987). Flavor improvement of soybean preparations by genetic removal of lipoxygenase-2. *Journal of the American Oil Chemists Society*, *64*(10), 1428–1433. http://doi.org/10.1007/BF02636994

Dierking, E. C., & Bilyeu, K. D. (2008). Association of a Soybean Raffinose Synthase Gene with Low Raffinose and Stachyose Seed Phenotype. *The Plant Genome Journal*, *1*(2), 135. http://doi.org/10.3835/plantgenome2008.06.0321

Iassonova, D. R., Johnson, L. A., Hammond, E. G., & Beattie, S. E. (2009). Evidence of an Enzymatic Source of Off Flavors in "Lipoxygenase-Null" Soybeans. *Journal of the American Oil Chemists' Society*, *86*(1), 59–64. http://doi.org/10.1007/s11746-008-1314-y

King, J. M., Svendsen, L. K., Fehr, W. R., Narvel, J. M., & White, R. J. (1998). Oxidative and flavor stability of oil from lipoxygenase-free soybeans. *Journal of the American Oil Chemists' Society*, *75*(12), 1121–1126. http://doi.org/10.1007/s11746-998-0300-8

Kitamura, K. (1984). Biochemical Characterization of Lipoxygenase Lacking Mutants, L-l-less, L-2-less, and L-3-less Soybeans. *Agricultural and Biological Chemistry*, *48*(9), 2339–2346. http://doi.org/10.1080/00021369.1984.10866488

Kris-Etherton, P. M., Pearson, T. A., Wan, Y., Hargrove, R. L., Moriarty, K., Fishell, V., & Etherton, T. D. (1999). High-monounsaturated fatty acid diets lower both plasma cholesterol and triacylglycerol concentrations. *The American Journal of Clinical Nutrition*, *70*(6), 1009–1015.

Kulkarni, D. S., Kapanoor, S. S., Girigouda, K., Kote, N. V., & Mulimani, V. H. (2006). Reduction of flatus-inducing factors in soymilk by immobilized alpha-galactosidase. *Biotechnology and Applied Biochemistry*, *45*(Pt 2), 51–57. http://doi.org/10.1042/BA20060027

Ku, S., Wei, L. S., Steinberg, M. P., Nelson, A. I., & Hymowitz, T. (1976). Extraction of oligosaccharides during cooking of whole soybeans. *Journal of Food Science*, *41*(2), 361–364. http://doi.org/10.1111/j.1365-2621.1976.tb00619.x

Kwok, K.-C., Liang, H.-H., & Niranjan, K. (2002). Optimizing conditions for thermal processes of soy milk. *Journal of Agricultural and Food Chemistry*, *50*(17), 4834–4838.

Lawrence, S. E., Lopetcharat, K., & Drake, M. A. (2016). Preference Mapping of Soymilk with Different U.S. Consumers: Preference mapping of soymilk…. *Journal of Food Science*, *81*(2), S463–S476. http://doi.org/10.1111/1750-3841.13182

Lin, F. M., & Wilkens, W. F. (1970). Volatile Flavor Components of Coconut Meat. *Journal of Food Science*, *35*(5), 538–539. http://doi.org/10.1111/j.1365-2621.1970.tb04802.x

Månsson, H. L. (2008). Fatty acids in bovine milk fat. *Food & Nutrition Research*, *52*(0). http://doi.org/10.3402/fnr.v52i0.1821

Matsuura, M., Obata, A., & Fukushima, D. (1989). Objectionable Flavor of Soy Milk Developed during the Soaking of Soybeans and its Control. *Journal of Food Science*, *54*(3), 602–605. http://doi.org/10.1111/j.1365-2621.1989.tb04662.x

Murugkar, D. A. (2014). Effect of sprouting of soybean on the chemical composition and quality of soymilk and tofu. *Journal of Food Science and Technology*, *51*(5), 915–921. http://doi.org/10.1007/s13197-011-0576-9

Naczk, M., Amarowicz, R., & Shahidi, F. (1997). α-Galactosides of Sucrose in Foods: Composition, Flatulence-Causing Effects, and Removal. In F. Shahidi (Ed.), *Antinutrients and Phytochemicals in Food* (Vol. 662, pp. 127–151). Washington, DC: American Chemical Society. Retrieved from http://pubs.acs.org/doi/abs/10.1021/bk-1997-0662.ch008

Næs, T., Brockhoff, P. B., & Tomić, O. (2010). *Statistics for sensory and consumer science*. Chichester, West Sussex: Wiley.

N'Kouka, K. D., Klein, B. P., & Lee, S.-Y. (2004). Developing a Lexicon for Descriptive Analysis of Soymilks. *Journal of Food Science*, *69*(7), 259–263. http://doi.org/10.1111/j.1365-2621.2004.tb13625.x

Palacios, O. M., Badran, J., Drake, M. A., Reisner, M., & Moskowitz, H. R. (2009). Consumer acceptance of cow's milk versus soy beverages: impact of ethnicity, lactose tolerance and sensory preference segmentation. *Journal of Sensory Studies*, *24*(5), 731–748. http://doi.org/10.1111/j.1745-459X.2009.00236.x

Pangborn, R. M., Bos, K. E., & Stern, J. S. (1985). Dietary fat intake and taste responses to fat in milk by under-, normal, and overweight women. *Appetite*, *6*(1), 25–40.

Pham, A.-T., Lee, J.-D., Shannon, J. G., & Bilyeu, K. D. (2010). Mutant alleles of FAD2-1A and FAD2-1B combine to produce soybeans with the high oleic acid seed oil trait. *BMC Plant Biology*, *10*(1), 195. http://doi.org/10.1186/1471-2229-10-195

Phillips, L. G., Mcgiff, M. L., Barbano, D. M., & Lawless, H. T. (1995). The Influence of Fat on the Sensory Properties, Viscosity, and Color of Lowfat Milk. *Journal of Dairy Science*, *78*(6), 1258–1266. http://doi.org/10.3168/jds.S0022-0302(95)76746-7

Powers, J.J. (1988). Current practices and application of descriptive methods. In *Sensory Analysis of Foods* (J.R. Pigott, ed., pp. 187–266). Barking, UK: Elsevier.

Rao, M.S.S, Mullinix, B.G., Rangappa, M., Cebert, E., Bhag-sari, A.S., Sapra, V.T., … Dadson, R.B. (2002). Genotype × environment interactions and yield stability of food-grade soybean genotypes., *Agronomy Journal*(94), 72–80.

Rohm, H., Kovac, A., & Kneifel, W. (1994). Effects of starter cultures on sensory properties of set-style yoghurt determined by quantitative descriptive analysis. *Journal of Sensory Studies*, *9*(2), 171–186. http://doi.org/10.1111/j.1745-459X.1994.tb00239.x

Schlueter, J. A., Vasylenko-Sanders, I. F., Deshpande, S., Yi, J., Siegfried, M., Roe, B. A., … Shoemaker, R. C. (2007). The FAD2 Gene Family of Soybean. *Crop Science*, *47*(Supplement_1), S–14. http://doi.org/10.2135/cropsci2006.06.0382tpg

Schyver, T., & Smith, C. (2005). Reported Attitudes and Beliefs toward Soy Food Consumption of Soy Consumers versus Nonconsumers in Natural Foods or Mainstream Grocery Stores. *Journal of Nutrition Education and Behavior*, *37*(6), 292–299. http://doi.org/10.1016/S1499-4046(06)60159-0

Torres-Penaranda, A. V., Reitmeier, C. A., Wilson, L. A., Fehr, W. R., & Narvel, J. M. (2006). Sensory Characteristics of Soymilk and Tofu Made from Lipoxygenase-Free and Normal

Soybeans. *Journal of Food Science*, *63*(6), 1084–1087. http://doi.org/10.1111/j.1365-2621.1998.tb15860.x

Tu, V. P., Valentin, D., Husson, F., & Dacremont, C. (2010). Cultural differences in food description and preference: Contrasting Vietnamese and French panellists on soy yogurts. *Food Quality and Preference*, *21*(6), 602–610. http://doi.org/10.1016/j.foodqual.2010.03.009

van der Ven, C., Matser, A. M., & van den Berg, R. W. (2005). Inactivation of soybean trypsin inhibitors and lipoxygenase by high-pressure processing. *Journal of Agricultural and Food Chemistry*, *53*(4), 1087–1092. http://doi.org/10.1021/jf048577d

Villegas, B., Carbonell, I., & Costell, E. (2009). Acceptability of Milk and Soymilk Vanilla Beverages: Demographics Consumption Frequency and Sensory Aspects. *Food Science and Technology International*, *15*(2), 203–210. http://doi.org/10.1177/1082013208105166

Wilcox, J. R., Cavins, J. F., & Nielsen, N. C. (1984). Genetic alteration of soybean oil composition by a chemical mutagen. *Journal of the American Oil Chemists' Society*, *61*(1), 97–100. http://doi.org/10.1007/BF02672055

Wu, Y., Fontenot Molaison, E., Pope, J. F., & Reagan, S. (2005). Attitudes and acceptability of soy-based yogurt by college students. *Nutrition & Food Science*, *35*(4), 253–257. http://doi.org/10.1108/00346650510605649

Xu, B., Chang, S. K. C., Liu, Z., Yuan, S., Zou, Y., & Tan, Y. (2010). Comparative studies on the chemical and cell-based antioxidant activities and antitumor cell proliferation properties of soy milk manufactured by conventional and commercial UHT methods.

Journal of Agricultural and Food Chemistry, *58*(6), 3558–3566. http://doi.org/10.1021/jf903796c

Yang, A., Smyth, H., Chaliha, M., & James, A. (2016). Sensory quality of soymilk and tofu from soybeans lacking lipoxygenases. *Food Science & Nutrition*, *4*(2), 207–215. http://doi.org/10.1002/fsn3.274

Yuan, S., & Chang, S. K.-C. (2007). Selected Odor Compounds in Soymilk As Affected by Chemical Composition and Lipoxygenases in Five Soybean Materials. *Journal of Agricultural and Food Chemistry*, *55*(2), 426–431. http://doi.org/10.1021/jf062274x

Yuan, S., Chang, S. K. C., Liu, Z., & Xu, B. (2008). Elimination of Trypsin Inhibitor Activity and Beany Flavor in Soy Milk by Consecutive Blanching and Ultrahigh-Temperature (UHT) Processing. *Journal of Agricultural and Food Chemistry*, *56*(17), 7957–7963. http://doi.org/10.1021/jf801039h

Zhang, Y., & Chang, S. K. C. (2016). Isoflavone Profiles and Kinetic Changes during Ultra-High Temperature Processing of Soymilk: Isoflavone profiles and kinetic changes …. *Journal of Food Science*, *81*(3), C593–C599. http://doi.org/10.1111/1750-3841.13236

Zhang, Y., Guo, S., Liu, Z., & Chang, S. K. C. (2012). Off-Flavor Related Volatiles in Soymilk As Affected by Soybean Variety, Grinding, and Heat-Processing Methods. *Journal of Agricultural and Food Chemistry*, *60*(30), 7457–7462. http://doi.org/10.1021/jf3016199

Chapter 4

Digestibility of full-fat, reduced oligosaccharide soybean determined with the precision-fed cecectomized rooster assay

Abstract

Soybeans are an important protein and oil crop, and are a main component of poultry diets. However, soy has several anti-nutritional factors that limit its use in diet formulations, namely the content of indigestible oligosaccharides such as raffinose and stachyose and low true metabolizable energy. Reducing raffinose and stachyose, and increasing sucrose levels is hypothesized to have a positive effect on the true metabolizable energy of full-fat soybean meal. This work used two sources of soybean meal for comparison: a wild-type line, and a genetically enhanced low-raffinose, lipoxygenase-null line developed at the University of Missouri. The objective of this study was to evaluate the true metabolizable energy of the two sources of soybean meal. Although there was no significant difference in passage rate or dry matter digestibility for the two soybean sources, the excreta from birds fed low-oligosaccharide soy meal had less gross energy than wild-type-fed birds. Also, the low-oligosaccharide soybean meal had more gross energy than wild-type. Therefore, there is more digestible energy in low- oligosaccharide soybean than wild-type soybean. Low- oligosaccharide soy had a measured TME of 2,797 kcal/kg, compared with 2,330 kcal/kg for wild-type. As low digestible energy is a major limiting factor in the percent of soybean meal that can be used in poultry diets, these results substantiate the use of higher concentrations of low-oligosaccharide, full-fat soy in the formulated diets.

Introduction

Soybean meal is one of the primary components of poultry diets, and supplies most of the protein requirement for the animal. Soy is a superior source of inexpensive, easily digested protein, and has a both high quantity and high quality amino acid profile. However, although soybean meal has about 10% more gross energy than corn, the metabolizable energy in soy is only about 72% that of corn (Dale, 2000). A large reason for the relatively low metabolizable energy content of soy is the presence of indigestible oligosaccharides in the carbohydrate fraction, namely raffinose and stachyose. Poultry are monogastric animals, and cannot digest the α-1,6-glycosidic bond in raffinose and stachyose because they lack the α-galactosidase enzyme in the foregut (Gitzelmann and Auricchio, 1965). Bacterial fermentation of raffinose and stachyose in the hindgut are known to cause diarrhea in pigs and decreased weight gain for poultry (Zhang et al, 2001; Perryman et al, 2013). Further, oligosaccharides have been shown to increase passage rate of the diet through the digestive system, causing a reduction in digestion and absorption of other essential nutrients (Wiggins, 1984; Leske et al, 1991).

Several methods have been tested to reduce the negative effects of soy oligosaccharides on poultry nutrition. Exogenous α-galactosidase enzyme could be added to the diet to aid in digestion, but thus far no studies using the method have been successful (Irish et al, 1995; Graham et al, 2002; Waldroup et al, 2006). Ethanol extraction of the meal has been shown to significantly increase TME and digestibility in at least one study (Coon et al, 1990), however in another study, TME was not affected by ethanol extraction (Irish et al, 1995). Removal of oligosaccharides from the soybean seed using plant breeding has been

effective; lines of soybean have been developed with dramatically reduced oligosaccharide content and increased sucrose content (Dierking and Bilyeu, 2008).

The genetically improved, reduced-oligosaccharide soybean varieties were tested in Parsons et al (2000) and Chen et al (2013) and they saw a 5 to 7.5% increase in true metabolizable energy for the reduce-oligosaccharide meals. Other studies have evaluated the effects of adding raffinose and stachyose to poultry diets, and found an overall negative impact of these compounds on animal health and nutrition (Leske, 1993).

Properly heated full-fat soybean meal has a higher metabolizable energy value than conventional soybean meal (SBM) – (3,300 kcal/kg for full-fat vs 2,440 kcal/kg for regular SBM (NRC, 1994). Also, studies have shown replacement of some of the diet's regular SBM with full-fat meal has a positive impact on turkey (MacIssac et al, 2005) and chicken (Hamilton and McNiven, 2000) growth. Producing full-fat soybean meal requires less processing than regular soybean meal, and as US poultry production is centered in regions that are also high in soybean production (Waldroup, 1982), the use of on-farm produced whole-bean soybean meal may be a viable economic option. Reduction of oligosaccharides in the soybean using plant breeding could make full-fat soybean meal an even more viable option for poultry producers. The objective of this research was to evaluate the metabolizable energy of two sources of full-fat soybean meal: a conventional variety with wild-type levels of oligosaccharides, and a genetically improved low-oligosaccharide variety.

Materials and Methods

Soybean composition analysis

Soybean sucrose content was measured by enzymatic assay using the Megazyme® kit (Megazyme, Bray, Ireland). Briefly, total carbohydrate was extracted from 12.5mg of lyophilized, ground soybean with 1mL of 50% ethanol for 30 minutes at 70°C. After centrifugation to remove solids, 100 uL of supernatant was diluted with 100uL of H_2O. 10uL of the dilution was mixed with 10uL of Megazyme Solution II in a new 96-well plate and incubated at 50°C for 20 minutes. 20uL of standards were prepared using the supplied 1mg/mL glucose standard, and 150uL of GOPOD reagent was added to all wells. After incubation at 50°C for 20 minutes, 125uL of the sample was analyzed on a spectrophotometer at 510nm. The assay measures glucose concentration, and percent sucrose was calculated assuming no endogenous glucose in the seed. Galactinol, raffinose, and stachyose content in seeds were determined by HPLC. 12.5mg of lyophilized, ground seed was extracted with 1mL of 50% ethanol at 70°C for 30 minutes with occasional agitation. 50 uL of the extraction supernatant vacuum-dried to remove ethanol, and then resusupended in 250uL of ddH_2O. 10uL of the resuspension was used for injection for HPLC.

Soybean meal preparation

Whole soybeans were roasted at 55ºC for one hour and allowed to cool, then ground in a Wiley laboratory mill fitted with a 3mm screen one day prior to feeding. The ground

soybean was weighed into 30 gram portions and transported to the poultry research farm for feeding.

Precision-feeding assay

All experimental protocols were reviewed and approved by the University of Missouri Animal Care and Use Committee, and the experiment operated under permit #8477. Birds were housed in individual cages with raised wire floors in a well-ventilated barn, and given free access to water. 14-week-old Single Comb White Leghorn roosters were obtained from a local pullet farm, and at around 18 weeks of age, cecectomy surgeries were performed following the procedures described in Parsons (1985). After four weeks of recovery, the precision feeding assay was conducted using 14 total roosters. Roosters were weighed and separated into two groups with each group having approximately equal total bird weight. Roosters were fasted for 24 hours to empty the digestive tract and then tube-fed 30 grams each of wild-type or low-raffinose soybean. Excreta was collected every four hours for the first 24 hours, and then again at the 48-hour mark. A second experiment was conducted without feeding to measure the endogenous excreta loss during a 48-hour period following 24 hours of fast.

Sample preparation and analysis

After drying the excreta samples for 24 hours at 65ªC, weights were recorded and then all samples from each bird were pooled. Dried excreta were ground with a 1mm screen and gross energy was measured by a Parr model 1341 oxygen bomb calorimeter and model 1108 oxygen combustion vessel calibrated with benzoic acid (Parr Instrument Company,

Moline, IL). Dry matter digestibility and true metabolizable energy were calculated as in Sibbald (1976):

$$TME\ (kcal/kg) = \frac{(GE_f \times F_i) - (GE_e \times (Y_f - Y_e))}{F_i}$$

Where, GE_e = gross energy of the excreta (kcal/kg)

GE_f = gross energy of the feed (kcal/kg)

F_i = feed input (g)

Y_f = excreta from fed bird (g)

Y_e = excreta from unfed bird (g)

Statistical analysis

SAS software (Version 9.4) was used for all data analysis. Means were separated by PROC ANOVA and results were considered significant at p=0.05.

Results

Soybean composition

While wild-type seeds had more total soluble carbohydrates than low-oligosaccharide seed, 11.4g and 9.34g, respectively, the proportion of digestible carbohydrates was much higher for low-oligosaccharide seed. The low-oligosaccharide soybean had about half the raffinose, and less than 5% of the stachyose of wild-type. Also, low-oligosaccharide seed had about 20% more sucrose than wild-type (Table 4.1). The major reduction in indigestible oligosaccharides and increase in sucrose are expected knowing the genotypes of these soybean varieties. Also, the gross energy of the low-oligosaccharide soybean meal as measured by bomb calorimetry was higher (5080 kcal/kg) than wild-type soybean meal (4690 kcal/kg).

Passage rate of feedstuff

Because there is evidence that oligosaccharide content in poultry diets increases passage rate, samples were collected every four hours for the first 24 hours after feeding. The results are summarized in Figure 4.1, and no significant difference in passage rate was noted between the two soybean sources.

Dry matter digestibility

First, endogenous loss was calculated to correct for the amount of endogenous urine and fecal energy lost during the 48-hour period the test was conducted. Total endogenous loss for 48 hours of fasting averaged 3.482g of dry excreta per bird, and that value was subtracted from the after-feeding excreta weight for each bird. The dry matter digestibility

of the soybean meals was not significantly different between the two sources, wild-type soybeans were digested at 38.975% and low-oligosaccharide at 37.710% on a dry matter basis.

True metabolizable energy

True metabolizable energy was calculated, and although dry matter digestibility was not significantly different between the soybean sources, the higher gross energy of low-oligosaccharide soybean meal and lower gross energy of excreta from low-oligosaccharide-fed birds gave a higher TME value to low-oligosaccharide soy. The TME for low-oligosaccharide soybeans was 2797 kcal/kg, compared with 2330 kcal/kg for wild-type beans, for an increase of 467 kcal or 20% (Table 4.2). The means are significantly different, with a p value of 0.0307.

Discussion

The true metabolizable energy of full-fat, low-oligosaccharide soybean meal was significantly higher than full-fat conventional soybean meal at 2797 and 2330 kcal/kg, respectively. The increase in TME for low-oligosaccharide soybean meal was 466 kcal, or 20% increase compared to wild-type. These results are partially explained because the gross energy of the low-oligosaccharide soybean meal was higher than wild-type by 8.3%. The difference in gross energy between the soybean meals could be due to low sensitivity of the bomb calorimeter, as the model used only measured temperature increase in increments of 0.1ªC. If a more reliable value for gross energy of the soybean meal was obtained, TME for the meals could be more accurately estimated.

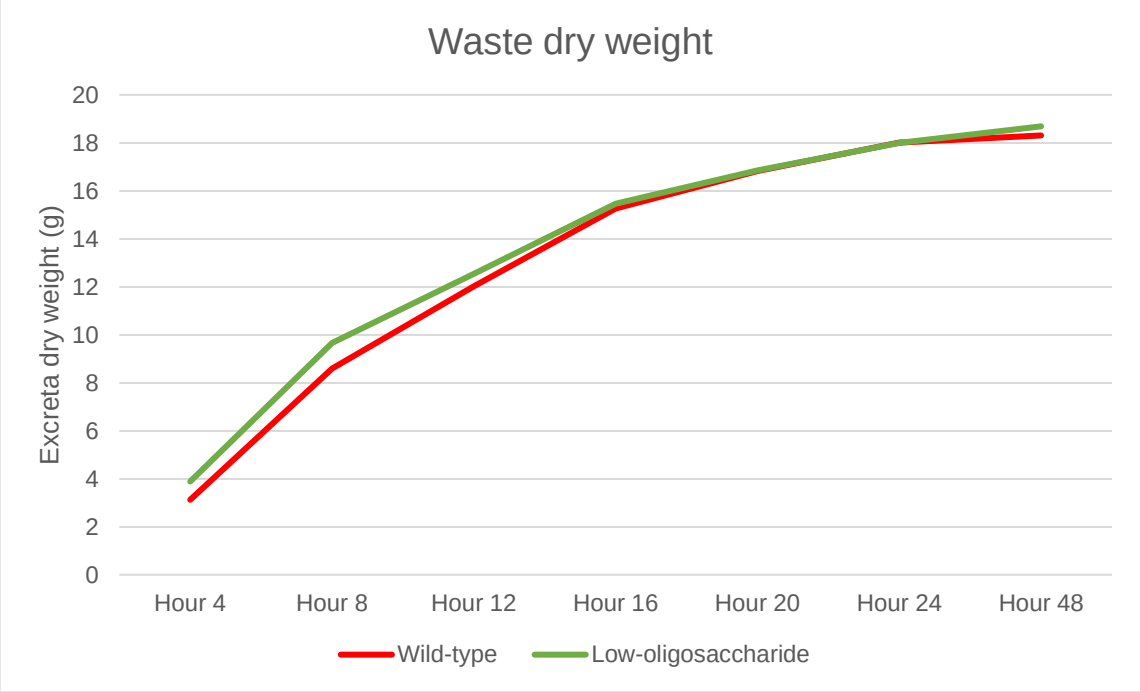
Values obtained from this study were similar to TME values listed in the 9th Edition of the NRC for "soybean seeds, heat processed" at 2,990 kcal/kg. The gross energy values measured in this study are comparable to full-fat gross energy values measured in Longo, et al (2004). Further, Chen et al, 2013 measured gross energy of processed conventional soybean meal at 4643 kg/kcal and reduced-oligosaccharide at 4,690 kcal/kg. Full-fat soybean meal retains its oil content, so it has more energy than processed soybean meal (NRC, 1994), and the TME values of the current study are greater than Chen's results, as expected.

This study provides a baseline for the use of full-fat, reduced-oligosaccharide soybean meal in chicken diets. However future experiments should be done using soybeans grown side-

by-side in the field during a growing season to better measure the contribution of only carbohydrate profile on true metabolizable energy. Oligosaccharide content and overall low metabolizable energy is a deterrent of using high percentages of soybean meal in poultry diets, but the use of genetically improved, low-oligosaccharide soy as a whole-bean protein, oil, and carbohydrate source may increase metabolizable energy in the diet while maintaining a low formulation cost.

Figures

Figure 4.1
Passage rate of 30g samples of soybean meal. No significant differences between wild-type and low-oligosaccharide soybean were observed for any time point.



Tables

Table 4.1
Carbohydrate composition of soybean lines used in the study

Soybean Line	*Galactinol*	*Sucrose*	*Raffinose*	*Stachyose*
Low-raffinose	0.54	8.31	0.34	0.15
Wild-type	0.03	6.95	0.63	3.79

Table 4.2
Calculations to estimate true metabolizable energy (TME)
GE = gross energy

Feed calculations

	GE of feed (cal/g)	Weight of feed (g)	kcal in feed per bird
Wild-type	4690.388	30	140.712
Low-oligo	5079.803	30	152.394

Excreta calculations

	GE of excreta (cal/g)	Weight of excreta (g)	Endogenous loss (g)	kcal in waste per bird
Wild-type	4771.262	18.307	3.482	70.799
Low-oligo	4503.950	18.687	3.482	68.491

TME calculations

	kcal absorbed by bird from 30 g feed	TME of feed (kcal/g)	**TME (kcal/kg)**
Wild-type	69.913	2.330	**2330.427[b]**
Low-oligo	83.903	2.797	**2796.770[a]**

p=0.0307

References

Chen, X., Parsons, C. M., & Bajjalieh, N. (2013). Nutritional evaluation of new reduced oligosaccharide soybean meal in poultry. *Poultry Science*, *92*(7), 1830–1836. http://doi.org/10.3382/ps.2012-02856

Coon, C. N., Leske, K. L., Akavanichan, O., & Cheng, T. K. (1990). Effect of Oligosaccharide-Free Soybean Meal on True Metabolizable Energy and Fiber Digestion in Adult Roosters. *Poultry Science*, *69*(5), 787–793. http://doi.org/10.3382/ps.0690787

Dale, N. M. (200AD). Soy products as protein sources in poultry diets. In *Animal Nutrition* (J. K. Drackley, pp. 283–288). Savoy, Il: Federation of Animal Science Societies.

Dale, N. M. (2006). *Feedstuffs Ingredient Analysis Table.* Minnetonka, MN: Miller Publishing Co.

Dierking, E. C., & Bilyeu, K. D. (2008). Association of a Soybean Raffinose Synthase Gene with Low Raffinose and Stachyose Seed Phenotype. *The Plant Genome Journal*, *1*(2), 135. http://doi.org/10.3835/plantgenome2008.06.0321

Gitzelmann, R., & Auricchio, S. (1965). THE HANDLING OF SOYA ALPHA-GALACTOSIDES BY A NORMAL AND A GALACTOSEMIC CHILD. *Pediatrics*, *36*, 231–235.

Graham, K., Kerley, M., Firman, J., & Allee, G. (2002). The effect of enzyme treatment of soybean meal on oligosaccharide disappearance and chick growth performance. *Poultry Science*, *81*(7), 1014–1019. http://doi.org/10.1093/ps/81.7.1014

Hamilton, R. M. G., & McNiven, M. A. (2000). Replacement of soybean meal with roasted full-fat soybeans from high-protein or conventional cultivars in diets for broiler chickens. *Canadian Journal of Animal Science*, *80*(3), 483–488. http://doi.org/10.4141/A99-064

Irish, G. G., Barbour, G. W., Classen, H. L., Tyler, R. T., & Bedford, M. R. (1995). Removal of the alpha-galactosides of sucrose from soybean meal using either ethanol extraction or exogenous alpha-galactosidase and broiler performance. *Poultry Science*, *74*(9), 1484–1494.

Leske, K. L., Akavanichan, O., Cheng, T. K., & Coon, C. N. (1991). Effect of Ethanol Extract on Nitrogen-Corrected True Metabolizable Energy for Soybean Meal with Broilers and Roosters. *Poultry Science*, *70*(4), 892–895. http://doi.org/10.3382/ps.0700892

Leske, K. L., Jevne, C. J., & Coon, C. N. (1993). Effect of oligosaccharide additions on nitrogen-corrected true metabolizable energy of soy protein concentrate. *Poultry Science*, *72*(4), 664–668.

Longo, F., Menten, J., Pedroso, A., Figueiredo, A., Racanicci, A., Gaiotto, J., & Sorbara, J. (2004). Determination of the energetic value of corn, soybean meal and micronized full fat soybean for newly hatched chicks. *Revista Brasileira de Ciência Avícola*, *6*(3), 147–151. http://doi.org/10.1590/S1516-635X2004000300003

MacIsaac, J. L., Burgoyne, K. L., Anderson, D. M., & Rathgeber, B. R. (2005). Roasted Full-Fat Soybeans in Starter, Grower, and Finisher Diets for Female Broiler Turkeys. *The Journal of Applied Poultry Research*, *14*(1), 116–121. http://doi.org/10.1093/japr/14.1.116

NRC. (1994). *Nutrient Requirements of Poultry. 9th rev. ed.* Washington, DC: National Academy Press.

Parsons, C. M. (1985). Influence of caecectomy on digestibility of amino acids by roosters fed distillers' dried grains with solubles. *The Journal of Agricultural Science*, *104*(02), 469. http://doi.org/10.1017/S0021859600044178

Parsons, C. M., Zhang, Y., & Araba, M. (2000). Nutritional evaluation of soybean meals varying in oligosaccharide content. *Poultry Science*, *79*(8), 1127–1131.

Perryman, K. R., Olanrewaju, H., & Dozier, W. A. (2013). Growth performance and meat yields of broiler chickens fed diets containing low and ultra-low oligosaccharide soybean meals during a 6-week production period. *Poultry Science*, *92*(5), 1292–1304. http://doi.org/10.3382/ps.2012-02723

Sibbald, I. R. (1976). A bioassay for true metabolizable energy in feedingstuffs. *Poultry Science*, *55*(1), 303–308.

Waldroup, P. W. (1982). Whole soybeans for poultry feeds. *World's Poultry Science Journal*, *38*(1), 28–35.

Waldroup, P. W., Keen, C. A., Yan, F., & Zhang, K. (2006). The Effect of Levels of -Galactosidase Enzyme on Performance of Broilers Fed Diets Based on Corn and Soybean Meal1. *The Journal of Applied Poultry Research*, *15*(1), 48–57. http://doi.org/10.1093/japr/15.1.48

Wiggins, H. S. (1984). Nutritional value of sugars and related compounds undigested in the small gut. *Proceedings of the Nutrition Society*, *43*(01), 69–75. http://doi.org/10.1079/PNS19840029

Zhang, L., Li, D., Qiao, S., Wang, J., Bai, L., Wang, Z., & Han, I. K. (2001). The Effect of Soybean Galactooligosaccharides on Nutrient and Energy Digestibility and Digesta

Transit Time in Weanling Piglets. *Asian-Australasian Journal of Animal Sciences*, *14*(11), 1598–1604. http://doi.org/10.5713/ajas.2001.1598

Vita

Michelle Folta is the daughter of Jerry and Zenda Folta of Hawk Point, Missouri. She graduated top of her class from Troy Buchanan High School in 2008, and then came to the University of Missouri to pursue higher education. During undergraduate work, she worked in the MU Crop Physiology Lab, advised by Dr. Felix Fritschi, where she discovered a passion for plant science research and crop improvement. She also worked one summer as a Trait Development Intern at Monsanto in Jerseyville, Illinois and another summer as a Field Intern for DuPont-Pioneer in Macomb, Illinois. In 2011, she graduated summa cum laude from MU with a bachelor's degree in Plant Science and a minor in Chemistry.

After undergraduate work, she chose to pursue a doctoral degree also from MU. Her primary work has been in the lab of Dr. Zhanyuan Zhang and the MU Plant Transformation Core Facility working with soybean seed trait improvements using transgenic strategies. She also had a summer rotation experience in the lab of Dr. Kristin Bilyeu learning techniques of soybean breeding and gene mapping. Michelle's graduate work has been supported by the MU Life Sciences Fellowship.