Improving Breeding Selection of Seed Quality Traits for Food-Grade Soybeans

Diana Marcela Escamilla Sanchez

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

Master of Science
In
Crop and Soil Environmental Sciences

Bo Zhang, Chair
M.A. Saghai Maroof
Laura K. Strawn
Qijian Song

December 7, 2017
Blacksburg, VA

Keywords: Soyfood, fungi, sprouts, natto, seed coat deficiency, seed quality, small-seeded, mycotoxin, disinfection treatment, QTL, Molecular assisted selection, breeding.

Improving Breeding Selection of Seed Quality Traits for Food-Grade Soybeans

Diana Marcela Escamilla Sanchez

ACADEMIC ABSTRACT

Natto and sprout soybeans are produced using small-seeded soybeans and their production is a high value alternative to grow grain soybeans for food in U.S. The development of soybean cultivars with improved natto and sprout quality is crucial for maintaining and increasing the soyfood market. However, there is insufficient information on sprout soybean characteristics. Therefore, the first objective of this study was to evaluate seed and sprout traits as potential selection criteria and study the storage effect on sprout quality. Seeds can be a vehicle for transmission of pathogens capable of causing human illness. That is why, the second objective was to identify seed-borne pathogens on a commercial soybean cultivar and to evaluate different seed decontamination treatments. Finally, seed coat deficiency is an undesirable trait for natto soybean seeds because it causes inferior appearance of the product. Thus, the third objective was to identify quantitative trait loci (QTL) underlying seed coat deficiency (SCD) and associated markers. Results showed that seed size, high-, average- and low-quality sprout percentage, hypocotyl thickness and length and sprout yield are the most important variables for breeding sprout cultivars; and one-year seed storage at room temperature reduced sprout quality. *Fusarium, Alternaria* and *Diaphorte* were the most frequent genera isolated from soybean seeds, and 2% calcium hypochlorite and 5% acetic acid were promising seed disinfection treatments. A stable QTL, *qSCD20_1,* was identified across two years explaining up to 25% of the variation of SCD; and eight molecular markers tightly linked and nearby *qSCD20_1* were identified. Information presented will be helpful for sprout and natto soybean cultivar development.

Improving Breeding Selection of Seed Quality Traits for Food-Grade Soybeans

Diana Marcela Escamilla Sanchez

GENERAL ABSTRACT

Natto and sprout soybeans are produced using by small-seeded soybeans and their production is a high value alternative to grow grain soybeans for feed in U.S. The development of soybean cultivars with improved natto and sprout quality is crucial for maintaining and increasing the soyfood market. However, there is insufficient information on sprout soybean characteristics. Therefore, the first objective of this study was to evaluate important seed and sprout traits as potential selection criteria of soybean cultivars and study the storage effect on sprout quality. Seeds can be contaminated with fungi capable of causing human illness. So, the second objective was to identify fungus species associated with seeds of a sprout soybean cultivar and evaluate different seed disinfection treatments. Finally, seed coat deficiency is an undesirable trait for natto soybean seeds because it causes inferior appearance of the product. Thus, the third objective was to identify molecular markers associated with seed coat deficiency. Results showed that seed size, high-, average- and low-quality sprout percentage, hypocotyl thickness and length and sprout yield are the most important variables for breeding sprout cultivars; and one-year seed storage at room temperature reduced sprout quality. Fungal species that may cause human illness were isolated from sprout soybean seeds and calcium hypochlorite and acetic acid are promising seed disinfection treatment for reducing fungus incidence. Finally, eight molecular markers associated with seed coat deficiency were identified which may be potentially used in selecting natto soybean cultivars with low seed coat deficiency. Information presented will be helpful for sprout and natto soybean cultivar development.

Dedication

I would like to dedicate this thesis to my parents, brother and grandparents for all their love, encouragement and support. A special thanks to my mom and my dad for putting me through the best education possible, they are the great inspiration of my life and I appreciate enormously their sacrifices. I wouldn't have been able to get to this stage without them.

Attributions

Many collaborators and colleagues contributed significantly to the research results presented and a description of their roles is described below.

Chapter 2: Evaluation of important seed and sprout traits as potential selection criteria in breeding varieties for sprout soybeans

Diana Escamilla, Master's Degree Candidate in the Department of Crop and Soil Environmental Science, Virginia Tech. Ms. Escamilla performed the research exposed in this chapter, analyzed the data and prepared the manuscript presented in this chapter.

Bo Zhang, Assistant Professor in the Department of Crop and Soil Environmental Science, Virginia Tech, and committee chair. Dr. Zhang aided in the experimental design development, data analysis for this research and editing the manuscript presented in this chapter.

Luciana Rosso, Research Associate in the Department of Crop and Soil Environmental Sciences, Virginia Tech. Dr. Rosso provided advice to the experimental design and fungal identification. She also edited the manuscript that is presented in this chapter.

Laura Strawn, Assistant professor of food science at the department of Food Science, Virginia Tech. Dr. Strawn provided advice and help about identification of fungal species isolated from soybean seeds.

Chapter 2: Fungi associated with soybean seeds of a commercial sprout soybean cultivar and potential seed disinfection treatments

Diana Escamilla, performed the research exposed in this chapter, analyzed the data and prepared the manuscript presented in this chapter.

Luciana Rosso, aided with the fungal isolation and identification protocols, fungal DNA extraction, experimental design and edition of the manuscript presented in this chapter.

Bo Zhang, contributed to the experimental design, data analysis and editing the manuscript presented in this chapter.

Chapter 3: QTL analysis of seed coat deficiency of soybeans (*Glycine max* [L.] Merr.)

Diana Escamilla, performed the research exposed in this chapter, analyzed the data and prepared the manuscript presented in this chapter.

Song Qijian, Research Geneticist at Beltsville Agricultural Research Center, USDA-ARS. Dr. Song provided advice on mapping population development and software for QTL analysis.

Bo Zhang, contributed to the experimental design, mapping population development, data analysis and editing the manuscript presented in this chapter.

Table of Contents

List of Tables

List of Figures

Introduction

Soybean, among other legumes, has the highest protein content (40%) with about 20% oil and 35% carbohydrates (Liu, 1997). It is the dominant oilseed crop in US accounting for about 90% of total US oilseed production (Pompelli, 2009). During the past seven years, the United States planted over 30.0 million hectares per year with 36.2 million hectares in 2017 (USDA-NASS, 2017). A great part of soybeans planted in USA are processed to extract oil for food and industrial use and high protein meal for animal feed; and a small amount of soybeans are used in human foods (Pompelli, 2009).

Soyfoods have been part of the daily life in Asia for over 5000 years and a great diversity of soyfood such as tofu, natto, edamame, soymilk, soy sauce and soy sprouts are now available (Cui et al., 2004; Liu, 1997). Over the past years, soy-foods have increased in popularity outside of Asia because of changes in consumer acceptance (Hartman et al., 2011; Mayta et al., 2014). Soybeans used to produce various soy-foods of Asian origin such as soybean sprouts and natto, assorted health food snacks, energy foods and cereals, are known as specialty or novel soybeans (Kentucky, 2013); and their production is a high value alternative to growing grain soybean for feed (Cui et al., 2004; Taira, 1990). Specialty soybeans in U.S. are grown to meet the demand from two markets, Asian countries with Japan as a major buyer and U.S consumers that realize the health benefits of soybeans in their diets. Food grade soybean cultivars are generally grown under contract agreements and processors specify quality requirements, price premiums and the number of acres to be planted (Kentucky, 2013). Additionally, cultivar selection is mainly based on seed-yield performance, disease resistance, and value-increasing seed attributes (Mayta et al.,

1

2014). Therefore, breeding programs of food grade soybeans in the United States are focused on developing varieties with specific seed characteristics required by soyfood buyers.

Natto and sprout soybeans are produced by small-seeded soybeans and their production is a high-value alternative for growing grain soybeans for feed in the Mid-Atlantic region of U.S. Therefore, the development of soybean cultivars with improved food grade characteristics is crucial for maintaining and increasing the soy-food market. In order to develop superior food grade cultivar, it is important to generate information about quality traits for natto and sprout soybeans, which lead to the development of the research conducted in this thesis.

References

Cui Z., James A., Miyazaki S., Wilson R., Carter T. (2004) Breeding Specialty Soybeans for Traditional and New Soyfoods, in: L. K (Ed.), Soybeans as Functional Foods and Ingredients, AOCS Press, Urbana, IL. pp. 264-322.

Hartman G.L., West E.D., Herman T.K. (2011) Crops that Feed the World 2. Soybean—Worldwide Production, Use, and Constraints Caused by Pathogens and Pests. Food Security 3:5-17. DOI: 10.1007/s12571-010-0108-x.

Kentucky Universtiy. (2013) Specialty Soybeans, in: U. K. Coperative Extension Service (Ed.), http://www.uky.edu/Ag/NewCrops/introsheets/specialtysoy.pdf.

Liu k. (1997) Soybeans: Chemistry, technology, and utilization Chapman & Hall, New York.

Mayta J., Chen P., Popp M.P., Dong D., Wu C.J., Zhang B., Smith S.F., Scaboo A.M. (2014) Break-Even Profitability for Food-Grade Specialty Soybeans. Agricultural Science 2:1-11. DOI: 10.12735/as.v2i2p01.

Pompelli G.K. (2009) Crop Production Capacity in North America, in: R. Lal (Ed.), Agricultural Sciences, Encyclopedia of Life Support Systems (EOLSS), Eolss Publishers, Oxford, United Kingdom. pp. 14.

Taira H. (1990) Quality of Soybeans for Processed Food. JARQ 24:224-230.

USDA-NASS. (2017) National Statistics for Soybeans, United States Department of Agriculture-National Agricultural Statistics Service.

Zhang B., Chen P., Florez-Palacios S.L., Shi A., Hou A., Ishibashi T. (2010) Seed Quality Attributes of Food-Grade Soybeans from the U.S. and Asia. Euphytica 173:387-396. DOI: 10.1007/s10681-010-0126-y.

1. Literature Review

Natto Soybeans

Consumption, preparation and processing conditions

Natto is a fermented soyfood originated in Japan about 1000 years ago and it is served for breakfast and dinner with rice (Hosoi and Kiuchi, 2003; Hu et al., 2010). Fermentation, a process used in the food industry, improves sensory characteristics of the product, makes nutrients more accessible and eliminates undesirable constituents (Hu et al., 2010). Natto is prepared by cooking presoaked small-seeded soybean seeds for about 30 min until tender. Then, beans are dried, cooled to 40 °C, inoculated with a water suspension of *Bacillus subtillis*, packed in a wooden box or polyethylene bag, and incubated at 40-43 °C for 12-20 h (Hosoi and Kiuchi, 2003; Hu et al., 2010; Liu, 1997; Wei et al., 2001). Around 5-10 ml of *B. subtilis* suspension are needed for 60 kg of raw soybeans (Liu, 1997). *B. subtilis* utilizes the proteins, peptides, and amino acids in soybeans for its growth; and the kinds and quantities of peptides and amino acids produced by the bacteria during fermentation affect the flavor of natto. Hence, soybean with high sucrose and protein content are preferred (Hosoi and Kiuchi, 2003). Natto is ready to be served after the beans are covered with white sticky coating, and sweet taste, light yellow color, unique aroma, and soft and sticky texture are common characteristics of good quality natto (Wei and Chang, 2004). The sticky coating in natto is mainly composed by γ-Polyglutamic acid produced during fermentation (Ho et al., 2006). To obtain better quality natto product, it is recommended to keep the package of fermented natto at a refrigerating temperature for one to two days to allow maturation before consumption or retailing (Liu, 1997). The degree and speed of the reactions, and substances, enzymes, amino acids and peptides released during natto fermentation, depend on the conditions of soaking, steaming, fermentation and bacteria strain (Wei et al., 2001).

Contamination of natto products by other microorganisms must be avoided, as it affects the quality of the natto and may cause food poisoning; however, normal tap water has shown to be adequate for the production of natto (Hosoi and Kiuchi, 2003). Over-fermentation leads to the release of ammonia that spoils natto flavor, destroys *B. subtilis* and promotes spoilage by other organisms. Natto is not accepted by some Chinese because of the ammonia smell (Ma et al., 2015). Ma et al (2015) found that a 1:4 ratio of black soybean to yellow soybean, steamed for 30 min under high pressure, inoculated with 5 mL of 10% solution of *B. subtilis* and fermented at 38°C for 13 h, produced natto with a good taste and better smell because black soybeans produce less nattokinase (enzyme), resulting in less ammoniacal smell (Ma et al., 2015). However, natto made from black soybeans has not traditionally been well accepted because of its unappetizing appearance and lack of characteristic aroma; but it has been recently more accepted and its high content of polyphenols has been emphasized, mainly for their health benefits (Hosoi and Kiuchi, 2003). A longer steaming time (121 °C for 40 min) during natto production also reduced fermentation time and ammonia content in the final product (Wei et al., 2001). As a result, natto with different texture, aroma and flavor could be produced by combining different processing conditions and raw materials (black and yellow soybeans).

Natto quality characteristics

In the industry, acceptability of small-seeded cultivars for natto production is assessed by measuring water absorption of the seeds, water loss of the seeds after steaming, hardness and darkness of the seeds in the natto after fermentation, and sensory evaluation of natto texture and flavor. High water absorption and low water loss are preferred by natto manufactures due to high natto yield, while the standards for hardness and color are driven by the consumer preferences (Geater et al., 2000). Water absorption is evaluated in natto production through the weight

increase ratio (WIR) after soaking, which is determined by soaking seeds at 22-25 ºC for 16 h (Wei and Chang, 2004). The suitability of natto texture and flavor for the natto market is usually determined by sensory panels (Zhang et al., 2008). Good quality natto must have colored mucous substance, characteristic flavor, light yellow color and soft and sticky texture (Hu et al., 2010; Liu, 1997; Wei and Chang, 2004). Other desirable qualities of natto soybeans are uniform seed shape, slight changes in constituents during storage, sweet flavor of natto, low cracking ratio of the soybeans after steaming, and low ammonia content (Cober et al., 1997; Hosoi and Kiuchi, 2003; Taira, 1990; Wei and Chang, 2004).

Nutritional content and health benefits

Soybean seeds contain approximately 21% oil, 40% protein and 34% carbohydrate (Burton, 1997). Thus, soyfood provides high contents of proteins, fiber, linoleic acid, all essential amino acids, and high content of isoflavones that decrease the loss of bone mass in postmenopausal women (Weng and Chen, 2012; Zhang et al., 2011). In addition, isoflavones are more abundant in natto than in other soybean products (Ikeda et al., 2006; Somekawa et al., 2001). Natto also contains functional compounds such as poly (γ--glutamic acid), bioactive peptides, nattokinase, milk clotting enzymes, and lipases produced by *B. subtilis* that have been applied in food processing, cosmetics, and synthesis of chemicals and pharmaceuticals (Dabbagh et al., 2014; Ma et al., 2006; Shieh et al., 2009; Shih and Yu, 2005; Weng and Chen, 2012). Bioactive peptides are produced as a result of hydrolysis of soybean proteins (Glycinin and β-conglycinin) during fermentation, and they may act like regulatory compounds and exhibit bioactive properties such as anti-hypertensive, antimicrobial, antioxidant, anti-diabetic, and prevention of cancer and gastrointestinal disorders among others (Gibbs et al., 2004; Sanjukta and Rai, 2016; Tamang et al., 2016). Nattokinase, an extracellular enzyme produced by *B.*

subtilis during fermentation, is stable enough in the gastrointestinal track and is a useful agent for the oral thrombolytic therapy because it helps to reduce blood pressure to the arteries and heart valves (Dabbagh et al., 2014; Ma et al., 2015; Tuan et al., 2015). All these characteristics make natto a very nutritional and healthy food.

Breeding of natto soybean cultivars

Seed morphological and physiological traits influencing natto quality

Small-seeded genotypes (< 9 g/100 seeds) with uniform seed size, round shape, easily washable, clear hilum, and yellow and smooth seed coat are preferred by the industry; because small seeds have high water absorption capacity and good to combine with rice in size and texture (Cui et al., 2004; Geater et al., 2000; Liu, 1997; Salas et al., 2006). In Japan, soybean seed size is classified by diameter into four groups: extra small seeds of less than 5.5 mm, small ranges from 5.5 mm to 7.3 mm, medium ranges from 7.3 mm to 7.9 mm and large of 7.9 mm or more. Extra small and small soybeans are considered the most suitable for natto as they tend to have distinctive natto flavor and stronger taste due to excess fermentation and higher degradation rate of proteins (Hosoi and Kiuchi, 2003).

Hardness of cooked seed is another critical trait for soyfood that requires processing and cooking because seed texture is an important factor in sensory evaluation of soyfood. Harder seeds decrease consumer acceptances of soyfood products. However, reducing seed hardness can be challenging since it is affected by many factors such as seed size, seed coat permeability, imbibition, dormancy, seed coat structure and chemical composition (Orazaly et al., 2015; Qutob et al., 2008; Zhang et al., 2008).

The initial water absorption step in natto processing where soybean seeds are softened and soluble sugars are released is important (Cook and Rainey, 2010). It is desirable for natto soybean cultivars to at least double their dry weight through water absorption due to the high water content (60%) of finished natto (Cook and Rainey, 2010; Wei et al., 2001). Cook and Rainey (2010) reported water absorptions of 215% to 225% in natto soybean germplasm adapted to mid-Atlantic region in the US. In soybean seeds, without physiological dormancy, germination is determined by the seed coat (Qutob et al., 2008); which regulates water absorption to prevent imbibitional damage (Copeland and McDonald, 2001). Koizumi et al., (2008) reported that intact seed coat of soybean seeds regulated water incorporation into the radicle, hypocotyl and cotyledons and it prevented the destruction of the seed tissues at the beginning of imbibition. It is ideal for natto soybean to quickly absorb as much water as possible while avoiding seed coat cracking, which is an undesirable trait of natto soybeans due to a resulting inferior appearance of the products and clogging of the production lines (Yasui et al., 2017). However, natto manufacturers observed that some natto germplasm adapted to mid-Atlantic region in US had poor seed coat integrity (Cook and Rainey, 2010). There are several seed tests that have been developed to evaluate the integrity of the seed coat such as sodium hypochlorite, fast green, indoxyl acetate, tetrazolium and Chowdhury methods (Copeland and McDonald, 2001; VanUtrecht et al., 2000). Rodda et al. (1973) reported a method based on the fact that cracked soybean seeds absorb moisture faster and swell in size more than undamaged seeds. Based on the same principle, a seed coat deficiency (SCD) test, was developed by Cook and Rainey (2010). Seed with cracks or blisters around the hilum or detached seed coat from the hull are counted as deficient seeds. In natto soybeans, the SCD scores ranged from 23 to 70%, which indicated that breeding selection is feasible to efficiently reduce SCD. SCD was positively

correlated to seed size because in small seeds the seed coat represents a greater part of the total seed mass, and seed coat is likely responsible of a high resistance to cracking. Natto lines with high water absorption also showed undesirable high SCD scores suggesting that selection for water absorption alone may have compromised seed coat strength (Cook and Rainey, 2010). Therefore, selection of soybeans with low seed coat deficiency and high water absorption must be carried out simultaneously.

Seed composition traits related to natto quality

Protein and isoflavones contents in soybean seeds play an important role in the nutritional and health value of natto soybeans (Ikeda et al., 2006; Somekawa et al., 2001; Weng and Chen, 2012; Zhang et al., 2011). Therefore, cultivars with high isoflavones and protein content are preferred. In order to produce natto with good quality and flavor, it is also desirable that seeds have high carbohydrate content, mainly soluble sugars, that promotes microbial growth and makes the finished product sweeter (Hosoi and Kiuchi, 2003; Wei and Chang, 2004). Soybean seeds have approximately 34% carbohydrate, and up to 1.6% of carbohydrates are soluble sugars (Burton, 1997; Hymowitz and Collins, 1974). Soybean seeds contain the monosaccharides glucose, fructose, and galactose; the disaccharide sucrose; and the oligosaccharides raffinosse and stachyose. Sucrose was found in greatest quantity in soybean seeds followed by stachyose and raffinosse representing 4.9%, 3% and 1.5% of mature seeds; small amounts of glucose and fructose were also found (<1%) (Hymowitz and Collins, 1974; Liu, 1997; Yazdi-Samadi et al., 1977). Of the soluble sugars, sucrose is one of the most important flavor components in soyfood because it contributes sweetness in soybean products such as soymilk and tofu. Raffinosse and stachyose, are considered undesirable sugars because they are not readily digestible, and cause flatulence or diarrhea (Wang et al., 2014).

Low content of calcium, magnesium and protein content are desired as well (Wei and Chang, 2004). Yoshikawa et al. (2014) reported that good sensory quality of natto soybeans was positively associated with sucrose, but negatively correlated with seed hardness, protein, protein plus oil, calcium, manganese, and boron contents. Therefore, selection of soybeans lines with low protein, oil, calcium, manganese, and boron content as well as high sucrose can be an effective strategy for natto soybean breeding.

Genetic control of soybean seed traits

Many important agronomic traits are controlled by multiple genes and are known as quantitative traits. Because the seed traits are quantitative and ambiguous (Cook and Rainey, 2010; Cui et al., 2004; Liu, 1997), it is a challenging and long process to improve those traits by conventional breeding. Therefore, the identification of QTL and markers associated with natto seed -related traits plays an important role in natto soybean breeding through marker assisted selection (MAS). SoyBase (http://www.soybase.org/ accessed 6 Sep. 2017) reported 240 QTL for protein, 322 for oil, 120 for seed weight, 14 for seed hardness, 37 for sucrose content, 14 for seed coat cracking and 73 for isoflavones content. However, genetic positions of some of these reported QTL are very close (< 10 cM) so they may not be truly different QTL (Pathan et al., 2013).

Numerous studies have identified QTL associated with seed compositions influencing the quality of natto soybeans. For example, Maughan et al. (2000) reported seventeen markers significantly associated with sucrose variation in a segregating F_2 population; and suggested that protein, oil and sucrose content are inherited as clusters of linked loci or that 'major' QTL with pleiotropic effects may control all three traits. A negative correlation (-0.68) between sucrose and protein content showed that selection for high sucrose would likely be accompanied by a

reduction in protein (Jaureguy et al., 2011). Several QTL for sucrose, raffinosse and stachyose have been identified previously (Kim et al., 2006; Lee et al., 2015; TM et al., 2005; Zeng et al., 2014). Recently, a SSR marker (Satt359) on a QTL located on chromosome 11 [Linkage group (LG) B1] showed to be significantly associated with sucrose (LOD = 5.192; R^2= 0.134), raffinosse (LOD = 3.95; R^2 = 0.104) and stachyose (LOD = 5.192; R^2 = 0.314); thus it can be used to assist breeding selection for sucrose, raffinosse and stachyose contents simultaneously (Wang et al., 2014). Seed composition traits such as protein, oil, sugar and isoflavones contents have complex genetic relationships that must be studied further to facilitate breeding of superior natto cultivars with specific profiles of protein, oil, sugar and isoflavones content.

Seed morphological and physiological traits such as seed shape and weight, seed coat cracking, water uptake and seed hardness are also quantitative (Liang et al., 2010; Molnar et al., 2012; Orazaly et al., 2015; Yasui et al., 2017). Hundred seed weight and seed shape are complex, polygenic traits; therefore, use of molecular markers for indirect selection of these traits may be helpful for breeders (Liang et al., 2010). Nineteen QTL were identified (LOD≥3.7) in ten LGs for seed shape (height, width, length, volume) and only one QTL was stable across populations and environments (Salas et al., 2006). Liang et al. (2010) identified 24 QTL for seed length, seed width, seed thickness and hundred seed weight; distributed also on several LGs. At least two common QTL among seed length, width and thickness were detected, indicating their similar genetic basis. In addition, eight gene candidates were also predicted including Glyma06g08860 (X90–X92) that may be involved in the process of seed development, and Glyma06g08800 (X83–X85) that encodes ARM repeat superfamily protein, which is involved in the ubiquitination pathway, regulating the development of seed size (Xie et al., 2014). Narrow sense heritability of 0.42 to 0.88 for seed shape traits (height, width, length, volume), and the number

and distribution of QTL for seed shape and hundred seed weigh traits confirmed the complex genetic control of these traits (Liang et al., 2010; Salas et al., 2006).

Several QTL have also been reported for seed coat cracking. Oyoo et al. (2010), reported two types of genetically controlled seed coat cracking, type I with irregular cracks and type II with net-like cracks. They also detected two QTL for seed coat cracking type II in molecular linkage group C1 with a distance of about 62 cM accounting for 16% and 32.4% of phenotypic variation, respectively. QTL for seed coat cracking after soaking (qSCAS1, qSCAS2 and qSCAS3) and after cooking (qSCAC1 and qSCAC2) located on chromosomes 4 (LG: C1), 6 (LG C2) and 8 (LG A2), were detected on a recombinant inbred line (RIL) population; and the interactions between QTL were also observed (Yasui et al., 2017). In addition, seed coat wrinkling, produced by high temperatures and alternate periods of wet and dry conditions, can also predispose seeds to mechanical damage at harvest, reducing germination and the quality of soybeans. A gene controlling seed coat wrinkling located on chromosome 5 (LG A1) was identified (Kebede et al., 2012). Thus, genetic control of these traits exist and molecular markers can be developed to allow a faster selection of cultivars that produce seeds with low seed coat wrinkling and cracking during soaking and cooking of soybean seeds.

There is limited information of QTL associated with water absorption in natto soybean. Five different QTL and associated microsatellite markers Sat_220, Satt651, Satt333, Satt456, and Sat_244, were detected for water uptake. QTL were located on linkage groups D2, E, A2, J and M and mapped to similar or nearby regions as known QTL for seed weight, seed yield, seed fill, flowering time and maturity (Molnar et al., 2012). Although Cook and Rainey (2010) reported that natto lines with high water absorption have shown undesirable seed coat cracking, there is no information about the genetic relationship of seed coat cracking and water uptake on

soybean seeds. As a result, further studies would be required to identify and understand their genetic relationship and facilitate breeding selection of these traits simultaneously.

Calcium concentrations and seed coat impermeability (thick seed coat) are associated with seed hardness (Orazaly et al., 2015; Zhang et al., 2008). Jang et al., (2015) studied the molecular basis of *qHS1,* a quantitative trait locus for seed hardness located in linkage group D1b (chromosome 2), and revealed that the hard-seed allele *qHS1* encodes an endo-1, 4-β glucanase, which seems to be involved in accumulation of β-1, 4-glucan derivatives such as xyloglucan and/or β-(1, 3), (1, 4)-glucan that reinforce the impermeability of seed coats in soybean. A single nucleotide polymorphism (SNP) in permeable cultivars introduced an amino acid substitution reducing or eliminating enzyme affinity for substrates (Jang et al., 2015). Two QTL were identified for seed hardness *Ha1* and *Ha2,* explaining 7.9% of the phenotypic variance (Zhang et al., 2008). Orazaly et al. (2015) revealed a new putative QTL for seed hardness (*Ha3)* in an approximate 46-cM region and linked to markers Satt547 and Satt414 on chromosome 16, and confirmed the chromosomal region of previously reported hardness QTL (*Ha2*) (Orazaly et al., 2015). A stable and already confirmed QTL (*qHbs3-1*) for seed harness was also detected on chromosomes 3. Therefore, SSR markers closely linked to *qHbs3-1* such as BARCSOYSSR_03_0165 and BARCSOYSSR_03_0185 as well as markers of the other reported QTL could be useful for marker-assisted selection (Hirata et al., 2014).

Despite the huge number of studies that have identified QTL associated with seed traits, confirmation of these QTL has rarely been carried out. For instance, out of 300 and 325 QTL identified for protein and oil seed content, only 16 QTL have been confirmed for each protein and oil content (available online: http://soybase.org). A QTL is confirmed when it is evaluated in an independent set of meiotic events (population) sharing at least one parent with the original

population, in independent environments and at an experiment-wise error rate of less than 0.01 (available online: http://soybase.org). On the other side, the identification of common QTL regions for specific traits is not always simple due to different marker sets, statistical methods, parents, population size and generation, environments, etc. (Van and McHale, 2017). Therefore, several factors such as mapping populations, scoring methods, replications and environments, appropriate quantitative genetic analysis, various genetic backgrounds and independent verification, must be considered when selecting markers previously reported for successful marker assisted selection.

Cultivar selection and genetic diversity of food-grade germplasms

The improvement of natto quality characteristics is important for maintaining and increasing its market (Cober et al., 1997). Cultivar selection for specialty soybeans is mainly based on seed-yield performance, disease resistance, and value-increasing seed attributes (Mayta et al., 2014). Natto manufacturers prefer certain soybean cultivars, and their preferences can vary among regions because of consumer preferences (Hosoi and Kiuchi, 2003). ARK1, ARK2 and ARK3 are superior natto cultivars used in commercial natto production. ARK1 was developed by Nakeko Seed Company, Gunma Japan; and ARK2 and ARK3 were developed by Takano Foods Corporation, Ibaraki, Japan (Yoshikawa et al., 2014). Danatto, Minnatto, Natto King and MN-468 are commercial natto soybean cultivars grown in Northern America, which had seed diameter that ranged from 5.61 to 6.57 mm with round shapes, except for Natto King, which had slightly oval shape. Their 100 seed weights varied from 10.47 to 11.65 g and seed coat cracking ratios from 0.78 to 5.80%. Among these cultivars Danatto, Minnatto and MN91-468 have shown more suitable than Natto King for natto manufacturing or processing because Natto King exhibited cracked hulls and separation of skins from cotyledons after cooking (Wei and Chang,

2004). SS-516 is a small-seeded cultivar marketed by Southern States seed company, having an intermediate marketing value (Cook and Rainey, 2010; Yoshikawa et al., 2014). MFS series such as MFS-561 developed in Virginia are also popular natto varieties. Other Japanese natto cultivars include Canatto, TNS, Nattosan, AC T2653, AC Pinson and Nattawa, which have mean protein values on a dry weight basis that range from 30.4% in AC T2653 to 31.0% in Nattawa, 32.1% in TNS, 32.3% in Canatto and 34.2% in Nattosan. All these Japanese cultivars had similar contents of essential amino acid (EAA= 45.0 to 47.0%) (Zarkadas et al., 1997). Several Natto cultivars have also been released in Canada. Examples of them are Natto 3 and DH3604 (Cober et al., 2006; Poysa and Voldeng, 2007). Therefore, there are available soybean cultivars that have been developed for natto production not just in Japan, but also in countries such as United States and Canada. Generally, natto cultivars are developed to fulfill specific characteristics for the markets where they are going to be commercialized; and continuous breeding is needed to maintain and improve natto quality through natto market evolution.

Diversity of food-grade soybeans is critical for utilization of genetic resources in cultivar development, germplasm enhancement, and end product commercialization (Zhang et al., 2010). Seed sugar content in soybean plant introductions from different origins and regions exhibited a great variation in individual and total sugars, where high sucrose (ranging from 1.6 to 95.4 mg/g) and low stachyose (ranged from 0.2 to 69.6 mg/g) types are the most valuable for breeding specialty soybeans (Hou et al., 2009). Soybean germplasms with unique sugar profiles are a useful source of diversity for future breeding and genetic research of sugar content in soybeans. A genetic diversity study among 105 food-grade soybean genotypes from USA and Asia detected different genetic backgrounds among US and, Japan and South Korea food-grade lines; which may serve as valuable gene pools to improve the genetic diversity of US lines. A negative

correlation between protein and oil content (r = -0.67) and several markers for oil and protein content were also identified among the U.S and Asia germplasm (Shi et al., 2010a). Zhang et al. (2010) also studied the genetic diversity among US (54) and Asian (51) food-grade lines, identifying higher genetic diversity for seed traits such as protein content, calcium content, seed hardness and seed uniformity than other quality traits. The same study suggested that small-seeded soybean genotypes from U.S. are desirable for natto production because of their softer texture with higher water absorption capacity and lower stone seed ratio. Stone seeds do not absorb water during soaking causing serious problems for food processing. However, seed size uniformity of both small and large-seeded types was low, indicating the importance of developing cultivar with uniform seed size for food-grade soybean market (Zhang et al., 2010). Heritability estimates for important seed traits for soyfood quality ranged from 0.46 to 0.97. Hence, selection for genotypes with specific combination of these traits should be reliable (Jaureguy et al., 2011; Yoshikawa et al., 2014). Available information showed that the indirect selection of natto soybean cultivars through selection of seed quality traits such as sugar content, seed size (hundred seed weight), water absorption, seed coat cracking among other traits is possible. The genetic base of U.S. and Asia germplasms is relatively broad for cultivar development and germplasm enhancement in food-grade soybean programs in US.

Sprout Soybeans

Germination of soybean seed for sprout production is another growing market for use of soybeans (Ghani et al., 2016). Soybean sprouts can be consumed in soups, salads and side dishes and it is very important in Asian countries such as Korea, China, and Japan (Choi and Bajpai, 2010 and Lee et al., 2007). During cooking, it is desirable to minimize heating to maintain the crisp texture and minimize the destruction of vitamins (Liu, 1997). Traditionally, soybean

sprouts are prepared by soaking small seeded soybeans for 3-4 hours. Those are then, placed at 23 °C in a deep container with holes at the bottom. The container is covered for keeping seeds under dark conditions and seeds are sprinkled with water 3-4 times a day. Most of the sprouts reach a length of about 8 cm in less than a week and about 7-9 lb of sprouts are produced by 1 lb of dry soybeans (Liu, 1997).

Germination of soybean seeds is a health beneficial process for soy food preparation due to the increased nutrition protein, phytosterols and tocopherol (vitamin E) levels during germination while keeping unchanged isoflavones levels (Shi et al., 2010b). Dietary intake of isoflavones has been shown to reduce the risk of several major diseases in humans (Primomo et al., 2005). While, Phytosterols can lower intestinal cholesterol absorption thus reducing the risk of atherosclerosis (Wang et al., 2011). In addition soybeans provide high contents of fiber, linoleic acid and all essential amino acids (Weng and Chen, 2012; Zhang et al., 2011).

Soybean sprouts must have a crispy texture with yellow cotyledons and long and bright white hypocotyls (Liu, 1997). Soybean cultivars used for sprouts should have small seed size, 40–150 mg seed^{-1}, and good seedling vigor for sprout production (Lee et al., 2001; Park et al., 1995). Uniform seed size and high and rapid water uptake are also desired characteristics for sprout cultivars (Pietrzak et al., 2002). Sprout length of 8-12 cm, sprout thickness of 2.0-2.2 mm and sprout yield of 4-6 g fresh weight/g dry weight have been previously reported for small seeded soybeans (Lee et al., 2001; Park et al., 1995). Other sprouts characteristics include fresh weight of sprouts, percentage of abnormal seedlings (decayed seeds and seedlings, stunted seedlings, less than 2-cm growth) and percentage of high-quality sprouts (vigorous and straight sprouts) (Lee et al., 2007; Lee et al., 2001).

To date, there is limited information about quality traits of soybean sprouts. However, the genetic relationship of some sprout traits has been studied. Quantitative trait loci (QTL) conditioning sprout yield and seed weight were on the same genomic region, indicating that sprout yield was genetically linked to seed weight (Lee et al., 2001). Seed weight was negatively correlated to sprout yield (r = -0.763***), so using small seeds was more profitable than bigger seeds. Sprout yield also showed a negative correlation with a higher incidence of abnormal seedlings because more abnormal seedlings produced less sprouts (Lee et al., 2001).

The main traits used for breeding of soybean sprouts include seed weight of about 10 g/ 100 seed, good germination (90%), seed vigor, bright hypocotyl color, yellow cotyledons, hypocotyl length and sprout yield (Ghani et al., 2016). However, breeders in U.S. use natto soybeans as sprout genotypes because natto cultivars and sprout soybean cultivars share some characteristics (Zhang et al., 2010). Natto, as well as sprout soybeans, requires cultivars with small, uniform seed size (< 9 g 100 seeds^{-1}), which is the primary characteristic used to identify natto cultivars, smooth and light colored seed coat, clear hilum and high water absorption (Cui et al., 2004; Liu, 1997). Therefore, there is a need for improving the understanding of seed quality and sprout traits to facilitate breeding programs in U.S.

References

Burton J.W. (1997) Soyabean (Glycine max(L.) Merr.).   Field Crops Research   53:171-186.

Cober E.R., Fregeau-Reid J.A., Pietrzak L.N., McElroy A.R., Voldeng H.D. (1997) Genotype and Environmental Effects on Natto Soybean Quality Traits.   Crop Science   37:1151-1154.

Cober E.R., Voldeng H.D., Frégeau-Reid J.A., Guillemette R.J.D. (2006) DH3604 Soybean. Canadian Journal of Plant Science   86:483-484.   DOI: 10.4141/P05-188.

Cook D.E., Rainey K.M. (2010) Seed Coat Deficiency, Trait Stability, and Other Soybean Seed Quality Traits for Natto Cultivar Development.   Crop Science   50:1244-1249.   DOI: 10.2135/cropsci2009.06.0300.

Copeland L.O., McDonald M.B. (2001) Principles of Seed Science and Technology. Springer Sceince+Business Media, LLC, New York.

Cui Z., James A., Miyazaki S., Wilson R., Carter T. (2004) Breeding Specialty Soybeans for Traditional and New Soyfoods, in: L. K (Ed.), Soybeans as Functional Foods and Ingredients, AOCS Press, Urbana, IL. pp. 264-322.

Dabbagh F., Negahdaripour M., Berenjian A., Behfar A., Mohammadi F., Zamani M., Irajie C., Ghasemi Y. (2014) Nattokinase: Production and Application.   Applied Microbiology and Biotechnology   98:9199-9206.   DOI: 10.1007/s00253-014-6135-3.

Geater C.W., Fehr W.R., L A.W. (2000) Association of Soybean Seed Traits with Physical Properties of Natto.   Crop Science   40:1529-1534.

Ghani M., Kulkarni K.P., Song J.T., Shannon J.G., Lee J.-D. (2016) Soybean Sprouts: A Review of Nutrient Composition, Health Benefits and Genetic Variation.   Plant Breeding and Biotechnology   4:398-412.   DOI: 10.9787/pbb.2016.4.4.398.

Gibbs B.F., Zougman A., Masse R., Mulligan C. (2004) Production and Characterization of Bioactive Peptides from Soy Hydrolysate and Soy-Fermented Food. Food Research International 37:123-131. DOI: 10.1016/j.foodres.2003.09.010.

Hirata K., Masuda R., Tsubokura Y., Yasui T., Yamada T., Takahashi K., Nagaya T., Sayama T., Ishimoto M., Hajika M. (2014) Identification of Quantitative Trait Loci Associated with Boiled Seed Hardness in Soybean. Breed Sci 64:362-70. DOI: 10.1270/jsbbs.64.362.

Ho G.-H., Ho T.-I., Hsieh K.-H., Su Y.-C., Lin P.-Y., Yang J., Yang K.-H., Yang S.-C. (2006) -Polyglutamic Acid Produced by Bacillus subtilis (natto): Structural Characteristics, Chemical Properties and Biological Functionalities. Journal of the Chinese Chemical Society 53:1363-1384.

Hosoi T., Kiuchi K. (2003) Fermenting Cooked Soybeans with Bacillus Subtilis (natto). , in: E. R. Farnworth (Ed.), Handbook of Fermented Functional Foods, CRC Press, USA. pp. 227-250.

Hou A., Chen P., Alloatti J., Li D., Mozzoni L., Zhang B., Shi A. (2009) Genetic Variability of Seed Sugar Content in Worldwide Soybean Germplasm Collections. Crop Science 49. DOI: 10.2135/cropsci2008.05.0256.

Hu Y., Ge C., Yuan W., Zhu R., Zhang W., Du L., Xue J. (2010) Characterization of Fermented Black Soybean Natto Inoculated with Bacillus Natto During Fermentation. J Sci Food Agric 90:1194-202. DOI: 10.1002/jsfa.3947.

Hymowitz T., Collins F.I. (1974) Variability of Sugar Content in Seed of Glycine ruax (L.) Merrill and G. soja Sieb. and Zucc. Agronomy Journal 66:239-240.

Ikeda Y., Iki M., Morita A., Kajita E., Kagamimori S., Kagawa Y., Yoneshimayy H. (2006) Intake of Fermented Soybeans, Natto, Is Associated with Reduced Bone Loss in Postmenopausal Women: Japanese Population-Based Osteoporosis (JPOS) Study.   J. Nutr.   136:1323–1328.

Jang S.J., Sato M., Sato K., Jitsuyama Y., Fujino K., Mori H., Takahashi R., Benitez E.R., Liu B., Yamada T., Abe J. (2015) A Single-Nucleotide Polymorphism in an Endo-1,4-beta-Glucanase Gene Controls Seed Coat Permeability in Soybean.   PLoS One   10:e0128527. DOI: 10.1371/journal.pone.0128527.

Jaureguy L.M., Chen P., Scaboo A.M. (2011) Heritability and Correlations Among Food-Grade Traits in Soybean.   Plant Breeding   130:647-652.   DOI: 10.1111/j.1439-0523.2011.01887.x.

Kebede H., Smith J.R., Ray J.D. (2012) A New Gene that Controls Seed Coat Wrinkling in Soybean.   Euphytica   189:309-320.   DOI: 10.1007/s10681-012-0818-6.

Kim H.K., Kang S.T., Oh K.W. (2006) Mapping of Putative Quantitative Trait Loci Controlling the Total Oligosaccharide and Sucrose Content of *Glycine max* Seeds.   J Plant Res   119:533-8. DOI: 10.1007/s10265-006-0004-9.

Koizumi M., Kikuchi K., Isobe S., Ishida N., Naito S., Kano H. (2008) Role of Seed Coat in Imbibing Soybean Seeds Observed by Micro-Magnetic Resonance Imaging.   Ann Bot 102:343-52.   DOI: 10.1093/aob/mcn095.

Lee J.-D., Shannon J.G., Jeong Y.-S., Lee J.-M., Hwang Y.-H. (2007) A Simple Method for Evaluation of Sprout Characters in Soybean.   Euphytica   153:171-180.   DOI: 10.1007/s10681-006-9252-y.

21

Lee J.S., Kim S.-M., Kang S. (2015) Fine Mapping of Quantitative Trait Loci for Sucrose and Oligosaccharide Contents in Soybean [*Glycine max* (L.) Merr.] using 180 K Axiom® SoyaSNP Genotyping Platform.   Euphytica   208:195-203.   DOI: 10.1007/s10681-015-1622-x.

Lee S.H., Park K.Y., Lee H.S., Park E.H., Boerma H.R. (2001) Genetic Mapping of QTL Conditioning Soybean Sprout Yield and Quality.   Theor Appl Genet 103:702–709.

Liang H., Xu L., Yu Y., Yang H., Dong W., Zhang H. (2010) Identification of QTL with Main, Epistatic and QTL by Environment Interaction Effects for Seed Shape and Hundred-Seed Weight in Soybean Across Multiple Years.   Journal of Genetics   95:475-477.

Liu k. (1997) Soybeans: Chemistry, Technology, and Utilization Chapman & Hall, New York.

Ma J., Zhang Z., Wang B., Kong X., Wang Y., Cao S., Feng Y. (2006) Overexpression and Characterization of a Lipase from *Bacillus subtilis*.   Protein Expression and Purification 45:22-29.   DOI: http://dx.doi.org/10.1016/j.pep.2005.06.004.

Ma Y.-Y., Liu Q.-B., Yang H.-W., Tang Y.-Y., Yang J.-G. (2015) Study on the Improvement of Natto-production Process.   Advance Journal of Food Science and Technology   7:704-708. DOI: 10.19026/ajfst.7.1631.

Maughan P.J., Maroof M.A.S., Buss G.R. (2000) Identification of Quantitative Trait Loci Controlling Sucrose Content in Soybean (Glycine max).   Molecular Breeding   6:105–111.

Mayta J., Chen P., Popp M.P., Dong D., Wu C.J., Zhang B., Smith S.F., Scaboo A.M. (2014) Break-Even Profitability for Food-Grade Specialty Soybeans.   Agricultural Science   2:1-11. DOI: 10.12735/as.v2i2p01.

Molnar S.J., Charette M., Cober E.R. (2012) Mapping Quantitative Trait Loci for Water Uptake in a Recombinant Inbred Line Population of Natto Soybean. Canadian Journal of Plant Science 92:257-266. DOI: 10.4141/cjps2011-122.

Orazaly M., Chen P., Zeng A., Zhang B. (2015) Identification and Confirmation of Quantitative Trait Loci Associated with Soybean Seed Hardness. Crop Science 55:688. DOI: 10.2135/cropsci2014.03.0219.

Oyoo M.E., Githiri S.M., Benitez E.R., Takahashi R. (2010) QTL Analysis of Net-Like Cracking in Soybean Seed Coats. Breeding Science 60:28-33.

Park M.H., Kim D.C., Kim B.S., Nahmgoong B. (1995) Studies on Pollution-Free Soybean Sprout Production and Circulation Market Improvement Korea Soybean Dig 12:51-67.

Pathan S.M., Vuong T., Clark K., Lee J.-D., Shannon J.G., Roberts C.A., Ellersieck M.R., Burton J.W., Cregan P.B., Hyten D.L., Nguyen H.T., Sleper D.A. (2013) Genetic Mapping and Confirmation of Quantitative Trait Loci for Seed Protein and Oil Contents and Seed Weight in Soybean. Crop Science 53:765. DOI: 10.2135/cropsci2012.03.0153.

Pietrzak L.N., Frégeau-Reid J., Chatson B., Blackwell B. (2002) Observations on Water Distribution in Soybean Seed During Hydration Processes Using Nuclear Magnetic Resonance Imaging. Can. J. Plant Sci.:513-519.

Poysa V., Voldeng H. (2007) Natto 3 Soybean. Canadian Journal of Plant Science 87:97-98. DOI: 10.4141/P06-142.

Primomo V.S., Poysa V., Ablett G.R., Jackson C.-J., Gijzen M., Rajcan I. (2005) Mapping QTL for Individual and Total Isoflavone Coentent in Soybean Seeds. Crop Science 45:2454-2463.

Qutob D., Ma F., Peterson C.A., Bernards M.A., Gijzen M. (2008) Structural and Permeability Properties of the Soybean Seed Coat.   Botany   86:219-227.

Rodda E.D., Steinberg M.P., Wei L.S. (1973) Soybean Damage Detection and Evaluation for Food Use.   Transactions of the ASAE   16:365.   DOI: https://doi.org/10.13031/2013.37521.

Salas P., Oyarzo-Llaipen J.C., Wang D., Chase K., Mansur L. (2006) Genetic Mapping of Seed Shape inThree Populations of Recombinant Inbred Lines of Soybean (*Glycine max* L. Merr.). Theor Appl Genet   113:1459-66.   DOI: 10.1007/s00122-006-0392-1.

Sanjukta S., Rai A.K. (2016) Production of Bioactive Peptides During Soybean Fermentation and their Potential Health Benefits.   Trends in Food Science & Technology   50:1-10.   DOI: http://dx.doi.org/10.1016/j.tifs.2016.01.010.

Shi A., Chen P., Zhang B., Hou A. (2010a) Genetic Diversity and Association Analysis of Protein and Oil Content in Food-Grade Soybeans from Asia and the United States.   Plant Breeding   129:250-256.   DOI: 10.1111/j.1439-0523.2010.01766.x.

Shi H., Nam P.K., Ma Y. (2010b) Comprehensive Profiling of Isoflavones, Phytosterols, Tocopherols, Minerals, Crude Protein, Lipid, and Sugar During Soybean (*Glycine max*) Germination.   J Agric Food Chem   58:4970-6.   DOI: 10.1021/jf100335j.

Shieh C.-J., Phan Thi L.-A., Shih I.-L. (2009) Milk-clotting Enzymes Produced by Culture of *Bacillus subtilis natto.*   Biochemical Engineering Journal   43:85-91.   DOI: http://dx.doi.org/10.1016/j.bej.2008.09.003.

Shih I.-L., Yu Y.-T. (2005) Simultaneous and Selective Production of Levan and Poly(γ-glutamic acid) by *Bacillus subtilis.*   Biotechnology Letters   27:103-106.   DOI: 10.1007/s10529-004-6936-z.

Somekawa Y., Chiguchi M., Ishibashi T., Aso T. (2001) Soy Intake Related to Menopausal Symptoms, Serum Lipids, and Bone Mineral Density in Postmenopausal Japanese women. Obstetrics & Gynecology 97:109-115. DOI: http://dx.doi.org/10.1016/S0029-7844(00)01080-2.

Taira H. (1990) Quality of Soybeans for Processed Food. JARQ 24:224-230.

Tamang J.P., Shin D.H., Jung S.J., Chae S.W. (2016) Functional Properties of Microorganisms in Fermented Foods. Front Microbiol 7:578. DOI: 10.3389/fmicb.2016.00578.

TM H.-K.K., Kang S.-T., Cho J.-H., Choung M.-G., Suh D.-Y. (2005) Quantitative Trait Loci Associated with Oligosaccharide and Sucrose Contents in Soybean (Glycine max L.). Journal of Plant Biology 48:106-112.

Tuan N.A., Thuan D.T.H., Tam T.T.M., Huong N.T. (2015) Determination the Optimum Fermentation in Obtaining Nattokinase by *Bacillus subtilis natto*. Journal of Innovation and Applied Studies 13:pp. 663-668.

Van K., McHale L.K. (2017) Meta-Analyses of QTL Associated with Protein and Oil Contents and Compositions in Soybean [Glycine max (L.) Merr.] Seed. Int J Mol Sci 18. DOI: 10.3390/ijms18061180.

VanUtrecht D., Bern C.J., Rukunudin I.H. (2000) Soybean Mechanical Damage Detection. Applied Engineering in Agriculture 16:137-141.

Wang F.-Q., Yao K., Wei D.-Z. (2011) From Soybean Phytosterols to Steroid Hormones, Soybean and Health Prof., in: H. El-Shemy (Ed.), Soybean and Health, In Tech, https://www.intechopen.com/books/soybean-and-health/from-soybean-phytosterols-to-steroid-hormones.

Wang Y., Chen P., Zhang B., Singh R. (2014) Quantitative Trait Loci Analysis of Soluble Sugar Contents in Soybean. Plant Breeding 133:493-498. DOI: 10.1111/pbr.12178.

Wei Q., Chang S.K.C. (2004) Characteristics of Fermented Natto Products as Affected by Soybean Cultivars. Journal of food processing preservation 28:251-273.

Wei Q., Wolf-Hall C., Chang K.C. (2001) Natto Characteristics as Affected by Steaming Time, Bacillus Strain, and Fermentation Time. Journal of food Science 86:167-173.

Weng T.M., Chen M.T. (2012) Effect of Drying Methods on γ-PGA Isoflavone Contents and ACE Inhibitory Activity of Natto (A Fermented Soybean Food). Journal of Food Processing and Preservation 36:483-488. DOI: 10.1111/j.1745-4549.2011.00591.x.

Xie F.-T., Niu Y., Zhang J., Bu S.-H., Zhang H.-Z., Geng Q.-C., Feng J.-Y., Zhang Y.-M. (2014) Fine Mapping of Quantitative Trait Loci for Seed Size Traits in Soybean. Molecular Breeding 34:2165-2178. DOI: 10.1007/s11032-014-0171-7.

Yasui T., Toda K., Yamada T., Yumoto S., Takahashi K., Takahashi M., Hajika M. (2017) Seed Coat Cracking of Soybean (Glycine max [L.] Merr.) After Soaking and Cooking. Cereal Chemistry Journal 94:717-722. DOI: 10.1094/cchem-12-16-0284-r.

Yazdi-Samadi B., Rinne R.W., SeiP R.D. (1977) Components of Developing Soybean Seeds: Oil, Protein, Sugars, Starch, Organic Acids, and Amino Acids'. Agronomy Journal 69:481-485.

Yoshikawa Y., Chen P., Zhang B., Scaboo A., Orazaly M. (2014) Evaluation of Seed Chemical Quality Traits and Sensory Properties of Natto Soybean. Food Chem 153:186-92. DOI: 10.1016/j.foodchem.2013.12.027.

Zarkadas C.G., Voldeng H.D., Yu Z.R., Shang K., Pattison P.L. (1997) Comparison of the Protein Quality of Five New Northern Adapted Natto Soybean Cultivars by Amino Acid Analysis. Journal of Agricultural and Food Chemistry 45:2013-2019. DOI: 10.1021/jf9604697.

Zeng A., Chen P., Shi A., Wang D., Zhang B., Orazaly M., Florez-Palacios L., Brye K., Song Q., Cregan P. (2014) Identification of Quantitative Trait Loci for Sucrose Content in Soybean Seed. Crop Science 54. DOI: 10.2135/cropsci2013.01.0036.

Zhang B., Chen P., Chen C.Y., Wang D., Shi A., Hou A., Ishibashi T. (2008) Quantitative Trait Loci Mapping of Seed Hardness in Soybean. Crop Science 48:1341. DOI: 10.2135/cropsci2007.10.0544.

Zhang B., Chen P., Florez-Palacios S.L., Shi A., Hou A., Ishibashi T. (2010) Seed Quality Attributes of Food-Grade Soybeans from the U.S. and Asia. Euphytica 173:387-396. DOI: 10.1007/s10681-010-0126-y.

Zhang B., Jaureguy L., Florez-Palacios L., Chen P. (2011) Food-Grade Soybean: Consumption, Nutrition, and Health, in: E. P. a. I. Golubev (Ed.), Beans: Consumption, Nutrition, and Health, Nova Science Publishers, Inc., Hauppauge, NY. pp. 18.

2. Evaluation of important seed and sprout traits as potential selection criteria in breeding varieties for sprout soybeans

Diana M. Escamilla[1], Luciana Rosso[1], Laura Strawn[2], and Bo Zhang[1]

[1] Dep. of Crop and Soil Environmental Sciences, [2] Dep. of Food Science, Virginia Tech, Blacksburg, VA 24060

Abbreviations: HQS, high quality sprout; AQS, average quality sprout; LQS, low quality sprout; W, Warsaw VA; MH, Mount Holly VA.

Abstract

Soybean sprouts, a traditional vegetable in Asia, are gaining popularity in the United States. Soybean sprout demand has been supplied by natto (a Japanese soyfood) cultivars that share some seed characteristics with sprout cultivars. However, natto seeds do not meet all requirements of sprouts and are rejected by sprout manufacturers. The objectives of this study were to evaluate important seed and sprout traits as potential selection criteria in breeding sprout soybeans and to study the storage effect on soybean sprout quality. Almost all genotypes produced thicker and longer hypocotyls and higher fresh-sprouts than 'MFS-561', a commercial soybean sprout variety. Hypocotyl length ranged from 13.8 to 16.2 cm. Four fungi genera *Bipolaris* sp., *Cercospora* sp., *Botrytis* sp. and *Chaetomium* sp. were isolated from seeds. Cracked cotyledons and abnormal seedlings were the two main constraints affecting soybean sprout quality. Correlation coefficients among all traits indicated that percentage and weight of high- and average-quality sprouts would determine sprout yield. Acceptable yield and several traits were recommended to be used simultaneously while breeding superior sprout soybean cultivars. Good sprout varieties should produce high-quality sprouts > 48%, average-quality sprouts < 38%, low-quality sprouts < 14%, sprout yield > 5.7 g/ g seed, hypocotyl thickness > 1.6 cm and hypocotyl length > 13 cm. One-year seed storage at room temperature reduced sprout quality. V09-3876 and V12-1939 had superior seed and sprout traits and are promising lines for further evaluation for sprout production. Seed storage over time affects seed germination and seedling vigor, and fungi on seed can cause reduced seed quality.

Introduction

Soybeans average 20% oil and 40% protein content and are a major source of protein and fatty acids in human and animal nutrition (Burton, 1997). The United States plants over 30.0 million hectares per year with 33.6 million hectares in 2014 and is the world's largest soybean producer and exporter (USDA-NASS, 2015). Changes in consumer acceptance, growing global demand, and improved processing technologies have promoted the development of the specialty soybean industry (Mayta et al., 2014). Soyfood made from large- seeded soybean varieties (> 20 g/100 seeds) include tofu (soybean curd), edamame (green vegetable soybeans), miso (fermented soup-based paste), and soymilk (soybeans, soaked, fine-ground and strained). On the other hand, soyfood made from small seeded soybean varieties (< 12 g/ 100 seeds) include natto (fermented whole soybeans), soy sauce (tamari, shoyu, and teriyaki), tempeh (made of whole cooked soybeans), and bean sprouts (Zhang et al., 2010).

Soybean sprouts are a very important vegetable consumed in many Asian countries such as Korea, China, and Japan. More than 500,000 tons of soybean sprouts are consumed annually as a vegetable in soups, salads and side dishes in Korea (Hwang et al., 2004; Lee et al., 2007). There is not a universal approach for sprouting soybeans and numerous factors can affect the sprouting process (Ghani et al., 2016). Air and water temperature of 20 °C, relative humidity of 80% and several applications of water during soybean sprouting are recommended for sprouts with good quality (Ghani et al., 2016; Lee et al., 2007). The sprout harvest time is determined by seed quality and temperature (air/water) during sprouting; however, under most conditions sprouts can be ready for harvest 5-7 days after germination (Ghani et al., 2016; Silva et al., 2013). Light affects root elongation and initiates photosynthesis producing long roots and green

cotyledons, which are undesirable traits for soybean sprouts (Liu, 1997; Shi et al., 2010). Consequently, soybean sprouts in Asia are generally prepared by soaking soybean seeds for 4-5 h before the seeds are placed in a dark growth chamber at room temperature (20 °C) and watered several times per day. After four or more days, the sprouts that reach a length of about 8 cm would be ready for serving (Liu, 1997).

Several criteria such as seed weight, high germination, seed vigor, bright hypocotyl color, yellow cotyledons, hypocotyl length and sprout yield defined as sprout fresh weight produced from 50 g dry seeds divided by 50 g have been used by breeders to select sprout soybeans (Lee et al., 2007; Lee et al., 2001). Sprout soybean cultivars must have small seed size, 40–150 mg seed[-1] with high germination (> 90%) and good seedling vigor that produce sprouts with bright yellow cotyledons, good sprout length (8-12 cm) and thick hypocotyls (2.0-2.2 mm) for high sprout yield (4-6 g fresh weight/g dry weight) (Lee et al., 2001; Park et al., 1995). In addition, seeds need to absorb a considerable amount of water before germination and a uniform hydration is important for high quality soyfood, thus uniform seeds that have high and rapid water uptake are preferred (Pietrzak et al., 2002). Other sprout characteristics reported by the sprout industry and supported by previous scientific studies were percentage of abnormal seedlings (decayed seeds and seedlings, stunted seedlings, less than 2 cm growth) and percentage of high-quality sprouts (vigorous and straight sprouts). All sprout characters were measured 4-5 days after initial germination (Lee et al., 2007; Lee et al., 2001). Sprout trait determination generally is made only on advanced lines in soybean breeding programs because the evaluation requires large amount of seeds (Lee et al., 2001).

The genetic relationship among some sprout traits have been studied. For example, quantitative trait loci (QTLs) associated with sprout yield and seed weight were in the same

31

genome region, indicating that sprout yield was genetically linked to seed weight (Lee et al., 2001). Seed weight was negatively correlated with sprout yield (r = -0.763***), so using smaller seed was more acceptable than bigger seeds. Sprout yield was negatively correlated with a higher incidence of abnormal seedlings because abnormal seedlings produce fewer sprouts (Lee et al., 2001).

Sprouting conditions could favor the propagation of various microorganisms including plant pathogenic bacteria and fungi. Eighteen and seven species of fungi were isolated from spoiled *Vigna* spp. and soybean sprouts in Japan, respectively, and approximately 70% of the isolates were plant pathogens and at least 14 species were known to be seed-borne (Sato et al., 2014). The fungus species isolated from soybean sprouts were *Cercospora kikuchii* (Tak. Matsumoto & Tomoy.) M.W. Gardner, *Diaporthe phaseolorum* var. caulivora Athow and Caldwell, *Fusarium graminearum* Schwabe, *Fusarium oxysporum* Schltdl., *Penicillium oxalicum* Currie and Thom, *Phoma medicaginis* Malbr. and Roum, *Phomopsis phaseoli* var. sojae (Lehman) Sacc. and *Syncephalastrum racemosum* Cohn ex J. Schröt (Sato et al., 2014). In addition, three bacteria species, *Bacillus cereus, Aeromonas hydrophila,* and *Pseudomonas aeruginosa,* found in raw soybean sprouts were tested for pathogenesis or food spoilage (Kim et al., 2002). However, seed-borne fungi incidence has not been included as a trait to evaluate soybean sprout quality.

Breeders use natto soybeans as sprout genotypes because natto cultivars and sprout soybean cultivars share some characteristics (Zhang et al., 2010). Natto, as well as sprout soybeans, require cultivars with small, uniform seed size (< 9 g 100 seeds^{-1}), which is the primary characteristic used to identify natto cultivars. Other important characteristics are smooth seeds, light colored seed coat, clear hilum and high water absorption (Cui et al., 2004; Liu,

1997). In addition, natto seed coat should remain intact despite weakening and expanding during water absorption, which is defined as low seed coat deficiency (Cook and Rainey, 2010). A high content of carbohydrates is also important for quick conversion to simple sugars, an essential step in the fermentation process which is the basis of natto preparation (Liu, 1997; Taira, 1990). However, there is no information about whether or not the natto quality traits affect soybean sprout acceptance.

There is insufficient information on sprout soybean characteristics for breeding purposes. Seed quality and sprout quality studies are needed to facilitate breeding programs in the United States. Bad sprout quality causes seed rejection by sprout manufacturers, but soybean sprout quality traits, such as seed-borne fungi incidence, abnormal seedlings, and high-quality sprouts, have not been well characterized. In the sprout industry, seeds are stored for usually up to one year before a new crop can be used to produce sprouts. Sometimes, seeds are stored even longer due to the cleaning and shipping process of newly harvested seeds. However, germination of stored seeds is often lower than freshly harvested seeds (Mbofung et al., 2013). Therefore, sprout quality may also be reduced due to long-term storage. Nevertheless, no information is available on how sprout quality change after seeds are stored after a period. Therefore, the objectives of the current study were to measure seed and sprout traits related to good quality sprouts for selection of sprout soybeans, and to study the storage effect on soybean sprout quality.

Materials and Methods

Materials

A total of 15 small-seeded genotypes including 'MFS-561', a commercial cultivar for soybean sprouts and one conventional yield check, 'Glenn', were evaluated for sprouting characteristics in order to improve sprout varieties (Table 1). All genotypes had white flowers to produce white hypocotyls to meet sprout requirements. To study storage effects, eight genotypes were tested for sprout quality after they were stored at room temperature for seven and 14 months.

Experimental field procedures

The genotypes were planted as full-season soybean in Warsaw, VA (W) and as double-cropped soybean after small-grain was harvested in Mount Holly, VA (MH) with three replications in a complete randomized design in 2014. Eight seeds were planted per foot. The plots in W were 4-rows, 5.5 m long spaced 0.76 m and the plots in MH were 5-rows, 5.5 m long spaced 0.18 m, and all plots were end-trimmed to 3.7 m to reduce the end effects of each plot. W has kempsville loam soil and MH has state soil series. Fertilizer was applied before planting according to soil test results at both locations, pre-herbicide Dual Magnum, 1.66 pints per acre, was applied before planting to reduce weed pressure. No irrigation or insecticides were applied.

Measurement of sprout characteristics

A total of 350 unbroken and undamaged seeds of each genotype were selected and grown in a bean sprouter (Cheong Si Ru, SC-9000TS, Korea) at room temperature for five days with two replications. The 350 dry seeds were weighed before sprouting. Water in the sprouters

was changed twice per day to keep sprouts fresh and to reduce mold. Sprout characteristics were measured on the fifth day.

Sprouts were categorized into three levels: high-quality (HQ) sprouts were those with vigorous, straight and long (> 7 cm) with bright yellow cotyledons and white hypocotyls (Fig. 1); average (AQ) sprouts were those with short hypocotyls (less than 7 cm), curled seedlings, damage cotyledons or cracked and open cotyledons (Fig. 2), and the low-quality (LQ) group included the decayed seeds and seedlings, stunted seedlings (Fig. 3). Percentage of high-, average- and low-quality sprouts out of the 350 evaluating seeds was calculated for each sample. The fresh weight of high- , average-, and low-quality sprouts was also recorded. Sprout yield was calculated as total fresh weight of high- and average- quality sprouts divided by the dry weight of the sample**.** Hypocotyl length was measured from the point of initiation of the first root to the cotyledons, and the thickness was measured at the middle point of the sprout hypocotyl using a digital caliper (Fraction Plus 6 in. 3-Mode Digital Caliper, Model # 147). Hypocotyl length and thickness were taken from 20 random seedlings from the high-quality category of each sample.

Measurement of mold incidence

Ten seeds of each genotype were washed in sterile water for one minute and then placed on PDA plates (Fig. 4). PDA is a relatively rich medium for growing a wide range of fungi. The plates were incubated at room temperature. Fungi were identified at the genera level by asexual structures using a microscope with a 440 x magnification (Olympus, model # 920806, made in Japan). The number of infected seeds was recorded, and the percentage of molding was calculated as the percentage of infected seeds out of the ten evaluated seeds.

Measurement of water absorption

A 20 g seed sample of each line was placed in a plastic container and immersed in 100 ml of deionized water at room temperature for 16 h. After soaking, the samples were drained for 15 s and then placed on a paper towel to remove any excess surface water. These samples were then reweighed. The procedure was also repeated. Water absorption was calculated as (weight after total water absorption/initial weight) × 100 (Geater et al., 2000).

Seed storage effect on sprout characteristics

Seeds of eight genotypes were harvested in Warsaw on November 2014, and stored at room temperature for seven and 14 months before they were tested for seed and sprout traits. The eight genotypes included 'MFS-561', a commercial cultivar for soybean sprouts, 'Glenn', one conventional yield check, and six small-seeded breeding lines (Table 1).

Statistical Analysis

Statistical analysis was computed in JMP statistical version 11.0 (SAS Institute, Raleigh, NC). Prior to conducting statistical analyses, assumptions for two-way analysis of variance model were checked for all response variables. Assumptions are normal distribution of data and homogeneous variance (McKillup 2012). Normality assumption was assessed by Shapiro-Wilk test and normal probability plots; and homogeneity of variance was evaluated by residuals vs. predicted values plots and plots of residuals by groups (McKillup 2012). Variables that violated the assumptions of normal distribution and equal variance were properly transformed and analyzed. All variables met assumptions after transformation.

Soybean genotypes were evaluated under a two factor completely randomized design with genotypes (15 small-seeded lines) and locations (MH and W) as factors for studying each

genotype performance and environmental effects on seed and sprout quality traits. A two factor completely randomized design with genotypes (8 small-seeded lines) and seed storage time (7 and 14 months) as factors was used to assess the effect of seed storage on sprout traits. The sprouting performance of genotypes, the environmental effect, genotype x environment (G x E) interaction effect and seed storage effect were evaluated using a two-way analysis of variance (ANOVA). Tukey's HSD test was used to show the variables that differed significantly at P=0.05. Pearson correlation coefficients among seed and sprout traits were calculated.

Results and Discussion

Fungi incidence

Results showed that four fungus genera were widely distributed among genotypes. They were *Bipolaris* sp., *Cercospora* sp., *Botrytis* sp., *Chaetomium* sp. (brown and yellow) and one unknown mold (Fig. 5). Fungi incidence average across genotypes and environments was 63.33% and there was no significant difference among genotypes or environments (Tables 2, 3). Although mold percentage was high on seeds in this experiment, the fungi development on actual sprout production was low if environmental conditions for fungi growth (such as high temperature) were not favorable. *Cercospora kikuchii* (Tak. Matsumoto & Tomoy.) was reported as a fungus spoiling soybean sprouts, while fungus species of *Chaetomium* sp. were isolated from *Vigna* spp. bean sprouts (Sato et al., 2014). Some *Chaetomium* species have antagonistic properties that make them interesting for biological hazard studies (Prokhorov and Linnik, 2011). To our knowledge, *Botrytis* sp. and *Bipolaris* sp. have not been reported on soybean sprouts before. *Botrytis* species are important pathogens in many crops; the most common specie, *Botrytis cinerea,* infects more than 200 plant species and it is the causal agent of gray mold in many economically important crops (Elad et al., 2004). The genus *Bipolaris* also includes a significant number of pathogens on rice, wheat and other host plants (Manamgoda et al., 2014). Thus, all fungus genera isolated from this study, are plant pathogens, except *Chaetomium* sp. that is a biological control agent against several plant pathogens. In this study, seeds were only treated with water before the mold test to allow seedborne fungi to develop and grow from environmental contamination. Soybean sprout consumption has been associated with numerous outbreaks of foodborne illness, which represents a risk to human health and the soybean sprout industry in the United States. Therefore, further studies are needed to identify fungal species in

38

the Virginia growth environment and to establish risk management strategies to prevent foodborne diseases and develop food safety and sanitation polices within the soybean sprout industry. Due to the non-significant difference among genotypes and the low fungi development on actual sprout production, mold incidence may not be an important trait in the breeding selection of sprout cultivars.

Seed size

Small to medium-sized seeds are preferred for sprouting because of their uniform germination and it is one of the factors considered while choosing soybean cultivars for sprout production (Ghani et al., 2016). Seed size of small-seeded lines ranged from 8.3 to 12.4 g/ 100 seed, these seed sizes met the requirement for sprout soybeans where a seed weight of <12 g/ 100 seed is preferred (Kwon et al., 1972). Glenn, the commercial check, had the biggest seed size of 14.9 g/ 100 seeds (Table 4). Environment and G x E interaction did not show significant effect on seed size (Table 3). However, significant differences among genotypes were presented indicating seed size varies even within a 4 g/ 100 seeds range in the mid-Atlantic small-seeded germplasm (Table 3).

Water absorption

A high and rapid water uptake by soybean seeds is important for quality soyfood (Pietrzak et al., 2002). Average water absorption of all genotypes across environments was 226% (Table 2). There was no difference among the mean of water absorption for the genotypes and environments tested (Table 3). A study of natto soybean traits reported that water absorption rate ranged from 215 and 225% among natto soybeans which was in agreement with our results (Cook and Rainey, 2010). The lack of phenotypic differences in this study indicated that water

39

absorption may not be an informative trait for selecting superior sprout cultivars because small seeds tend to have similar water absorption rates.

Hypocotyl thickness and length

Soybean lines had an average of 14.95 cm long and 1.74 mm thick hypocotyl (Table 2). There was significant phenotypic variation among genotypes for both traits (Table 3). The hypocotyl length of most lines was around 15 cm except for V12-1764 at 13.8 cm. The hypocotyl thickness of all genotypes ranged from 1.6 to 1.9 mm where V12-1725 and V12-1818 had significant higher hypocotyl thickness than the commercial sprout variety 'MFS-561' (Table 4). The average hypocotyl length and thickness in our study were higher than those previously reported by Lee et al. (2007), where they averaged 10 cm and 1.6 mm, respectively. The sprouting temperature might affect the hypocotyl growth because previous sprout studies were conducted at a temperature of 20 °C while this study was conducted at 22 °C (Lee et al., 2007; Lee et al., 2001). It was reported that sprout quality such as whole sprout length and hypocotyl length increased at a higher germination temperature (25 vs 20 °C). Hypocotyl length at 20 °C was around 7 cm and 16 cm at 25 °C (Koo et al., 2015). In addition, the genetic differences among varieties and the sprouter we used that provides favorable sprout growing condition would also influence hypocotyl length and thickness.

High-, average- and low- quality sprouts

Significant differences among genotypes were shown in high-quality sprout percentage (HQS%), average-quality sprout percentage (AQS%), low-quality sprout percentage (LQS %), HQS, AQS and LQS fresh weight; meaning that high quality sprout varieties can be developed through effective selection for specific sprout traits. HQS%, AQS% and fresh weight of HQS and AQS were also significantly affected by environment, so it is important to choose the right

locations and agronomic practice to produce sprout seeds. In addition, AQS% and AQS fresh weight were affected by genotype x environment interaction (Table 3), which would make it difficult to identify the best genotypes for these specific traits and environments during breeding selection.

The percentage of HQ, AQ and LQ sprouts ranged from 33.3 to 62.7%, 28.0 to 49.2%, and 7.7 to 29.7%, respectively (Table 5). V12-1818 (62.7%), V09-3876 (59.7%), V12-1827 (57.9%) and V12-1789 (57.5%) had the highest percentage of HQ sprouts, the least percentage of AQ sprouts and small levels of LQ sprouts. They also had the highest fresh weight of HQ sprouts among all lines. V12-1939 had the highest AQ sprout fresh weight and percentage among all lines. V12-1764 and V09-3771 had the least HQ sprouts, so they were not good soybean varieties for sprout use (Table 5).

One of the main reasons for reduced sprout quality was cracked cotyledons (Fig. 3), which generated a high AQ sprout percentage and fresh weight. Soaking injury or soaking damage, when seeds are severely desiccated, cracked, or lacking cotyledons or overall damage to the hypocotyl, of leguminous seeds is a classic problem in agriculture and food processing. Increasing moisture content of seeds before imbibition may alleviate soaking damage since seeds with the water content of 12-17% did not suffer from soaking effects (Ashworth and Oberdorf, 1980; Koizumi et al., 2008; Pollock, 1969). It was reported that environmental stresses (chilling temperatures about 15 °C) during flowering induced seed coat cracking as well as mottling around the hilum region (Sunada and Ito, 1982). Cracked cotyledons were also associated with harvest injuries and excess seed loss during seed cleaning (Liu and Pappelis, 1971). In addition, seeds with a semi-hard and tightly adhering seed coat do not crack during imbibition at a low temperature because of the slow initial water uptake regulated by the seed coat (Sorrels and

Pappelis, 1976). Koizumi et al. (2008) reported that intact seed coat of soybean seeds regulated water incorporation into the radicle, hypocotyl, and cotyledons, and prevented the destruction of seed tissues at the beginning of imbibition. Thus, seed coat deficiency could be associated with susceptibility to cracked cotyledons. Low seed coat deficiency is also an important trait for natto soybeans (Cook and Rainey, 2010). Therefore, conditions during seed development, mechanical damage, high seed coat deficiency, and low moisture content of the seeds may have caused cracked cotyledons in this study. The average over locations for low-quality sprout percentage was 13.8% (Table 2). Similar results were found in a previous sprout trait study with abnormal seedlings averaging from 3 to 34% with the heritability of abnormal seedlings of only 0.13 (Lee et al., 2001). Results showed that cracked cotyledons and abnormal seedlings were the two main constraints in soybean sprout quality. Hence, reduction in the number of cracked cotyledons and abnormal seedlings should be studied further.

Sprout yield and promising sprout lines

Small-seeded lines in this study showed significant phenotypic variation in sprout yield (Table 3). V09-3876 had higher sprout yield (7.0 g/ g seeds) than the two checks 'MFS-561' (5.24 g/ g seeds) and 'Glenn' (5.24 g/ g seeds) (Table 4). The average sprout yield in our study was 6.1 g/ g seed, which was consistent with the results of a previous study that sprout yield of small- seeded genotypes (11.3 and 16.4 g/100 seeds) ranged from 4.8 and 6.3 g/ g seed (Table 2) (Lee et al., 2001). V09-3876 was the best sprout line among all genotypes tested because it had small seed size of 9.2 g/ 100 seeds, high percentage of HQ sprouts, HQS fresh weight, long sprouts and the highest sprout yield than the other genotypes (Fig. 6). Although V12-1818, V12-1827 and V12-1789 showed desirable sprout characteristics, their seed size was slightly larger than optimum seed size, of ≤ 10 g/ 100 seeds. V12-1939 could also be considered as a sprout

variety candidate because it was 9.9 g/ 100 seeds and showed competitive sprout characteristics compared to the checks.

Seed and sprout trait correlations

Correlation coefficients between all traits (Table 6) indicated low-quality sprout percentage was significantly and negatively correlated to high-quality sprout percentage (-0.72), high-quality sprout fresh weight (-0.62) and sprout yield (-0.61). High-quality sprout percentage was positively correlated to sprout yield (0.59). High-, average-, and low-quality sprout percentages were positively correlated to high- (0.90), average- (0.61) and low- (0.93) quality sprout fresh weight, respectively. Seed size had a positive and significant correlation with HQS (0.48) and AQS (0.44), and hypocotyl thickness (0.46), so smaller seeds within acceptable sprout seed size (8.3 g/ 100 seed) produced sprouts with lower quality. Seed size was not significantly correlated to sprout yield. However, a study of soybean sprouts reported that seed weight was negatively associated with sprout yield (r= -0.763***), and sprout producers prefer cultivars with a small seed weight (Lee et al. 2001). In our results, seed size was not correlated to sprout yield probably because all lines tested had small seeds (<12.4 g/ 100 seeds), while in the previous sprout soybean study large and small seeded genotypes were tested. Small seed size also explained why there was no significant difference in water absorption among genotypes in our study (Lee et al., 2001). Water absorption showed a positive correlation with LQS% (Table 5). Higher water absorption may also be as a result of cracked seed coats associated with bad quality seed. Hypocotyl length had a significant positive correlation with HQS% (0.33) while hypocotyl thickness had a significant correlation with HQS (0.47). Mold incidence was not significantly correlated with any trait. Correlation coefficients indicated that several traits must be selected simultaneously while breeding superior sprout soybean cultivars with high quality and sprout

yield. Those traits were recommended: seed size, hypocotyl length and thickness, high- and average- quality sprout percentage and weight, and sprout yield.

Seed storage effects

There was a significant effect of seed storage on most seed and sprout traits except for seed size, AQS%, AQS fresh weight and hypocotyl thickness (Table 7). Soybean seeds usually have an extended period of postharvest storage; stored grains are biologically active and they may deteriorate under long storage periods, high temperature and relative humidity (Yousif, 2014). Studies about seed storage effects on sprout traits have not been reported. Sprout traits depend on high seed viability, and a reduction in sprout quality was expected as time in seed storage increased. A high LQS% represents a reduction in seed germination and an increase of abnormal seedlings. Thus, our results showed a significant reduction in seed germination, fungi incidence, hypocotyl length, sprout yield and high-quality sprouts (HQS% and HQS fresh weight) after seeds were stored at room temperature (20 °C) for more than a year. There was no significant difference of AQS% and AQS fresh between the storage times. This indicated that seed storage may mainly affect seed germination (LQS %) and seedling vigor (HQS %), but it did not have a significant effect on the number of curled seedlings or cracked and open cotyledons. Studies have shown that seedling vigor and germination were affected during seed storage. Fluctuation of temperature and relative humidity were determinants of the seed deterioration rates (Mbofung et al., 2013). Seeds that were stored at a controlled temperature of 15 or 20 °C had higher germination rate than that of those stored at fluctuating environmental temperatures (Mbofung et al., 2013). It was also determined that seed viability remained high (> 92%) in seeds under cold storage (10 °C) and seed viability was moderately high (> 78%) in seeds under warm storage (20 °C) after 20 months (Mbofung et al., 2013). In this study, seeds

were stored at room temperature (20 °C), which may explain the significant reduction in seed germination (high LQS%), HQS%, sprout yield and hypocotyl length after 14 months of storage. Storage fungi are a major cause of reduced seed quality in stored seed (Mbofung et al., 2013). Our results showed that fungi incidence was reduced after a longer storage. It has been reported a reduction in the percentage of field fungi with longer storage due to the absence of high relative humidity in the seeds (Bhattacharya and Raha, 2002). Results showed that water absorption increased with length of seed storage because seeds are drier, and it may also be as a result of seed coat deterioration that allows water to enter more freely. Thus, unfavorable soybean storage will result in a significant loss of sprout quality that could lower commercial value and reduce acceptability of soybean sprouts.

Conclusions

Among seed and sprout traits, mold incidence and water absorption of small seeds might not be informative traits for selecting superior sprout soybean cultivars because they were similar among all small-seeded genotypes. Fresh weight of HQS and AQS were significantly correlated to the percentage of HQS and AQS, respectively. High- and average- quality sprout percentage are currently used to evaluate sprouts in the sprout industry, so the fresh weight of HQS and AQS were not recommended as sprout breeding selection traits. Hypocotyl length and thickness are important traits for sprout quality and yield. Thus, seed size, high-, average- and low-quality sprout percentage, hypocotyl thickness and length and sprout yield should be considered as the most important variables to evaluate soybean sprout quality in breeding programs. Using 'MFS-561' and the average value of traits as references, we suggested that soybean sprout varieties should have a seed size of about 10 g/ 100 seed, high-quality sprouts > 48%, average-quality sprouts < 38%, low-quality < 14%, sprout yield of 5.7 g/ g seed, hypocotyl thickness of 1.6 mm and hypocotyl length > 13 cm. V09-3876 and V12-1939 are considered sprout variety candidates because they were < 10 g/ 100 seeds and showed competitive sprout characteristics compared to the checks. Seed storage can affect seed germination and seedling vigor, and storage fungi can be a major cause of reduced seed quality.

References

Ashworth E.N., Oberdorf R.L. (1980) Imbibitional Chilling Injury in Soybean Axes: Relationship to Stelar Lesions and Seasonal Environments. Agron J 72:923-928.

Bhattacharya K., Raha S. (2002) Deteriorative Changes of Maize, Groundnut and Soybean Seeds by Fungi in Storage. Mycopathologia 155:135-141.

Burton J.W. (1997) Soyabean (Glycine Max(L.) Merr.). Field Crops Research 53:171-186.

Cook D.E., Rainey K.M. (2010) Seed Coat Deficiency, Trait Stability, and Other Soybean Seed Quality Traits for Natto Cultivar Development. Crop Science 50:1244-1249. DOI: 10.2135/cropsci2009.06.0300.

Cui Z., James A., Miyazaki S., Wilson R., Carter T. (2004) Breeding Specialty Soybeans for Traditional and New Soyfoods, in: L. K (Ed.), Soybeans as Functional Foods and Ingredients, AOCS Press, Urbana, IL. pp. 264-322.

Elad Y., Williamson B., Tudzynski P., Delen N. (2004) Botrytis Spp and Diseases They Cause in Agriculture Systems-an Introduction, in: E. Y, et al. (Eds.), Botrytis: Biology, Pathology and Control, Kluwer Academic Publishers, Dordrecht, the Netherlands pp. 1-6.

Geater C.W., Fehr W.R., L A.W. (2000) Association of Soybean Seed Traits with Physical Properties of Natto. Crop Science 40:1529-1534.

Ghani M., Kulkarni K.P., Song J.T., Shannon J.G., Lee J.-D. (2016) Soybean Sprouts: A Review of Nutrient Composition, Health Benefits and Genetic Variation. Plant Breeding and Biotechnology 4:398-412. DOI: 10.9787/pbb.2016.4.4.398.

Hwang Y.H., Jeong Y.S., Lee J.D. (2004) Present Status and Future Developmental Direction of Soy-Related Industries in Korea. Korea Soybean Digest 21:28-44.

Kim H.-J., Koo K.-M., Kim G.-N., Lee D.S., Paik H.-D. (2002) Isolation and Identification of Pathogenic Microorganims from Soybean Sprouts.   Nutraceuticals & Food   7:305-309.

Koizumi M., Kikuchi K., Isobe S., Ishida N., Naito S., Kano H. (2008) Role of Seed Coat in Imbibing Soybean Seeds Observed by Micro-Magnetic Resonance Imaging.   Ann Bot 102:343-52.   DOI: 10.1093/aob/mcn095.

Koo S.C., Kim S.G., Bae D.W., Kim H.Y., Kim H.T., Lee Y.H., Kang B.K., Baek S.B., Baek I.Y., Yun H.T., Choi M.S. (2015) Biochemical and Proteomic Analysis of Soybean Sprouts at Different Germination Temperatures.   J Korean Soc Appl Biol Chem. 58:397-407.   DOI: 10.1007/s13765-015-0053-7).

Kwon S.H., Im K.H., Kim J.R. (1972) Studies on Diversity of Seed Weight in Korean Soybean Land Races and Wild Soybean.   Korean J Breed   4:70-75.

Lee J.-D., Shannon J.G., Jeong Y.-S., Lee J.-M., Hwang Y.-H. (2007) A Simple Method for Evaluation of Sprout Characters in Soybean.   Euphytica   153:171-180.   DOI: 10.1007/s10681-006-9252-y.

Lee S.H., Park K.Y., Lee H.S., Park E.H., Boerma H.R. (2001) Genetic Mapping of Qtls Conditioning Soybean Sprout Yield and Quality.   Theor Appl Genet 103:702–709.

Liu K. (1997) Soybeans: Chemistry, Technology, and Utilization Chapman & Hall, New York.

Liu K.C., Pappelis J. (1971) Transverse and Crescent Cracks in Soybean Cotyledons Associated with Imbibition.   Trans Ill Acad Sci   64:43-45

Manamgoda D.S., Rossman A.Y., Castlebury L.A., Crous P.W., Madrid H., Chukeatirote E., Hyde K.D. (2014) The Genus Bipolaris.   Stud Mycol   79:221-88.   DOI: 10.1016/j.simyco.2014.10.002.

Mayta J., Chen P., Popp M.P., Dong D., Wu C.J., Zhang B., Smith S.F., Scaboo A.M. (2014) Break-Even Profitability for Food-Grade Specialty Soybeans. Agricultural Science 2:1-11. DOI: 10.12735/as.v2i2p01.

Mbofung G.C.Y., Goggi A.S., Leandro L.F.S., Mullen R.E. (2013) Effects of Storage Temperature and Relative Humidity on Viability and Vigor of Treated Soybean Seeds. Crop Science 53:1086. DOI: 10.2135/cropsci2012.09.0530.

Park M.H., Kim D.C., Kim B.S., Nahmgoong B. (1995) Studies on Pollution-Free Soybean Sprout Production and Circulation Market Improvement Korea Soybean Dig 12:51-67.

Pietrzak L.N., Frégeau-Reid J., Chatson B., Blackwell B. (2002) Observations on Water Distribution in Soybean Seed During Hydration Processes Using Nuclear Magnetic Resonance Imaging. Can. J. Plant Sci.:513-519.

Pollock B.M. (1969) Imbibition Temperature Sensitivity of Lima Bean Seeds Controlled Y Initial Seed Moisture. Plant Physiol 44.

Prokhorov V.P., Linnik M.A. (2011) Morphological, Cultural, and Biodestructive Peculiarities of Chaetomium Species. Moscow University Biological Sciences Bulletin 66:95-101. DOI: 10.3103/s0096392511030072.

Sato T., Aoki M., Aoki T., Kubota M., Yaguchi T., Uzuhashi S., Tomioka K. (2014) Fungi Isolated from Spoiled Bean Sprouts in Japan. JARQ 48:317-329.

Shi H., Nam P.K., Ma Y. (2010) Comprehensive Profiling of Isoflavones, Phytosterols, Tocopherols, Minerals, Crude Protein, Lipid, and Sugar During Soybean (Glycine Max) Germination. J Agric Food Chem 58:4970-6. DOI: 10.1021/jf100335j.

Silva L.R., Pereira M.J., Azevedo J., Gonçalves R.F., Valentão P., de Pinho P.G., Andrade P.B. (2013) Glycine Max (L.) Merr., Vigna Radiata L. And Medicago Sativa L. Sprouts: A

Natural Source of Bioactive Compounds. Food Research International 50:167-175. DOI: 10.1016/j.foodres.2012.10.025.

Sorrels M.E., Pappelis A.J. (1976) Effect of Temperature and Osmotic Concentration on Cotyledon Cracking During Imbibition of Soybean. Crop Soil Sci 16:413–415.

Sunada K., Ito T. (1982) Soybean Grain Quality as Affected by Low Temperature Treatments in Plants (Color of Hilum, Seed Coat Cracking). Rep Hokkaido Branch Crop Sci Soc Jpn and Hokkaido Branch Jpn Soc Breed 22:34.

Taira H. (1990) Quality of Soybeans for Processed Food. JARQ 24:224-230.

USDA-NASS. (2015) National Statistics for Soybeans, United States Department of Agriculture-National Agricultural Statistics Service.

Yousif A.M. (2014) Soybean Grain Storage Affects Grain Testa Color, Texture and Cooking Quality. Journal of Food Quality 37 18-28.

Zhang B., Chen P., Florez-Palacios S.L., Shi A., Hou A., Ishibashi T. (2010) Seed Quality Attributes of Food-Grade Soybeans from the U.S. And Asia. Euphytica 173:387-396. DOI: 10.1007/s10681-010-0126-y.

Tables and Figures

Table 1: Fifteen soybean genotypes tested for sprout traits across Warsaw and Mount Holly, VA

Name	Pedigree
V09-3876	V00-3650 x V01-4937
V10-3916*	V00-3488 x V03-1154
MFS-561*	Check
V09-3854*	V00-3650 x V01-4937
V09-3771	V00-3650 x S99-3181
V11-0815	V00-3636 x V03-1154
V05-5973W	S98-3212 X V97-3835
V12-1939	V03-0986 x R04-198
V12-2249	V01-4937 x R04-198
V12-1725	V03-0986 x V01-2245
V12-1789*	V03-0986 x V01-2245
Glenn*	Check
V12-1827*	V03-0986 x V01-2245
V12-1818*	V03-0986 x V01-2245
V12-1764	V03-0986 x V01-2245

* Tested for seed storage effect on sprout traits

Table 2: Mean of seed and sprout traits across Warsaw and Mount Holly, VA

Seed and sprout traits[1]	Units	Mean	SE
SS	g/100 seed	10.46	0.21
HQS %	%	47.84	1.52
AQS %	%	38.31	1.05
LQS %	%	13.85	1.10
HQS	g	131.64	5.02
AQS	g	89.09	2.79
LQS	g	15.49	1.31
Hlgth	cm	14.95	0.11
Hthk	mm	1.74	0.02
Wabs	%	225.89	0.49
Syld	g/g seeds	6.03	0.10
Finc	%	63.33	28.38

[1] The seed and sprout trait names are abbreviated as follows: SS is seed size, HQS% is high quality sprout percentage, AQS% is average quality sprout percentage, LQS% is low quality sprout percentage, HQS is high quality sprout fresh weight, AQS is average quality sprout fresh weigh, LQS is low quality sprout fresh weight, Hlgth is hypocotyl length, Hthk is hypocotyl thickness, Wabs is water absorption, Syld is sprout yield, Finc is fungi incidence and SE is standard error

Table 3: Analysis of variance results of 15 soybean genotypes grown in Warsaw and Mount Holly, VA

Source	DF	Seed and Sprout Traits[1]											
		Finc	SS	Wabs	HQS%	AQS%	LQS%	HQSg	AQS	LQS	Hlght	Hthk	Syld
Genotype	14	0.1925	<.0001*	0.115	<.0001*	<.0001*	0.0006*	<.0001*	<.0001*	0.0002*	0.0043*	0.005*	0.0094*
Location	1	0.7856	0.2921	0.439	0.0065*	<.0001*	0.4745	0.0984	<.0001*	0.8156	0.8144	0.673	0.7391
Genotype* Location	14	0.7867	0.4226	0.617	0.0727	0.007*	0.6704	0.2135	0.001*	0.6205	0.3470	0.738	0.1241

[1] The seed and sprout trait names are abbreviated as follows: SS is seed size, HQS% is high quality sprout percentage, AQS% is average quality sprout percentage, LQS% is low quality sprout percentage, HQS is high quality sprout fresh weight, AQS is average quality sprout fresh weigh, LQS is low quality sprout fresh weight, Hlgth is hypocotyl length, Hthk is hypocotyl thickness, Wabs is water absorption, Syld is sprout yield, Finc is fungi incidence and DF is degrees of freedom
*Significant at P≤ 0.01 level

Table 4: Mean seed and sprout traits of 15 soybean genotypes grown in Warsaw and Mount Holly, VA

Genotype	SS (g/ 100 seeds)[1]		Hlght (cm)[1]		Hthk (mm)[1]		Syld (g/ g seeds)[1]	
	Mean	SE[†]	Mean	SE	Mean	SE	Mean	SE
Glenn	14.9 a	0.35	14.6 abc	0.44	1.83 ab	0.04	5.24 b	0.22
V12-1827	12.5 b	0.19	16.2 a	0.27	1.85 ab	0.01	6.11 ab	0.13
V12-1818	11.6 bc	0.14	14.8 abc	0.31	1.88 a	0.03	6.57 ab	0.19
V12-1789	11.4 bcd	0.17	15.4 abc	0.33	1.74 ab	0.06	6.04 ab	0.24
V12-1725	10.9 cde	0.24	14.7abc	0.16	1.91 a	0.07	6.02 ab	0.27
V12-2249	10.6 cdef	0.20	14.9 abc	0.35	1.76 ab	0.02	6.2 ab	0.40
V10-3916	10.2 defg	0.65	14.7 abc	0.70	1.69 ab	0.05	5.73 ab	0.31
V12-1939	9.9 efg	0.09	15.7 ab	0.17	1.73 ab	0.02	6.81 ab	0.11
V11-0815	9.8 efg	0.14	15.6 abc	0.23	1.64 ab	0.02	6.11 ab	0.45
V05-5973W	9.7 fg	0.27	14.1 bc	0.27	1.75 ab	0.02	6.16 ab	0.39
V12-1764	9.4 gh	0.31	13.8 c	0.26	1.68 ab	0.13	5.6 ab	0.39
V09-3854	9.2 gh	0.23	14.9 abc	0.23	1.69 ab	0.04	6.29 ab	0.36
V09-3876	9.2 gh	0.05	15.4 abc	0.34	1.69 ab	0.05	7.01 a	0.23
MFS-561	9.2 gh	0.02	14.7abc	0.32	1.57 b	0.03	5.24 b	0.31
V09-3771	8.3 h	0.04	14.8 abc	0.50	1.63 ab	0.06	5.33 b	0.75

[1] The seed and sprout trait names are abbreviated as follows: SS is seed size, Hlgth is hypocotyl length, Hthk is hypocotyl thickness and Syld is sprout yield

[†]SE is standard error

*Different letters in the same column indicated the significant difference at the 0.05 probability level

Table 5: Sprout quality of 15 soybean genotypes grown in Warsaw and Mount Holly, VA

Genotype	HQS%[1]		AQS%[1]		LQS%[1]		HQS (g)[1]		AQS (g)[1]		LQS (g)[1]	
	Mean	SE[†]	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Glenn	51.1 abcd	2.89	35.1 bc	2.97	13.9 abc	3.06	146.9 abcd	17.99	125.7 a	9.59	9.5 bcd	0.55
V12-1827	57.9 abc	4.42	32.1 bc	4.63	10.1 abc	0.89	181.5 ab	18.28	85.6 cd	9.70	11.7 abcd	0.74
V12-1818	62.7 a	2.54	28.0 c	1.56	9.3 c	1.79	193.1 a	9.67	72.4 d	3.37	11.5 abcd	1.72
V12-1789	57.5 abc	4.02	33.7 bc	3.63	8.8 c	1.06	162.5 abc	10.17	79.2 d	9.32	9.8 cd	2.24
V12-1725	39.9 cd	3.61	43.4 ab	1.38	16.7 abc	2.81	121.4 bcd	9.04	110.2 abc	7.13	23.8 abc	3.67
V12-2249	51.0 abcd	4.99	38.6 abc	4.99	10.4 abc	0.92	136.7 abcd	13.99	93.1 bcd	12.94	10.9 bcd	1.62
V10-3916	43.1 bcd	6.95	41.1 ab	6.26	15.7 abc	6.34	121.5 bcd	17.20	83.7 cd	9.08	20.4 abcd	8.79
V12-1939	43.1 bcd	3.82	49.2 a	3.86	7.7 c	0.85	116.4 bcd	10.78	118.7 ab	10.08	7.1 d	0.84
V11-0815	45.8 ab	3.76	43.9 ab	2.10	10.3 abc	1.76	120.3 bcd	10.36	89.4 cd	4.04	11.8 abcd	2.57
V05-5973W	46.6 ab	4.67	39.1 abc	1.29	14.3 abc	4.08	124.1 bcd	10.63	83.1 d	1.99	16.8 abcd	4.56
V12-1764	32.9 d	3.03	37.4 abc	1.83	29.6 ab	3.58	90.8 d	7.65	91.5 cd	4.63	30.5 ab	5.56
V09-3854	45.8 abcd	7.12	42.2 ab	5.28	11.9 bc	3.82	114.6 cd	18.74	88.6 cd	12.56	12.4 bcd	4.35
V09-3876	59.7 ab	2.35	33.1 bc	3.21	7.1 c	1.06	157.7 abc	11.52	68.8 d	5.46	7.2 d	0.30
MFS-561	47.2 abcd	1.36	40.7 ab	1.67	12.1b	1.08	103.8 cd	5.64	75.3 d	2.32	14.7 abcd	0.49
V09-3771	33.3 d	6.83	36.9 bc	3.63	29.7a	3.56	83.4 d	22.13	71.1 d	4.78	34.3 a	2.97

[1] The seed and sprout trait names are abbreviated as follows: HQS% is high quality sprout percentage, AQS% is average quality sprout percentage, LQS% is low quality sprout percentage, HQS is high quality sprout fresh weight, AQS is average quality sprout fresh weigh and LQS is low quality sprout fresh weight

[†] SE is standard error. *Different letters in the same column indicated the significant difference at the 0.05 probability level

Table 6: Correlation coefficient among sprout traits

Traits[1]	SS	HQS%	AQS%	LQS%	HQS	AQS	LQS	Hlght	Hthk	Wabs	Syld	Finc
SS	1.00											
HQS%	0.32*	1.00										
AQS%	-0.18	-0.69***	1.00									
LQS%	-0.27	-0.72***	-0.01	1.00								
HQS	0.48***	0.90***	-0.65***	-0.62***	1.00							
AQS	0.44**	-0.37*	0.61***	-0.07	-0.28	1.00						
LQS	-0.34*	-0.72***	0.07	0.93***	-0.62***	-0.12	1.00					
Hlght	0.10	0.33*	-0.01	-0.44**	0.24	-0.01	-0.46**	1.00				
Hthk	0.46**	0.25	-0.22	-0.13	0.47***	0.22	-0.12	-0.09	1.00			
Wabs	-0.15	-0.31*	0.18	0.26	-0.29	0.12	0.33**	-0.22	-0.17	1.00		
Syld	-0.13	0.59***	-0.22	-0.61***	0.61***	-0.08	-0.58***	0.19	0.30	-0.12	1.00	
Finc	0.04	-0.06	-0.04	0.12	-0.04	-0.05	0.15	-0.03	-0.001	0.11	-0.19	1.00

[1] The seed and sprout trait names are abbreviated as follows: SS is seed size, HQS% is high quality sprout percentage, AQS% is average quality sprout percentage, LQS% is low quality sprout percentage, HQS is high quality sprout fresh weight, AQS is average quality sprout fresh weigh, LQS is low quality sprout fresh weight, Hlgth is hypocotyl length, Hthk is hypocotyl thickness, Wabs is water absorption, Syld is sprout yield and Finc is fungi incidence
*Significant at the 0.01 significance level
**Significant at the 0.001 significance level
***Significant at the 0.0001 significance level

Table 7: Storage effect on sprout traits of eight soybean genotypes produced in Warsaw, VA

Seeds Storage time	SS[1] (g/ 100 seeds)	HQS[1] (%)	AQS[1] (%)	LQS[1] (%)	HQS[1] (g)	AQS[1] (g)	Hlgth[1] (cm)	Finc[1] (%)	Hthk[1] (mm)	Wabs[1] (%)	Syld[1] (g/ g seeds)
7 months	10.9	52.9 a	34.5	12.6 b	147.7 a	80.9	14.8 a	64.4 a	1.8	226.21 b	6.0 a
14 months	10.8	20.3 b	32.8	46.9 a	44.8 b	74.5	10.7 b	39.4 b	1.8	232.6 a	3.2 b

[1] The seed and sprout trait names are abbreviated as follows: SS is seed size, HQS% is high quality sprout percentage, AQS% is average quality sprout percentage, LQS% is low quality sprout percentage, HQS is high quality sprout fresh weight, AQS is average quality sprout fresh weigh, LQS is low quality sprout fresh weight, Hlgth is hypocotyl length, Hthk is hypocotyl thickness, Wabs is water absorption, Syld is sprout yield and Finc is fungi incidence
*Different letters in the same column indicated the significant difference at the 0.05 probability level
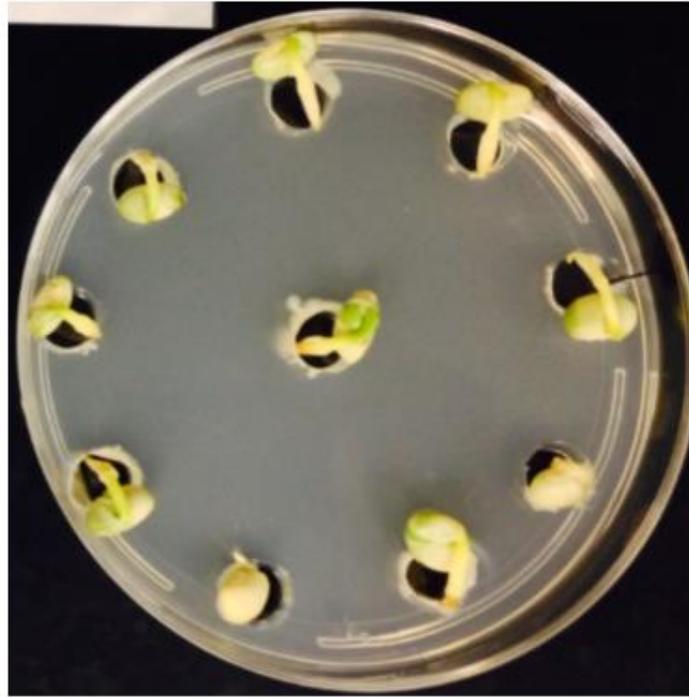
Figure 1. High-quality sprouts with vigorous, straight and long (> 7 cm) hypocotyls with bright yellow cotyledons
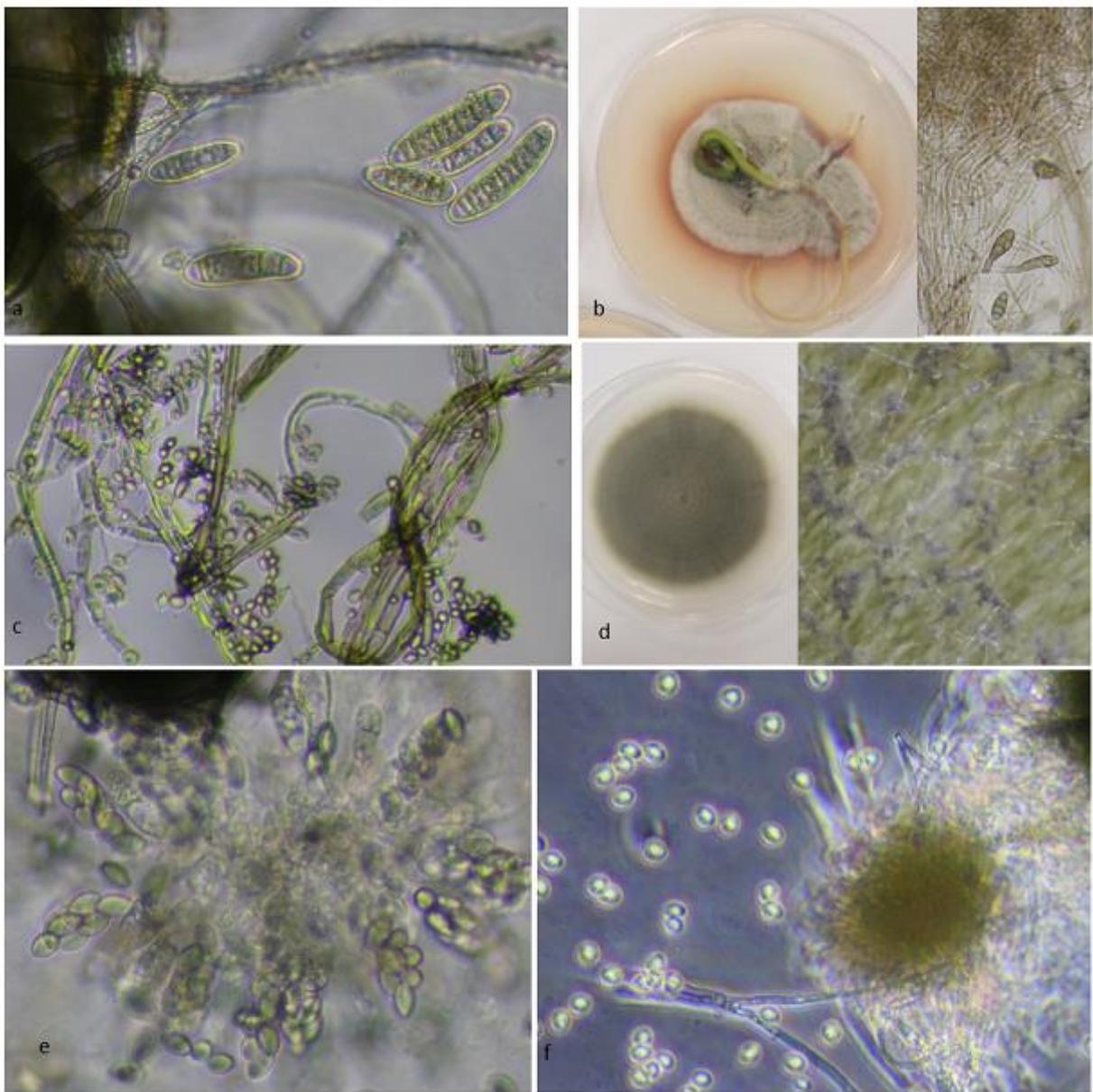
Figure 2. Average-quality sprouts with short hypocotyls (less than 7 cm) with cracked and open cotyledons (left) and curling seedlings (right)

Figure 3. Low-quality sprouts with abnormal seedlings and non-germinated seeds

Figure 4. Soybean seeds on potato dextrose agar plates (PDA) to establish the mold test on seeds

Figure 5. Four fungus genera detected on the soybean genotypes germinated on potato dextrose agar. **a** Asexual structures of *Bipolari*s sp. **b** Asexual structures and colony of *Cercospora* sp. **c** Asexual structures of *Botrytis* sp. **d** Colony of *Botrytis* sp. **e** Asexual structures of *Caethomium* sp. (brown color) and **f** Asexual structures of *Caethomium* sp. (yellow color), isolated from soybean seeds

Figure 6. High-quality sprouts of V09-3876 from Mount Holly, VA in 2014

3. Identification of fungi associated with soybeans and effective seed disinfection treatments

Diana M. Escamilla[1], Luciana Rosso[1] an Bo Zhang[1*]

[1]Department of Crop and Soil Environmental Sciences, Virginia Tech, Blacksburg, VA 24060

Abbreviations: HQS, high quality sprout; AVQ, average quality sprout; LQS, low quality sprout, EV, eastern Virginia; SV, southern Virginia, NC, northeastern North Carolina, PDA, potato dextrose agar; PD, potato dextrose broth, NCBI, National Center for Biotechnology Information; BLASTN, nucleotide basic local alignment search tool, PCR, polymerase chain reaction, ITS, internal transcribed spacer; DMSO, dimethyl sulfoxide.

Abstract

Soybean sprouts, a traditional soyfood vegetable in Asia, are gaining popularity in the U.S. Sprouts can be a vehicle for the transmission of several pathogens capable of causing human illness and the potential source of contamination is seed used for sprouting. The limited information about seed-borne pathogens as well as their incidence on soybean seeds for the sprout industry led the objective of this study that was to identify seed-borne pathogens on commercial sprout soybean seeds and to evaluate different decontamination treatments on disinfection effectiveness and sprout quality. Seeds of 'MFS-561', a sprout soybean cultivar, from three production regions in Virginia and North Carolina were used in this study. The internal transcribed spacer (ITS1 and ITS2) DNA sequences of the isolated fungi from MFS-561 seeds were used for species identification. Seven disinfection treatments were evaluated on their effectiveness on reducing fungal incidence and impact on sprout characteristics. Out of 55 fungal isolates obtained from the soybean seeds, seven species and six genera were identified. The most frequent genera across regions were *Alternaria*, *Diaphorte* and *Fusarium.* Most of the species isolated have been previously reported to be associated with mycotoxin production, especially *Fusarium* spp., *Alternaria alternata* and *Penicillium citrinum.* Due to their toxicological risk, it was necessary to discover prevention strategies for the seed-borne pathogens. The results showed that soaking seeds in 2% calcium hypochlorite for 10 minutes and 5% acetic acid for 2 minutes before sprouting were promising seed disinfection treatments as they significantly reduced fungi incidence without any negative effects on sprout quality.

Introduction

In recent years, the development of specialty soybeans has been promoted due to changes in consumer acceptance, growing global demand, and improved processing technologies (Mayta et al., 2014). Specialty soybeans are used to produce various soy-foods of Asian origin, assorted health food snacks, energy foods and cereals; and are generally grown under contract agreements. Sprout soybeans, a soyfood made from small seeded varieties (< 12 g/100 seeds), is an important vegetable consumed in many Asian countries such as Korea, China and Japan. In Korea, more than 500,000 tons of soybean sprouts are consumed annually as a vegetable in soups, salads and side dishes (Hwang et al., 2004; Lee et al., 2007; Zhang et al., 2010). In the united states, 10% of the population eats sprouts (alfalfa, mung bean, and soybean) regularly, resulting in a $250 million market (Sikin et al., 2013).

Soybean sprouts in Asia are generally prepared by soaking the seeds for about 4 to 5 h followed by placing them in a dark growth chamber (sprouter) at room temperature and watering them several times per day; under most conditions, sprouts can be ready for harvest 5-7 days after germination (Ghani et al., 2016; Liu, 1997; Silva et al., 2013). Darkness was required to avoid root elongation and green cotyledons (Liu 1997; Shi et al. 2010); and air and water temperature of 20 ℃, relative humidity of 80% and several applications of water during soybean sprouting are recommended to produce good quality (Lee et al. 2007; Ghani et al. 2016). Sprouting conditions such as temperature, high relative humidity, a pH close to neutral and nutrient availability are also favorable for the growth of microorganism, if present on the soybean seeds. Mesophilic bacteria and fungi grow well at 15-40 ℃ and a pH close to neutral (5-9) (Griffin, 1994; Pitt and Hocking, 2009; Stanga, 2010). Bacterial growth decreases with a

lower pH, while fungal growth increases (Rousk and Bååth, 2011). However, each specie has its own particular tolerance and some microorganism can survive even under extreme temperatures and pH (Pitt and Hocking, 2009; Stanga, 2010).

Fresh produce, such as sprouts, can be a vehicle for the transmission of several bacterial, protozoan and viral pathogens capable of causing human illness (FDA, 1999). This is the reason why sprouts represent a special food safety concern (NACMCF, 1999). From 1998 to 2010, 33 outbreaks from seed and bean sprouts were reported in the United States (Dechet et al., 2014). The most popular foodborne diseases are those caused by *Escherichia coli* O157:H7, *Salmonella, Campylobacter,* and *Cryptosporidium* (Johnston et al., 2005). Kim et al., (2002) tested raw soybean sprouts for contamination agents and detected the presence of *Bacillus cereus, Aeromonas hydrophila,* and *Pseudomonas aeruginosa*, but *Salmonella* spp., *E. coli* O157:H7, and *Yersinia enterocolitica* that are associated with serious diseases in humans were not detected. Fungi also have the potential, less than bacteria, to cause food spoilage and pathogenesis. However, the information of mycological identification of fungi species isolated from soybean sprouts was very limited. Eighteen and seven species of fungi were isolated from spoiled *Vigna* spp. and soybean sprouts in Japan, respectively; and approximately 70% of the isolates were plant pathogens and at least 14 species were known to be seed-borne (Sato et al., 2014). The fungus species isolated from soybean sprouts were *Cercospora kikuchii* (Tak. Matsumoto & Tomoy.) M.W. Gardner, *Diaporthe phaseolorum* var. *caulivora* Athow and Caldwell, *Fusarium graminearum* Schwabe, *Fusarium oxysporum* Schltdl., *Penicillium oxalicum* Currie and Thom, *Phoma medicaginis* Malbr. and Roum, *Phomopsis phaseoli* var. *sojae* (Lehman) Sacc., and *Syncephalastrum racemosum* Cohn ex J. Schröt. Some strains of *F. graminearum* have already been reported to produce high concentrations of deoxynivalenol (DON) (Sato et al., 2014). DON

is a mycotoxins produced by certain *Fusarium* species, affecting animal and human health by causing acute temporary nausea, vomiting, diarrhea, abdominal pain, headache, dizziness, and fever (Sobrova et al., 2010). Mycotoxins are compounds produced by fungi that can contaminate food and feed commodities during both pre- and post-harvest and when ingested, it causes mycotoxicoses, which can have acute or toxic effects on human and animals (Milićević et al., 2010). Mycotoxins effects are diverse and include carcinogenic, nephrotoxic, hepatotoxic, dermotoxic and neurotoxic thus affecting human and animal health (Bovdisova et al., 2016; de Souza et al., 2013; Desjardins et al., 2000; Dong et al., 1987; Milićević et al., 2010; Ostry, 2008; Stankovic et al., 2007; Zain, 2011).

In soybean sprout outbreaks, the seeds are the potential source of contamination, although poor sanitation and unhygienic practices at the sprout operation could also produce contamination (NACMCF, 1999). As a result, seed disinfection treatments were considered the major intervention in a multi-step approach to reduce the risk of illness associated with contaminated sprouts. Calcium hypochlorite (20,000 ppm) has been considered the standard reference of seed disinfection treatment for over a decade (Ding et al., 2013; FDA, 1999). In addition, sodium hypochlorite solution has also been widely used. Solution of 1-5% sodium hypochlorite for 1 minute have shown to be good seed surface disinfectants (Sauer and Burroughs, 1986). However, one of the main limitations of the current chlorine-based washes was the possible hazards associated with production, transporting, and handling large amounts of chlorines; thus making organic acids as a potential chemical alternative to disinfect seeds and sprouts (Sikin et al., 2013). Treatments of acetic acid, lactic acid and a combination of lactic/acetic acid followed by calcium hypochlorite treatment have shown to be effective on reducing *E. coli* *O*157:H7 populations. Reductions of 2-4 $\log_{10}$ CFU/g of *E. coli* were achieved

by 5% acetic and lactic acid treatments (Lang et al., 2000). Acetic acid, as a disinfectant treatment, has been studied on radish seeds, alfalfa seeds and sprouts, and mung bean and cowpea sprouts; whereas lactic acid has been studied on alfalfa seeds (Lang et al., 2000; Nei et al., 2011; Oms-Oliu et al., 2010; Singh et al., 2005). Chemical treatment effectiveness may be limited by inaccessibility to pathogens sheltered in scarified surfaces and the interior of the seeds (Ding et al., 2013).

Physical treatments are more environmentally friendly and may have better penetration characteristics than chemical treatments (Ding et al., 2013). Heat, radiation, electric and magnetic fields, high pressure, and ionization produce changes in the physical conditions, which can interfere with microorganism life and cause death. Some physical disinfection treatments can be applied to foods without causing deterioration, unlike chemical treatments (Stanga, 2010). Hot water treatments (85 °C for 9 s) were equally or more effective than 20,000 ppm calcium hypochlorite treatments on disinfecting alfalfa seeds inoculated with *E. coli* (Enomoto et al., 2002). Hot water and dry heat treatments have also been studied on radish, alfalfa and mung bean seeds, and mung bean sprouts to eliminate *E. coli* and/or *Salmonella* (Bari et al., 2009a; Bari et al., 2003; Bari et al., 2009b; Weiss and Hammes, 2005a). In addition, seed biological interventions also have shown to be similar to 20,000 ppm calcium hypochlorite to eliminate pathogen populations (Ding et al., 2013). For instance, a combination of *Enterobacter asburiae* (strain "JX1") and lytic bacteriophages, isolated from pig or cattle manure, showed being effective on controlling *Salmonella* on sprouting mung beans and alfalfa seeds (Ye et al., 2017). Combination of chemical and physical treatments has also shown to be an effective strategy. However, identifying the optimal conditions might be challenging due to the complexity introduced by the application of several treatments, but once stablished they might provide the

69

best control (Ding et al., 2013). Sikin et al. (2013) reported that the thermal inactivation of seeds and irradiation of sprouts were the most practical and safe interventions for sprout production from all current decontamination technologies for seeds and sprouts. However, the reported studies focused on eliminating bacteria such as *E. coli* and *Salmonella* due to the high risk they represented for human health.   Therefore, additional studies of alternative seed disinfection treatments against other microorganisms, such as fungi, were needed to improve management strategies and sprout safety.

A successful seed treatment should reduce microbial pathogens while preserving seed viability, germination (95%), seed vigor and the sensorial attributes of the final product (FDA, 2017; Sikin et al., 2013). Additionally, an antimicrobial treatment that is effective for one type of seed may not be as appropriate for other types because surface features of the seeds can influence the access of a treatment to inactivate pathogens on or in the seed (FDA, 2017). Despite the fact that soybean sprouts are in high demand in the edible market due to their numerous health benefits (Ghani et al., 2016), limited information was available on fungus species present on soybean seeds used for sprouting and effective seed disinfection treatments. Therefore, soybean seeds from a commercial sprout soybean cultivar, MFS-561, were evaluated for the presence of fungi; which allowed us to identify potential risks for food spoilage and pathogenesis in soybean sprout production in VA, US. In addition, effectiveness of several seed disinfection treatments and their effect on important sprout quality traits were also evaluated to establish effective and practical sanitation procedures of soybean seeds.

Materials and Methods

Materials

Montague Farms, Inc. provided composite seed samples of MFS-561, a sprout commercial variety from three different growing regions: Southern Virginia (SV), Eastern Virginia (EV) and Northeastern North Carolina (NC) in US. Seeds were produced by a network of growers who supply seeds to Montague Farm, Inc.

Fungal isolation from soybean sprout seeds

MFS-561 seeds from SV, EV and NC harvested in 2015, were used for fungal isolation and identification. In order to determine the fungal incidence, seed surfaces were rinsed with sterile water. Afterwards, 10 seeds were placed in holes punched on 4% potato a dextrose agar (PDA) plate that is a relatively rich medium for growing a wide range of fungi. Three plates from each region were incubated at room temperature (20 °C) for five days. Then, each fungal colony that grew on the seeds was isolated by cutting a small piece from the edge of the mycelial of each individual colony. The mycelial pieces were transferred to fresh PDA plates and incubated at 20°C for 7 days. The diameter of each fungal colony was recorded. After the incubation period, a piece of the mycelial from each fungal colony was transferred to 250 mL Erlenmeyer flask containing potato dextrose broth (PD; Difco). In order to produce enough fresh fungal mycelium for molecular identification, liquid cultures were grown for 10-14 days on a Lab-line incubator shaker Orbit (400 rpm) at room temperature before DNA extraction. Fungi colonies on PDA were characterized morphologically and a total of 55 isolations from Southern Virginia (23), North Eastern North Carolina (13) and Eastern Virginia (19) were sent for sequencing.

Molecular identification of fungal species by DNA extraction, PCR and sequencing

71

For extraction of genomic DNA of fungal isolations, mycelium from axenic cultures grown in PD broth was ground to a fine powder in liquid N_2 using a mortar and a pestle. Genomic DNA was extracted following the CTAB method as described by Gontia-Mishra et al. (2014). DNA was eluted in 100 µl of distilled H_2O and diluted to a concentration of 30 ng/µl prior to PCR. The complete ITS 1 and 2 regions, along with the short structural gene (5.8S), were amplified with ITS5 (5'- GGAAGTAAAAGTCGTAACAAGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') by PCR (Fig. 7) (White et al., 1990). The ITS regions were chosen to identify fungal species, as they have been widely used as a DNA barcoding regions for molecular identification of fungi (Schoch et al., 2012). Amplifications were carried out in 25µl reaction mixtures containing 15.39µl of sterile water, 2.4µl of 10x PCR buffer, 0.8µl of 50 mM $MgCl_2$, 0.25µl of each primer (100µM), 0.5µl of 10µM dNTPs, 1.26 µl of 10% DMSO, 0.4 µl of Taq DNA polymerase (Apex Taq, 5U/µl) and 4 µl of DNA template (30 ng/µl). PCRs were repeated twice to obtain 50 µl of PCR product for each fungus isolated. PCRs were performed as follows: an initial denaturation for 5 min at 95° C, 40 thermal cycles (1 min at 94° C, 1 min at 55° C and 1 min at 72° C), and a final 10 min extension at 72° C. PCR products were run on 1% agarose gel to verify amplification. The amplified DNA (50 µl) was purified using QIAquick PCR purification kit following manufacturer's instructions (QIAGEN, Germany); and sequenced (eurofins genomics). The sequence data were searched with "Standard Nucleotide BLAST" in the NCBI website (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) to identify the species; results were compared and confirmed based on the fungal colonies morphology on PDA.

Disinfection seed treatments

MFS-561 seeds from SV, EV and NC harvested in 2015 were used for testing seed disinfection treatments. Seed treatments included physical intervention methods of hot water at 60 °C for 2 minutes and dry heat at 50°C for 1 hour; and chemical intervention methods of calcium hypochlorite (2% w/v Ca [ClO]$_2$) for 10 minutes, sodium hypochlorite (2% v/v NaOCl) for 10 minutes, acetic acid (5% v/v) for 2 minutes and lactic acid (5% v/v) for 10 minutes. Commercial bleach with 8.25% NaOCl (Clorox), granulated calcium hypochlorite with 65% content of chlorine (Pfaltz & Bauer), 99% acetic acid (Fisher Chemical) and 10% lactic acid (RICCA Chemical Company) were diluted in sterile water for preparing the treatment solutions of calcium hypochlorite, sodium hypochlorite, acetic acid and lactic acid, respectively. The pH of the treatment solutions was 11.97 for calcium hypochlorite, 8.7 for sodium hypochlorite, and 2.9 for acetic and lactic acid. Ten seeds from each growing region were placed in a 50-ml beaker and covered with 40 mL of the respective disinfection solutions during the times specified above. For the physical interventions, ten seeds were placed in a strainer and submerged in the water bath for 2 minutes for the hot water treatment; and ten seeds were placed in a container and put in the oven for 1 hour for the dry heat treatment; temperatures were set at 60 and 50°C, respectively. After chemical and physical treatments, seeds were rinsed with sterile water for 1 minute and placed in holes punched on 4% PDA (Potato dextrose agar) plates; each plate had 10 seeds. Plates were incubated at room temperature (20 °C) for five days and the number of infected seeds out of the ten initial seeds was recorded as the fungus incidence. Sterile water treatment was used as a control and each treatment was repeated five times.

Effect of seed treatments on sprout traits

A total of four composite seed samples of MFS-561 were used to study the effect of seed disinfection treatments on sprout traits. Seed source including growing region and year were provided in Table 8. A total of 350 unbroken and undamaged seeds of MFS-561 from each composite sample were selected and grown in a bean sprouter (Cheong Si Ru, SC-9000TS, made in Korea) at room temperature for five days. Before sprouting, seeds were treated with 2% calcium hypochlorite for 10 minutes, 5% acetic acid for 2 minutes, hot water (60°C) for 2 minutes and sterile water for 1 minute (control) as described previously; each treatment was repeated three times for one seed sample. Sprouting conditions and sprout trait evaluation were performed as described by Escamilla et al. (2017). Sprout traits evaluated include percentage of high-, average- and low- quality sprouts, hypocotyl length and thickness, and sprout yield.

Statistical analysis

Statistical analysis was computed in JMP statistical version 11.0 (SAS Institute, Raleigh, NC). Prior to conducting statistical analyses, assumptions for two-way analysis of variance model were checked for all response variables. The normality assumption was assessed by Shapiro-Wilk test and normal probability plots; while the assumption of homogeneity of variance was evaluated by residuals vs. predicted values plots and plots of residuals by groups (McKillup 2012). Variables that violated the assumptions of normal distribution and equal variance were properly transformed and analyzed. All variables met assumptions after transformation. Seed disinfection treatment performances were evaluated under a two factor completely randomized design with seed treatments (6 treatments + control) and locations (SV, EV and NC) as factors. A two factor completely randomized design with seed treatments (3 treatments + control) and composite seed sample (CS1, CS2, CS3 and CS4) as factors was used to assess the effect of seed

disinfection treatments on sprout traits. The seed treatment effects on fungi incidence and sprout

quality traits were analyzed using a two-way analysis of variance (ANOVA). Tukey's HSD test

was used to show the treatments that differed significantly at $P = 0.05$.

Results and Discussion

Fungi species isolated from MFS-561, a commercial soybean sprout cultivar in VA, US.

A total of 55 isolates identified as 7 species and 6 genera were obtained from MFS-561 seeds across the three regions. Fungal species, source locations and most representative accessions with high identity hit in each BLASTN search were listed in Table 9. The most common genera across three growing regions were *Alternaria* (29.1%)*, Diaphorte* (29.1%) and *Fusarium* (23.6%). While *Phoma* (1.8%), *Penicillium* (3.6%)*, and *Cladosporium* (7.3%) were identified only on one or two locations. *Fusarium equiseti* (30.77%) was the most common species in NC followed by *Diaphorte longicolla* (23.1%) and *Penicillium citrinum* (14.4%). In SV, the most common species were *Alternaria alternata* and *Diaphorte* sp. with frequency of 34.8%. *A. alternata* (15.8%) and *D. longicolla* (15.8%) were also the most common species in EV together with *Fusarium* sp. (15.8%) (Table 10).

Alternaria sp. and *A. alternata*

A. alternata was isolated from NC, SV and EV with a frequency of 7.7%, 34.8% and 15.8%, respectively. Several isolations were identified as *Alternaria* sp. on SV (8.7%) and EV (10.53%) seeds when searching on NCBI databases (Table 10). When growing in PDA plates, *A. alternata* species were characterized by colonies that could either have 50-70 mm diameter or cover the whole petri dish with a grey brown mycelium and a brown to nearly black color on the reverse side (Pitt and Hocking, 2009). In this study, colonies of *A. alternata* and *Alternaria* sp. had a grey brown mycelium covering the whole plate and a black color on the reverse (Fig. 8 A-B). There were not noticeable differences on colony morphology among *A. alternata* and *Alternaria* spp. isolations.

Alternaria was one of the most common genera across three regions, and *A. alternata* was the most frequent species in SV and EV. Most *Alternaria* species are saprophytes, others well known post-harvest pathogens, are opportunistic plant pathogens causing several diseases on important agronomic crops. Some *Alternaria* spp. are of clinical importance because they produce toxic secondary metabolites and some of these mycotoxins were associated with the development of cancer in mammals; where *A. Alternata* particularly is gaining distinction as an emerging human pathogen (Thomma, 2003). Some of the mycotoxins produced by *A. alternata* are alternariol, altenusin, hydroxyalternariol monomethyl ether, alternariol monomethyl ether, tenuazonic acid and alternuene (Bottalico and Logrieco, 1988; de Souza et al., 2013). Previously, *A. alternata* was detected at high frequencies in soybean seeds in Argentina, where some strains were associated with tenuazonic acid production (Broggi et al., 2007; Oviedo et al., 2009). Tenuazonic acid (TA) is toxic to several animal species (mice, chicken, dogs) while alternariol, alternariol monomethyl ether and alternuene are not highly toxic; and many of these mycotoxins are frequently detected in food commodities (Bottalico and Logrieco, 1988; Ostry, 2008). Based on the fungal contamination results for soybean seeds, occurrence of toxic metabolites such as TA could be a potential risk for sprout safety. However, further studies are necessary to identify the toxigenic ability of *A. alternata* strains isolated from this study and the occurrence of mycotoxins on sprouts, if present.

Diaphorte sp. and *D. longicolla*

Diaporthe sp. represented a 7.7%, 34.8% and 5.3% from the total isolations of NC, SV and EV, respectively; and D. *longicolla* represented a 23.1% and 15.8% from the total isolations of NC and EV, respectively (Table 10). *D. longicolla* (anamorph = *P.* longicolla) colonies on PDA were described for the first time by Hobbs et al. (1985) as floccose, dense and white with

occasional green-ish-yellow areas and colorless in the reverse with large black stomata; similar characteristics were observed on *Diaphorte* sp. and *D. longicolla* isolations from this study (Fig. 8 K-L). There were not significant morphological differences among colonies of *D. longicolla* and *Diaphorte* sp.

Diaphorte also was one of the most common genus across the regions with *D. longicolla* (NC and EV) and *Diaporthe* sp. (SV) as the most frequent species. Soybeans harbor a complex of *Diaphorte* (anamorph = *Phomopsis*) species that are associated with several soybean diseases including seed decay, pod and stem blight and stem canker; thus being responsible of considerable crop production losses (Lehman, 1923; Santos et al., 2011; Sun et al., 2013). The most recognized species of the *Diaphorte/Phomopsis* complex are *D. phaseolorum* var. *caulivora, D. Phaseolorum* var. *sojae,* and *Phomopsis longicolla* (Xue et al., 2007). *Diaporthe longicolla* (anamorph = *Phomopsis longicolla*) is the primarily cause of seed decay of soybeans and it is the major cause of poor-quality seed (Darwish et al., 2016; Li et al., 2010; Mengistu et al., 2014; Santos et al., 2011). Mycotoxin production has just been reported for *Phomopsis leptostromiformis* that infects lupins and produces phomopsins. This mycotoxin causes liver disease of sheep and cattle that have been fed with infected lupin (Battilani et al., 2011; Ding et al., 2016). However, no mycotoxins were reported to be associated with *D. longicolla*. As a result, the presence of *D. longicolla* and *Diaphorte* sp. in soybean seeds was expected because of its high pathogenicity to soybeans and could represent a potential risk for seed and sprout quality.

Fusarium sp.; F. proliferatum; F. equiseti and *F. chlamydosporum*

Fusarium species were found in significant frequencies on soybean seeds of NC (38.46%) and EV (26.32%). A percentage of 7.7 from NC and 15.8 from EV of the total

isolations were identified at genus level as *Fusarium* sp. when searching on NCBI databases. Three *Fusarium* species were isolated from soybean seeds. One was *F. chlamydosporum* detected at low frequency in EV (10.53%) (Table 10). Its colonies presented a powdery appearance (felty mycelium) from profuse microconidial production on PDA, with pale salmon color and paler margins as described by Pitt and Hocking (2009) (Fig. 8 G-H). Other was *F. proliferatum* found at low frequency (4.35%) in SV (Table 10). Its colonies were floccose with a pale orange to white mycelium and a light orange in the reverse in accordance to Pitt and Hocking (2009) (Fig. 8 E-F). Lastly, *F. equisete* isolated in high frequency (30.77%) on soybean seeds from NC and, in low frequency (10.53%) on seeds from EV (Table 10). On PDA *F. equiseti* colonies had floccose mycelium with white to pale salmon color that became brown with age, a central mass of orange sporodochia sometimes surrounded by poorly defined sporodochial rings and a pale salmon color on the reverse often flecked with brown (Pitt and Hocking, 2009). Colonies of *F. equiseti* isolated in this study showed several of these characteristics (Fig. 8 C-D).

Fusarium was the second most common genus across regions after *Alternaria* and D*iaphorte. Fusarium* species are associated with several plant diseases such as vascular wilts, root and stem rots, pre- and post- emergence blight among others (Pitt and Hocking, 2009). These pathogens are widely distributed in cultivated soils and participate in the decomposition of cellulosic plant materials. In addition, *Fusarium* species are also a major source of storage rots of fruits and vegetables; and are commonly associated with cereals (Pitt and Hocking, 2009). *Fusarium* species are of special concern of public health because *Fusarium* is one of the three major fungal genera producing toxins in infected plants and/or in plant products, where the most widespread *Fusarium* mycotoxins are the trichothecenes. A high number of compounds (> 50) are known to be produced by *Fusarium* species and some are highly toxic (Ding et al., 2013; Pitt

and Hocking, 2009). Three *Fusarium* species were isolated from sprout soybean seeds in this study. *F. equiseti* was the most common specie on NC seeds and it was also isolated from EV seeds but at low frequency. This specie is frequently found on soybean seeds and it is associated with the root rot complex, a major soybean disease (Ivic, 2014; Zhang et al., 2013). Ivic et al. (2009) also reported F. *equiseti* as one of the most common species isolated from soybean seeds. Besides, it has also been reported in other crops as beans and kidney beans (Pitt et al., 1994). *F. equiseti* is well known for producing mycotoxins such as type A trichothecenes (diacetoxyscirpenol and related compounds, T-2 toxin, HT-2 toxin and neosolaniol), type B trichothecenes (deoxynivalenol, 15-acetyldeoxynivalenol, fusarenone and nivalenol), zearalenone, moniliformin, beauvericin, fusarochromanone and related compounds (Desjardins, 2006; Hestbjerg et al., 2002; Kosiak et al., 2005; Langseth, 1998; Logrieco et al., 1998). *F. chlamydosporum* was also isolated from sprout soybean seeds of EV but in very low frequency. *F. chlamydosporum* is a seed-borne fungus in several crops including soybeans, pea, mung bean and sorghum, which plays important roles as plant pathogen causing root rot and wilt diseases, and bio controller of plant diseases such as groundnut rust (Desjardins et al., 2000; Ivic, 2014; Ivic et al., 2009; Mathivanan et al., 1998; Pitt et al., 1994; Singh et al., 2013). *F. chlamydosporum* has been associated with production of mycotoxins type A trichothecenes including T-2 toxin, HT-2 toxin, monoacetoxyscirphenol, neosolaniol and iso-neosolaniol (Park and Chu, 1993). However, there is no recent evidence of trichothecene production and moniliformin seems to be the major mycotoxin produced by *F. chlamydosporum* (Desjardins, 2006). The toxicology of moniliformin is not well understood and it's typically associated with myocardiac damage (Peltonen et al., 2010). Lastly, *F. proliferatum* was only isolated from SV seeds at low frequency. It is also an important plant pathogen affecting several plants (Bernardi-

Wenzel et al., 2016) and has been found in rice, soybeans, peas, wheat, maize and sorghum (Desjardins et al., 2000; Ivic, 2014; Ivic et al., 2009; Pitt et al., 1994). In soybeans, it was reported as a causal agent of root rot (Chang et al., 2015). *F. proliferatum* is a major producer of fumonisins B1, B2 and B3, moniliformin, beauvericin, enniatins and fusarin (Desjardins, 2006; Desjardins et al., 2000; Ivic, 2014; Ivic et al., 2009; Marasas et al., 1986; Ross et al., 1990). *Fusarium* species were common on soybean seeds and they have been associated with mycotoxins with a wide range of toxicity for plants, animals and humans in the past. Thus, seed-borne *Fusarium* species, especially *F. equiseti,* could represent a potential risk for sprout production and safety.

Cladosporium cladosporioides

C. cladosporioides was detected on soybean seeds from SV (8.7%) and EV (10.5%) at low frequencies (Table 10). Its colonies were low and dense, lightly wrinkled surface, lightly floccose and olive color with a dark grey on the reverse as described by Pitt and Hocking (2009) (Fig. 8 M-N). *C. cladosporioides* is an endophyte that has shown potential as biological control agent of weeds and chrysanthemum white rust (Torres et al., 2017; Waqas et al., 2013). It also was the most commonly isolated species (49%) on soybean seeds from Thailand and it is considered a significant storage fungi of soybean seeds (Pitt et al., 1994). In addition, the genus *Cladosporium* has also been isolated at high frequency (54.6%) from soybean seed samples used for meju preparation, an important food ingredient in Korean cuisine (Kim et al., 2013). However, *C. cladosporioides* is not known to produce mycotoxins (Pitt and Hocking, 2009). Despite the low frequency on soybean seeds and the low risk of mycotoxin production, *C. cladosporioides* could reduce sprout quality, if growth. Therefore, good seed storage conditions

and adequate seed disinfection treatments are required to eliminate *C. cladosporioides* from soybean seeds.

Penicillium citrinum

P. citrinum was only isolated from NC soybean seeds representing 15.4% of total isolations from North Carolina (Table 10). Colonies of *P. citrinum* are floccose with a yellow brown or olive color, white mycelium in peripheral areas, radially sulcate and dull brown color on the reverse (Pitt and Hocking, 2009). In this study, *P. citrinum* colonies exhibited similar characteristics but with a dull yellowish brown color in the reverse (Fig. 8 I-J). *P. citrinum* has been widely isolated from food and it is considered a storage fungi, which has been isolated from mung beans, soybeans, rice, sorghum and other beans (red beans) (Pitt and Hocking, 2009; Pitt et al., 1994). Food spoilage caused by *P. citrinum* is rare, however it can't be considered a mere contaminant because of its mesophilic nature, worldwide distribution and its ability to grow down to 0.80 water activity (a_w), which helps to secure its niche in a very wide range of habitats (Domsch et al., 1980; Pitt, 1979; Pitt and Hocking, 2009). *P. citrinum* is known for producing the mycotoxin citrinin and cellulose digesting enzymes like cellulose and endoglucanase (Khan et al., 2008). Citrinin is a common contaminant of various food and feed materials that has shown clear evidence for reproductive toxicity, teratogenic, nephrotoxic, hepatotoxic and embryotoxic; however the mechanism of its action are not known properly (Bovdisova et al., 2016). Therefore, *P. citrinum* might be a risk for sprout spoilage because of its high ability to grow in wide range of conditions and mycotoxin production. However, its presence was very low.

Phoma sp.

A low percentage (4.3%) of isolations from SV was identified as *Phoma* sp. (Table 10). Its colonies have a color that varies from dark to black, olive-gray or grayish-brown and a diffusible pigment either reddish or brown is sometimes released into the media. The colonies don't wrinkle but they may form concentric circles with darker shades of color in the center (Sciortino, 2017). In this study, colonies of *Phoma* sp. had an olive-gray color with a reddish pigment clearly visible in the front and reverse of the PDA plate and concentric circles (Fig. 8 O-P).

Phoma genus was rare on soybean seeds; it was only detected on SV representing a low percentage of isolations. This genus is part of a group of rarely observed pathogens that have been isolated from spots on leaves and pods of soybeans, and it is a cosmopolitan, ubiquitous specie on diseased and dead plant materials that is defined frequently as opportunistic parasite (Kövics et al., 2014). *Phoma* sp. has been isolated from seeds of mung beans, soybeans, rice, sorghum, red beans and other beans (Pitt et al., 1994), and it has also been found on spoiled sprouts of mung bean and soybeans (Sato et al., 2014). Tenuazonic acid, viridactol, alternariol and alternariol monomethyl ether production by *Phoma* sp. has been previously reported (Mousa et al., 2015; Steyn and Rabie, 1976). However, there is limited information about mycotoxin production by *Phoma* sp. and the significance of these toxins. Presence of *Phoma* sp. on seeds was rare representing a lower risk for sprout production compared with other fungal species isolated in this study.

From all the fungal species isolated, *Penicillium citrinum* and *Cladosporium cladosporioides* are widely recognized as storage fungi, while the other species are found either as storage or field fungi. In this study, most fungal species isolated from soybean seeds have

been previously associated with mycotoxin production, especially *Fusarium* species, *A. alternata* and *P. citrinum.* However, presence of mycotoxins in food, seeds or animal feed depends on many factors such as environmental conditions (storage conditions), fungal strain specificity, strain variation and instability of toxigenic properties. Environmental conditions that favor fungus growth not necessarily favor mycotoxin production (Langseth, 1998; Oviedo et al., 2009; Zain, 2011). Identification of the specific mycotoxin is necessary to make a good diagnosis, which is generally done by liquid chromatography of extracts from the fugal source and/or food (Bennett and Klich, 2003; Plumlee and Galey, 1994). For that reason, the identification of *A. alternata, Fusarium* species and *P. citrinum* on sprout soybean seeds doesn't necessarily mean mycotoxin production and/or sprout contamination even when the species has already been associated with mycotoxins.

Nowadays, modern agricultural practices and the legislatively regulated food processing and marketing system have helped to reduce the population exposure to mycotoxins. A lot effort is being directed toward the evaluation of current contamination status of commodity foods and the impact of mycotoxins on humans and animal health (Milićević et al., 2010). Thus, further studies are needed to evaluate the toxigenic ability of fungal species associated with soybean seeds in Virginia, U.S., optimal conditions for growth and mycotoxin production, and the occurrence of their toxins in soybean sprouts, if produced. However, the frequent appearance of fungal species that could produce mycotoxins represents a potential risk of multiple mycotoxin contamination on soybean seeds and possibly on soybean sprouts. Besides, fungi presence on soybean seeds could also reduce sprout quality and produce economic losses for soybean sprout growers and industry. As a result, soybean seed disinfection treatments play an important role in

reducing the risk of mycotoxin contamination among all other approaches at different critical control points that are being developed to avoid mycotoxin contamination in food chains.

Disinfection seed treatments

Fungi incidence among treatments were significantly different, but not among growing regions. Non-treated seeds had a fungi incidence of 59%. Hot water at 60 ℃ for 2 minutes was the most effective treatment for reducing fungal soybean seed contamination with a fungi incidence of 5.2% (Fig. 9). Previous reports showed that a wide range of hot water treatments at different temperatures and exposure times have successfully disinfected seeds of radish, alfalfa and mung beans contaminated with *E. coli* and *Salmonella* (Enomoto et al., 2002; Jaquette et al., 1996; Pao et al., 2008; Weiss and Hammes, 2005b). Some of these treatments were hot water at 57 to 60℃ for 5 minutes, 70 ℃ for 10 seconds, 80 ℃ for 5 seconds, and 90 ℃ and 100 ℃ for 3 seconds that eliminated *Salmonella* from previously inoculated seeds (Jaquette et al., 1996; Pao et al., 2008). Hot water treatment at 85 ℃ for 9 seconds was effective on disinfecting alfalfa seeds inoculated with *E. coli* (Enomoto et al., 2002). Reductions of *Salmonella* and *E. coli* populations by more than 5log CFU/g have been achieved by hot water treatments (Weiss and Hammes, 2005b). Hot water treatment has been widely used against bacteria such as *Salmonella* and *E. coli* because they are the most common causes of foodborne illness. Our results showed that hot water treatment was also effective on reducing field and storage fungi incidence on soybean seeds. Further studies at several temperatures and times might be needed to identify the conditions (temperature and time) that allow total elimination of fungi from soybean seeds.

Calcium hypochlorite (20,000 ppm for 10 minutes) was the second most effective seed treatment after hot water with fungi incidence of 12.8% (Fig. 9). Its effectiveness in reducing microbial pathogens has been evaluated in the past (Damron et al., 2005; Ding et al., 2013). The

85

U.S. Food and Drug Administration (FDA) has recommended widely pre-germination treatments with 20,000 ppm Ca(OCl)$_2$ in sprout production (Christopher et al., 2003; Jaquette et al., 1996). The chlorine molecule was the molecule that actually disinfected seeds by disrupting the cell membrane of microorganisms (Damron et al., 2005). For that reason, several studies reported the free chlorine content of treatment solutions instead of calcium or sodium hypochlorite concentrations when referring to seed disinfection treatments. Chlorine concentrations of 20,000 and 2,000 ppm have shown to be effective controlling *E. coli* and *Salmonella*, respectively, in alfalfa seeds (Jaquette et al., 1996; Lang et al., 2000). Bari et al. (2003) reported that combination treatments of 200 ppm of active chlorine with sonication, dry heat and other chemical sanitizers have also showed to be effective in reducing *E. coli* from alfalfa, radish and mung bean seeds. The effect of sodium and calcium hypochlorite has been studied mainly on bacteria as *E. coli* and *Salmonella* due to their importance as human pathogens. In this study, 2% calcium hypochlorite was more effective on reducing fungi incidence than 2% sodium hypochlorite treatment with 39.3% fungi incidence (Fig. 9). Leonardo et al. (2016) found that solutions of 2.5% calcium and sodium hypochlorite (~ pH 12) provided approximately 2.2% and 1.5% of chlorine content, respectively. Hence, when diluted at the same concentration, chlorine was expected to have a higher content in calcium hypochlorite solutions than in sodium hypochlorite solutions, making calcium hypochlorite a more effective seed treatment. In addition, chlorine was unstable at acidic pH with a limit temperature of 50 °C to prevent risk of corrosion (Stanga, 2010). Both calcium and sodium hypochlorite solutions at pH~12 have shown to be stable for 30 days of storage when kept at 4 C or at 25 °C (Leonardo et al., 2016). While, 1-5% sodium hypochlorite solutions at pH 8 have shown to be more effective on reducing pathogen populations but unstable (Sauer and Burroughs, 1986). However, freshly prepared

sodium hypochlorite solutions tended to have lower pH when compared with calcium hypochlorite solutions (Leonardo et al., 2016). In this study, the pH of calcium hypochlorite solution (11.97) was higher than the pH of sodium hypochlorite solution (8.7). Therefore, higher chlorine content and a more stable solution might be the reason why calcium hypochlorite treatment had high effectiveness as seed disinfectant. In general, low concentrations (1-5%) of sodium hypochlorite are used for eliminating seed surface fungi but not internal fungi (Sauer and Burroughs, 1986). Therefore, higher concentrations of sodium hypochlorite may eliminate pathogen populations in the seeds.

The third most effective treatment was 5% acetic acid followed by 5% lactic acid with seed fungi incidence of 16.8 and 32%, respectively (Fig. 9). Similar treatment conditions were evaluated by Lang et al. (2000), including combination of 5% acetic or lactic acid for 10 minutes followed by 2,000 ppm of active chlorine achieved reductions of 2-4 $\log_{10}$ CFU/g of *E. coli* in alfalfa seeds. However, none of the treatments prevented regrowth of surviving *E. coli* (Lang et al., 2000). Lactic and acetic acid treatments at several concentrations and times have successfully controlled bacteria populations on other seed types with longer exposure time than those evaluated in this study. Treatment of 8.7% acetic acid at 55 °C for 2-3 hour reduced populations of *E. coli* and *Salmonella* on alfalfa and radish seeds by more than 5.0 log CFU/g and longer treatment completely eliminated *E.coli* populations (Nei et al., 2011). Solutions of 5% acetic acid eliminated *Salmonella* from alfalfa and mung bean sprouts after 4 and 16 h, respectively. While solutions of 2% acetic acid eliminated *Salmonella* after 24 and 48 hours for alfalfa and mung bean, respectively (Pao et al., 2008). Thus, higher concentration of acetic acid required shorter time to eliminate pathogen populations. In this study, seeds were treated with acetic acid for 2 minutes because longer times showed to affect adversely seed viability when observed on PDA

87

plates. As a result, other combinations of exposure times and acetic/lactic acid concentrations might increase acetic and lactic acid effectiveness on reducing fungal populations on soybean seeds. The antibacterial activity of acetic acid has been proven in several studies (Halstead et al., 2015; Lang et al., 2000; Nei et al., 2011; Pao et al., 2008). However, its antifungal activity hasn't been widely explored in the sprout industry. Doran (1928) reported that soils treated with 1 to 1.2% acetic acid protected tobacco plants against black root rot, brown root rot, and bed rot or damping off caused by several fungal species. A more recent study identified the antifungal properties of acetic acid and lactic acid produced by lactic acid bacteria (Cabo et al., 2002). However, the available studies about antifungal activity of acetic or lactic acid were very limited. In this study, acetic acid treatment and lactic acid treatment significantly reduced fungi incidence on soybean seeds compared with control treatment, where a higher reduction in fungal population was achieved by acetic acid, being a potential seed disinfection treatment for sprout soybean seeds.

Dry heat at 50 °C for 1 hour was the most ineffective seed treatment probably because the temperature and/or exposure time were not enough for killing fungal populations. Seeds treated with dry heat had a fungi incidence of 61.6% (Fig. 9). Dry heat treatments at 50 °C for 1 and 17 hour have shown to be effective in eliminating *E. coli* and other pathogen populations from alfalfa, radish, mung bean and broccoli seeds when combined with irradiation and chemical sanitizers, respectively (Bari et al., 2003; Bari et al., 2009b). Exposure to dry heat at higher temperatures and longer times (60 °C for 7 days or 65 °C for 4 days) reduced lupin seed infection with *Colletotrichum lupini* (Thomas and Adcock, 2004). In addition, it showed that a wrong procedure of disinfection by heating could result in higher fungi growth because temperature could have a positive effect on spores and make them sprout (Stanga, 2010). Exposure time and

temperature to adequately sterilize seeds depend on pathogen populations (marginal growth temperatures) and seed type. Therefore, determination of marginal growth temperatures of fungal populations present on soybean seeds and the effect of high temperatures on soybean seed germination will allow stablishing more effective dry heat treatments.

Consequently, a reduction of 93.2%, 74.9% and 71.6% from the initial fungal levels can be achieved by treating soybean seeds with hot water at 60 °C for 2 minutes, 5% calcium hypochlorite for 10 minutes and 5% acetic acid for 2 minutes, respectively. However, seed treatments could have negative effects on sprout quality. For this reason, it was important to evaluate their effect on sprout traits before making a seed treatment recommendation.

Effect of decontamination treatments on sprout traits

Calcium hypochlorite and acetic acid treatments did not produce any adverse effect on sprout traits compared with the control. Seeds treated with calcium hypochlorite, acetic acid and sterile water had High Quality Sprout (HQS) % and Average Quality (AVQ) % of 40.21% and 24.05%, 37.52% and 25.41%, and 38.81% and 21.5%, respectively. Low Quality Sprout (LQS) % and sprout yield ranged from 35.7% to 39.7% and 3.05 to 4.07 g/100 g seeds, respectively (Table 11). Similar results were found in a previous sprout trait study with HQS% and AQS% that ranged from 33.3 to 62.7% and 28 to 48.2%, respectively (Escamilla et al., 2017). Sprouts in this study had a lower sprout yield and higher LQS% than those previously reported by Escamilla et al. (2017), where they ranged from 5.24 to 7.01 g/100 g seeds and 7.7 to 29.7%, respectively. It suggested a lower seed quality in this study, which resulted in lower quality sprouts and sprout yield. These differences were likely caused by the environments because seed quality largely depended upon environmental conditions.

Calcium hypochlorite and acetic acid treatment effects on seed germination have been reported previously in other plants. A study on alfalfa seeds treated with acetic acid, calcium hypochlorite and a combination of both showed that germination of seeds was not adversely affected by any of the treatments obtaining germinations higher than 90% (Lang et al., 2000). Similarly, a treatment of 8.7% of acetic acid at 55 ℃ for 2-3 hour didn't affect germination rates of alfalfa and radish seeds (Nei et al., 2011). Treatments of up to 5,000 ppm $Ca(OCl)_2$ haven't showed negative effects on germination rate and seed viability of alfalfa seeds (Damron et al., 2005; Ding et al., 2013; Jaquette et al., 1996). However other studies showed that treatments of 200 ppm of active chlorine for 5 minutes reduced germination rate of alfalfa, broccoli, kohlrabi, kyona, mustard, red kohlrabi, red young radish and violet radish when combined with dry heat (80℃ for 24 h) (Choi et al., 2016). Seed treatments of cowpea seeds with 5% acetic acid for 45 minutes also significantly deteriorated seed germination (Singh et al., 2005). Therefore, seeds tolerance to chlorine and acetic acid varied among crops, exposure time and concentrations. As many studies showed that Calcium hypochlorite and acetic acid treatment has not negative effect on seed germination, treatments of 2% calcium hypochlorite for 10 minutes and 5% acetic acid for 2 minutes were also recommended as soybean seed disinfectants in this study because they significantly reduced fungal populations on soybean seeds without affecting seed germination and sprout quality traits.

On the contrary, hot water treatment produced sprouts with lower HQS% (6.76) and sprout yield (3.05) compared with the control and other treatments, while producing sprouts with higher AQS% and LQS% (Table 11). Despite this study showed that hot water was the best disinfectant seed treatment; hot water treatment significantly reduced sprouts quality and seed germination. It is well known that high and fluctuant temperatures speed up seed deterioration

rates reducing seed germination and viability (Mbofung et al., 2013). Similarly, hot water treatments (85℃ for 9 s) on alfalfa seed reduced germination (73%) and yield (78.4%) compared with the control (Enomoto et al., 2002). However, some other studies with dry heat at several temperature and times have shown to be effective as seed disinfectants without adverse effects on germination rates (Bari et al., 2009a; Jaquette et al., 1996). Hypocotyl length and thickness were not adversely affected by any of the treatments. As a result, despite its high effectiveness on reducing fungal populations hot water treatment at 60℃ for 2 minutes was not recommended for soybean seeds because it adversely affected germination rate and sprout quality by producing more sprouts with low and average quality. Further studies, exploring other temperatures and shorter times might produce satisfactory results in disinfecting soybean seeds without compromising seed viability and sprout quality.

Conclusion

Most fungal species isolated in this study were associated with mycotoxin production, especially *Fusarium* spp., *P. citrinum* and *A. alternata.* But their presence didn't necessarily mean sprout and seed contamination by mycotoxins. In order to make a complete diagnosis the specific mycotoxin must be detected from the fungal source and/or food. Further studies to evaluate the toxicological risk of isolated fungal strains will allow determining the actual risk that these species represent in sprout production in Virginia. Moreover, fungal contamination and frequent appearance indicated a potential risk for multiple mycotoxin contaminations of soybean seeds and sprouts, which could produce significant economic losses for soybean sprout growers and sprout industry. A significant reduction on initial levels of fungal populations could be achieved by treating soybean seeds with hot water, calcium hypochlorite and acetic acid, respectively. However, seeds treated with hot water had a lower germination rate and poor sprout quality. As a result, seed disinfection treatments of 2 % calcium hypochlorite for 10 minutes and 5% acetic acid for 2 minutes were recommended as potential seed disinfection treatments for soybean seeds because they reduced fungi incidence without reducing seed germination and sprout quality. None of the treatments completely eliminated the fungal populations, so optimization of relevant treatment parameters may be needed (temperature, time and dosage) to achieve a higher control. Exploring other temperature types and exposure times for hot water treatment might produce satisfactory results in disinfecting soybean seeds without reducing seed germination. Antimicrobial seed treatments are specific for seed type and microbe; and they must be evaluated for the specific seed type under normal production conditions. Thus, calcium hypochlorite and acetic acid treatment were recommended specifically for soybean seeds to control fungal populations.

A complete elimination of microbial contamination by seed disinfection treatments in sprout production is very unlikely. Food contamination is possible even if seed treatments are successful because treated pathogens may survive. Thus, seed disinfection must be implemented together with other preventative strategies such good agricultural practices (seed cleaning, seed storage, and handling), good manufacturing practices and hazard analysis.

References

Bari L., Sugiyama J., Kawamoto S. (2009a) Repeated Quick Hot-and-Chilling Treatments for the Inactivation of *Escherichia coli* O157:H7 in Mung Bean and Radish Seeds. Foodborne Pathogens and Disease 6:137.

Bari M.L., Nazuka E., Sabina Y., Todoriki S., Isshiki K. (2003) Chemical and Irradiation Treatments for Killing *Escherichia coli* O157:H7 on Alfalfa, Radish, and Mung Bean Seeds. Journal of Food Protection 66:767-774. DOI: 10.4315/0362-028x-66.5.767.

Bari M.L., Nei D., Enomoto K., Todoriki S., Kawamoto S. (2009b) Combination Treatments for Killing *Escherichia coli* O157:H7 on Alfalfa, Radish, Broccoli, and Mung Bean Seeds. Journal of Food Protection 72:631–636.

Battilani P., Gualla A., Dall'Asta C., Pellacani C., Galaverna G., Giorni P., Costa L.G. (2011) Phomopsins: An Overview of Phytopathological and Chemical Aspects, Toxicity, Analysis and Occurrence. World Mycotoxin Journal 4:345-359. DOI: doi:http://dx.doi.org.ezproxy.lib.vt.edu/10.3920/WMJ2011.1302.

Bennett J.W., Klich M. (2003) Mycotoxins. Clinical Microbiology Reviews 16:497-516. DOI: 10.1128/cmr.16.3.497-516.2003.

Bernardi-Wenzel J., Quecine M.C., Azevedo J.L., Pamphile J.A. (2016) *Agrobacterium*-Mediated Transformation of *Fusarium proliferatum*. Genet Mol Res 15. DOI: 10.4238/gmr.15027944.

Bottalico A., Logrieco A. (1988) Toxigenic Alternaria Species of Economic Importance, in: S. KK and B. D (Eds.), Mycotoxins in Agriculture and Food Safety, Marcel Dekker, New York. pp. 65-108.

Bovdisova I., Zbynovska K., Kalafova A., Capcarova M. (2016) Toxicological Properties of Mycotoxin Citrinin. Journal of Microbiology, Biotechnology and Food Sciences 05:10-13. DOI: 10.15414/jmbfs.2016.5.special1.10-13.

Broggi L.E., González H.H.L., Resnik S.L., Pacin A. (2007) *Alternaria alternata* Prevalence in Cereal Grains and Soybean Seeds from Entre Ríos, Argentina. Rev Iberoam Micol 24:47-51.

Cabo M.L., Braber A.F., Koenraad P.M.F.J. (2002) Apparent Antifungal Activity of Several Lactic Acid Bacteria against Penicillium discolor Is Due to Acetic Acid in the Medium. Journal of Food Protection 65:1309-1316. DOI: 10.4315/0362-028X-65.8.1309.

Chang K.F., Hwang S.F., Conner R.L., Ahmed H.U., Zhou Q., Turnbull G.D., Strelkov S.E., McLaren D.L., Gossen B.D. (2015) First report of *Fusarium proliferatum* Causing Root Rot in Soybean (Glycine max L.) in Canada. Crop Protection 67:52-58. DOI: https://doi.org/10.1016/j.cropro.2014.09.020.

Choi S., Beuchat L.R., Kim H., Ryu J.-H. (2016) Viability of Sprout Seeds as Affected by Treatment with Aqueous Chlorine Dioxide and Dry Heat, and Reduction of *Escherichia coli* O157:H7 and *Salmonella enterica* on Pak Choi Seeds by Sequential Treatment with Chlorine Dioxide, Drying, and Dry Heat. Food Microbiology 54:127-132. DOI: 10.1016/j.fm.2015.10.007.

Christopher J.G., William E.K., Janet C.M.-B., Jeff A.F., Patti L.W., Christine G.H., Paul R.C. (2003) Alfalfa Seed Decontamination in a Salmonella Outbreak. Emerging Infectious Disease journal 9:474. DOI: 10.3201/eid0904.020519.

Damron E., Klein C., Leach M., Mourot J., Murphy T., Seamans A., Wilson R. (2005) Use of Calcium Hypochlorite as a Sanitizer for Seeds Used for Sprouting: Task #2 impact: Improved Alfalfa Decontamination Technologies.   INQUIRY   6:73-85.

Darwish O., Li S., Matthews B., Alkharouf N. (2016) Genome-Wide Functional Annotation of *Phomopsis longicolla* Isolate MSPL 10-6.   Genomics   Data   8:67-69.   DOI: https://doi.org/10.1016/j.gdata.2016.03.006.

de Souza G.D., Mithofer A., Daolio C., Schneider B., Rodrigues-Filho E. (2013) Identification of *Alternaria alternata* Mycotoxins by LC-SPE-NMR and their Cytotoxic Effects to Soybean (*Glycine max*) Cell Suspension Culture.   Molecules   18:2528-38.   DOI: 10.3390/molecules18032528.

Dechet A.M., Herman K.M., Chen P.C., Taormina P., Johanson J., Tauxe R.V., Mahon B.E. (2014) Outbreaks Caused by Sprouts, United States, 1998–2010: Lessons Learned and Solutions Needed.   Foodborne   Pathogens   and   Disease   11:635-644.   DOI: https://doi-org.ezproxy.lib.vt.edu/10.1089/fpd.2013.1705.

Desjardins A.E. (2006) *Fusarium* Mycotoxins. Chemistry, Genetics and Biology The American Phytopathological Society, St Paul, MN.

Desjardins A.E., Manandhar H.K., Plattner R.D., Manandhar G.G., Poling S.M., Maragos C.M. (2000) Fusarium Species from Nepalese Rice and Production of Mycotoxins and Gibberellic Acid by Selected Species.   Applied and Environmental Microbiology   66:1020-1025.

Ding H., Fu T.J., Smith M.A. (2013) Microbial Contamination in Sprouts: How Effective is Seed Disinfection Treatment?   Journal of Food Science   78:R495-R501.   DOI: 10.1111/1750-3841.12064.

Ding W., Liu W.-Q., Jia Y., Li Y., van der Donk W.A., Zhang Q. (2016) Biosynthetic Investigation of *Phomopsins* Reveals a Widespread Pathway for Ribosomal Natural Products in Ascomycetes. Proceedings of the National Academy of Sciences 113:3521-3526. DOI: 10.1073/pnas.1522907113.

Domsch K.H., Gams W., Anderson T.H. (1980) Compendium of Soil Fungi Academic Press, London.

Dong Z., Liu G., Dong Z., Qian Y., An Y., Miao J., Zhen Y. (1987) Induction of Mutagenesis and Transformation by the Extract of *Alernaria alternata* Isolated from Grains in Linxian, China. Carcinogenesis 8:989-991. DOI: 10.1093/carcin/8.7.989.

Doran W.L. (1928) Acetic Acid as a Soil Disinfectant Journal of Agricultural Research 36:269-280.

Enomoto K., Takizawa T., Ishikawa N., Suzuki T. (2002) Hot-Water Treatments for Disinfecting Alfalfa Seeds Inoculated with *Escherichia coli* ATCC 25922. Food Science and Technology Research 8:247-251. DOI: 10.3136/fstr.8.247.

Escamilla D.M., Rosso M.L., Strawn L.K., Zhang B. (2017) Evaluation of Important Seed and Sprout Traits as Potential Selection Criteria in Breeding Varieties for Sprout soybeans. Euphytica 213. DOI: 10.1007/s10681-017-2007-0.

FDA. (1999) Microbiological Safety Evaluations and Recommendations on Sprouted Seed, Food and Drug Administration

FDA. (2017) Compliance with and Recommendations for Implementation of the Standards for the Growing, Harvesting, Packing, and Holding of Produce for Human Consumption for Sprout Operations: Guidance for Industry, FDA, U. S Food and Drug Administration.

Ghani M., Kulkarni K.P., Song J.T., Shannon J.G., Lee J.-D. (2016) Soybean Sprouts: a Review of Nutrient Composition, Health Benefits and Genetic Variation.   Plant Breeding and Biotechnology   4:398-412.   DOI: 10.9787/pbb.2016.4.4.398.

Gontia-Mishra I., Tripathi N., Tiwari S. (2014) A Simple and Rapid DNA Extraction Protocol for Filamentous Fungi Efficient for Molecular Studies.   Indian Journal of Biotechnology 13:536-539.

Griffin D.H. (1994) Fungal physiology. 2nd ed. ed. Wiley-Liss, New York :.

Halstead F.D., Rauf M., Moiemen N.S., Bamford A., Wearn C.M., Fraise A.P., Lund P.A., Oppenheim B.A., Webber M.A. (2015) The Antibacterial Activity of Acetic Acid Against Biofilm-Producing Pathogens of Relevance to Burns Patients.   PLoS One   10:e0136190. DOI: 10.1371/journal.pone.0136190.

Hestbjerg H., Nielsen K.F., Thrane U., Elmholt S. (2002) Production of Trichothecenes and other Secondary Metabolites by *Fusarium culmorum* and *Fusarium equiseti* on Common Laboratory Media and a Soil Organic Matter Agar: An Ecological Interpretation.   Journal of Agricultural and Food Chemistry   50:7593-7599.   DOI: 10.1021/jf020432o.

Hobbs T.W., Schmitthenner A.F., Kuter G.A. (1985) A new Phomopsis Species from Soybean. Mycologia   77:535-544.   DOI: 10.2307/3793352.

Hwang Y.H., Jeong Y.S., Lee J.D. (2004) Present Status and Future Developmental Direction of Soy-Related Industries in Korea.   Korea Soybean Digest   21:28-44.

Ivic D. (2014) Pathogenicity and Potential Toxigenicity of Seed-Borne *Fusarium* Species on Soybean and Pea.   Journal of Plant Pathology 96:541-551.

Ivic D., Domijan A.M., Peraica M., Milicevic T., Cvjetkovic B. (2009) *Fusarium* spp. Contamination of Wheat, Maize, Soybean, and Pea Grain in Croatia. Arh Hig Rada Toksikol 60:435-42. DOI: 10.2478/10004-1254-60-2009-1963.

Jaquette C.B., Beuchat L.R., Mahon B.E. (1996) Efficacy of Chlorine and Heat Treatment in killing *Salmonella stanley* Inoculated onto Alfalfa Seeds and Growth and Survival of the Pathogen during Sprouting and Storage. Applied and Environmental Microbiology 62:2212-2215.

Johnston L.M., Moe C.L., Moll D., Jaykus L.-A. (2005) The Epidemiology of Produce-Associated Outbreaks of Foodborne Disease, Microbial Hazard Identification in Fresh Fruit and vegetables, John Wiley & Sons, Inc. pp. 37-72.

Khan S.A., Hamayun M., Yoon H., Kim H.Y., Suh S.J., Hwang S.K., Kim J.M., Lee I.J., Choo Y.S., Yoon U.H., Kong W.S., Lee B.M., Kim J.G. (2008) Plant Growth Promotion and *Penicillium citrinum*. BMC Microbiol 8:231. DOI: 10.1186/1471-2180-8-231.

Kim D.H., Kim S.H., Kwon S.W., Lee J.K., Hong S.B. (2013) Mycoflora of Soybeans Used for Meju Fermentation. Mycobiology 41:100-7. DOI: 10.5941/MYCO.2013.41.2.100.

Kim H.-J., Koo K.-M., Kim G.-N., Lee D.S., Paik H.-D. (2002) Isolation and Identification of Pathogenic Microorganims from Soybean Sprouts. Nutraceuticals & Food 7:305-309.

Kosiak E.B., Holst-Jensen A., Rundberget T., Gonzalez Jaen M.T., Torp M. (2005) Morphological, Chemical and Molecular Differentiation of *Fusarium equiseti* Isolated from Norwegian Cereals. International Journal of Food Microbiology 99:195-206. DOI: https://doi.org/10.1016/j.ijfoodmicro.2004.08.015.

Kövics G.J., Sándor E., Rai M.K., Irinyi L. (2014) Phoma-Like Fungi on Soybeans. Critical Reviews in Microbiology 40:49-62. DOI: 10.3109/1040841X.2012.755948.

Lang M.M., Ingham B.H., Ingham S.C. (2000) Efficacy of Novel Organic Acid and Hypochlorite Treatments for Eliminating *Escherichia coli* O157:H7 from Alfalfa Seeds Prior to Sprouting. International Journal of Food Microbiology 58:73-82. DOI: https://doi.org/10.1016/S0168-1605(00)00297-X.

Langseth W. (1998) Mycotoxin Production and Cytotoxicity of *Fusarium* Strains Isolated from Norwegian Cereals. Mycopathologia 144:103-113. DOI: 10.1023/a:1007016820879.

Lee J.-D., Shannon J.G., Jeong Y.-S., Lee J.-M., Hwang Y.-H. (2007) A Simple Method for Evaluation of Sprout Characters in Soybean. Euphytica 153:171-180. DOI: 10.1007/s10681-006-9252-y.

Lehman S.G. (1923) Pod and Stem Blight of Soybean. Annals of the missouri botanical garden 10:111-178. DOI: 10.2307/2394035.

Leonardo N.G.e.S., Carlotto I.B., Luisi S.B., Kopper P.M.P., Grecca F.S., Montagner F. (2016) Calcium Hypochlorite Solutions: Evaluation of Surface Tension and Effect of Different Storage Conditions and Time Periods Over pH and Available Chlorine Content. Journal of Endodontics 42:641-645. DOI: https://doi.org/10.1016/j.joen.2016.01.006.

Li S., Hartman G.L., Boykin D.L. (2010) Aggressiveness of *Phomopsis longicolla* and other *Phomopsis* spp. on Soybean. Plant Disease 94:1035-1040. DOI: 10.1094/PDIS-94-8-1035.

Liu K. (1997) Soybeans: Chemistry, Technology, and Utilization Chapman & Hall, New York.

Logrieco A., Moretti A., Castella G., Kostecki M., Golinski P., Ritieni A., Chelkowski J. (1998) Beauvericin Production by *Fusarium* Species. Applied and Environmental Microbiology 64:3084-3088.

Marasas W.F.O., Thiel P.G., Rabie C.J., Nelson P.E., Toussoun T.A. (1986) Moniliformin Production in *Fusarium* Section Liseola. Mycologia 78:242-247. DOI: 10.2307/3793169.

Mathivanan N., Kabilan V., Murugesan K. (1998) Purification, Characterization, and Antifungal Activity of chitinase from *Fusarium chlamydosporum,* a Mycoparasite to Groundnut Rust, *Puccinia arachidis*. Can. J. Microbiol. 44:646-651.

Mayta J., Chen P., Popp M.P., Dong D., Wu C.J., Zhang B., Smith S.F., Scaboo A.M. (2014) Break-Even Profitability for Food-Grade Specialty Soybeans. Agricultural Science 2:1-11. DOI: 10.12735/as.v2i2p01.

Mbofung G.C.Y., Goggi A.S., Leandro L.F.S., Mullen R.E. (2013) Effects of Storage Temperature and Relative Humidity on Viability and Vgor of Treated Soybean Seeds. Crop Science 53:1086. DOI: 10.2135/cropsci2012.09.0530.

Mengistu A., Castlebury L.A., Morel W., Ray J.D., Smith J.R. (2014) Pathogenicity of *Diaporthe* spp. Isolates Recovered from Soybean (Glycine max) seeds in Paraguay. Canadian Journal of Plant Pathology 36:470-474. DOI: 10.1080/07060661.2014.966151.

Milićević D.R., Škrinjar M., Baltić T. (2010) Real and Perceived Risks for Mycotoxin Contamination in Foods and Feeds: Challenges for Food Safety Control. Toxins 2:572-592. DOI: 10.3390/toxins2040572.

Mousa W.K., Schwan A., Davidson J., Strange P., Liu H., Zhou T., Auzanneau F.-I., Raizada M.N. (2015) An Endophytic Fungus Isolated from Finger Millet (Eleusine coracana) Produces

101

Anti-Fungal Natural Products. Frontiers in Microbiology 6. DOI: 10.3389/fmicb.2015.01157.

NACMCF. (1999) Microbiological Safety Evaluations and Recommendations on Sprouted Seeds. International Journal of Food Microbiology 52:123-153. DOI: https://doi.org/10.1016/S0168-1605(99)00135-X.

Nei D., Latiful B.M., Enomoto K., Inatsu Y., Kawamoto S. (2011) Disinfection of Radish and Alfalfa Seeds Inoculated with *Escherichia coli* O157:H7 and *Salmonella* by a Gaseous Acetic Acid Treatment. Foodborne Pathogens and Disease 8:1089+.

Oms-Oliu G., Rojas-Graü M.A., González L.A., Varela P., Soliva-Fortuny R., Hernando M.I.H., Munuera I.P., Fiszman S., Martín-Belloso O. (2010) Recent Approaches Using Chemical Treatments to Preserve Quality of Fresh-Cut Fruit: A review. Postharvest Biology and Technology 57:139-148. DOI: https://doi.org/10.1016/j.postharvbio.2010.04.001.

Ostry V. (2008) Alternaria Mycotoxins: an Overview of Chemical Characterization, Producers, Toxicity, Analysis and Occurrence in Foodstuffs. World Mycotoxin Journal 1:175-188. DOI: 10.3920/WMJ2008.x013.

Oviedo M.S., Ramirez M.L., Barros G.G., Chulze S.N. (2009) Effect of Environmental Factors on Tenuazonic Acid Production by *Alternaria alternata* on Soybean-Based Media. J Appl Microbiol 107:1186-92. DOI: 10.1111/j.1365-2672.2009.04301.x.

Pao S., Kalantari A., Khalid M.F. (2008) Eliminating *Salmonella enterica* in Alfalfa and Mung Bean Sprouts by Organic Acid and Hot Water Immersions. Journal of Food Processing and Preservation 32:335-342. DOI: 10.1111/j.1745-4549.2008.00182.x.

Park J., Chu F.S. (1993) Immunochemical Analysis of Trichothecenes Produced by Various Fusaria.   Mycopathologia 121:179-192.

Peltonen K., Jestoi M., Eriksen G. (2010) Health Effects of Moniliformin: a Poorly Understood *Fusarium* Mycotoxin World Mycotoxin Journal 3:403-414.   DOI: https://doi.org/10.3920/WMJ2010.1232

Pitt J.I. (1979) The Genus *Penicillium* and its Teleomorphic States *Eupenicillium* and *Talaromyces* Academic Press., London.

Pitt J.l., Hocking A.D. (2009) Fungi and Food Spoilage. Third ed. Springer US, New York.

Pitt J.l., Hocking A.D., Bhudhasamai K., Miscamble B.F., Wheeler K.A., Tanboon-Ek p. (1994) The Normal Mycoflora of Commodities from Thailand. 2. Beans, Rice, Small Grains and other Commodities.   International Journal of Food Microbiology 23:35-53.

Plumlee K.H., Galey F.D. (1994) Neurotoxic Mycotoxins: a Review of FungalToxins that Cause Neurological Disease in Large Animals.   Journal of Veterinary Internal Medicine   8:49-54. DOI: 10.1111/j.1939-1676.1994.tb03195.x.

Ross P.F., Nelson P.E., Richard J.L., Osweiler G.D., Rice L.G., Plattner R.D., Wilson T.M. (1990) Production of Fumonisins by *Fusarium moniliforme* and *Fusarium proliferatum* isolates Associated with Equine Leukoencephalomalacia and a Pulmonary Edema Syndrome in swine. Applied and Environmental Microbiology   56:3225-3226.

Rousk J., Bååth E. (2011) Growth of Saprotrophic Fungi and Bacteria in Soil.   FEMS Microbiology Ecology   78:17-30.   DOI: 10.1111/j.1574-6941.2011.01106.x.

Santos J.M., Vrandecic K., Cosic J., Duvnjak T., Phillips A.J. (2011) Resolving the *Diaporthe* Species Occurring on Soybean in Croatia. Persoonia 27:9-19. DOI: 10.3767/003158511X603719.

Sato T., Aoki M., Aoki T., Kubota M., Yaguchi T., Uzuhashi S., Tomioka K. (2014) Fungi Isolated from Spoiled Bean Sprouts in Japan. JARQ 48:317-329.

Sauer D.B., Burroughs R. (1986) Disinfection of Seed Surfaces with Sodium Hypochlorite. Phytopathology 76:745-749.

Schoch C.L., Seifert K.A., Huhndorf S., Robert V., Spouge J.L., Levesque C.A., Chen W., Fungal Barcoding C., Fungal Barcoding Consortium Author L. (2012) Nuclear Ribosomal Internal Transcribed Spacer (ITS) Region as a Universal DNA Barcode Marker for Fungi. Proc Natl Acad Sci U S A 109:6241-6. DOI: 10.1073/pnas.1117018109.

Sciortino C.V. (2017) Atlas of Clinically Important Fungi Wiley.

Sikin A.M., Zoellner C., Rizvi S.S. (2013) Current Intervention Strategies for the Microbial Safety of Sprouts. J Food Prot 76:2099-123. DOI: 10.4315/0362-028X.JFP-12-437.

Silva L.R., Pereira M.J., Azevedo J., Gonçalves R.F., Valentão P., de Pinho P.G., Andrade P.B. (2013) Glycine max (L.) Merr., Vigna radiata L. and Medicago sativa L. Sprouts: A Natural Source of Bioactive Compounds. Food Research International 50:167-175. DOI: 10.1016/j.foodres.2012.10.025.

Singh B., Chandra M., Agrawal R., Babu N. (2005) Curing of *Salmonella enterica*, Serovar Typhimurium-Contaminated Cowpea Seeds and Sprouts with Vinegar and Chlorination.

Singh R., Soni S.K., Kalra A. (2013) Synergy between *Glomus fasciculatum* and a Beneficial *Pseudomonas* in Reducing Root Diseases and Improving Yield and Forskolin Content in Coleus

forskohlii Briq. under organic field conditions. Mycorrhiza 23:35-44. DOI: 10.1007/s00572-012-0447-x.

Sobrova P., Adam V., Vasatkova A., Beklova M., Zeman L., Kizek R. (2010) Deoxynivalenol and its Toxicity. Interdiscip Toxicol 3:94-9. DOI: 10.2478/v10102-010-0019-x.

Stanga M. (2010) Disinfectants and Sanitation Technology, Sanitation, Wiley-VCH Verlag GmbH & Co. KGaA. pp. 499-554.

Stankovic S., Levic J., Petrovic T., Logrieco A., Moretti A. (2007) Pathogenicity and Mycotoxin Production by *Fusarium proliferatum* Isolated from Onion and Garlic in Serbia. European Journal of Plant Pathology 118:165-172. DOI: 10.1007/s10658-007-9126-8.

Steyn P.S., Rabie C.J. (1976) Characterization of Magnesium and Calcium Tenuazonate from *Phoma sorghina.* Phytochemistry 15:1977-1979. DOI: https://doi.org/10.1016/S0031-9422(00)88860-3.

Sun S., Kim M.Y., Chaisan T., Lee Y.W., Van K., Lee S.H. (2013) Phomopsis (Diaporthe) Species as the Cause of Soybean Seed Decay in Korea. Journal of Phytopathology 161:131-134. DOI: 10.1111/jph.12034.

Thomas G.J., Adcock K.G. (2004) Exposure to Dry Heat Reduces *Anthracnose* Infection of Lupin Seed. Australasian Plant Pathology 33:537–540.

Thomma B.P.H.J. (2003) Alternaria spp.: from General Saprophyte to Specific Parasite. Molecular Plant Pathology 4:225-236. DOI: 10.1046/j.1364-3703.2003.00173.x.

Torres D.E., Rojas-Martínez R.I., Zavaleta-Mejía E., Guevara-Fefer P., Márquez-Guzmán G.J., Pérez-Martínez C. (2017) *Cladosporium cladosporioides* and *Cladosporium pseudocladosporioides* as Potential New Fungal Antagonists of *Puccinia horiana* Henn., the

Causal Agent of Chrysanthemum White Rust.   PLoS ONE   12:e0170782.   DOI: 10.1371/journal.pone.0170782.

Waqas M., Khan A.L., Ali L., Kang S.M., Kim Y.H., Lee I.J. (2013) Seed Germination-Influencing Bioactive Secondary Metabolites Secreted by the Endophyte *Cladosporium cladosporioides* LWL5.   Molecules   18:15519-30.   DOI: 10.3390/molecules181215519.

Weiss A., Hammes W.P. (2005a) Efficacy of Heat treatment in the Reduction of s*almonella* and *Escherichia coli* O157:H– on Alfalfa, Mung Bean and Radish Seeds Used for Sprout Production. European Food Research and Technology   221:187-191.   DOI: 10.1007/s00217-004-1125-9.

Weiss A., Hammes W.P. (2005b) Efficacy of Heat Treatment in the Reduction of *Salmonella* and *Escherichia coli* O157:H– on Alfalfa, Mung Bean and Radish Seeds Used for Sprout Production. European Food Research and Technology   221:187-191.   DOI: 10.1007/s00217-004-1125-9.

White, Bruns T., Lee S., Taylor J. (1990) White, T. J., T. D. Bruns, S. B. Lee, and J. W. Taylor. Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for phylogenetics.

Xue A.G., Morrison M.J., Cober E., Anderson T.R., Rioux S., Ablett G.R., Rajcan I., Hall R., Zhang J.X. (2007) Frequency of isolation of species of Diaporthe and Phomopsis from Soybean Plants in Ontario and Benefits of Seed Treatments.   Canadian Journal of Plant Pathology 29:354-364.   DOI: 10.1080/07060660709507481.

Ye J., Kostrzynska M., Dunfield K., Warriner K. (2017) Control of Salmonella on Sprouting Mung Bean and Alfalfa Seeds by Using a Biocontrol Preparation Based on Antagonistic Bacteria and Lytic Bacteriophages.   Journal of Food Protection   73:9–17.

Zain M.E. (2011) Impact of Mycotoxins on Humans and Animals.   Journal of Saudi Chemical Society   15:129-144.   DOI: 10.1016/j.jscs.2010.06.006.

Zhang B., Chen P., Florez-Palacios S.L., Shi A., Hou A., Ishibashi T. (2010) Seed Quality Attributes of Food-Grade Soybeans from the U.S. and Asia. Euphytica 173:387-396. DOI: 10.1007/s10681-010-0126-y.

Zhang J.X., Xue A.G., Cober E.R., Morrison M.J., Zhang H.J., Zhang S.Z., Gregorich E. (2013) Prevalence, Pathogenicity and Cultivar Resistance of Fusarium and Rhizoctonia Species Causing Soybean Root Rot. Canadian Journal of Plant Science 93:221-236. DOI: 10.4141/cjps2012-223.

Tables and Figures

Table 8: Composite seed samples of MFS-561used to evaluate seed disinfection treatment effect on sprout traits.

Sample ID [a]	Growing Region	Year	Amount (g)
CS1	Southern Virginia	2015	453.6
CS2	Eastern Virginia	2015	453.6
CS3	Southern Virginia	2016	453.6
CS4	North Eastern North Carolina	2016	453.6

[a]CS: Composite seed sample.

Table 9: Fungi isolated from seeds of a sprout soybean cultivar grown in southern Virginia, Eastern Virginia and Northeastern North Carolina in 2015.

Division[a]	Species	Isolation source	Location source[b]	Year	BLAST hit (accessions)[c]	Homology[g]
A	*Alternaria alternata*	Seeds	NC, SV and EV	2015	KX783406.1	100/95
A	*Alternaria* sp.	Seeds	EV and SV	2015	KX878965.1	100/92
A	*Cladosporium cladosporioides*	Seeds	EV and SV	2015	KX258800.1	99/100
A	*Diaporthe longicolla*	Seeds	NC and EV	2015	KX878969.1	98/99
A	*Diaporthe* sp.	Seeds	NC, SVand EV	2015	MF435146.1	90/97
A	*Fusarium chlamydosporum*	Seeds	EV	2015	KU516827.1	82/94
A	*Fusarium proliferatum*	Seeds	SV	2015	JQ957846.1	90/77
A	*Fusarium equiseti*	Seeds	EV and NC	2015	KX878922.1	93/95
A	*Fusarium* sp.	Seeds	EV and NC	2015	KU886151.1	97/93
A	*Penicillium citrinum*	Seeds	NC	2015	KX867539.1	99/99
A	*Phoma* sp.	Seeds	SV	2015	KM387394.1	90/80

[a] A: Ascomycota. Divisions of fungi species were found at Mycobank webpage.
[b] SV: Southern Virginia, EV: Eastern Virginia, NC: North Eastern North Carolina.
[c] Strains identification preserved at the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI.
[d] Identity (%)/query coverage (%).

Table 10: Relative frequency of fungal species by growing regions southern Virginia, Eastern Virginia and Northeastern North Carolina.

Species	% Frequency			
	NC[a]	SV[a]	EV[a]	Total[b]
Alternaria alternata	7.69	34.78	15.79	21.82
Alternaria sp.	0.00	8.70	10.53	7.27
Cladosporium cladosporioides	0.00	8.70	10.53	7.27
Diaphorte longicolla	23.08	0.00	15.79	10.91
Diaphorte spp.	7.69	34.78	5.26	18.18
Fusarium Chlamydosporum	0.00	0.00	10.53	3.64
Fusarium equiseti	30.77	0.00	10.53	10.91
Fusarium proliferatum	0.00	4.35	0.00	1.82
Fusarium spp.	7.69	0.00	15.79	7.27
Penicillium citrinum	15.38	0.00	0.00	3.64
Phoma spp.	0.00	4.35	0.00	1.82
Uncultured fungus	7.69	4.35	5.26	5.45

[a] The growing region names are abbreviated as follows: NC is North Eastern North Carolina, SV is Southern Virginia and EV is Eastern Virginia.
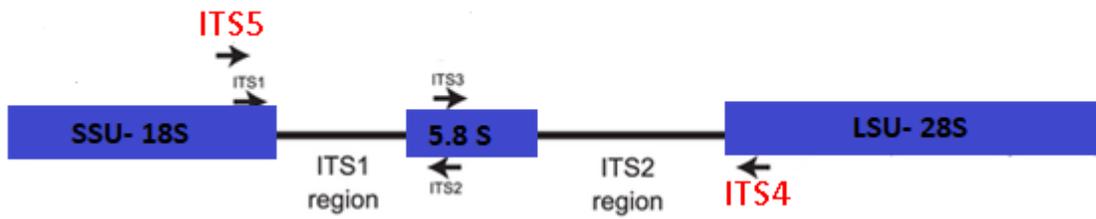[b] Total frequency across all regions.

Table 11: Mean of sprout quality traits of seed disinfection treatments.

Treatments	HQS%[a]		AQS%[a]		LQS%[a]		Syld (g/g seed)[a]		Hlght (cm)[a]		Hthk (mm)[a]	
	Mean	SE[b]	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
Ca(ClO)2 2%, 10 min	40.21 a *	2.88	24.05 b	2.45	35.74 b	2.74	4.07 a	0.23	9.64 a	0.43	1.74 a	0.05
Acetic acid 5%, 2 min	37.52 a	3.02	25.41 b	2.32	37.07 b	1.88	3.74 ab	0.15	9.20 a	0.54	1.71 a	0.02
Hot water 60 °C, 2 min	6.76 b	1.20	42.43 a	3.6	50.81 a	3.65	3.05 b	0.22	9.46 a	0.37	1.73 a	0.01
Sterile Water 1 min	38.81 a	4.27	21.50 b	2.06	39.69 b	3.37	3.46 ab	0.35	9.20 a	0.55	1.74 a	0.05

[a] The sprout trait names are abbreviated as follows: HQS% is high quality sprout percentage, AQS% is average quality sprout percentage, LQS% is low quality sprout percentage, Hlgth is hypocotyl length, Hthk is hypocotyl thickness, Syld is sprout yield and DF is degrees of freedom.
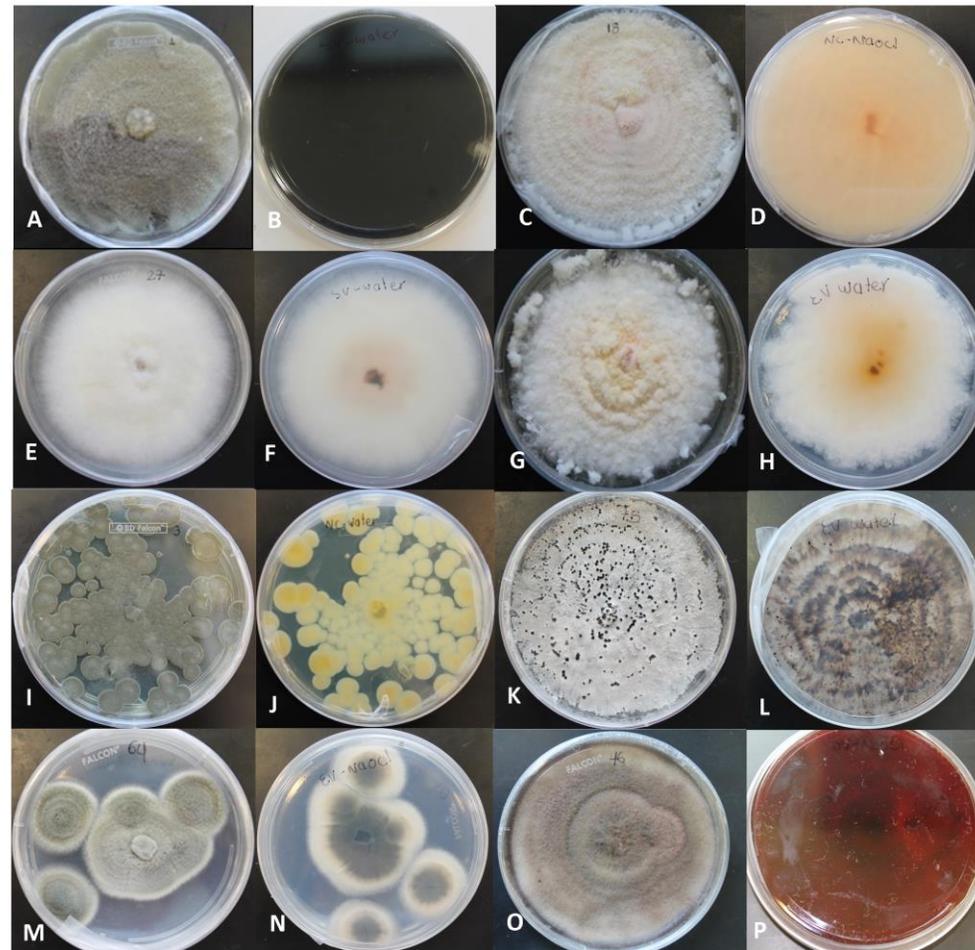
[b] SE is standard error.
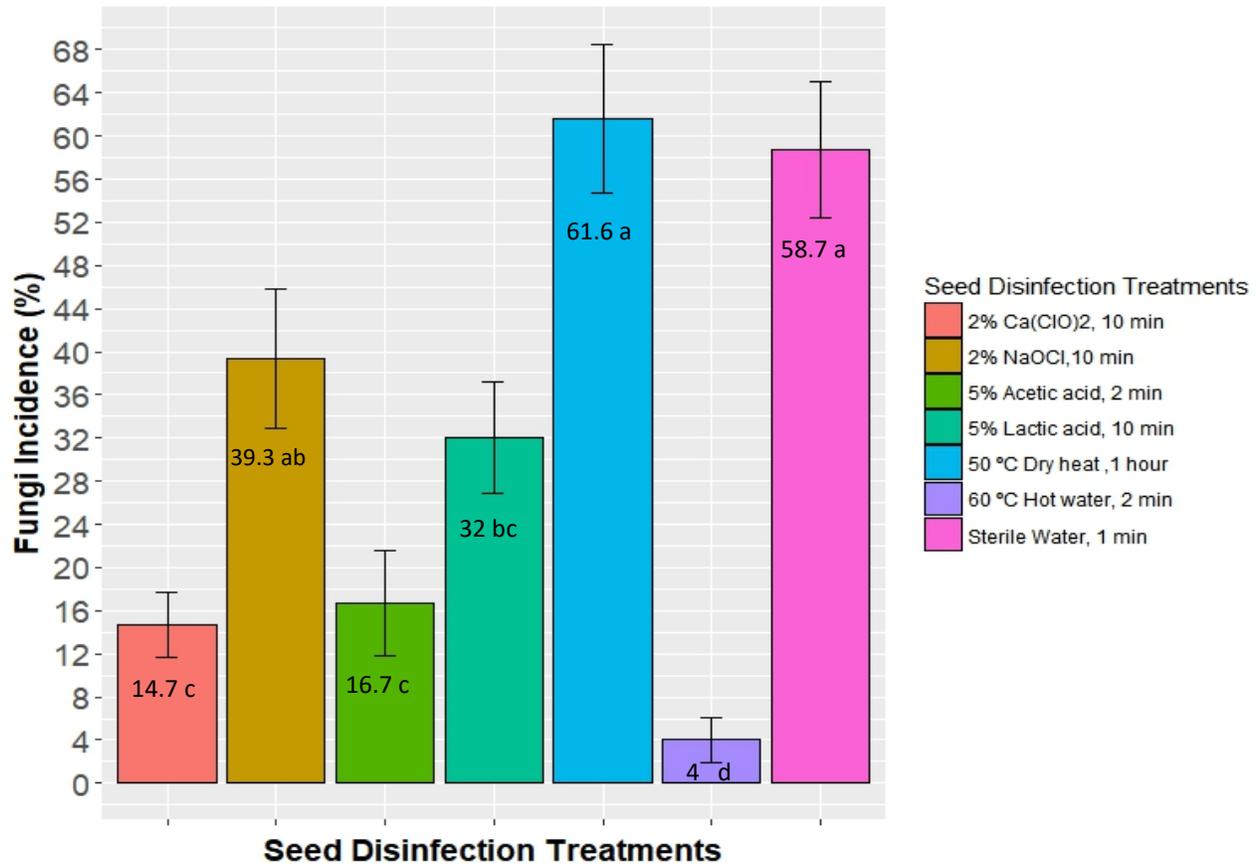
Figure 7. Nuclear ribosomal RNA genes and their ITS regions. Positions of forward (right-pointing arrow) and reverse (left-pointing arrow) primers are shown on the map of ITS regions and surrounding ribosomal RNA genes. LSU is large subunit and SSU is small subunit.

Figure 8. Fungi isolated from seeds of a commercial sprout cultivar in VA, United States.
A: surface side of a colony on a PDA plate of *Alternaria alternata*, B: reverse side of *A. alternata* colony, C: surface side of a colony on a PDA plate of *Fusarium equiseti*, D: reverse side of *F. equiseti* colony, E: surface side of a colony on PDA plate of *Fusarium proliferatum*, F: reverse side of *F. proliferatum* colony, G: surface side of a colony on PDA plate of *Fusarium chlamydosporum,* H: reverse side of *F. chlamydosporum* colony, I: surface side of a colony on a PDA plate of *Penicillium citrinum,* J: reverse side of *P. citrinum* colony, K: surface side of a colony on a PDA plate of *Diaphorte longicolla,* L: reverse side of *D. longicolla* colony, M: surface side of a colony on PDA plate of *Cladosporium cladosporioides,* N: reverse side of *C. cladosporioides* colony, O: surface side of a colony on PDA plate of *Phoma* sp., P: reverse side of *Phoma* sp. colony.

Figure 9. Average fungus incidence on soybean seeds by seed treatments across three different seed production regions. Different letters represent significant differences at P≤0.05 level.

4. QTL analysis of seed coat deficiency of soybean seeds (*Glycine max* [L.] Merr.)

Diana M. Escamilla[1], Song Qijian[2] and Bo Zhang[1*]

[1]Department of Crop and Soil Environmental Sciences, Virginia Tech, Blacksburg, VA 24060;

[2]Beltsville Agricultural Research Center, USDA-ARS.

Abbreviations: SCD, seed coat deficiency; SNP, single nucleotide polymorphism; LG, linkage group; LOD, logarithmic odd score; MAS, marker assisted selection.

Abstract

Seed coat deficiency is an undesirable trait of soybean seeds for food use such as natto soybeans because they cause an inferior appearance of the products and clogging of the production lines. Development of soybean cultivars with improved natto quality characteristics is crucial for maintaining and increasing the natto soybean market. However, the understanding of the genetic control of seed coat deficiency on soybean seeds was very limited. Thus, the objective of this study was to identify and validate quantitative trait loci (QTL) underlying seed coat deficiency. Mapping population derived from V11-0883 x V12-1626 consisted of 240 F_3 plants and 240 $F_{4:5}$ lines grown in Blacksburg, VA in 2016 and 2017, respectively. Validation population derived from V13-1687 x V12-1885 with 166 F_3 plants was grown in Blacksburg, VA in 2016. A total of 1,318 and 1,605 SNPs out of 6,000 SNPs tested were polymorphic between parents in the mapping and validation population, respectively. Linkage maps consisted of 20 linkage groups with a total length of 1,826 cM and average marker interval of 1.6 cM for the mapping population, and a total length of 1,970 cM and average marker interval of 1.4 cM for the validation population. Broad sense heritability was estimated to be 0.40 for seed coat deficiency. One stable QTL across years (*qSCD20_1*) was identified near Gm20_34881595_C_T and Gm20_36002148_T_C on chromosome 20 (LG I) with logarithmic odds score of 8.43 and 15.15, accounting for 11.3 and 24.3% of the variation of the trait in 2016 and 2017, respectively. In the validation population, a QTL (*qSCD20_2)* was identified 1.2 Mb apart from *qSCD20_1* with 8.56 of LOD value and explaining 22.6% of the total variation of the trait. Further confirmation experiments are needed to validate *qSCD20_1* and *qSCD20_2*, which is essential for further utilization in MAS.

Introduction

Improved processing technologies have made it possible to use soybeans to make foods that are familiar to consumers, which have prompted soy-food popularity outside of Asia (Hartman et al., 2011, Mayta et al., 2014). Specialty soybeans are used to produce various soy-foods of Asian origin; they are generally grown under contracts and receive premium prices (Kentucky, 2013). Therefore, specialty soybeans is a high value alternative to growing grain soybean for feed (Cui et al., 2004, Taira, 1990). Natto is a traditional Japanese soyfood fermented by *Bacillus subtillis* (*natto*) that is served for breakfast and dinner with rice (Hosoi and Kiuchi, 2003, Hu et al., 2010, Wei and Chang, 2004, Liu, 1997). It is considered a very nutritional and healthy food because of its large amount of protein, fiber, amino acids and isoflavones (Zhang et al., 2011, Weng and Chen, 2012, Ikeda et al., 2006, Somekawa et al., 2001). In addition, bioactive peptides produced during natto fermentation have shown a wide range of benefits for human health (Sanjukta and Rai, 2016, Tamang et al., 2016, Gibbs et al., 2004).

Natto is made from small seeded soybean varieties (<9 g/100 seeds) (Zhang et al., 2010, Liu, 1997, Cui et al., 2004). It is prepared by soaking and steaming/cooking soybean seeds until tender. Once, seeds are dried and cooled to 40 °C they are inoculated with *B. subtillis*, packed in a wooden box or polyethylene bag and incubated at 40-43 °C for 12-20 h (Hosoi and Kiuchi, 2003, Wei et al., 2001, Liu, 1997, Hu et al., 2010). Natto quality is determined by soybean cultivars, processing conditions (soaking, cooking and fermentation) and bacteria strain (Wei and Chang, 2004); and the suitability of potential cultivars or seed product for the natto market is usually determined by professional testers based on sensory evaluations (Zhang et al., 2008). Good quality natto must have a characteristic flavor (sweet), a white colored mucous substance,

117

light yellow color, low ammonia content and soft and sticky texture (Wei and Chang, 2004, Hu et al., 2010, Liu, 1997, Hosoi and Kiuchi, 2003). However, evaluation of natto cultivars based on sensory panels is time consuming, so indirect selection for natto quality based on seed traits would be desirable. Natto soybean industry prefers soybean seeds that are easily washable, and have uniform size, round shape, clear hilum, slight changes in constituents during storage, low cracking ratio after steaming, high water absorption, and yellow and smooth seed coat (Cober et al., 1997, Wei and Chang, 2004, Taira, 1990, Hosoi and Kiuchi, 2003, Geater et al., 2000, Cui et al., 2004, Liu, 1997).

The initial water absorption, where seeds are softened and soluble sugars are released, is an important step in natto fermentation (Cook and Rainey, 2010). Finished natto has high water content (60%) and it is desirable for natto soybean cultivars to at least double dry weight through water absorption (Wei et al., 2001, Cook and Rainey, 2010). Thus, small-seeded genotypes are preferred because small seeds have high water absorption capacity (Geater et al., 2000, Cui et al., 2004, Liu, 1997). It is well known that seed germination is determined by the seed coat (Qutob et al., 2008); which regulates water absorption and prevent the destruction of seed tissue at the beginning of imbibition (Copeland and McDonald, 2001, Koizumi et al., 2008). Therefore, it is also ideal for natto soybean to quickly absorb as much water as possible, but the seed coat remains intact despite weakening and expanding during absorption (Cook and Rainey, 2010). In addition, seed coat cracking after soaking is an undesirable trait for natto soybeans because they cause an inferior appearance of the product and clogging of the production lines (Yasui et al., 2017). Because natto manufacturers expressed that some natto germplasm adapted to mid-Atlantic region in US had poor seed coat integrity (Cook and Rainey, 2010), there was a need to develop soybean cultivars with improved seed coat strength.

118

Several seed tests have been developed to evaluate seed coat integrity in soybeans. Rodda et al. (1973) soaked soybean seeds on 0.1% sodium hypochlorite for 5 minutes to detect damage seeds based on the fact that cracked soybean seeds absorb moisture faster and swell in size more than undamaged seeds. Based on the same principle, a seed coat deficiency test (SCD) was developed by Cook and Rainey (2010) for evaluating the seed coat strength of natto soybeans. Seeds were soaked in water for 10 minutes and cracked seed or severely blistered around the hilum or detached seed coat from the hull, were counted as deficient seeds. Small-seeded soybeans have showed large differences in seed coat strength with SCD percentages ranging from 23 to 70%; and a higher resistance to seed coat cracking than large-seeded soybeans (Cook and Rainey, 2010). In a recent study, seed coat cracking during water absorption of natto soybeans was evaluated by soaking soybean seeds for 18 h at 10 °C and seed coat cracking ratios ranged from 4 to 100% (Yasui et al., 2017). Other seed tests used to evaluate the integrity of the seed coat that could be use on natto soybeans included sodium hypochlorite, fast green, indoxyl acetate, tetrazolium and Chowdhury methods (VanUtrecht et al., 2000, Copeland and McDonald, 2001). As a result, several tests were available for evaluating the seed coat strength of natto soybeans during water absorption referred in previous studies as seed coat deficiency and seed coat cracking after soaking (Cook and Rainey, 2010, Yasui et al., 2017). In this study, we adopted the name given by Cook and Rainey (2010) and used seed coat deficiency to refer to seed coat integrity during water absorption.

Soybean (*Glycine max* L. Merr.) has 20 chromosomes and a genome size of 1.1-1.15 Gb (Arumuganathan and Earle, 1991). Reports of the high quality whole genome sequence for *Glycine max* var. Williams 82 are available, representing about 85% of the predicted genome. Around 75% of the genes are present in multiple copies due to duplication events occurred 59 to

119

13 million years ago, and 57% of the genomic sequence occurs in repeat- rich, low recombination heterochromatic regions surrounding the centromeres. The average ratio of genetic-to physical distance is 1cM per 197 kb in euchromatic regions, and 1 cM per 3.5 Mb in heterochromatic regions (Schmutz et al., 2010). To date, the availability of accurate soybean genome sequences has facilitated the identification of the genetic base of many soybean traits and accelerates the breeding process of improved soybean varieties (Schmutz et al., 2010).

Soyfood quality traits are generally affected by genetically controlled seed traits that are quantitative and ambiguous, where genes or loci associated with a quantitative trait are known as quantitative trait loci (QTL)   (Liu, 1997, Cui et al., 2004, Cook and Rainey, 2010, Gupta et al., 2005). Breeding varieties resistant to seed coat cracking is challenging due to the complicated genetic behavior and environmental interactions (Ha et al., 2012). Thus, QTL and markers associated with seed coat deficiency play an important role in natto soybean breeding through marker assisted selection (MAS). Currently, SoyBase (http://www.soybase.org/ accessed 16 Oct. 2017) reported 14 QTL for seed coat cracking, however they haven't been confirmed yet. Oyoo et al. (2010b) reported that there were two types of genetically controlled seed coat cracking, type I with irregular cracks and type II with net-like cracks; and detected two QTL for seed coat cracking type II in molecular linkage group C1 (chromosome 4) with a distance of about 62 cM accounting for 16% and 32.4% of phenotypic variation. Two other QTL were identified for average cracking index (not cracked, 0, to severely cracked, 4), cr1 in linkage group D1b (chromosome 2) with LOD score of 4.77 in an F_7 population and cr2 in linkage group M (chromosome 7) with LOD score of 4.10 in an F_6 population accounting for 23.5% and 16.2% of phenotypic variation, respectively (Oyoo et al., 2010a). In addition, a total of 10 QTL with additive effects were identified for seed coat cracking, explaining 48.0% of the phenotypic

variation, with additive effects of 33.8%, epistatic effect of 7.6% and additive-by-environment interaction effect and additive-by-epistatic interaction effect of 6% (Ha et al., 2012). However, the phenotype of above mentioned QTL have been obtained by determining physical seed coat cracking after harvest. While seed coat deficiency, the trait of interest for natto soybeans evaluates the resistance to seed coat cracking of soybean seeds during water absorption. In a recent study, QTL for seed coat cracking after soaking (qSCAS1, qSCAS2 and qSCAS3) and after cooking (qSCAC1 and qSCAC2) located on chromosomes 4 (LG: C1), 6 (LG: C2) and 8 (LG: A2), were detected on a recombinant inbred line (RIL) population; and interactions between QTL were also observed (Yasui et al., 2017). To date, there was limited information about genetic control of seed coat deficiency on soybeans.

Development of soybean cultivars with improved natto quality characteristics is crucial for maintaining and increasing the natto soybean market. Although several studies have been conducted to identify QTL associated with seed coat cracking in soybeans, only one study detected QTL associated with seed coat deficiency also referred as seed coat cracking after soaking. In addition, the understanding of the genetic control of seed coat deficiency on soybean seeds was very limited. Therefore, the aims of the present study were to identify quantitative trait loci (QTL) underlying seed coat deficiency and associated markers in the V11-0883 x V12-1626 population and verify them through a validation population derived from V13-1687 x V12-1885; which would provide useful information for researchers and breeders to improve natto cultivars.

Materials and Methods

Plant Material

Parental soybean lines were chosen based on the severity of seed coat deficiency. Two high SCD soybean lines, V11-0883 and V13-1687, were crossed with two low SCD soybean lines, V12-1626 and V12-1885 for developing two populations (Table 12). Crosses were made in summer of 2015 at the Virginia Tech farm in Blacksburg, VA. F_1 plants were spaced-planted and harvested in a Puerto Rico winter nursery in winter of 2015; the hybrid plants were harvested to form F_2 populations. Six SSR markers, Satt449, Satt197, Satt281, Satt268, Satt431 and Satt345, and flower color were used to verify true hybrids (Table 12). F_2 populations were spaced-planted in Puerto Rico winter nursery. Since F_2 generation, populations were advanced to F_4 generation using single seed descent method to minimize the amount of genetic sampling as described by Fehr (1987). F_3 and F_4 populations were spaced planted at the Virginia Tech farm in Blacksburg (2016), VA and winter nursery in Puerto Rico (2016), respectively. F_4 plants were harvested individually to derive $F_{4:5}$ lines and each line were kept as a separate entry for replicated trials. $F_{4:5}$ lines were grown at the Virginia Tech farm in Blacksburg (2017), VA. Flower and pubescent color were taken to control population's purity and avoid contaminations. A total of 240 $F_{4:5}$ individuals from the cross of V11-0883xV12-1626 were used as a mapping population and 166 F_3 individuals obtained from the cross of V13-1687xV12-1885 were also analyzed; which are referred in this study as mapping and validation population, respectively.

Experimental field procedures

Generations F_1, F_2 and F_4, were grown in Puerto Rico under winter nursery field conditions. F_3 plants were spaced planted (0.15 m) in single rows of 1.83 m in Blacksburg, VA; and $F_{4:5}$ lines were grown in a complete randomized design (RCD), in individual rows of 1.83 m

long spaced 0.76 m with two replications at Blacksburg VA. Fertilizer was applied before planting according to soil test results, pre-herbicide Dual Magnum, 1.66 pints per acre, was applied before planting to reduce weed pressure. No irrigation or insecticides were applied. When 95% of the plants in a row reached R8 maturity and dried for 5 to 10 days with less than 15% moisture, they were harvested.

Seed coat deficiency phenotyping

Two hundred seed samples from each F_3 plant and a hundred seed sample from each replication of $F_{4:5}$ lines were used for phenotyping of the mapping population; and two hundred seed samples from each F_3 plants were used for phenotyping of the validation population. To quantify the seed coat deficiency (SCD) we used a modified method from protocols described previously (Cook and Rainey, 2010, Rodda et al., 1973) for easier observation. The hundred unbroken seed samples were soaked in 1% commercial bleach solution for ten minutes. After soaking, seeds were immediately evaluated and cracked seed or severely blistered around the hilum or detached seed coat from the hull, were counted as deficient seeds.

DNA isolation and genotyping

A bulk sample of young leaves from each F_3 plant and a bulk sample of young leaves from 4-5 plants of $F_{4:5}$ lines, were collected at the V5 growth stage and stored at -80 °C from the mapping and validation population, respectively. Then, leaves were freeze dried using a FreeZone 6 Liter Console Freeze Dry System (-56°C and -0.220 mbar) and about 200 mg of leaf tissue were placed in a 2.5 mL tube. Tissue was ground to fine powder in liquid N_2 using glass stirring rods. Total genomic DNA was isolated from leaf tissue by a modified protocol from the CTAB method of Saghai-Maroof et al. (1984). Freeze-dried tissue (0.20 g, dry weight) was mixed with 750 µL of extraction buffer (0.1 M Tris, pH 8.0, 1.4 M NaCl, 20 mM EDTA, 1% β-

mercaptoethanol), and incubated at 65°C for 60 min with occasional mixing by gentle swirling. Phenol/chloroform/Isoamyl alcohol was added to each tube (750 µL) and the solution was mixed by inversion several times, and then placed on the rotating shaker for 45 minutes followed by a centrifugation step at 13,500 rpm for 10 minutes at room temperature. The upper phase (600 µL) was transferred into a new 1.5 mL tube and chloroform/octanol, 24:1 (v/v) (600 µL), was added to each tube. Samples were placed on the rotating shaker for 30 minutes and centrifuged at 13,500 rpm for 5 minutes at room temperature. Then, the upper phase (600 µL) was transferred into a new 1.5 mL tube and cold isopropanol was added (750 µL); samples were left overnight at 4 °C. Later, samples were centrifuged at 13,500 rpm for 10 minutes at 4 °C, where pellet remains at the bottom of the tube. Pellet was washed twice using 75% cold ethanol followed by centrifugation at 13,500 rpm for 10 minutes. Genomic DNA was suspended in 100 mL TE buffer (VWR, pH 8.0) and 3 µL of RNAase (10 mg/ml) (Qiagen). Fifty microliters of DNA of each genotype were sent to the USDA–ARS Beltsville Agricultural Research Center laboratory, Beltsville, MD, where SNP analysis was run. The parental lines and F_3 and $F_{4:5}$ Populations were genotyped using the Illumina 6000 SNP INF-BARCSoySNP6K_v2 Beadchip. Single-nucleotide polymorphism genotyping was conducted at the USDA–ARS Soybean Genomics and Improvement Laboratory, Beltsville, MD, on the Illumina platform following the Infinium HD Assay Ultra Protocol (Illumina, Inc.). Single-nucleotide polymorphism allele calling was done using the GenomeStudio Module v2.0.3 (Illumina, Inc.). Where, low seed coat deficiency parents were scored as A and high seed coat deficiency parents were scored as B.

Linkage map construction and quantitative trait loci analysis

Raw SNP data was manually inspected in Genome Studio v2.0.3 (Illumina, Inc.) and SNP data with no call, with severe segregation distortion, minor allele frequency-MAF<10%,

missing data (<5%) and monomorphic SNPs between parents were discarded. Linkage maps were constructed by Joinmap 4.0 (Van Ooijen, 2006) using a regression approach with a minimum logarithm of odds (LOD) threshold of 3 for linkage group construction. Recombination frequencies were converted to centimorgan (cM) using Kosambi's mapping function (Kosambi, 1943). QTL for seed coat deficiency were analyzed separately in each year for the mapping population. QTL analysis was performed by single marker analysis, interval mapping and composite interval mapping implemented in Windows Cartographer (WinQTLCart version 2.5) (Wang et al., 2007), R/QTL (Broman and Sen, 2009) and ICI mapping v 4.1 (Wang et al., 2016). For SMA, p<0.001 was used as a threshold for significant markers. In the CIM and SIM, the empirical significance threshold was determined by 1000-time permutation with a walk speed of 1 cM and significance level of 0.05. MapChart (Voorrips, 2002) was used to create the LOD plots based on JoinMap 3.0, R/QTL and ICI mapping v 4.1 data.

Statistical Data Analysis

Phenotypic data for seed coat deficiency, collected in 2016 and 2017, were analyzed using JMP statistical version 11.0 (SAS Institute). Normality assumption was assessed by Shapiro-Wilk test and normal probability plots. SCD trait was analyzed by one-way analysis of variance (ANOVA) for the 2016 and 2017 data individually, and by two-way ANOVA for the combined data. Histograms of SCD distributions were elaborated in R software. Variance-component heritability estimates were calculated by Analyses of variance using R software. Broad sense heritability of seed coat deficiency was estimated using the equation: $H^2 = s^2{}_g / [\, s^2{}_g + (s^2{}_{ge}/e) + (s^2/re)]$, where H^2 is heritability, $s^2{}_g$ is genotypic variance, $s^2{}_{ge}/e$ is genotype x environment interaction variance, s^2 is error variance, r is the number of replications, and e is the number of environments (Nyquist and Baker, 1991).

Results and Discussions

Seed coat deficiency

There was significant variation in seed coat deficiency among lines, years and the interaction of genotype and year. In the mapping population SCD of F_3 plants and $F_{4:5}$ lines ranged from 0 to 48% with a mean of 14.8% and from 3 to 94% with a mean of 43.3%, respectively. The two year average ranged from 6.7 to 61.7% with a mean of 29.09% (Table 13). Parental lines, V12-1626 and V11-0883, had SCD of 0.5 and 18.5%, and 40 and 68% in 2016 and 2017, respectively (Table 13). In the F_3 validation population there was also significant variation in SCD among plants, ranging from 2 to 51% with a mean of 17.6%. The parents, V12-1885 and V13-1687, had an average of 6.5 and 29%, respectively (Table 13). Thus, large phenotypic difference were observed among individuals in the populations, which was in consistency of previous studies with SCD values ranging from 23 to 70% (Cook and Rainey, 2010) and 4 to 100% (Yasui et al., 2017). Significant difference among years was also identified for the mapping population (Table 13). This difference was likely produced by the specific conditions of the season because seed quality is largely dependent upon environmental conditions. Several factors as chilling temperatures during flowering, harvest injuries and low moisture content can induce and favor seed coat cracking (Pollock, 1969, Ashworth and Oberdorf, 1980, Liu and Pappelis, 1971, Koizumi et al., 2008, Sunada and Ito, 1982). Results suggested that seed coat deficiency on soybeans had complex environmental interactions.

Progenies from the same populations usually present a continuous distribution of quantitative traits, because they are controlled by several genes and are strongly influenced by the environment (Kole, 2013). In this study, SCD showed continuous variation in both of the mapping and independent population, suggesting that seed coat deficiency is a quantitative trait

(Figs. 10 and 11). The mean value of each genotype fell between or outside the values of the two parents; and there were transgressive segregation in F_3 plants and $F_{4:5}$ lines (Figs. 10 and 11). Absolute values of skewness and kurtosis for SCD in each population and generation were less than 1, which indicated that SCD was normally distributed and was suitable for QTL mapping (Table 13).

The broad sense heritability (H^2) of SCD, 0.40, was estimated using variance components for SCD in 2-year average data of the mapping population. It was lower than heritability previously reported for seed coat cracking after soaking for 18 h (0.67) (Yasui et al., 2017), and after harvest (0.89 to 0.97) (Ha et al., 2012); and other food grade seed traits such as seed size, protein, sucrose, raffinosse and stachyose concentration where heritability estimates ranged from 0.45 to 0.86 (Jaureguy et al., 2011). Heritability estimates are indicators of potential success of breeders in improving genetic gain of specific traits (Jaureguy et al., 2011). According to Robinson et al. (1949), heritability is categorized as low (0-30%), medium (31-60%) and high (>60%). Therefore, moderate broad sense heritability and large SCD difference among years confirmed that complex genetic and environment interactions affected this trait. Thus, selection of seed coat deficiency through molecular markers would play an important role for reducing seed coat deficiency in natto soybeans together with proper seed coat deficiency determination.

Linkage map

A total of 1,318 and 1,605 polymorphic SNPs between parents out of the 6,000 evaluated SNPs in the mapping and validation populations were selected for further genetic linkage map, respectively; after filtering for monomorphic markers, SNPs with severe segregation distortion, SNPs with minor allele frequency-MAF<10% and missing data <5% (Table 14-15). The linkage map of the mapping population consisted of 20 chromosomes, which spanned 1,826 cM and

were defined by 1,258 SNP markers. The constructed map had average marker intervals ranging from 0.8 to 2.47 cM with an overall mean of 1.6 cM (Table 14). While the linkage map in the validation population consisted of 20 chromosomes and spanned 1,970 cM with a total of 1,505 SNPs mapped. The average marker intervals ranged from 0.90 to 1.97 cM with an overall mean of 1.4 cM (Table 15). The levels of polymorphism detected in the mapping and validation populations using BARCsoySNP6K were 22 and 27%, respectively. The polymorphism rate and coverage were lower than those (48% polymorphism, 2,346 cM map length and 1.1cM marker interval) from a RIL population that differ for many morphological traits (Lee et al., 2017, Lee et al., 2015). A shared parent (Essex) in the pedigree of the parents of the mapping and validation populations may contribute to the low map coverage and genetic diversity among parents (Fig. 12).

QTL analysis

Quantitative trait loci for seed coat deficiency were analyzed separately in each year for the mapping population and the QTL consistent across years were considered putative QTL for seed coat deficiency. In the CIM and IM, empirical threshold was computed as LOD of 3.28 after 1000 permutation in 2016 and 2017. Based on ICI mapping results, two and three QTL regions were significantly associated with seed coat deficiency in the mapping population in 2016 and 2017, respectively. QTL were located on chromosomes 15 (LG E), 20 (LG I) and 10 (LG O) explaining up to 24.29% of the phenotypic trait variation (Table 16). QTL analysis was congruent among IciMapping, R/QTL and QTL Cartographer. QTL detected in 2016 and 2017 explained 21.39 and 36.56% of the total variation of the trait, respectively. Variation that remained unexplained may be attributed to QTL not detected in this study because of incomplete genome coverage, environment or week linkage relationships among marker loci and QTL

(Primomo et al., 2005) and other minor QTL. However, QTL detected on chromosome 15 (LG E) and 10 (LG O) were not stable across years. Unstable QTL are not desirable for MAS because they can give unreliable data; which was why, it is recommended to conduct QTL mapping studies in as many environments and populations as feasible (Panthee et al., 2005, Fasoula et al., 2004).

The QTL region detected on chromosome 20 (LG I) was stable across years and significantly associated with SCD, accounting for 11.3% and 24.3% of the total variation of the trait in 2016 and 2017, respectively. We tentatively named this QTL for seed coat deficiency as *qSCD20_1* (Table 16). In 2016, *qSCD20_1* was positioned at 48 cM between flanking markers Gm20_35625615_C_T and Gm20_36002148_T_C with 8.43 of LOD value. In 2017, *qSCD20_1* was positioned at 45 cM between flanking markers Gm20_34881595_C_T and Gm20_35625615_C_T with 15.15 of LOD value (Table 16). These two regions being in close proximity and sharing one flanking marker were considered the same QTL (Fig. 13). Therefore, qSCD20_*1* was positioned in 44.5-50.5 cM of the linkage map of chromosome 20 flanked by markers Gm20_34881595_C_T and Gm20_36002148_T_C (Fig. 13). The additive effect (Add) of *qSCD20_1* was negative (-3.05 in 2016 and -9.26 in2017), indicating that the alleles from the low SCD parent (V12-1626) conferred reduced SCD at the QTL (Table 16).

The SNPs Gm20_34881595_C_T and Gm20_36002148_T_C were located at 53.7 and 58.9 cM, on Chromosome 20, respectively, in the most recently developed high density linkage map for soybeans (Song et al., 2016). Based on physical position information of the SNPs, the physical distance between these markers is 1.08 Mb; and the corresponding genome positions of Gm20_34881595_C_T and Gm20_36002148_T_C in the reference soybean genome (Wm82.a2.v1 assembly) are 36.02 Mb-37.09 Mb that encompass a total of 21 gene candidates

129

(Song et al., 2016) (Table 16). To date, there are seven published genes on this region that are associated with cell signaling responses, cellular metabolism and seed development (NCBI, 2017). In addition, several QTL for seed composition (protein, oil and isoflavones), pod maturity, seed weight and seed hardness have also been identified in this region of chromosome 20 of soybeans (Soybase https://soybase.org). Therefore, the chromosome 20 of the soybean genome harbor several genes associated with seed traits in soybeans and seed development that may be associated with the seed coat strength.

The F3 validation population was used to confirm the detected QTL. Empirical threshold for ICM and IM was computed as LOD 4.13 after 1000 permutation. Based on IciMapping results, two QTL positioned at chromosome 10 (LG O) and 20 (LG I) were significantly associated with SCD, accounting for 11.32% and 22.6% of phenotypic trait variance, respectively (Table 17). Congruent results were obtained by IciMapping, R/QTL and QTL cartographer. The QTL detected on chromosome 20 (LG I) was tentatively named as *qSCD20_2* and was positioned at 49 cM between flanking markers Gm20_33482879_C_T and Gm20_33676755_C_T with 8.56 of LOD value (Table 17, Fig. 14). The additive effect (Add) of *qSCD20_2* was also negative -3.78, indicating that the low SCD parent (V12-1885) is the carrier of the favorable allele for low seed coat deficiency (Table 17). The SNP markers Gm20_33482879_C_T and Gm20_33676755_C_T are positioned at 44 and 45 cM respectively, in the most recently developed high density linkage map for soybeans (Song et al., 2016). Based on SNPs physical position information, the physical distance between SNP flanking markers is 0.19 Mb (193,878 bp). The corresponding genome positions of Gm20_33482879_C_T and Gm20_33676755_C_T in the reference soybean genome (Wm82.a2.v1 assembly) are 34.62 and 34.82 Mb (Song et al., 2016) (Table 17). Thus, *qSCD20_2* was located 1.2 Mb apart from

qSCD20_1 identified on the original mapping population. Similar to *qSCD20_1*, QTL associated with seed composition, seed weight and seed size have also been identified in the same region that *qSCD20_2* was detected.

Student t-test was performed to compare trait mean values for individuals with and without *qSCD20_1* and *qSCD20_*2. The SCD mean of individuals carrying *qSCD20_1* (22.7%) and *qSCD20_2* (12.6%) were significant lower than SCD mean obtained from plants without those QTL that were 32.94 and 19.73%, respectively (Fig. 15). These results confirmed that *qSCD20_1* and *qSCD20_2* conferred lower seed coat cracking of soybeans during water absorption. Due to the proximity of *qSCD20_1* and *qSCD20_2* and the shared parent among the two low SCD parental lines we suggested that *qSCD20_1* and *qSCD20_2* may be the same QTL or that they may be part of a rich QTL region for seed coat deficiency on LG I in soybeans; and that favorable alleles are probably inherited from the common parent V03-0986. There are several explanations for the inability to detect *qSCD20_1* on the exactly same position in the validation population such as differences in the linkage maps, generation, population size and parental genetic diversity (Van and McHale, 2017).

Molecular markers flanking QTL regions on LG I, that were significantly associated (p-value <0.001) to seed coat deficiency across years and populations were identified by SMA. In chromosome 20 (LG I), a total of 29 SNPs in the validation population, and 9 and 14 SNPs in 2016 and 2017 from the mapping population showed to be significantly (p<0.001) associated with seed coat deficiency. From these markers, 8 SNPs showed to be significant across years in the mapping population. Where, two of these SNPs also showed to be significantly associated with SCD in the validation population (Gm20_34942502 and Gm20_34881595) (Table 18). All SNPs were tightly linked or nearby *qSCD20_1* and *qSCD20_2* position (Fig. 14). As a result,

QTL and SNPs reported in this study were potential candidates for MAS of superior natto soybean cultivars.

A recent study identified QTL associated with seed coat cracking after soaking (qSCAS1, qSCAS2 and qSCAS3) and after cooking (qSCAC1 and qSCAC2) located on chromosomes 4 (LG: C1), 6 (LG: C2) and 8 (LG: A2),(Yasui et al., 2017). However, the way to score the phenotype, soaking for 18 h was different from our study of soaking for only 10 minutes. Thus, different scoring methods, population size and generation, marker set and environments explained different QTL associated with seed coat strength identified in two studies. Although no other QTL reported for seed coat deficiency, several QTL have been detected for seed coat cracking after harvest. A QTL (qSCC20) associated with seed coat cracking after harvest was also identified in LG I. However, qSCC20 was not stable across environments because it was just detected in two out of the four locations evaluated (Ha et al., 2012). Other QTL associated with seed coat cracking after harvest have been detected on LG C1, D1b, M, K, N and M (Oyoo et al., 2010a, Oyoo et al., 2010b, Ha et al., 2012). Therefore, *qSCD20_1/qSCD2_2* is a novel QTL for seed coat deficiency.

Conclusions

Soybean populations showed large phenotypic differences on seed coat deficiency, and genotypes contributed significantly to the trait variation, which indicated that selection of SCD is feasible. Identification of QTL and the verification through a parallel population has allowed us to detect a stable QTL region on chromosome 20 (LG I) across years and populations. Results suggested that genetic improvement of SCD could be possible through MAS using 8 potential candidate SNPs. However, the populations were only evaluated across two years and *qSCD20_1* and *qSCD20_2* were not detected in the same position. Thus, further confirmation experiments are needed to validate *qSCD20_1* and *qSCD20_2*, which is essential for further utilization in MAS for natto breeding. Further studies should focus on verifying QTL across environments and in different genetic backgrounds, transferring favorable alleles into high yielding soybean varieties and identification of potential additional QTL and epistatic interactions.

qSCD20_1 was detected in the same region of genes associated with cell signaling responses, cellular metabolism and seed development; and QTL for seed composition, pod maturity, seed weight and seed hardness; thus, chromosome 20 of the soybean genome harbor several genes associated with seed traits in soybeans and seed development that may be associated with the seed coat strength.

References

Arumuganathan K., Earle E.D. (1991) Estimation of Nuclear DNA Content of Plants by Flow Cytometry.   Plant Mol. Biol. Rep. 9:229–241.

Ashworth E.N., Oberdorf R.L. (1980) Imbibitional Chilling Injury in Soybean Axes: Relationship to Stelar Lesions and Seasonal Environments.   Agron J 72:923-928.

Broman K.W., Sen S. (2009) A Guide to Qtl Mapping with R/Qtl. 1. Aufl. ed. Springer, Dordrecht.

Cober E.R., Fregeau-Reid J.A., Pietrzak L.N., McElroy A.R., Voldeng H.D. (1997) Genotype and Environmental Effects on Natto Soybean Quality Traits.   Crop Science   37:1151-1154.

Cook D.E., Rainey K.M. (2010) Seed Coat Deficiency, Trait Stability, and Other Soybean Seed Quality Traits for Natto Cultivar Development.   Crop Science   50:1244-1249.   DOI: 10.2135/cropsci2009.06.0300.

Copeland L.O., McDonald M.B. (2001) Principles of Seed Science and Technology SPRINGER SCIENCE+BUSINESS MEDIA, LLC, New York.

Cui Z., James A., Miyazaki S., Wilson R., Carter T. (2004) Breeding Specialty Soybeans for Traditional and New Soyfoods, in: L. K (Ed.), Soybeans as Functional Foods and Ingredients, AOCS Press, Urbana, IL. pp. 264-322.

Fasoula V.A., Harris D.K., Boerma H.R. (2004) Validation and Designation of Quantitative Trait Loci for Seed Protein, Seed Oil, and Seed Weight from Two Soybean Populations Crop Science   44:1218-1225.

Fehr W. (1987) Principles of Cultivar Development: Theory and Technique Macmillan, New York.

Geater C.W., Fehr W.R., L A.W. (2000) Association of Soybean Seed Traits with Physical Properties of Natto.   Crop Science   40:1529-1534.

Gibbs B.F., Zougman A., Masse R., Mulligan C. (2004) Production and Characterization of Bioactive Peptides from Soy Hydrolysate and Soy-Fermented Food.   Food Research International   37:123-131.   DOI: 10.1016/j.foodres.2003.09.010.

Gupta V.S., M. G., Y. K.N., P. R., N. S.M. (2005) Quantitative Trait Loci (Qtl) Mapping and Marker Assisted Selection (Mas) in Food Legumes.   Food Legumes for Nutritional Security and Sustainable Agriculture   1.

Ha B.-K., Kim H.-K., Kang S.-T. (2012) Mapping Qtls with Epistatic Effects and Qtl-by-Environment Interactions for Seed Coat Cracking in Soybeans.   Euphytica   186:933-942.   DOI: 10.1007/s10681-012-0719-8.

Hartman G.L., West E.D., Herman T.K. (2011) Crops That Feed the World 2. Soybean—Worldwide Production, Use, and Constraints Caused by Pathogens and Pests.   Food Security   3:5-17.   DOI: 10.1007/s12571-010-0108-x.

Hosoi T., Kiuchi K. (2003) Fermenting Cooked Soybeans with Bacillus Subtilis (Natto). , in: E. R. Farnworth (Ed.), Handbook of Fermented Functional Foods, CRC Press, USA. pp. 227-250.

Hu Y., Ge C., Yuan W., Zhu R., Zhang W., Du L., Xue J. (2010) Characterization of Fermented Black Soybean Natto Inoculated with Bacillus Natto During Fermentation.   J Sci Food Agric   90:1194-202.   DOI: 10.1002/jsfa.3947.

Ikeda Y., Iki M., Morita A., Kajita E., Kagamimori S., Kagawa Y., Yoneshimayy H. (2006) Intake of Fermented Soybeans, Natto, Is Associated with Reduced Bone Loss in

Postmenopausal Women: Japanese Population-Based Osteoporosis (Jpos) Study. J. Nutr.   136:1323–1328.

Jaureguy L.M., Chen P., Scaboo A.M. (2011) Heritability and Correlations among Food-Grade Traits in Soybean.   Plant Breeding   130:647-652.   DOI: 10.1111/j.1439-0523.2011.01887.x.

Kentucky U.o. (2013) Specialty Soybeans, in: U. C. E. Service (Ed.), http://www.uky.edu/Ag/NewCrops/introsheets/specialtysoy.pdf.

Koizumi M., Kikuchi K., Isobe S., Ishida N., Naito S., Kano H. (2008) Role of Seed Coat in Imbibing Soybean Seeds Observed by Micro-Magnetic Resonance Imaging.   Ann Bot 102:343-52.   DOI: 10.1093/aob/mcn095.

Kole C. (2013) Genomics and Breeding for Climate-Resilient Crops: Vol. 2,Target Traits. 2013;1; ed. Springer Verlag, Berlin, Heidelberg.

Kosambi D.D. (1943) The Estimation of Map Distances from Recombination Values.   Annals of Eugenics   12:172-175.   DOI: 10.1111/j.1469-1809.1943.tb02321.x.

Lee S., Freewalt K.R., McHale L.K., Song Q., Jun T.-H., Michel A.P., Dorrance A.E., Mian M.A.R. (2015) A High-Resolution Genetic Linkage Map of Soybean Based on 357 Recombinant Inbred Lines Genotyped with Barcsoysnp6k.   Molecular Breeding   35. DOI: 10.1007/s11032-015-0209-5.

Lee S., Jun T.-H., McHale L.K., Michel A.P., Dorrance A.E., Song Q., Mian M.A.R. (2017) Registration of Wyandot × Pi 567301b Soybean Recombinant Inbred Line Population. Journal of Plant Registrations   11:324-327.   DOI: 10.3198/jpr2016.09.0042crmp.

Liu k. (1997) Soybeans: Chemistry, Technology, and Utilization Chapman & Hall, New York.

Liu K.C., Pappelis J. (1971) Transverse and Crescent Cracks in Soybean Cotyledons Associated with Imbibition. Trans Ill Acad Sci 64:43-45

Mayta J., Chen P., Popp M.P., Dong D., Wu C.J., Zhang B., Smith S.F., Scaboo A.M. (2014) Break-Even Profitability for Food-Grade Specialty Soybeans. Agricultural Science 2:1-11. DOI: 10.12735/as.v2i2p01.

NCBI. (2017).

Nyquist W.E., Baker R.J. (1991) Estimation of Heritability and Prediction of Selection Response in Plant Populations. Critical Reviews in Plant Sciences 10:235-322. DOI: 10.1080/07352689109382313.

Oyoo M.E., Benitez E.R., Matsumura H., Takahashi R. (2010a) Qtl Analysis of Seed Coat Cracking in Soybean. Crop Science 50. DOI: 10.2135/cropsci2009.08.0440.

Oyoo M.E., Githiri S.M., Benitez E.R., Takahashi R. (2010b) Qtl Analysis of Net-Like Cracking in Soybean Seed Coats. Breeding Science 60:28-33.

Panthee D.R., Pantalone V.R., West D.R., Saxton A.M., Sams C.E. (2005) Quantitative Trait Loci for Seed Protein and Oil Concentration and Seed Size in Soybean. Crop Science 45:2015-2022.

Pollock B.M. (1969) Imbibition Temperature Sensitivity of Lima Bean Seeds Controlled Y Initial Seed Moisture. Plant Physiol 44.

Primomo V.S., Poysa V., Ablett G.R., Jackson C.-J., Gijzen M., Rajcan I. (2005) Mapping Qtl for Individual and Total Isoflavone Coentent in Soybean Seeds. Crop Science 45:2454-2463.

Qutob D., Ma F., Peterson C.A., Bernards M.A., Gijzen M. (2008) Structural and Permeability Properties of the Soybean Seed Coat.   Botany   86:219-227.

Robinson H.F., Comstock R.E., Harvey P.H. (1949) Estimation of Heritability and the Degree of Dominance in Corn.   Agronomy Journal   41:353-359.

Rodda E.D., Steinberg M.P., Wei L.S. (1973) Soybean Damage Detection and Evaluation for Food Use.   Transactions of the ASAE   16:365.   DOI: https://doi.org/10.13031/2013.37521.

Saghai-Maroof M.A., Soliman K.M., Jorgensen R.A., Allard R.W. (1984) Ribosomal DNA Spacer-Length Polymorphisms in Barley: Mendelian Inheritance, Chromosomal Location, and Population Dynamics.   Proc. Natl. Acad. Sci. USA   81:8014-8018.

Sanjukta S., Rai A.K. (2016) Production of Bioactive Peptides During Soybean Fermentation and Their Potential Health Benefits.   Trends in Food Science & Technology   50:1-10. DOI: http://dx.doi.org/10.1016/j.tifs.2016.01.010.

Schmutz J., Cannon S.B., Schlueter J., Ma J., Mitros T., Nelson W., Hyten D.L., Song Q., Thelen J.J., Cheng J., Xu D., Hellsten U., May G.D., Yu Y., Sakurai T., Umezawa T., Bhattacharyya M.K., Sandhu D., Valliyodan B., Lindquist E., Peto M., Grant D., Shu S., Goodstein D., Barry K., Futrell-Griggs M., Abernathy B., Du J., Tian Z., Zhu L., Gill N., Joshi T., Libault M., Sethuraman A., Zhang X.C., Shinozaki K., Nguyen H.T., Wing R.A., Cregan P., Specht J., Grimwood J., Rokhsar D., Stacey G., Shoemaker R.C., Jackson S.A. (2010) Genome Sequence of the Palaeopolyploid Soybean.   Nature 463:178-83.   DOI: 10.1038/nature08670.

Somekawa Y., Chiguchi M., Ishibashi T., Aso T. (2001) Soy Intake Related to Menopausal Symptoms, Serum Lipids, and Bone Mineral Density in Postmenopausal Japanese

Women. Obstetrics & Gynecology 97:109-115. DOI: http://dx.doi.org/10.1016/S0029-7844(00)01080-2.

Song Q., Jenkins J., Jia G., Hyten D.L., Pantalone V., Jackson S.A., Schmutz J., Cregan P.B. (2016) Construction of High Resolution Genetic Linkage Maps to Improve the Soybean Genome Sequence Assembly Glyma1.01. BMC Genomics 17:33. DOI: 10.1186/s12864-015-2344-0.

Sunada K., Ito T. (1982) Soybean Grain Quality as Affected by Low Temperature Treatments in Plants (Color of Hilum, Seed Coat Cracking). Rep Hokkaido Branch Crop Sci Soc Jpn and Hokkaido Branch Jpn Soc Breed 22:34.

Taira H. (1990) Quality of Soybeans for Processed Food. JARQ 24:224-230.

Tamang J.P., Shin D.H., Jung S.J., Chae S.W. (2016) Functional Properties of Microorganisms in Fermented Foods. Front Microbiol 7:578. DOI: 10.3389/fmicb.2016.00578.

Van K., McHale L.K. (2017) Meta-Analyses of Qtls Associated with Protein and Oil Contents and Compositions in Soybean [Glycine Max (L.) Merr.] Seed. Int J Mol Sci 18. DOI: 10.3390/ijms18061180.

Van Ooijen J.W. (2006) Joinmap 4 ®: Software for the Calculation of Genetic Linkage Maps in Experimental Populations Kyazma B. V, Wageningen, The Netherlands.

VanUtrecht D., Bern C.J., Rukunudin I.H. (2000) Soybean Mechanical Damage Detection. Applied Engineering in Agriculture 16:137-141.

Voorrips R.E. (2002) Mapchart: Software for the Graphical Presentation of Linkage Maps and Qtls. Journal of Heredity 93:77-78. DOI: 10.1093/jhered/93.1.77.

Wang J., Li H., Zhang L., Meng L. (2016) Users' Manual of Qtl Icimapping Institute of Crop Science, CAAS, Beijing, and Crop Research Informatics Lab, Mexico.

Wang S., Basten C.J., Zeng Z.B. (2007) Windows Qtl Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Wei Q., Chang S.K.C. (2004) Characteristics of Fermented Natto Products as Affected by Soybean Cultivars. Journal of food processing preservation 28:251-273.

Wei Q., Wolf-Hall C., Chang K.C. (2001) Natto Characteristics as Affected by Steaming Time, Bacillus Strain, and Fermentation Time. Journal of food Science 86:167-173.

Weng T.M., Chen M.T. (2012) Effect of Drying Methods on Γ-Pga, Isoflavone Contents and Ace Inhibitory Activity of Natto (a Fermented Soybean Food). Journal of Food Processing and Preservation 36:483-488. DOI: 10.1111/j.1745-4549.2011.00591.x.

Yasui T., Toda K., Yamada T., Yumoto S., Takahashi K., Takahashi M., Hajika M. (2017) Seed Coat Cracking of Soybean (Glycine Max [L.] Merr.) after Soaking and Cooking. Cereal Chemistry Journal 94:717-722. DOI: 10.1094/cchem-12-16-0284-r.

Zhang B., Chen P., Chen C.Y., Wang D., Shi A., Hou A., Ishibashi T. (2008) Quantitative Trait Loci Mapping of Seed Hardness in Soybean. Crop Science 48:1341. DOI: 10.2135/cropsci2007.10.0544.

Zhang B., Chen P., Florez-Palacios S.L., Shi A., Hou A., Ishibashi T. (2010) Seed Quality Attributes of Food-Grade Soybeans from the U.S. And Asia. Euphytica 173:387-396. DOI: 10.1007/s10681-010-0126-y.

Zhang B., Jaureguy L., Florez-Palacios L., Chen P. (2011) Food-Grade Soybean: Consumption, Nutrition, and Health, in: E. P. a. I. Golubev (Ed.), Beans: Consumption, Nutrition, and Health, Nova Science Publishers, Inc., Hauppauge, NY. pp. 18.

Tables and figures

Table 12: Pedigree information and characteristics of small-seeded parental genotypes

Parent[a]	Name	Pedigree	FC[b]	PC[b]	SS[b] (g/100 seed)	SCD[b] (%)	Note
P1	V12-1626	V03-0986 x G03-3113	P	G	10.6	9.0	Low SCD
P2	V11-0883	V00-3488 X R04-198	P	G	11.3	63.3	High SCD
P3	V12-1885	V03-0986 x V01-2245	P	G	10.8	10.0	Low SCD
P4	V13-1687	MFS-541 x R04-198	P	G	11.7	56.0	High SCD

[a] Parent 1 (P1), Parent 2 (P2), Parent 3 (P3) and Parent 4 (P4)
[b] FC is flower color, P is purpura; PC is pubescent color, G is gray, SS is seed size and SCD is seed coat deficiency

Table 13: Descriptive statistics for seed coat deficiency (SCD) in F3 Plants and F4:5 lines of mapping and validation population.

Population	Year	Generation	Max[a]	Min[b]	Mean	SE[c]	Skewness	Kurtosis	P1[d]	P2[d]	P3[d]	P4[d]
V11-0883 x V12-1626	2016*	F3 Plants	48	0	14.8	0.39	0.88	0.52	0.5	40	-	-
V11-0883 x V12-1626	2017*	F4:5 lines	94	3	43.29	1.02	0.18	-0.83	18.5	68	-	-
V11-0883 x V12-1626	2016/2017**		61.5	6.7	29.09	0.82	0.28	-0.58	9.5	54		
V13-1687 xV12-1885	2016*	F3 plants	51	2	17.57	0.46	0.47	0.09	-	-	6.5	29

[a] Max stands for maximum; [b] Min stands for minimum; [c] SE stands for standard error
[d] P1 is V12-1626, P2 is V11-0883, P3 is V12-1885 and P4 is V13-1687.
*Significant difference among genotypes at the P<.0001
**Significant difference among genotypes (G), years (Y) and G x Y interaction at the p <.0001

Table 14: Summary of single nucleotide polymorphism (SNP) markers used in genotyping F4-derived F5 mapping population from V11-0883xV12-1626.

Chr.[a]	LG[b]	Length (cM)[c]	No. of total SNPs tested	No. of polymorphic SNPs	No. of mapped SNPs	Average distance between SNP (cM)
1	D1a	91.42	263	69	68	1.34
2	D1b	82.20	328	45	42	1.96
3	N	128.61	280	79	78	1.65
4	C1	101.17	266	49	49	2.06
5	A1	98.28	294	81	81	1.21
6	C2	79.48	313	44	43	1.85
7	M	87.61	319	52	52	1.68
8	A2	108.52	380	54	54	2.01
9	K	79.09	261	49	32	2.47
10	O	102.48	317	58	57	1.80
11	B1	98.22	279	55	55	1.79
12	H	47.71	259	33	20	2.39
13	F	97.68	381	105	104	0.94
14	B2	95.84	275	110	109	0.88
15	E	78.11	305	68	68	1.15
16	J	44.87	328	31	30	1.50
17	D2	87.83	265	74	63	1.39
18	G	86.79	380	117	108	0.80
19	L	118.03	326	85	85	1.39
20	I	112.58	271	60	60	1.88

[a] Chromosome
[b] Linkage group
[c] Chromosome length in centimorgan

Table 15: Summary of single nucleotide polymorphism (SNP) markers used in genotyping F3 plants from V13-1687 x V12-1885

Chr.[a]	LG[b]	Length (cM)[c]	No. of total SNPs tested	No. of polymorphic SNPs	No. of mapped SNPs	Average distance between SNP
1	D1a	113.944	263	168	68	1.68
2	D1b	54.049	328	42	42	1.29
3	N	98.528	280	90	90	1.09
4	C1	96.066	266	79	79	1.22
5	A1	126.197	294	68	68	1.86
6	C2	99.897	313	90	90	1.11
7	M	72.975	319	37	37	1.97
8	A2	146.302	380	87	87	1.68
9	K	40.203	261	35	35	1.15
10	O	120.049	317	104	104	1.15
11	B1	90.357	279	51	51	1.77
12	H	85.238	259	86	86	0.99
13	F	155.906	381	120	120	1.30
14	B2	106.653	275	99	99	1.08
15	E	84.239	305	81	81	1.04
16	J	95.727	328	58	58	1.65
17	D2	124.444	265	62	62	2.01
18	G	86.238	380	96	96	0.90
19	L	105.517	326	111	111	0.95
20	I	67.562	271	41	41	1.65

[a] Chromosome
[b] Linkage group
[c] Chromosome length in centimorgan

Table 16: Quantitative trait loci for seed coat deficiency (SCD) in the mapping population (V11-0883xV12-1626)

Year	QTL name	G[a]	P[a] (cM)	LG[a]	Flanking markers	GI (cM)[a]	Glyma1.01 physical interval (bp)[b]	Wm82.a2.v1 Physical interval (bp)[c]	DBM (bp)[a]	LOD[a]	PVE (%)[a]	Add[a]
2016	qSCD15_1	F$_3$	12	15	Gm15_511387_T_C-Gm15_4751337_G_A	8.0-13.0	511387-4751337	511864-4770814	4239950	6.05	10.06	-1.94
2016	qSCD20_1	F$_3$	48	20	Gm20_35625615_C_T-Gm20_36002148_T_C	46.5-50.5	35625615-36002148	36720824-37097315	376533	8.43	11.33	-3.05
2017	qSCD10_1	F$_{4:5}$	25	10	Gm10_11518881_G_A-Gm10_26188429_A_G	22.5-25.5	11518881-26188429	11710604-16707334	14669548	3.4	5.1	-4.27
2017	qSCD15_2	F$_{4:5}$	23	15	Gm15_5312718_C_T-Gm15_6066709_C_A	22.5-23.5	5312718-6066709	5331364-6085794	753991	4.98	7.17	-5.04
2017	qSCD20_1	F$_{4:5}$	45	20	Gm20_34881595_C_T-Gm20_35625615_C_T	44.5-46.5	34881595-35625615	36021058-36720824	754020	15.15	24.29	-9.26

[a] G= generation, P= peak, LG= linkage group, GI= genetic interval, DBM= distance between markers, LOD= logarithm of odds, PVE= percentage of variance explained, Add= additivity

[b] Physical interval according to genome positions at Glyma 1.01.

[c] Physical interval according to genome positions at Wm82.a2.v

Table 17: Quantitative trait loci for seed coat deficiency (SCD) in the validation population (V13-1687 xV12-1885) on IciMapping

QTL name	G[a]	P (cM)[a]	LG[a]	Interval	GI (cM)[a]	Glyma1.01 Physical interval (bp)[b]	Wm82.a2.v1 Physical interval (bp)[c]	DBM (bp)[a]	LOD[a]	PVE (%)[a]	Add[a]
qSCD10_2	F_3	108	10	Gm10_44445941_A_G-Gm10_44972284_T_C	106.5-108.5	44445941-44972284	45023591-45550230	526343	4.52	11.32	-2.66
qSCD20_2	F_3	49	20	Gm20_33676755_C_T-Gm20_33482879_C_T	48.5-49.5	33482879-33676755	34622286-34816162	193876	8.56	22.59	-3.78

[a] G= generation, P= peak, LG= linkage group, GI= genetic interval, DBM= distance between markers, LOD= logarithm of odds, PVE= percentage of variance explained, Add= additivity
[b] Physical interval according to genome positions at Glyma 1.01.
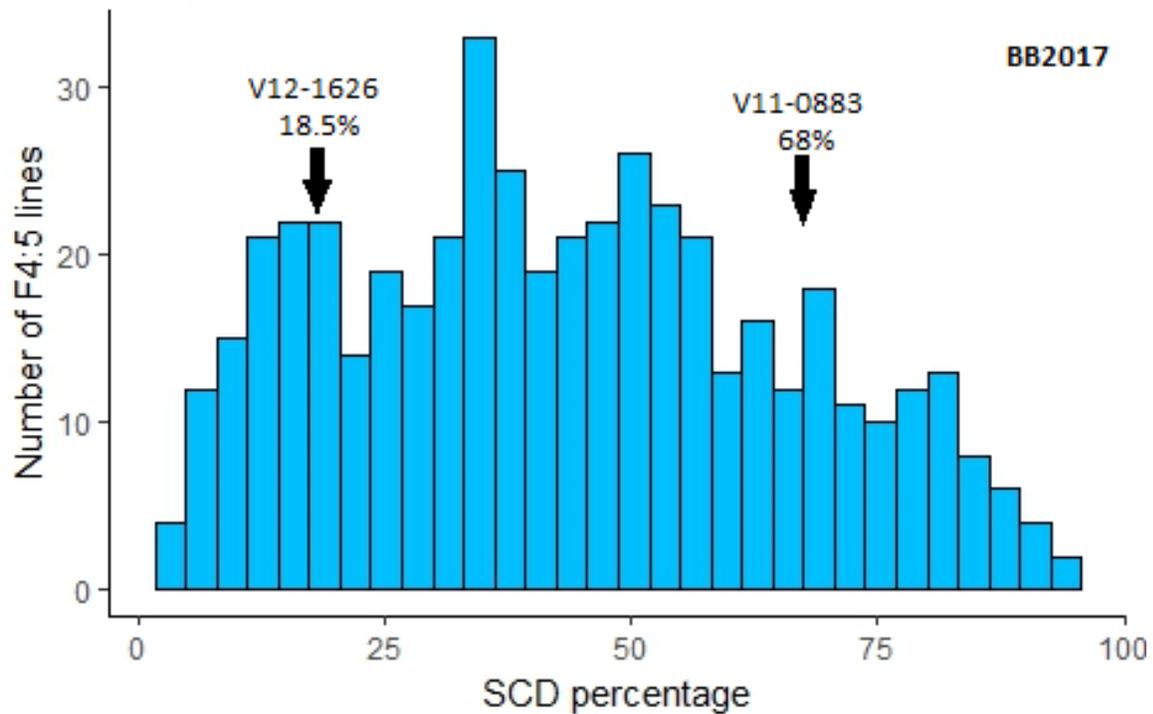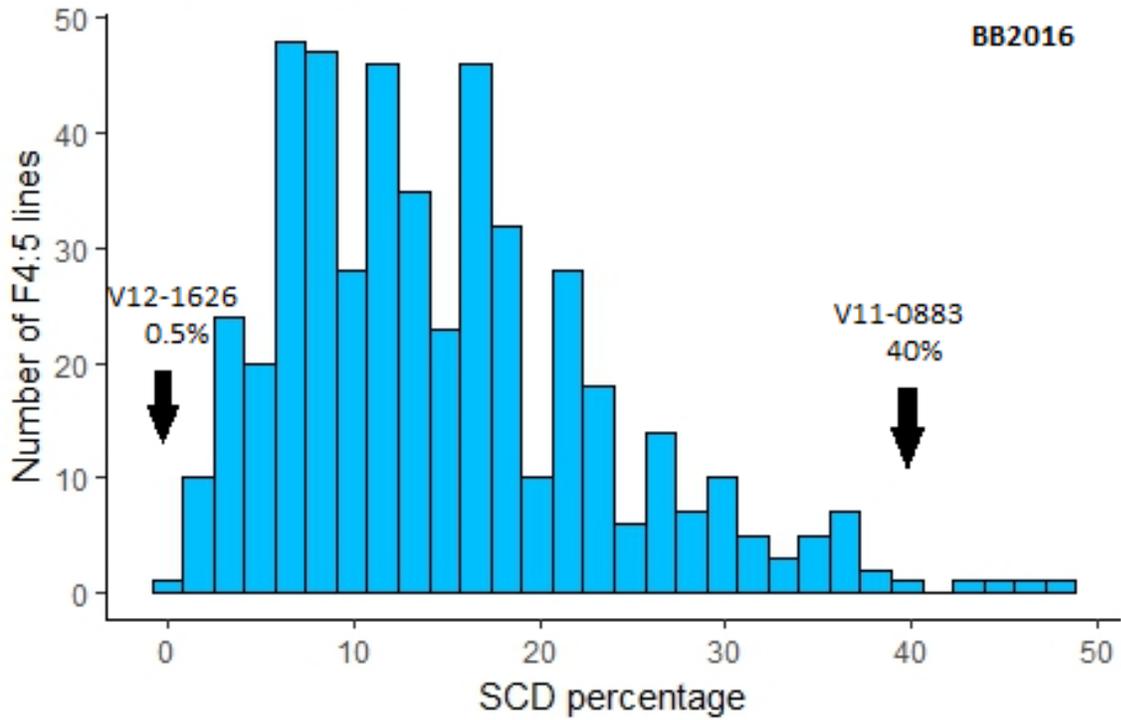[c] Physical interval according to genome positions at Wm82.a2.v

Table 18: SNP markers highly significant associated (p<0.001) with seed coat deficiency (SCD) in F3 plants and F4:5 lines from the mapping and validation populations.

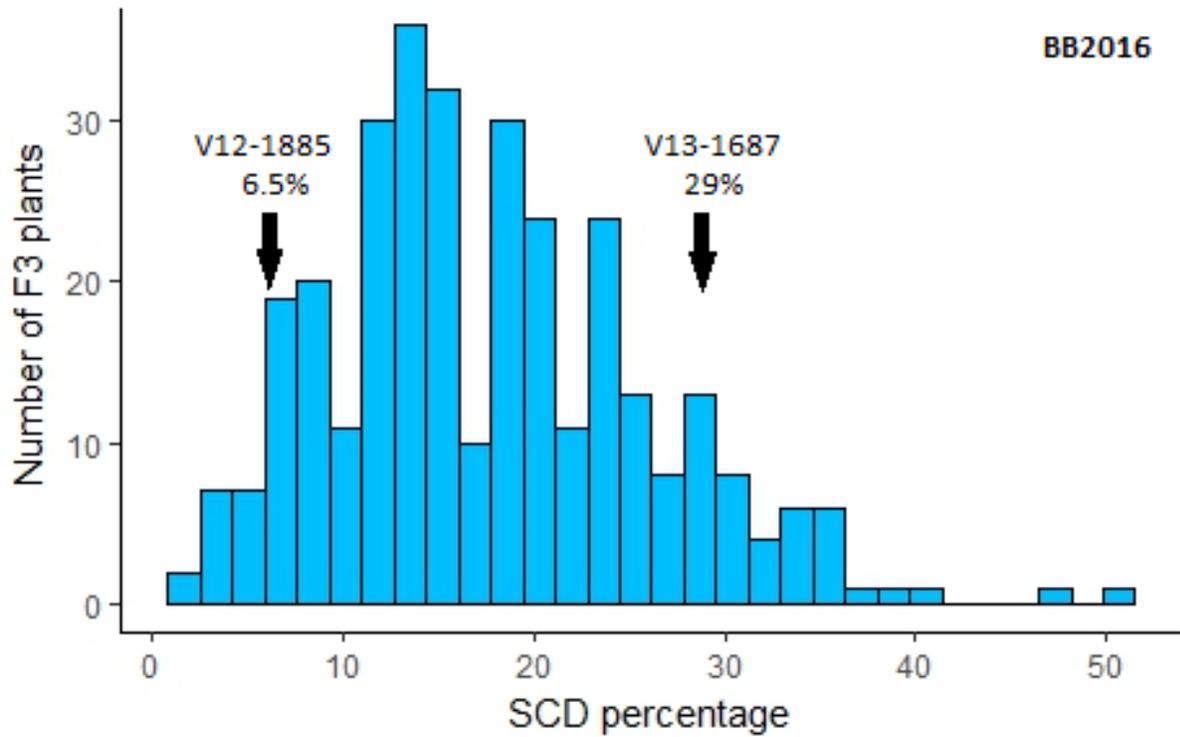| LG[a]	Marker ID	Position (cM)	Mapping population LOD[a]		Validation population F3 plants	
			F3 plants	F4:5 lines	Position (cM)	LOD
20	Gm20_34942502	44.51	5.48**	13.78***	53.75	8.94***
20	Gm20_34881595	44.72	5.84**	13.89***	53.78	8.94***
20	Gm20_35625615	47.91	7.25***	11.70***		
20	Gm20_36002148	51.28	6.27**	13.32***		
20	Gm20_36095037	52.02	6.44**	13.47***		
20	Gm20_36153048	52.80	6.34**	13.31***		
20	Gm20_36651429	55.36	5.91**	13.34***		
20	Gm20_36710448	55.80	6.41**	11.84***		

[a] LG is linkage group, LOD is = logarithm of odds
** SNPs significant associated with seed coat deficiency at a p-value<0.001
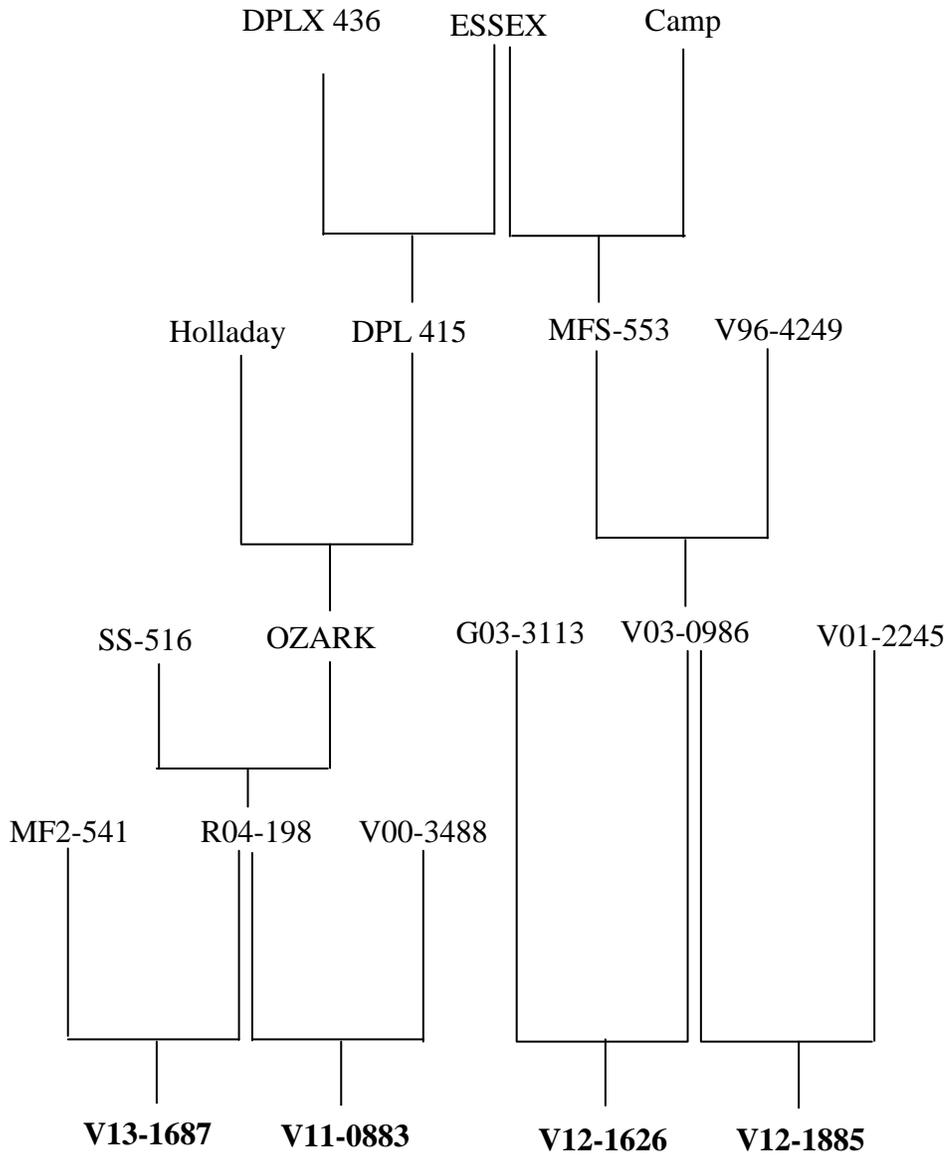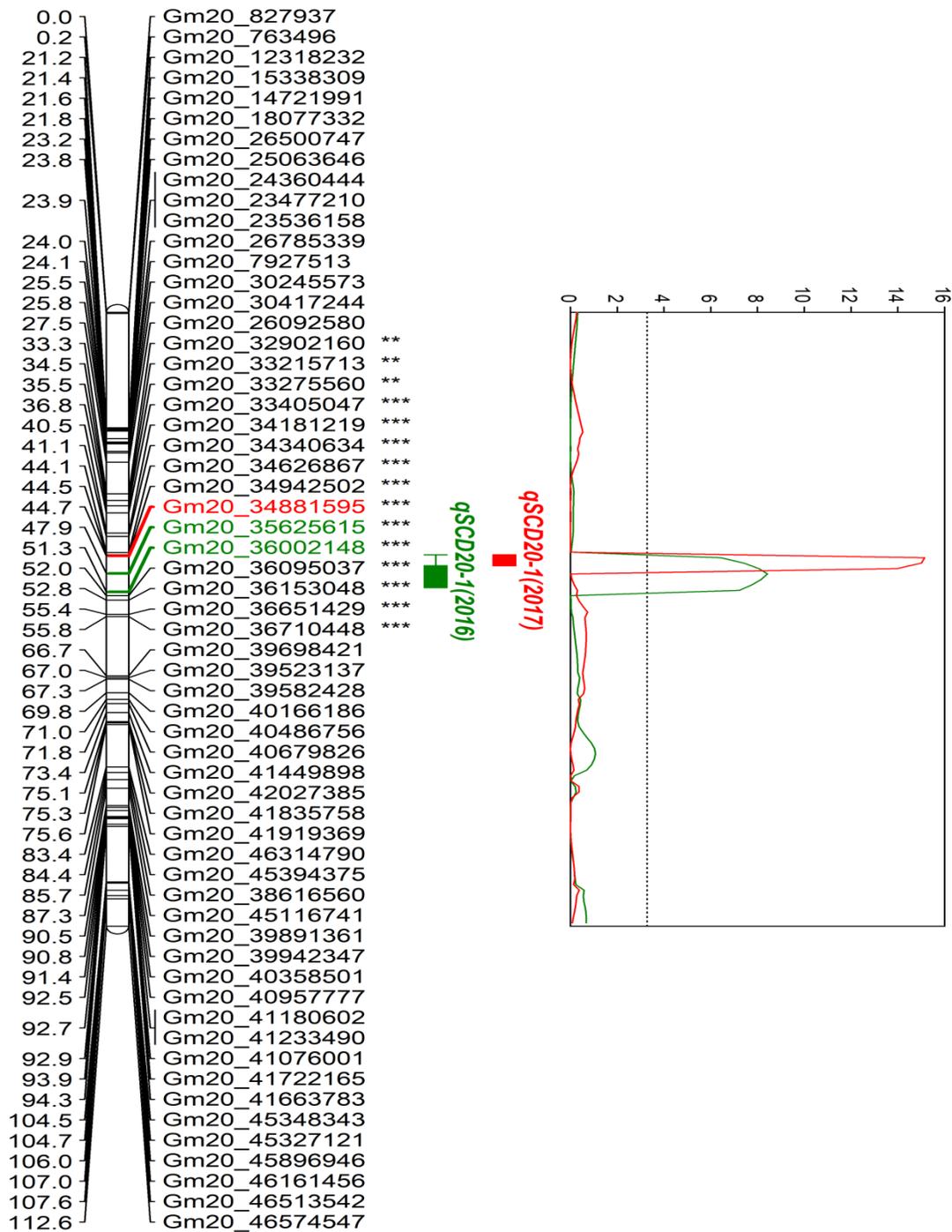*** SNPs significant associated with seed coat deficiency at a p-value<0.0001

Figure 10. Distribution of seed coat deficiency (SCD) in F4:5 lines developed from a cross between soybean lines [Glycine max (L.) Merr.], V12-1626 and V11-0883 in Blacksburg, VA in 2016 and 2017. Average SCD of parents V12-1626 and V11-0883 are shown.
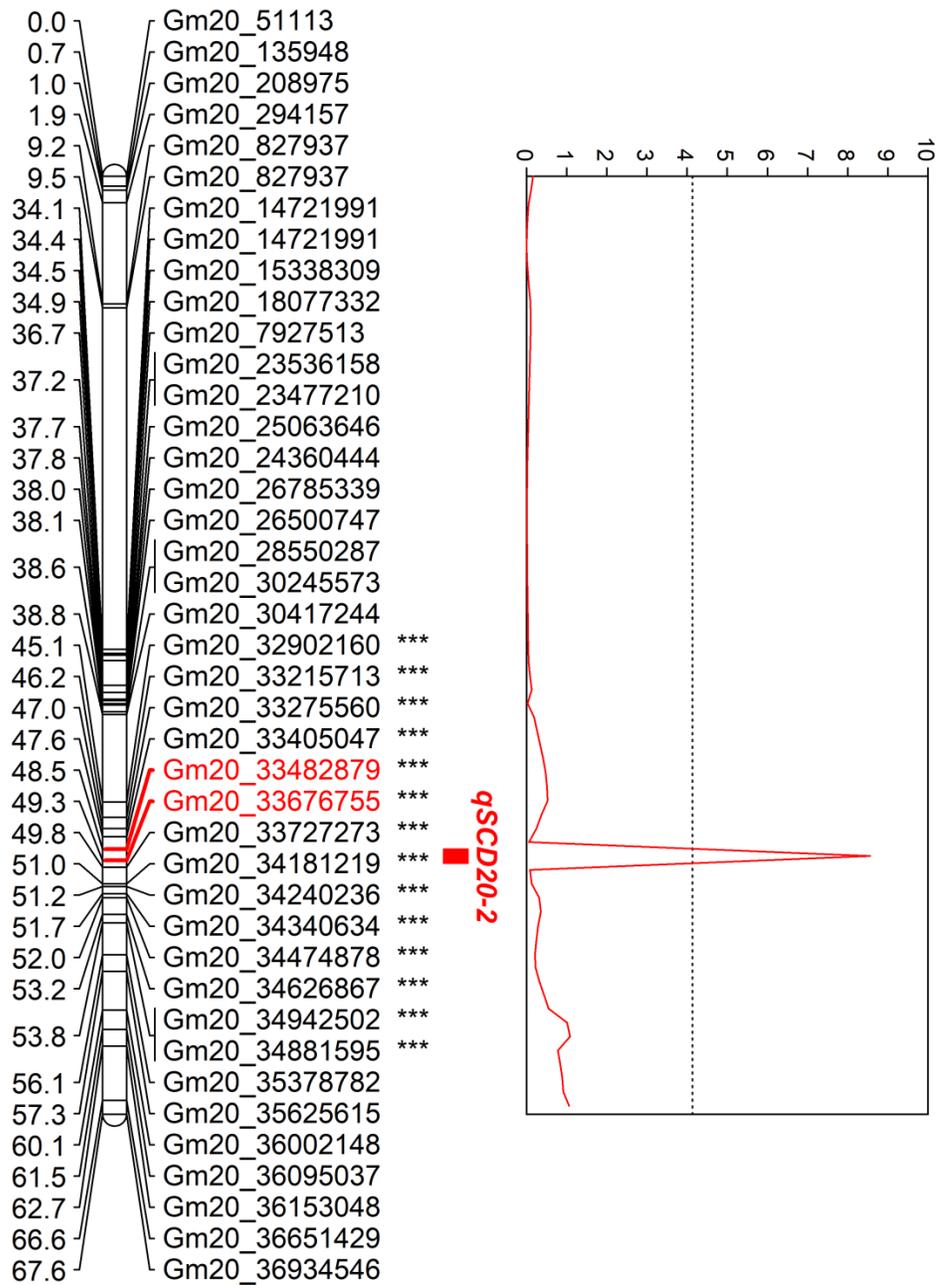
Figure11. Distribution of seed coat deficiency (SCD) in F3 plants developed from a cross between soybean [Glycine max (L.) Merr.] lines, V12-1885 and V13-1687 in Blacksburg, VA in 2016. Average SCD of parents V12-1885 and V13-1687 are shown.
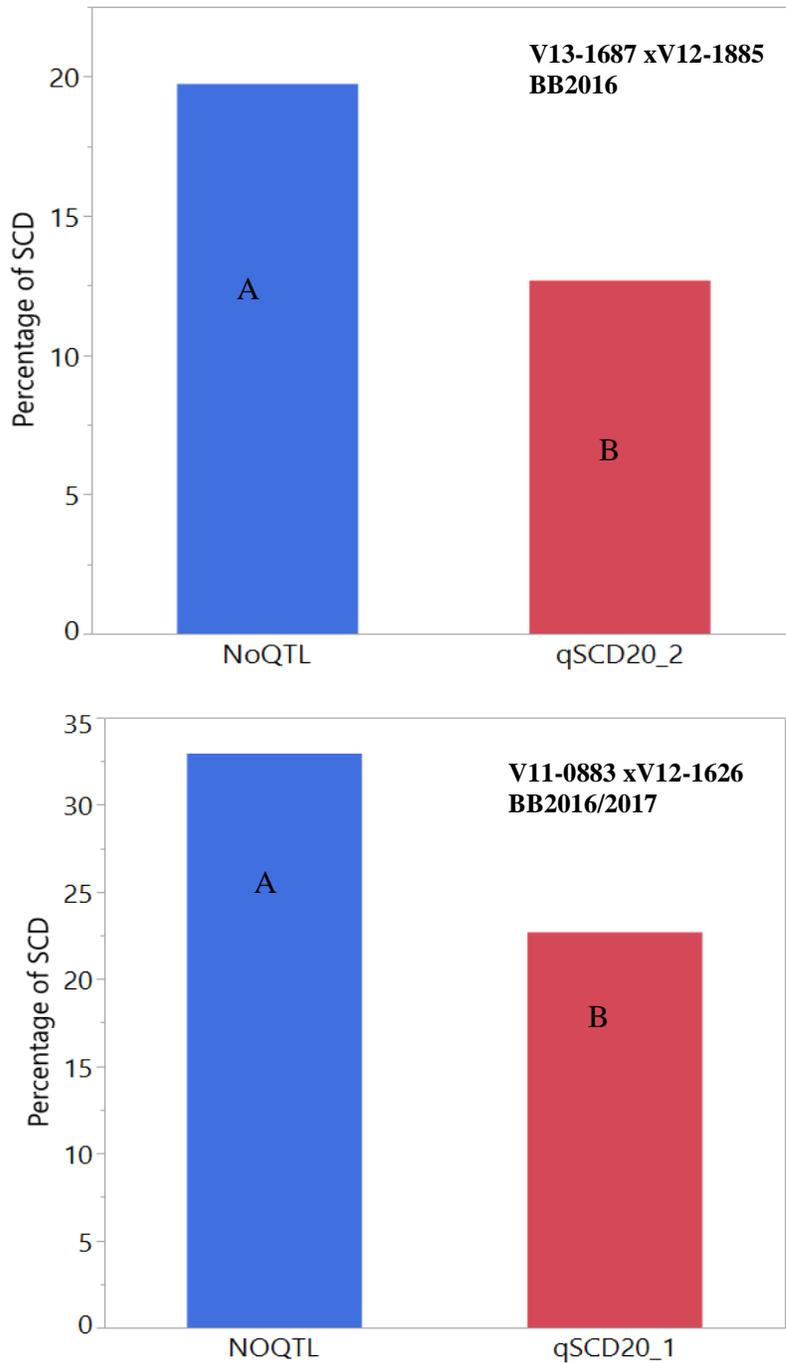
Figure 12. Pedigree of soybean [*Glycine max* (L.) Merr.] cultivars used in this study. Parental cultivars used in this study are shown in bold

Figure 13. Inclusive composite interval mapping for seed coat deficiency in the V11-0883xV12-1626 population in 2016 (F$_3$ plants) and 2017 (F$_{4:5}$ lines). QTL nomenclature is in the form of qTraitChr.number_#. ** SNPs significant associated with seed coat deficiency at a p-value<0.001. *** SNPs significant associated with seed coat deficiency at a p-value<0.0001

151

Figure 14. Inclusive composite interval mapping for seed coat deficiency in the V13-1687 xV12-1885 population in 2016 (F$_3$ plants). QTL nomenclature is in the form of qTraitChr.number_#. *** SNPs significant associated with seed coat deficiency at a p-value<0.0001

Figure 15. Average seed coat deficiency (SCD) of soybeans carrying qSCD20_1 and qSCD20_2. Above average SCD of 2016 from V11-0883 x V12-1626 population in 2016(F3 plants) and 2017 (F4:5 lines) with and without qSCD20_1. Below SCD of F3 plants of the V13-1687 x V12-1885 population with and without qSCD20_2. Different letters indicate significant difference at 0.05 probability level.

5. Conclusions

Among seed and sprout traits, seed size, high-, average- and low-quality sprout percentage, hypocotyl thickness and length and sprout yield should be considered as the most important variables to evaluate soybean sprout quality in breeding programs. Using 'MFS-561', current sprout cultivar and the average value of traits as references, we suggested that soybean sprout varieties should have a seed size of about 10 g/100 seed, high-quality sprouts > 48%, average-quality sprouts < 38%, low-quality < 14%, sprout yield of 5.7 g/ g seed, hypocotyl thickness of 1.6 mm and hypocotyl length > 13 cm. Where, seed storage can affect seed germination and seedling vigor, and storage fungi can be a major cause of reduced seed quality.

Most fungal species isolated from MFS-561, a commercial sprout soybean cultivar, are associated with mycotoxin production, but their presence don't necessarily mean sprout and seed contamination by mycotoxins. Seed disinfection treatments of 2 % calcium hypochlorite for 10 minutes and 5% acetic acid for 2 minutes are recommended as potential seed disinfection treatments for soybean seeds because they reduce fungi incidence without reducing seed germination and sprout quality. A complete elimination of microbial contamination by seed disinfection treatments in sprout production is very unlikely. Thus, seed disinfection must be implemented together with other preventative strategies such as good agricultural practices (seed cleaning, seed storage, and handling), good manufacturing practices and hazard analysis.

Soybean populations showed large phenotypic differences on seed coat deficiency, where genotypes contributed significantly to the trait variation, which indicates that selection of SCD is feasible. Identification of QTL and the independent verification through a parallel population have allowed detection of a stable QTL region on chromosome 20 (LG I) across years and

populations, indicating that genetic improvement of this unfavorable trait could be possible through marker assisted selection (MAS). For that, eight potential SNP markers were identified. However, the populations were only evaluated across two years and further confirmation studies are needed. *qSCD20_1* was detected on the same region as genes associated with seed metabolism and seed development; and QTL associated with several seed traits such as seed composition, seed hardness and seed weight; thus suggesting that chromosome 20 of the soybean genome harbor several genes associated with seed traits and soybean seed development that may be associated with the seed coat strength of soybeans.

6. Future directions

Sprout traits are highly influenced by the environment and sprouting conditions; and their evaluation is time consuming. Future studies should focus on studying sprout and seed traits stability and identifying seed composition and/or seed traits highly associated with good quality sprout that allow indirect and faster selection of soybean sprout cultivars. In addition, identifying QTL and molecular markers associated with sprout quality traits, will allow a better understanding of sprout quality and a faster selection of superior sprout cultivars.

In order to make a complete and accurate diagnosis for mycotoxin production, the specific mycotoxin must be detected from the fungal source and/or food (sprouts). Thus, further studies to evaluate the toxicological risk of the fungal strains isolated in this research will allow determining the actual risk these species represent for sprout production in Virginia.

None of the seed treatments eliminated completely the fungal populations and optimization of relevant treatment parameters may be needed (temperature, time, dosage) to achieve a higher control from calcium hypochlorite and acetic acid treatments. In addition, other studies exploring other specific temperatures and exposure times for hot water treatment might produce satisfactory results in disinfecting soybean seeds without reducing seed germination.

Finally, further confirmation experiments are needed to validate *qSCD20_1* and *qSCD20_2*, which is essential for further utilization in MAS for seed coat deficiency of specialty soybeans for natto production. Further studies should focus on verifying QTL across environments and in different genetic backgrounds, transferring of favorable alleles into high yielding soybean varieties and identification of potential additional of other QTL and epistatic interactions.