AN ABSTRACT OF THE THESIS OF

Chern Lin Koh for the degree of Master of Science in Food Science & Technology

presented on December 19, 2012

Title: Evaluation of DNA Recovery Methods for the Detection of Soy in Foods Using Real-Time PCR.

Abstract approved:

_____

Christina A. Mireles-DeWitt

Food allergies are an important health problem and affect up to 2% of the adult population and 8% of children worldwide. Under the Food Allergen Labeling and Consumer Protection Act (FALCPA) of 2004, foods that contain or derive from the "Big 8" allergens (milk, egg, finfish, crustacean shellfish, tree nuts, peanuts, wheat, and soybeans) must be declared and the "common or usual name" of the allergen source must be printed on the label of the food product.

Currently, the most common used detection methods for food allergens are enzyme-linked immunosorbent assay (ELISA) based. ELISA, a protein-based method, targets specific allergen(s) and detects by colorimetric reaction following binding with a specific-enzyme labeled antibody. However, studies have demonstrated that matrix interference and heat treatment can interfere with the detectability of commercial ELISA kits. An alternative approach to targeting the allergen in soy is to use deoxyribonucleic acid (DNA) as a unique marker that can be used to indicate the presence of soy in food. According to FALCPA the source of an allergen should be

declared on the label, therefore identifying an allergen, such as soy, by DNA detection could be a valid means of meeting FALCPA requirements. Real-time polymerase chain reaction (real-time PCR), a DNA-based method, can identify the presence of soy through amplification of specific sequences of DNA through the use of primers. However, the sensitivity of real-time PCR can be influenced by the amplification protocol, primer design and DNA extraction methods. Thus, the main objectives of this study were to 1) verify the specificity of primers designed to detect soy DNA from different soy products, 2) optimize the previously developed real-time PCR protocol to detect soy DNA, 3) investigate the application of two commercially available DNA recovery systems (column and magnetic beads) to recover soy DNA from different forms of soy products using real-time PCR and 4) determine the effect of food matrices and thermal processing on soy detection using DNA and ELISA methods.

In this study, Wizard Magnetic DNA Purification system kit (Promega, Madison, WI) was selected as the column DNA recovery system while DNeasy mericon Food Kit (Qiagen, Valencia, CA) was selected as the magnetic beads system. Neogen Veratox for soy allergen was selected as the ELISA system. The evaluations of both DNA recovery systems were conducted on soy protein isolates (SPI), powdered soybean and soymilk. The effect of thermal processing in soy detection was conducted on four different food matrices (protein, fat, carbohydrate and water). Each food matrix was spiked with 10% soy protein isolates and heated at 95℃ for an hour. Both DNA (column and magnetic beads DNA recovery system) and ELISA detection methods were used to detect soy in heated and non-heated food matrices.

The limit of detection for column DNA recovery method in soybean, SPI and soymilk can be as low as 20 ppm, while magnetic beads DNA method was matrix dependent. The magnetic beads methods demonstrated a lower detection for soybean sample (1.33 ppm) but higher for soymilk (133.3 ppm). The soy percent recovery for non-heated food matrices was higher in ELISA methods and lower in magnetic beads DNA method. For heated food matrices, percent recovery for both DNA methods was

higher than ELISA method. Overall, heat treatment can significantly reduce the ability of the ELISA method to detect soy in all food matrices. However, for DNA methods (column and magnetic beads), water and ranch matrices were the only two that were significantly affected by thermal processing. In terms of food matrices, water matrix (heated and non-heated) has the highest percent recovery of soy for all detection methods. However, percent recovery of soy in flour matrix (non-heated) was the lowest using both DNA methods.

Evaluation of DNA Recovery Methods for the Detection of Soy in Foods Using Real-Time PCR

by
Chern Lin Koh

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented December 19 2012

Commencement June 2013

Master of Science thesis of Chern Lin Koh presented on December 19, 2012.

APPROVED:

_____

Major Professor, representing Food Science & Technology

_____

Head of the Department of Food Science & Technology

_____

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

_____

Chern Lin Koh, Author

ACKNOWLEDGEMENTS

I would like to express my upmost gratitude to my major professor, Dr. Christina Dewitt, for her support, guidance and encouragement throughout the past eight years. Thank you for watching over me and the progression of all my research projects. I am very grateful for your continuous support, generosity and patience. I am lucky to enjoy the supervision of someone I admire as a thinker and leader.

For the rest of my committee member, Dr. Yi-Cheng, Su, Dr. Andrew Ross and Dr. Frank W.R. Chaplen, thank you for taking the time to participate in my defense and your help throughout my time at Oregon State University.

I would like to thank Dr. Gilliland for his guidance and sense of humor. I will always remember your generosity and witty comments. Thanks to Dustin Vann for being a great mentor and friend. Thank you for taking me under your wing and help me become a better person. I would like to thank Kris Novotny, Alisha Parsons, Akhila Vasan, Teresa Brown, Claudia Cerruto, Tyler Tate, Lina Ramirez and Jordan Denton for their support and friendship at Oklahoma State University.

I would like to thank Dr. Jae Park, Sue Hansell, Dr. Craig Holt, Angee Hunt, Dr. Reza Tahergorabi, Dr. Samanan Poowakanjana (Sand), Sureerat Phuvasate (Note), Jing Mou, Dunyu Xi, Yin Tao, Aoy Kokkaew, Zatil Athaillah, Sara Wu and all others at the OSU seafood Laboratory. Thank you for their support, friendship and sense of humor made my time in Astoria very enjoyable. I would also like to thank Austin Lowder for helping me out with all my SAS analysis and being a great lab mate.

I would also like to thank the support of Oregon State University and the Coastal Oregon Marine Experiment Station's Seafood Research & Education Center in Astoria, OR for funding this research project. It was also supported by the Oklahoma State University Robert M. Kerr Food & Agricultural Product Center and Department of Animal Sciences, Stillwater, OK

Special thanks to my family and my friends for their unconditional love and support. Thank you for believing in me and support my decision to move to Oregon.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

Chapter 1

INTRODUCTION

Soybean (*Glycine max*) is a popular crop due to its high nutritional values and potentials as a cash crop. Production and consumption of soy products in western countries has increased drastically in the last decade. As a major food ingredient, soybean is a great source of protein and provides numerous health benefits. Individuals who consume soybean-rich diets exhibit lower incidence of high plasma cholesterol, cancer (prostate and breast), diabetes mellitus, and increased bone density (2011; Desroches and others 2004; Li and others 2005b; Lovati and others 2000). Also, soybean has been widely used in processed foods for its functionality. However, soybean is among the most common food allergens.

Food allergies are an important health problem, affecting up to 2% of the adult population and up to 8% of children worldwide (Schubert-Ullrich and others 2009; Taylor and others 2009). Food allergies are caused by immunological responses to specific proteins (also known as food allergens) in foods. Currently, avoidance of allergenic food is the only prevention step recommended for this acute and potentially life-threatening adverse reaction.

Under the Food Allergen Labeling and Consumer Protection Act (FALCPA) of 2004, foods that contain or derive from the "Big 8" allergens (milk, egg, finfish, crustacean shellfish, tree nuts, peanuts, wheat, and soybeans) must be declared and the "common or usual name" of the allergen source must be printed on the label of the food product. Also, the European Union (E.U.) implemented 2 legislative act, Directives 2000/13/EC and 2003/89/EC, which enforced a mandatory declaration of allergenic foods. Other than the "Big 8" allergens, the legislative acts also included celery, mustard, sesame seeds and sulfites.

Reliable detection and quantification methods for soy allergens are necessary to ensure compliance with regulations. Since allergens in foods are often present in trace amounts and masked by the food matrix, detection limits for different food products would need to be low. Depending on the respective food, the general agreement for detection limits are somewhere around 1 to 100 ppm (mg allergic protein kg-1 food) (Poms and others 2004). Detection methods for allergens target either the allergen itself or a marker indicating the presence of the allergen source. The choice of methods is mainly dependent on the food of concerns and its processing conditions.

Currently, the prominent detection methods used for food allergens are enzyme-linked immunosorbent assay (ELISA) based. ELISA, a protein-based method, targets the allergen(s) and detect by colorimetric reaction following binding with a specific-enzyme labeled antibody. Simplicity and short reaction time of commercially available ELISA kit have been the main attraction. However, studies have demonstrated that matrix interference and heat treatment can interfere with the detectability of allergens by commercial ELISA kits (Platteau and others 2011; L'Hocine and others 2007).

Other than targeting the allergen in soybean (soy), deoxyribonucleic acid (DNA) can be used as marker to indicate the presence of soy. Real-time Polymerase Chain Reaction (PCR) has emerged as a robust and widely used methodology to detect and quantify DNA. The assay uses primers designed to target a specific region of DNA and amplified it through series of thermal cycle using thermostable polymerase (van Hengel 2007; Gatti and Ferretti 2009; Kirsch and others 2009).

Real-time PCR is highly specific, reproducible, sensitive and rapid. Taking into consideration the differences in product type and composition, an efficient DNA extraction step would be crucial for a PCR detection system (Gryson and others 2004; Olexova and others 2004). Commercially available DNA extraction kits and systems are becoming increasingly popular because of their ease of use, limited labor

requirements, and ability to consistently produce high-quality DNA (Gryson and others 2004; Torres and others 2002).

Thus, the main objectives of this study were to 1) verify the specificity of primers designed to detect soy DNA from different soy products, 2) optimize the previously developed real-time PCR protocol to detect soy DNA, 3) investigate the application of two commercially available DNA recovery systems (column and magnetic beads) to recover soy DNA from different forms of soy products using real-time PCR and 4) determine the effect of food matrices and thermal processing on soy detection using DNA and ELISA methods. .

Chapter 2

LITERATURE REVIEW

2.1 WHAT IS FOOD ALLERGY?

A food allergy is an abnormal immune response to specific protein components of foods (Wood 2003). Coombs and Gell (1975) classified allergic reactions as mostly type I reactions, or referred to them as acute or immediate hypersensitivity reactions. Thus, the reactions are typically characterized by rapid onset of symptoms and mediated by allergen-specific immunoglobulin E (IgE) (Taylor and others 1987; Poms and others 2004).

Adverse reactions to food after ingestion might also be due to other non-immune-mediated reactions (Sampson 2003). The non-immune-mediated food sensitivities are characterized by the absence of allergen-specific IgE and the delayed onset of symptoms. Those non immune-mediated sensitivities may be caused by pharmaceutical reactions, food intolerances, poisons, infectious agents, diseases in the digestive systems, idiosyncratic reactions or even psychological reasons (Nowak-Wegrzyn and Sampson 2006; Poms and others 2004). Among them, food intolerances account for a majority of the adverse reactions (Sampson 1999a).

2.1.1 *Epidemiology of food allergies*

According to the Center of Disease Control and Prevention (CDC), an estimated 3.9% of children under the age of 18 and 2% of adults have food allergies. Furthermore, the health council of the Netherlands estimates the prevalence of food allergies to be in the range of 1-3 % for children and 1-2% for adults (Taylor and others 2009). The literature supports a relative prevalence rate of 1-2% for adults (Niestijl Jansen and others 1994; Nowak-Wegrzyn and Sampson 2006; Ortolani and others 2004; Bock 1986; Sicherer and others 2001; Poms and others 2004). However,

for children, the rate of prevalence is not so clear cut and there is a general disagreement in the actual rate of allergy in children. Several studies have suggested that the actual prevalence of food allergies for children is higher at 6-8% (Wood 2003; Kagan 2003; Cordle 2004; Nowak-Wegrzyn and Sampson 2006; Poms and others 2004). In a report published by the CDC (2007) it was stated that from 2004 to 2006 there were approximately 9,500 hospital discharges per year with a diagnosis related to food allergy among children under 18. Food allergies are estimated to result in approximately 150 deaths each year in the United States alone (Taylor and others 2009). From 1997 to 2007, the prevalence of reported food allergies increased 18% among children under 18 years of age (Branum and Lukacs 2009). In another study of 480 newborns, 28% were reported to experience adverse food reactions, which mostly occur during the first year of life (Bock 1986). Over time most food allergy is lost or "overcome", but others, such as peanuts and tree nuts, can be long-lived (Wood 2003). Overall, there is general consensus that food allergies are more prevalent in children, especially young infants, and individuals with atopic diseases (Sampson 1999a).

2.1.2 How does a food allergy occur?

The development of an IgE-mediated response to an allergen is the result of a series of molecular and cellular interactions involving antigen- presenting cell (APCs), T cells, and B cells (Sampson 1999b). The allergic reactions involve a complex two-step process, sensitization and elicitation. Sensitization requires exposure to a food allergen, allergenic protein components in food, to initiate reactions. Taylor and others (2009) stated that inducing sensitization usually occurs through oral exposure, but it can also be induced through other routes. During sensitization, the allergen interacts with antigen-presenting cells and binds with allergen specific T cells. Newly formed T helper 2 cells associate with cytokines and induces the production of IgE towards allergens from B cells. The antigen-specific IgE in turn binds to the IgE receptor on mast cells. Afterwards, elicitation occurs when the allergen is re-encountered and cross-linked with IgE receptors on mast cells. Degranulation of mast

cells stimulates basophils and thus releases histamines and other inflammatory mediators (Sampson 1999b; Taylor and others 2009).

2.1.3 Symptoms of food allergy

Depending on the location of the mast cells, the released chemicals will exhibit either localized symptoms or systematic symptoms. Even though allergenic symptoms are typically related to the IgE antibody, studies have demonstrated that disorders can also be cell-mediated or from a mixture of IgE and non IgE mechanisms. Food allergenic disorders include acute, possible life-threatening reactions, as well as chronic debilitating diseases (Sicherer and Sampson 2009). Poms and others (2004) stated that intake of minute amount of allergens by highly sensitized individuals can provoke digestive disorders, respiratory symptoms, circulatory symptoms and skin reactions. Other studies also indicate that allergic reactions can result in symptoms such as anaphylactic shock, cardiac arrhythmia and laryngeal edema (Sampson 1999b; Taylor and others 2009).

Severity and onset time differ for individuals. Severity can be generally assumed to be proportional to the level of exposure to the allergen and the individual's sensitivity. Some individuals experience allergenic symptoms only if food is eaten before specific physical stimuli (exercise) or concomitant seasonal allergies. Individual sensitivity may also be affected by other factors such as stress, medication or even medical conditions (Kagan 2003).

2.1.4 Is oral tolerance equivalent to food allergy?

Oral tolerance is defined as a state of active inhibition of immune responses to an antigen prior to oral exposure. The specific suppression of cellular and/or humoral immune responses evolved to prevent hypersensitivity reactions to food proteins and bacterial antigens present in the mucosal flora (Chehade and Mayer 2005; Weiner

2000). Oral tolerance also provides mechanisms to suppress pathologic reactivity against self and, thus, to prevent or treat autoimmune diseases. Thus, a food allergy can be view as the breakdown in, or non-or-underdevelopment of oral tolerance in individuals. On the other hand, sensitization can be viewed as a spectrum ranging from complete oral tolerance to various degrees of non-tolerance (Taylor and others 2009).

Recent clinical research suggested that oral tolerance to peanut could be induced by intentional oral administration of very low, slowly escalating doses of peanuts (Jones and others 2009). Factors that influence the outcome of an immune response to orally administrated antigens includes antigen availability, the immune environment, type of the antigen presenting in Dendritic cells (DC), dendritic cell maturation and activation, the level and form of co-stimulation and the actions of regulatory T cells and their cytokines (van Wijk and Knippels 2007). Both food tolerance and sensitization to food are most likely to occur in infancy as new foods are introduced to the diet. However, with respect to sensitization, it is possible to become sensitized to food at any age.

2.1.5 Clinical assessment on food allergy

In most clinical literature, threshold has been defined as the lowest dose to elicit an allergic reaction. From a risk assessment perspective, the terms lowest observed adverse effect level (LOAEL) and the minimal eliciting dose (MED) were used to describe threshold. No observed adverse effect level (NOAEL) was used to describe the highest amount of the allergenic food which will not cause a reaction in individuals allergic to foods (Taylor and others 2002; Taylor and Hourihane 2008).

Presently, very little information exists on the basis of which to estimate threshold dose for specific allergenic foods. The study of sensitization in human subjects is complicated because most allergenic foods are normal components of the human diet and are likely to be consumed in reasonably large quantity until allergic

responses or symptoms occur. Therefore, most threshold dose studies have been obtained through elicitation from human clinical studies.

2.1.5.1 Threshold assessment

The inconsistency in clinical protocols and subject sizes complicate the estimation of threshold dosage. Most clinical trials are aimed at defining biological thresholds, immune therapy and low dose diagnostic DBPCFC. Among those studies, DBPCFC remains the gold standard for diagnosis of a food allergy (Bindslev-Jensen and others 2004; Allan Bock and others 1988). Allergen-specific IgE level in serum and the skin prick test wheal size were widely used in clinical practice to assess the progress of oral tolerance acquisition in pediatric patients and to determine appropriate times for confirmatory challenge trials (Sampson 2004; Roberts and Lack 2005). However, existing data are somewhat limited and correlation has not yet been found between patients who react on challenge to very low doses and patients with histories of severe reactions (Taylor and Hourihane 2008). Furthermore, most clinical trials using low dose challenge have mainly been conducted on peanut, milk, egg and hazelnut. Soybean research was mostly done in a smaller group and the result was too small to make a population based estimate. Overall, the MED for each specific allergen varied by study and it range from 1 mg to several grams (Taylor and Hourihane 2008).

2.1.5.2 Standardized clinical assessment

In 1999, a panel of 12 clinical allergists and other interested parties were invited to participate in a roundtable conference to discuss clinical approaches and share data on the threshold doses for major allergens. They have concluded that the most useful clinical data derived from the results of double-blind, placebo-controlled food challenges (DBPCFC) and is mostly available on peanuts, eggs and cow's milk. Some data were also available for fish, mustard seed, soybeans, wheat, sesame seed, tree nuts and crustaceans. Since most of the data were obtained by different protocols, estimation of the threshold dose was difficult. Thus, another roundtable conference

was convened to develop a consensus clinical protocol to determine the threshold dose. The panel developed a consensus protocol with the following key elements: selection criteria for patients, natural history of allergenic reactions to specific foods, medical and pharmacological status of patients immediately prior to challenge standardization of challenge materials and protocols, amount of doses for clinical, time interval between doses and interpretation of positive response (Taylor and others 2004; Taylor and others 2002).

2.2 WHAT ARE FOOD ALLERGENS?

In general, most food allergens are glycoproteins with acid isoelectric points, usually ranging between 10-70 kDa in size, stable and resistant against the effect of processing and digestive systems. They are capable of stimulating immune responses and are polyvalent with at least 2 IgE antibody binding sites. Sometimes, they are the major food proteins and the high dose can enhance the likelihood of producing an allergic response (Taylor and others 1987; Lehrer and others 2006). On the other hand, the allergenic reaction is highly dependent on the biochemical structure and source of the glycoproteins. For example, tropomysin is an allergen in shrimp and crustaceans but not in poultry and beef. Also, allergens in soy and peanut are mostly found in storage proteins and allergens in shrimp or other crustaceans are mainly found in the muscle protein (Daul and others 1994; Shanti and others 1993).

2.2.1 Common food allergens

So far, there are over 150 different foods worldwide that have been reported to cause allergic reactions. Among them, most are water-soluble glycoproteins that are between 10 to 70 kDa in size and relatively stable to heat, acid, and proteases (Sicherer and Sampson 2009; Hefle and others 1996). Depending on dietary habits, foods that are responsible for allergic reactions usually vary among countries. In most cases, certain types of allergen might be more prominent than others. For example, in

the U.S., the most common allergenic foods are peanuts, tree nuts, soy, eggs, milk, wheat, fish and crustacean shellfish. In addition, the European Union also considers sesame seed, mustard, celery, molluscan shellfish and lupine as allergens (Taylor 2009). Also, preparation of the food or degree of processing can affect the severity of the allergenic reaction. For example, peanuts are mostly roasted in the United State (U.S.), whereas in China, they are typically boiled or fried. Even though peanut consumption in the U.S. is similar to China, the prevalence for peanut allergy is higher in the U.S. (Sicherer and Sampson 2007). It has been suggested that the high temperatures used during roasting of peanuts may have led to increased stability and allergenicity of peanut allergens in the U.S. (Beyer and others 2001; Maleki and others 2003).

2.2.2 Current regulations for food allergens

In 2004, the U.S. Congress passed the Food Allergen Labeling and Consumer Protection Act (FALCPA) in an attempt to address the allergen problem. Under this Act, food labels on products manufactured after January 2006 need to clearly identify the source of ingredients derived from these eight major food allergens: peanuts, tree nuts, soy, eggs, milk, wheat, fish and crustacean shellfish. FALCPA required the presence of a major food allergen to be listed on the product label. A food allergen can be listed within the ingredient list (e.g. milk, soy or peanut), declaring the source after the ingredient (lecithin (soy)) or added on with the phrase "contains or may contains". Also, since the threshold level is still under investigation, any amount of a major food allergen that is an ingredient, or used in processing, must be labeled. Highly processed oil is the only exemption.

2.3 ROLE OF SOYBEAN IN FOOD

Soybean has been an integral part of the diet for people in the Far East for more than 5000 years. The nutrient-dense legume has been a great source of protein

and a component for drugs. In the U.S., the history of soybean production and soy protein products such as flour, concentrates, isolates and their derivatives are relatively short (Rhee 1994). However, soybean production has become one of the major agriculture crops and exports. According to the national soybean research laboratory, soybean can produce at least twice as much protein per acre than any other major grain crop; 5 to 10 times more protein than land set aside for grazing animals to make milk; and up to 15 times more protein per acre than land set aside for meat production. From 2005- 2009, U.S. soy exports set a new record with whole soy exports totaled 34.4 metric tons, soy oil exports for 1.2 metric tons and soybean meal exports equaled 8.6 metric tons(USDA 2012).

2.3.1 Benefits of soybean

Soybean has been identified as a good source of protein (approximately 37%) and its consumption has increased in recent years. The protein fractions of the crop are reported to be beneficial to human health (Lee and Brennand 2005). Studies indicated that soy protein reduces levels of cholesterol in blood plasma in individuals with high cholesterol levels thus reducing the risk of cardiac diseases. Also, the presence of isoflavones and other phytochemicals can reduce the risk of breast and prostate cancers and increase bone density (Desroches and others 2004; Li and others 2005b; Lovati and others 2000). Besides functioning as antioxidants, soy isoflavones interact with human estrogen receptors and other non-hormonal effects that influence signal transduction in cells (Maltas and others 2011). In October 1999, the FDA approved the use of the health claim: “Soybean can reduce the risk of cardiac disease” on the labels of soybeans and products containing soybeans. The following variation may also be used: “Diets with a low ingestion of saturated fat and cholesterol that include 25 g soy protein daily can reduce the risk of cardiac disease” (Li and others 2005b).

2.3.2 Soybean products

Prior to 1960s, soybean products were used primarily to meet nutritional needs and were mostly consumed as whole, ground into flour, soymilk or fermented products. With the advancement in science and technology, methods have been developed for extracting the proteins in soybean and transformed them into different processed products. Recently, soy protein is highly utilized in the food industry because of its functionality and potential health benefits. Processing condition variation can be used to differentiate soy into flours, concentrates and isolates.

Soy flours are typically made by grinding soybeans into a fine powder and come in 3 major forms: natural or full-fat (40% protein), defatted with 50-54% protein content, defatted with addition of lecithin and grits. Natural soy flours are the least refined and are commonly used in the baking industry and for the production of soymilks. Soy grits are similar to soy flour except the particles are larger (between 1-2.4 mm) and may be toasted.

Soy protein concentrates are defatted flours without the water-soluble carbohydrates. Normally, soy concentrates would contain at least 70% protein. Soy protein concentrates can either be processed by acid leaching at pH 4.5, extraction with 70-90% ethanol or extraction with water after heat-denaturation of protein. Furthermore, a more functional soy concentrate can be prepared by subjecting a low water-soy protein concentrate to steam injection. This type of treatment improves the solubility and functionality of the soy product so it can be used as an emulsifier.

Soy protein isolate is a highly refined or purified protein with at least 90% protein on a dry basis. There are different ways to achieve the isolation, but the common ways are acid and alkaline or pH shift treatment. Soy proteins are first extracted with mild alkaline water at pH 8.5-9.0 and centrifuged to remove the insoluble residue. Afterward, the supernatant is adjusted to pH 4.5 to precipitate the protein. The recovered proteins are then washed, spray dried to yield isolates and then neutralized

with NA-, K-, or Ca- proteinases. The conversion increases the solubility and functionality of the soy protein isolates (Rhee 1994; Kinsella 1979).

Furthermore, soy protein can also be partially hydrolyzed to improve functionality and solubility. Soy protein hydrolysates can be produced by acid, alkali or enzymatic hydrolysis. During enzymatic hydrolysis, proteolytic enzymes (such as pepsin, papain, bromelain) are used to reduce the molecular size of proteins to larger peptides and increased the numbers of charged groups. Depending on the resultant molecular size and hydrophobicity of the peptides, the functional properties would differ. Thus foaming, gelling and emulsifying properties of the product would differ too (Hrckova and others 2002; Rhee 1994; Kuipers and others 2005).

2.3.3 Functionality of soy protein

In general, the protein fractions are the principal functional components in soybean. Approximately 90% of the protein fractions are storage proteins known as globulins and the remaining proteins consist of intracellular enzymes, hemagglutinins, protein inhibitors and lipoproteins. Within the storage proteins, β-conglycinin (7S globulin fraction) and glycinin (11S globulin fraction) are the major components that constitute over 70% of soluble protein. The physicochemical and functional properties of soybean proteins can be modified by physical, chemical and enzymatic treatment. Through protein modification, the functional properties can be altered or extended for different usage.

2.3.3.1 Soy protein modification

Traditionally, thermal treatment has been used to reduce protease inhibitors, eliminate lipoxygenases and unfavorable volatile compounds. Also, thermal treatment induces dissociation and degradation of polypeptides that leads to increase in both indigestibility and solubility. On the other hand, chemical treatment usually refers to

acid-alkali denaturation. Similar to thermal treatment, chemical treatment can be used to reduce protease activity and to eliminate unfavorable odor. Such processing conditions can also reduce phytic acid content (increased bioavailability of minerals such as magnesium, zinc, iron and calcium) and decrease or increase solubility through deamination and mild hydrolysis (Barac and others 2004). It has also been suggested treatment with alkali can cause dissociation of storage proteins and results in increased viscosity and possible gelation (Kinsella 1979). Enzymatic treatment mostly refers to hydrolysis of 7S and 11S globulin fractions. As mentioned in the previous section, selective proteases can partially hydrolyze both storage proteins (7S and 11S globulin fraction) to form smaller polypeptides. Aside from proteases, oxidases can be added to irreversibly oxidize aldehydes to remove undesirable beany or green soy flavor that are largely caused by aldehydes. Depending on the specificity of the enzyme, the processing conditions and the extent of hydrolysis, a wide variety of peptides are generated. The resultant protein hydrolysate has been found to have improved nutritive properties, increased digestibility and better foaming and gel-forming ability (Barac and others 2004; Panyam and Kilara 1996). With the ability to provide soy proteins with different functional properties, the food industry has found numerous ways to incorporate soy protein in their products.

2.3.3.2 Soy protein in food products

Through protein modification, soy protein isolates or textured soy proteins are used in foods products to not only supplement protein content, but to act as binders and thus enhance the water absorption and retention properties of the final product. This use is prevalent in surimi-based products (such as fish blocks, fish patties, and fish cakes), processed meat products and meat analogs. Furthermore, addition of soy protein can also improve the sliceability, shelf life and yield of processed meat. Studies have demonstrated that soy protein can bind and hold the natural flavor and moisture within processed meat products while being able to withstand stresses caused

by different types of cooking methods and extended storage or holding time(Kinsella 1979; Rhee 1994; Riaz 1999). For extruded products, the addition of soy protein is not necessarily beneficial as product expansion is reduced which results in an increased product hardness (Li and others 2005; Veronica and others 2006).

2.3.3.3 Soybean as an allergen

Soy exhibits high functionality as a food ingredient. But its recognition as one of the major allergenic foods could be a drawback for the crop. Within the protein fraction, mostly in the globulins, at least sixteen IgE-reactive proteins have been reviewed and suggested as food allergens by the Allergome database (Ballmer-Weber and Vieths 2008).

Among them, Gly m 1 (hydrophobic protein), Gly m 2 (defensin), Gly m 3 (profilin), Gly m 4 (PR-10 protein), Gly m 5 (β-conglycinin), Gly m 6 (glycinin) and Gly m 7 (seed biotinylated protein) were officially identified as food allergens by International Union of Immunological Societies Allergen Nomenclature sub-Committee. Other proteins of interest are thiol-protease Gly m Bd 30k (Helm and others 1998; Ogawa and others 1995), Kunitz trypsin inhibitors (Burks and others 1994; Moroz and Yang 1980), an agglutinin/lectin precursor (Barnett and Howden 1987), the α-sub units of major storage protein β-conglycinin (Gly m 5) (Helm and others 1998; Holzhauser and others 2009; Ogawa and others 1995), acidic chain of major glycinin G1 subunits and the basic chain of G2 subunits (Beardslee and others 2000; Holzhauser and others 2009). Furthermore, Awazuhara and others (2006) detected ten IgE-binding and eight IgG4- binding proteins in soybean by immunoblotting with sera from 30 soybean- sensitive patients. Among these proteins, 3 of them reacted with both antibodies. Thus, they suggested that IgG4-binding proteins may also act anaphylactically in patients with soybean allergy.

2.4 DETECTION METHODS FOR FOOD ALLERGENS

Reliable detection and quantification methods for food allergens are necessary to ensure compliance with regulations. Since allergens in foods are often present in trace amounts and masked by the food matrix, detection limits for different food products would need to be low. Depending on the respective food, the general agreement for detection limits are somewhere around 1 to 100 ppm (mg allergic protein kg^{-1} food) (Poms and others 2004). Currently, detection methods for allergens target either the allergen itself or a marker indicating the presence of the allergen source. The choice of methods is mainly dependent on the food of concerns and its processing conditions.

2.4.1 Protein-based detection methods

Protein based detection methods target allergenic protein(s) and depend on immunochemical assays such as radio-allergosorbent test (RAST), enzyme-allergosorbent test (EAST), rocket-immuno-electrophoresis (RIE), immunoblotting and enzyme-linked immunosorbent assay (ELISA). Both RAST and EAST inhibition test are *in-vitro* tests and mainly used in clinical diagnosis of food allergen. Immunoblotting test such as sodium docecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) can be used for protein/allergen separation and identification. According to Scheibe and others (2001), SDS-PAGE can be used to identify potential allergens down to a detection limit of 5 ppm in food. However, this type of testing would be time consuming and human sera are needed for method development. On the other hand, RIE uses an antibody gel for electrophoresis and the rocket-shaped precipitates are the result of antigen-antibody complexes. In terms of detection limit, RIE was able to detect peanut allergens down to 2.5 ppm in various food products (Holzhauser and Vieths 1999). But, even though RIE has the ability to detect allergen, the gel preparation and staining protocol is still time consuming.

2.4.2 Enzyme-linked immunosorbent assay (ELISA)

ELISA test relies on the interaction between antigens and antibodies. Due to its simplicity, short reaction time and ability to detect allergen quantitatively, the test has been widely used in the food industry. Traditional ELISA typically involves an enzyme initiated reaction that produces observable color change to indicate presence of antigen. For food allergen testing, the common approach to quantify potential allergens in foods is a sandwich ELISA (Poms and others 2004; Schubert-Ullrich and others 2009). The sandwich ELISA measures the amount of antigen between two layers of antibodies. The capture antibodies (attached to a microtiter plate or a multiple well strips) were designed to attract specific antigens and detection antibodies were added to bind the antigens forming a "sandwich". Afterward, enzyme-labeled secondary antibodies were added and color changes occurred through chemical reaction with substrates. Depending on the desired sensitivity, either monoclonal or polyclonal antibodies can be used as capture and detection antibodies in the system. The rapid sandwich ELISA test kit can be found for detecting gluten, milk, soybean, peanut, hazelnut, almond, egg and crustaceans. Most tests can produce qualitative and/or semi-quantitative results within 30 – 60 minutes.

Other than sandwich ELISA, competitive ELISA has been used to quantify food allergens. The protocol for Competitive ELISA is slightly different than both traditional and sandwich ELISA. To start with, the targeted antigen was first bound to unlabeled antibodies to form antibody/antigen complexes. Afterward, the complexes were added to another antigen coated well. Similar to sandwich ELISA, secondary antibodies were added to couple with the enzyme (substrate). As suggested by the name, only the unlabeled antibodies will be bound to the second antigen and thus interact with the secondary antibodies. So unlike the sandwich ELISA, the higher the sample antigen concentration, the weaker the eventual signal. The major advantage of a competitive ELISA is the ability to use crude or impure samples and still selectively bind any antigen that may be present.

With the advancement in technology, Kirsch and others (2009) stated that a multiallergen immunoassay based on ELISA system has been developed. The assay can simultaneous determine detection of at least 1 ppm for each peanut and tree nut allergen in chocolate. They also stated that a combination ELISA and inductively coupled plasma MS (ICP-MS) technique has been developed and this new method can increase sensitivity and precision in detecting allergen. According to them, the new ELISA-ICP-MS method can detect peanut allergen down to 2 ppm in a cereal based matrix. The ELISA-ICP-MS uses secondary antibodies that are labeled with stable isotopes instead of enzyme, and can be used for quantification with a mass spectrometer (Careri and others 2007).

2.4.2.1 ELISA test for soy allergen detection

As mentioned earlier, the short reaction time and simplicity of ELISA test kits have attracted the attention of food industry and testing laboratory. Thus, commercial kits have been developed to help identify different type of allergens. As for the soy allergen kit, Pom and others (2004) listed a sandwich type ELISA that targets soy protein by Elisa System and a competitive ELISA that targets soy trypsin inhibitor. Furthermore, both kits were evaluated with soy protein isolates, chickpea isolates and food samples by L'Hocine and others (2007). They stated that both kits were highly sensitive with a limit of detection of 1-2 ppm. But they also discovered drawbacks related to food matrix interference and specificity. Protein modification such as hydrolysis or glycation may interfere with the detectability and accuracy of the assay.

Such observations were also seen in the study conducted by Platteau and others(2011). In this study, 3 different commercial kits (Veratox Soy Allergen, BioKits Soya Allergen Assay and Soy Residue by ELISA Systems) were used to evaluate the impact of the Maillard reaction on the detectability of soy proteins. The authors stated all 3 kits can be useful to give an indication of the presence of allergen, rather than an exact quantitative value of soy allergen. Furthermore, they also

concluded that Maillard reaction in soy proteins in a buffered system can strongly altered the detectability of each kits. Therefore a more sensitive analytical method for the detection of traces of allergens in processed foods is still needed.

2.4.3 *DNA-based detection methods*

DNA or deoxyribonucleic acid consists of 2 long polymers of nucleotides with backbones made of sugars and phosphate groups. It contains the genetic information (genes) needed in the development and functioning of all living organism. On the other hand, proteins consist of one or more polypeptides that are folded into a globular or fibrous form. A polypeptide is a single linear polymer chain of amino acids bonded together by peptide bonds at which the sequence of amino acid is defined by a gene. Thus, DNA and protein can both carried genetic information for and from a specific gene. Therefore, DNA-based detection methods can be used as a surrogate (or indirect) test to identify the source of the allergenic compound. Instead of targeting the proteins, the method targets a specific DNA fragments that is indicative to the allergenic foods and amplifies by PCR. This method was specific and sensitive in detecting food source but it does not detect the allergen itself or a comparable protein. DNA and proteins can also be affected differently during food processing and thus results from the detection method might vary. But DNA tends to remains intact for longer time under heat and pressure processing (Hird and others 2003). Thus it would provide the basis of a robust assay for the detection of allergen residues in foods. Furthermore, protein composition may vary between varieties, but the DNA of a gene will remain the same (van Hengel 2007). Thus, a single protocol would be sufficient to cover all different type of products (Poms and others 2004).

2.4.3. 1 Polymerase Chain Reaction (PCR)

The DNA-based detection method was initially developed using traditional PCR. A traditional PCR involves 2 main procedures, amplification of the DNA and agarose gel electrophoresis for size separation. Typically, a specific segment of DNA

is flanked by primers (2 oligonucleotides) and amplified by thermostable polymerase through a series of thermal cycles. The thermal cycle consists of 3 major functional steps: denaturation, annealing and elongation of DNA. Depending on the initial concentration of DNA, 25-45 cycles of reaction are need to established a detectable signal. Afterward, the amplified product can be visualized by staining with florescent dye or southern blotting in an agarose gel.

As an alternative approach, PCR application has been evaluated to detect specific wheat (gliadin) in 35 different food products (Allmann and others 1993). The results indicated that PCR technology was able to support and confirmed the analysis executed with an ELISA system. Furthermore, another study suggested that PCR detection can be used to detect hazelnut protein at 10 ppm in chocolate (Herman and others 2003). They also suggested that PCR method can be the future of food safety for detecting traces of hidden food allergens (Herman and others 2003; Holzhauser and Vieths 1999). On the other hand, PCR detection has been developed to detect trace of allergenic mackerel in foods(Hong 2007). The developed assay was designed to target the sequence for parvalbumin from mackerel (*S.Japonicus*). The assay was able to detect 5 ng of purified mackerel DNA in 50ng of surimi DNA.

As for soy, Meyer and others (1996) concluded that PCR detection can be used as an alternative method for detection of textured soy protein concentrated in processed meat products. The developed PCR assay was able to detect soy protein down to 0.01% in pork. Also, a more recent study concluded that the developed PCR assay has the detectability of 0.01% to 0.06% (w/w) in unprocessed and heat processed pork (Soares and others 2010). The assay was designed to target lectin gene and the assay was able to detect traces of soy DNA in products that were not labeled with soy. Samples were extracted using Promega Wizard DNA Purification kit and results were based on detections in agarose gel. The primers used were Lectin-F 5’ TCC ACC CCC ATC CAC ATT T 3’ and Lectin-R 5’GGC ATA GAA GGT GAA GTT GAA GGA 3’ along with a lectin probe, Lectin-TMP 5’ FAM-AAC CGG TAG CGT TGC CAG CTT CG--BHQ2 3’. Lastly, Wang and others (Wang and others

2012) developed a PCR method that targets one of the major allergen genes, Gly m Bd 28K. They concluded that the assay developed has a LOD of 10 to 100ppm depending on the food matrix.

But as PCR results are visualized by agarose electrophoresis, they are merely qualitative. Even by incorporating internal standards, the results are still semi quantitative (Studer and others 1998). Thus, to achieve a more superior quantification result, traditional PCR have been replaced by real-time PCR.

2.4.3.2 Real-time Polymerase Chain Reaction (Real-time PCR)

In recent years, real-time PCR has emerged as a robust and widely used methodology because it can detect and quantify very small amounts of specific nucleic acid sequences. Instead of visualizing with agarose gel electrophoresis, the new real time PCR machine can amplify and quantify DNA simultaneously as the reaction progresses in real time. Similar to traditional PCR, primers were designed to target a specific region of the DNA and amplified using thermostable polymerse. To obtain the "real time" results, 30-45 thermal cycles were repeated by either using a non-specific fluorescent dyes or by a sequence-specific DNA probes (Poms and others 2004).

Similar to traditional PCR, primer design is essential in target detection. Primer is a strand of nucleic acid that serves as the starting point for DNA synthesis. Thus a pair of primers is designed to target a double-strand DNA sequence. To start with, the primers need to have similar melting temperature and the region of DNA targeted needs to be unique to source (Dieffenbach and others 1993). Also, Basic Local Alignment Search Tool (BLAST) can be used to search through all possible known sequences to check for primer specificity.

2.4.3.2.1 Real-time PCR dyes and chemistries

Depending on the objective of the study, three main DNA florescence-monitoring systems (hydrolysis probes, hybridization probes and DNA-binding agents), shown in fig 2.1, can be used in real-time PCR detection (Dorak 2006). DNA hydrolysis probes (TaqMan) consist of oligonucleotides that have been labeled with fluorescent reporter dye and a quencher dye. As long as the probe is still intact, the fluorescence stays suppressed. But as Taq polymerase cleaves the hybridized probes, the dyes separate. With the quencher dye no longer in close proximity, the intensity of the fluorescence of the free reporter dye is able produce a measurable signal (Holland and others 1991). During each consecutive PCR cycle this fluorescence will further increase because of the progressive and exponential accumulation of fluorescent reporter dye.

The hybridizing probes (molecular beacon and scorpions) consist of two PCR primers and two sequence-specific probes (donor and receptor) that bind adjacent to each other in a head-to-tail arrangement. The probe is designed to hybridize to a region within the amplicon and is duel labeled with a reporter dye and a quenching dye (Heid and others 1996; Dorak 2006). After hybridization, the two probes are in close proximity and the donor probe, excited by an external light, will transfer energy to the receptor probe through Fluorescence Resonance Energy Transfer. Thus it results in the emission of fluorescence from the receptor probe. After each subsequent PCR cycle, more hybridization probes can anneal and result in higher fluorescence signals (Van der Velden and others 2003).

Lastly, the non-specific fluorescent DNA binding dyes (such as SYBR Green I) would intercalate with any double-stranded DNA (dsDNA) and caused florescence. Thus the fluorescence intensity increases proportionally to the concentration of dsDNA. During the extension phase, more and more SYBR Green I will bind to the PCR product, resulting in an increased fluorescence. Consequently, during each subsequent PCR cycle more fluorescence signal will be detected. The intensity of the florescence can be used to quantify the amount of DNA (Dragan and others ; Van der

Velden and others 2003). Quantification of DNA can be determined by either absolute quantification or relative quantification. Absolute quantification gives the exact number of target DNA molecules by comparison with DNA standards while relative quantification is based on internal reference genes to determine fold-differences in expression of the target gene (Poms and others 2004).

2.4.3.2.2 DNA extraction methods

Real-time PCR is highly specific, reproducible, sensitive and rapid in processing time. Taking into consideration the differences in type, composition and degree of products, an efficient DNA extraction step would be crucial for a PCR detection system (Olexova and others 2004; Gryson and others 2004). But the presence of PCR inhibitors, such as polysaccharides and polyphenols in foods, can interfere with the reaction and lead to decreasing or complete inhibition of DNA polymerase activity (Van Hoef and others 1998; Meyer and others 1996; Pinto and others 2007). Furthermore, DNA may degrade considerably during food processing, particularly in thermal treatment with soybean and maize (Olexova and others 2004; Querci and others 2006; Stephan and Vieths 2004).

Many different methods and technologies are available for extraction and purification of genomic DNA. Commercially available DNA extraction kits and systems are becoming increasingly popular because of their ease of use, limited labor, and ability to consistently produce high-quality DNA (Gryson and others 2004; Torres and others 2002). In general, all methods involved disruption or breaking down of tissue or cell, separation of nucleic acid from cellular proteins and debris, and purification of nucleic acid (Torres and others 2002). Extraction and purification of nucleic acids from different food samples often involve organic extraction and ethanol precipitation with a variable loss of non-negligible amounts of the original sample (Pinto and others 2007). Even so, application to isolate DNA from complex food matrices and processed food can be limited (Gryson and others 2004). Thus different types of DNA extraction and purification procedures have been studied for its application in food.

The objective of nucleic acid extraction methods is to obtain purified nucleic acids from various sources. According to Olexova and others (2004), cetyltrimethylammonium bromide solubilization with a subsequent liquid-phase extraction (CTABLPE) is accepted as a "gold standard". But this method is laborious and time consuming. Thus a chaotropic solid-phase extraction (SPE) method is often preferred. Column–based nucleic acid system is a SPE method that acts like a filtration system where DNA are bound to the silica membrane during washing and released through centrifugation and neutral ionic condition (Yang and others 1998; Wolfe and others 2002). The column based extraction method is mostly designed for small sampling size and is limited by the binding ability (space and conditions) of the silica-coated membranes layered within the column.

Recently, another SPE extraction approach, mobile solid phase system, has been used to compare against the column-based system. Magnetic beads technology was initially invented for the field of medicine and biological studies. Recent years, the method has been adapted to recover genomic DNA for PCR amplification. The magnetic beads procedures involve attracting DNA to magnetic beads, holding the beads in place using a magnetized source, and washing away other components of the sample. Unlike column-based systems, binding of nucleic acids to magnetic particles can occur in solution, resulting in increased binding kinetics and binding efficiency. Particles also can be completely re-suspended during the wash steps and thus enhancing removal of contaminants. The binding affinity of DNA can depend on either ionic strength (silica dioxide coated beads) or pH (surface charged beads).

Both extraction methods type have been suggested to produce high quality of DNA. To evaluate the application of the DNA recovery methods, soybean and SPI (soy protein isolates) were chosen to represent unprocessed and highly processed soy products. To compare the effectiveness of the approaches, Pinto and others (2007) decided to evaluate both systems with various type of food. They concluded that Promega Wizard Magnetic DNA Purification for food (mobile phase system) is highly efficient in a matrix rich with polysaccharides and polyphenols. The system was

stated to be less time-consuming and less technically demanding than the revised DNeasy Tissue kit (column-based system). However, the column-based system has higher binding affinity with animal tissues and complex matrices. Also, Datukishvili and others (2010) concluded that Promega Wizard Magnetic DNA Purification for Food gave the highest yield of genomic DNA, while DNeasy plant kit produced DNA with best purity and integrity.

2.4.3.3 Real-time PCR for allergen detection

Real time PCR methods have been largely used for the detection and quantification of GMO foods and have been used for peanut detection in food (Arlorio and others 2007). In a study, 33 commercial products were analyzed with both ELISA and real time PCR method in detecting trace of peanut in processed food. The results suggested that both assays were sensitive and specific to detect hidden allergen in processed food and could detect down to 10 ppm in semisweet chocolate. They also concluded the real-time PCR method they developed could only be considered as semiquantitative because of the high coefficient of variation (CV), influence of PCR inhibitors and thermal degradation that was not specifically addressed in the study (Stephan and Vieths 2004).

In addition, Arlorio and others (2007) suggested that real-time PCR method can be used as safety assessment for food potentially crossed contaminated by hazelnut. This study stated that the protocol developed is highly sensitive in spiked processed food and can detect hazelnut DNA down to 0.1ng. They also observed that the developed protocol can be limited by the extraction and clean up steps. Similar results were also observed in another study on detecting pea and walnut in food (Brezna and others 2006b; 2006a). The developed real-time PCR protocols were able to demonstrate a matrix-related detection limit of 50,000 ppm in pea and 10,000 ppm in walnut. The authors suggested that real-time PCR method was sensitive and selective in detecting walnut for bakery and confectionary products. Furthermore,

protocol developed for detection of pea can be quantitative with calibration for each product and processing conditions.

2.4.3.4 Real-time PCR for soy allergen detection

Soybean is commonly added as functional ingredients in food products. As soybean is a food allergen, possible cross contamination could be a problem in a plant that produced products without soy. Similar to other food commodities, the PCR method was first developed for detecting GMO soybean in foods. Both real-time PCR and competitive PCR have been largely used to quantify GMO and products derived from it (Foti and others 2006). As mentioned in the previous section, the PCR method can be used as surrogate test for allergen detection. Thus, the simplicity and high efficiency of real-time PCR would perhaps be a better option. Furthermore, there are at least 16 known soy allergens and the adverse reaction may differ between individuals. Thus, a more universal approach using detection of the gene as opposed to the protein might prove to be more beneficial.

Lectins are carbohydrate-binding proteins or glycoproteins that occur widely in plants, animals and microorganisms. The soy lectin gene represents one of the seed protein genes that are highly regulated during the soybean life. Also, even though advanced processing can degrade the lectin gene, it's still considered more stable than other exogenous genes found in soy (Chen and others 2005). Several studies have suggested that a PCR method targeting lectin gene could be used to detect soybean in foods (Siegel and others 2012; Soares and others 2010; Espineira and others 2010).

In a study conducted by Espineira and others (2010), both traditional and real-time PCR were evaluated for rapid detection of soy in processed food. Genomic DNA was extracted using either a silica based gel membrane method (NucleoSpin Tissue kit; Macherey-Nagel, Du ren, Germany) or a CTAB digestion with a silica-gel based purification system. In this study, a TaqMan probe targeting the lectin gene, 5' (FAM) CAC ATG CAG GTT ATC TTG GTC 3' (TAMRA), was developed and real

time PCR was performed with the following thermal cycling conditions: 95℃ for 30s followed by 45 cycles of 95℃ for 5s and 58℃ for 30s. The authors indicated that the detection limit of soy in cans of fish for real time PCR and traditional PCR developed were 500 ppm and 625 ppm. However, they suggested that PCR method developed could be used as a screening process, but quantification would need to be carried out with ELISA. Furthermore, they also tested 35 different processed products that were declared as free of soy. Within the samples, almost half of them were positive soy detection. Thus, they concluded that the PCR assay could be a valuable tool in verifying the implementation of labeling requirements to protect consumer rights.

Recently, an interlaboratory validation of two food allergen (soybean and mustard) was conducted using real-time PCR (Siegel and others 2012). The genomic DNA was extracted using a CTAB based method with a silica-gel membrane purification system. The primers used for soybean were Lectin probe 5'-FAM- AAC Cgg TAg CgT TgC CAg CTT Cg-TAMRA-3 or Lectin forward 5'-TCC ACC CCC ATC CAC ATT T-3 and Lectin-R 5'-GGC ATA GAA GGT GAA GTT GAA GGA-3'. The amplification was performed following the thermal protocol: Initial step for 10 min at 95 ℃, followed by 40 cycles of 15 s at 95 ℃ and 60 s at 60 ℃. The developed method was tested in 10 different laboratories and the results provided were consistent and able to detect a spike level of 10 ppm soybean in boiled sausages. Even so, the authors concluded that further matrix-based calibrants have to be produced to fully understand the accuracy of the results among foods.

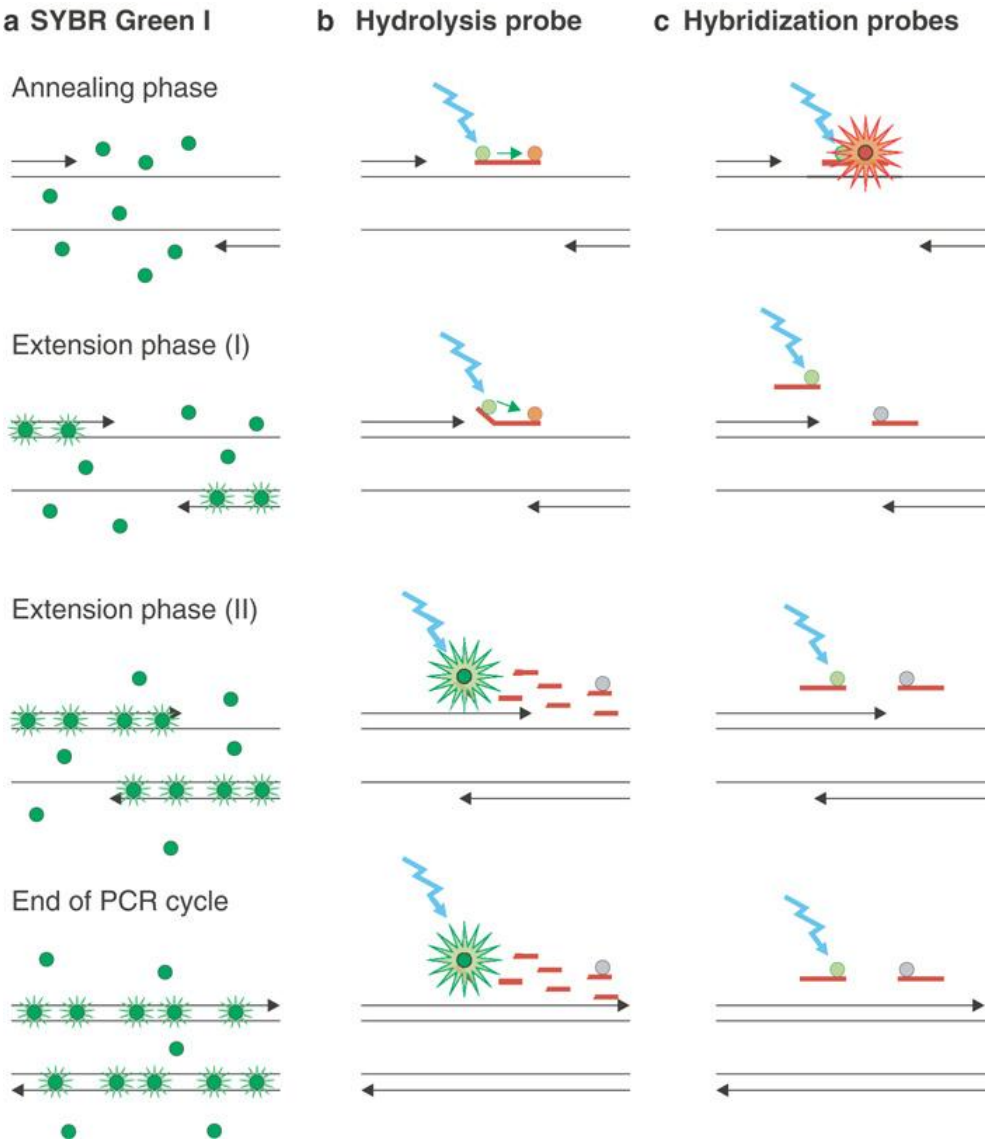


Figure 2.1: Principles of real-time PCR chemistries during annealing and extension phase, adapted from (Van der Velden and others 2003). (a) SYBR Green I technique, (b) Hydrolysis probe technique and (c) Hybridization probes technique**.**

Chapter 3

MATERIALS AND METHODS

3.1 PRIMER DESIGN

Previous research was conducted (Tate 2009) to identify a gene in which a primer could be developed for Real-time PCR identification of processed soy products. Initial investigations determined that the soy lectin 1 gene (Le1), unique to soy, would be a good gene to target for primer development. Lectins are proteins localized in the root hairs of the soybean plant. The gene sequence was obtained from the NCBI database Accession Number: K00821. Several primers were designed utilizing Primer3 software (Tate, 2009) and the sequences were screened through Basic Local Alignment Search Tool (BLAST). A Real-Time PCR protocol was developed for evaluation of the primers of varying size. The best performing primer (no production of primer dimer, ghost peaks, etc) was identified as Lectin Forward (5'-TCC ACC AAA TCC ACA CAT C-3') and Lectin Reverse (5'-GAA GCA AAA GAC CAA GAA AGC AC-3') that produced an amplification product of 291 bps when tested against whole, ground soybean. This primer was then further evaluated as described below.

3.2 Evaluation of Primer for Specificity of Soy Detection

For specificity verification (insuring the primer detects only soy products), DNA was extracted from raw ground soybean and green pea, confirmed with real time PCR (using the method developed by Tate, 2009) and gel electrophoresis was performed to check the purity of the amplified product. All DNA extractions were repeated twice for each sample type and PCR reactions were conducted in duplicates. In addition, an amplified product from each extraction of ground soybean and green pea was randomly selected for evaluation by gel electrophoresis. Description of DNA extraction conditions, PCR methods and gel electrophoresis are described below.

3.2.1 DNA extraction and real-time PCR

The primer set was previously screened through Basic Local Alignment Search Tool (BLAST) to verify the specificity of the primer design by Tate and others (2009). To further verify the specificity of the design, the primer set was evaluated on DNA extracted from a closely related legume, Green pea (*pisum sativum*). DNA was extracted from green pea and soybean (Qiagen DNeasy mericon Food Kit, Valencia, CA) and real time PCR was performed using Bio-Rad MyiQ Single Color Real-Time PCR Detection System (Hercules, CA) following the real-time PCR protocol mentioned below. The protocol used was adapted from previous research conducted by Tate and others (2009). PCR amplification was carried out in a 15 µl volumes with the following components: 7.5 µl of Bio-Rad iQ™ SYBR® Green Supermix, 5.1 µl of sterile nanopure DNase-free water, 0.1 µl of 150nM of forward primer, 0.1 µl of 150nM of reverse primer and 2.2 µl of DNA template. A master mix for all reaction components (except DNA template) was created and 12.8 µl of the mix was added to a well on Bio-rad iQ™ 96-well PCR plate (Hercules, CA). DNA template was then added and the plate was sealed with Bio-Rad Microseal® 'B' Adhesive Seals (Hercules, CA). The sealed plate was centrifuge for 1 min at 1000 rpm (J6-MI, Beckman, Brea, CA). The conditions for real-time PCR were: initial denaturation at 95 ℃ for 5 minutes, 40 cycle of amplification (denature at 98 ℃ for 30 seconds, annealing at 61℃ for 30 seconds and extension at 72℃ for 30 seconds) and melt curve from 55 to 95 ℃, at increment of 0.5℃/cycle for 5 seconds. Real-time PCR reactions were conducted in duplicate for each DNA extraction and results were expressed as quantification cycle (Cq). Cq value is determined as the cycle number at which fluorescence has increased above threshold. PCR reaction without DNA template or "no template control" (NTC), was used as blank.

3.2.2 Gel Electrophoresis

Gel electrophoresis was performed to verify the purity of PCR products. A 1.5% (w/v) agarose gel was made with Ultrapure DNA grade agarose in 1X Tris/Boric Acid/ EDTA (TBE) buffer (Cellgro, Manassas, VA). Ethidium bromide (EtBr, Bio-Rad) was added to the agarose gel mixture for staining nucleic acids and the final concentration in gel was 0.5µg/ml. For each gel, all extracted DNA and amplified products from both soybean and green pea were loaded. For loading DNA samples, 10 µl of extracted DNA was mixed with 8 µl of TBE and 2µl of 5x loading buffer before adding to the agarose well. A 100 bp Molecular Ruler (Bio Rad) was also loaded for verification of DNA sizes. Gel was ran using Bio-Rad Wide mini-sub cell GT system with Bio-Rad power PAC 300 power supply at 140 V for 45 minutes. Gel electrophoresis was ran twice and results were visualized and scanned using Biorad GelDoc^{TM} XR and Quantity one® software (Hercules, CA).

3.3 EVALUATION OF PRIMER ON A VARIETY OF SOY PRODUCTS

Soybean processing involves a series of steps to produce a wide variety of products for food uses. These steps typically involve traditional milling processes such as milling, rolling, and solvent extraction of oils. They also involve processing steps that remove soluble sugars through various leaching processes. The ability of the primer to produce an amplified product from a range of soy products, was evaluated with ground soybean, soy flour, soy protein concentrates, soy protein isolates and soy lecithin. The ground soybean was prepared for analysis by freezing with liquid nitrogen and then reducing it to a powder using a waring blender (blender jar was kept frozen until use, blending was conducted in a cold room). The resultant powder was transferred to a sterile sampling container and stored at room temperature. In addition, de-fatted soy flour (Cargill Prolia ^{TM} 200/70), soy protein concentrate (ADM Arcon F) , Soy protein fiber lecithin (ADM Protein Fiber Lecithin) and soy protein isolate (Solae ISP ADM, Ardex F Dispersible and Cargill Prolisse) were all

selected as examples of processed soy products. All DNA extractions were repeated three times for each sample type and PCR reactions were conducted in duplicates. DNA concentration and purity (A 260/280) were measure using NanoDrop 1000 spectrophotometer (Wilmington, DE) for extracted DNA. Gel electrophoresis was run on just the extracted DNA. Description of DNA extraction conditions, PCR methods and gel electrophoresis are described below:

3.3.1 DNA extraction and real-time PCR

DNA was extracted using Qiagen QIAamp® DNA Stool Mini Kit (Valencia, CA) following the protocol provided for isolation from stool for pathogen detection. DNA collected was stored at -20℃ after extraction. Detection of soybean DNA was performed using Bio-Rad MyiQ Single Color Real-Time PCR Detection System (Hercules, CA) with Bio-Rad iQ™ SYBR® Green Supermix. The real-time PCR protocol developed by Tate and others (2009) was followed. Description of the protocol was described in the previous section.

3.3.2 Gel electrophoresis

Soy samples selected for evaluation differ in degree of processing and thus the quality and concentration of the extracted DNA was variable from these products. Gel electrophoresis was performed on extracted DNA to visualize the differences among samples. A 1.5% (w/v) agarose gel was made with Ultrapure DNA grade agarose (sigma Aldrich) in 1X Tris/Acetic Acid/ EDTA (TAE) buffer. Ethidium bromide (EtBr, Bio-Rad) was added to the agarose gel mixture for staining nucleic acids and the final concentration in gel was 0.5µg/ml. The TAE buffer was diluted from a 50X stock solution made up of 242g of Tris base (sigma), 57.1mL of glacial acetic acid (fisher scientific), 100 ml of 500nM EDTA (pH 8.0) and brought up to volume (1L) with deionized water. The 500nM EDTA solution was prepared by dissolving 93.06 g of EDTA (sigma Aldrich) with 300 mL deionized water in a 500 mL volumetric flask,

adjust pH to 8.0 with 0.1N NaOH and bring it up to volume. The 50X stock solution was mixed, autoclaved and kept at room temperature. For each gel, all extracted DNA and a100 bp Molecular Ruler (Bio Rad) were loaded. For loading DNA samples, 6 µl of extracted DNA was mixed with 2µl of 2x loading buffer before adding to the agarose well. Gel was run using Mini-Sub® Cell GT Cell with Bio-Rad power PAC 300 power supply at 80 V for 1 hour. Gel electrophoresis was ran three times and results were visualized and scanned using Biorad GelDoc™ XR and Quantity one® software (Hercules, CA).

3.4 OPTIMIZATION OF REAL-TIME PCR

Due to the replacement of equipment, modification of the previously described protocol (developed by Tate and others 2009) was carried out and optimization of the new PCR protocol was needed. The newly acquired Bio-Rad CFX96 ™ Multiplex Real-Time PCR Detection System (Hercules, CA) has been proclaimed to have a better thermal control and faster scanning detection. Also, a new high-performance SsoAdvanced™ SYBR® Green Supermix was evaluated. The supermix has shown to increase sensitivity and efficiency in detection with shorter running time. As the reagent of the real-time PCR reaction changes, optimization of the new protocol is essential for accurate and reproducible quantification. Modification for the protocol was needed to incorporate the suggested protocol for the new supermix. Since the new supermix has a faster reaction time, a higher concentration of primer pairs was required. Also, the annealing and extension steps are combined and thus annealing time and temperature need to be investigated. The suggested reaction time was between 10-30 seconds and a temperature gradient close the adapted protocol (61 ℃) was chosen. DNA extracted from soybean (Qiagen DNeasy mericon Food Kit) was used as sample. A temperature gradient of 60-64 ℃ was conducted with primer concentration (300-400 nM with increment of 50nM) to verify the annealing temperature. Afterward, 2 different reaction times (10 and 30 seconds) were evaluated with primer concentration of (300-400 nM with increment of 50nM). Furthermore, a

series of primer concentrations (ranging from 300nM to 500nM with increment of 50nM) were tested. Lastly, the protocol was verified with SPI and soybean. DNA extracted from SPI and soybean was serial diluted (200,000 to 2 ppm) with deionized water and the real time PCR reaction was carried out as describe below. A standard curve was constructed with soybean DNA and SPI DNA to verify the optimized protocol.

3.4.1 Real-Time PCR protocol

The real-time PCR amplification was carried out in a 20 µl reaction mixture with the following components: 10 µl Bio-Rad SsoAdvanced™ SYBR® Green Supermix, 5.2 µl of sterile nanopure water, 1.4 µl of 350 nM of forward primer, 1.4 µl of 350nM of reverse primer and 2 µl of DNA template. A master mix was created to include all components (except DNA template) and 18 µl of the mix was added to a well on Multiplate™ Low-Profile 96-Well unskirted PCR Plate (Hercules, CA). DNA template was then added and the plate was sealed with Bio-Rad Microseal® 'B' Adhesive Seals (Hercules, CA). The sealed plate was centrifuge for 1 min at 1000 rpm (J6-MI, Beckman, Brea, CA). Real-time PCR was performed under the following conditions: Initial Denaturation at 98 ℃ for 2 minutes, 40 cycle of Amplification (denature 98 ℃ for 5 seconds, annealing and extension at 61℃ for 20 seconds) and melt curve from 55 to 95 ℃, at Increment of 0.5℃/cycle for 5 seconds. Real-time PCR results were expressed in amplification graph and recorded as quantification cycle (Cq). PCR reaction without DNA template, No template control (NTC), was used as blank.

3.5 EVALUATION OF COLUMN VS. MAGNETIC BEADS DNA EXTRACTION METHODS FOR SOY

After verifying the primer design and real-time PCR protocol, two different types of DNA extraction methods were evaluated for soy. Research has shown that

both column and magnetic bead extraction method can be used to extract DNA from food samples. The column based extraction method followed a bind-wash-elute procedure to extract and recover DNA from food sample. Silica based membranes are normally used in the column to bind DNA. Real time PCR was performed and Limit of Detection (LOD) was determined for both kits. Percent PCR recovery was obtained through dilutions performed for LOD.

SPI and soybean were used as sample and DNA extractions for each method were repeated three times. During sample extraction, samples were serially diluted with lysis buffer to cover concentration range of 2000 ppm to 2 ppb (for column) and 1333 to 1.3 ppb (for magnetic beads). Afterward, DNA recovery was performed on all samples and dilutions using both extraction methods. DNA concentration and purity (A 260/280) were measure using Eppendorf BioPhotometer Plus with Hellma Traycell (Hauppauge, NY) for extracted DNA. PCR reactions were conducted in duplicates for all recovered DNA. Standard curves were generated for both samples and extractions. Description of the both DNA recovery kits with serial dilution for LOD and PCR methods are described below.

3.5.1 Column type DNA recovery method

The protocol for a 200 mg sample was followed with modifications that were needed to determine the limit of detection (LOD). A 200 mg of sample was digested with 1mL of lysis buffer and 2.5 µL Proteinase K solutions. Sample was incubated for 30 min at 60 ℃ with constant shaking at 150 rpm (Precision Reciprocal shaking bath 25, Thermo scientific, Waltham, MA). After incubation, the sample was cool down to room temperature, centrifuged at 4000 rpm for 15 min (J6-MI, Beckman, Brea, CA) and supernatant was collected. The supernatant was first diluted (1:100) and then serially diluted (1:10) with lysis buffer in order to create a sample dilution range from 2000 ppm to 2 ppb sample. Afterward, a 700 µL aliquot from each dilution was obtained and combined with 500 µL chloroform, vortexed for 15 seconds, and

centrifuged at 12000 rpm for 15 min. A 350 µL aliquot from the upper aqueous phase was combined with 350 µL PB buffer, vortexed and transferred to QIAquick spin column with collection tub. The column was centrifuged at 13,000 rpm for 1.5 min and flow-through was discarded. After that, 500 µL buffer AW2 was added to the column and re-centrifuged at 1300 rpm for 1.5 min. Flow-through was discarded, and the spin column was centrifuged again at 13000 rpm for 1.5 min to dry the membrane. A fresh collection tube was placed on the spin column and 150 µL EB buffer was added directly to the membrane in the column. The spin column was allowed to incubate at room temperature for 1 min and then centrifuged at 14000 rpm for 1 min to release the DNA. Recovered DNA was then stored at -20 ℃.

3.5.2 Magnetic bead DNA recovery method

The Promega Wizard Magnetic DNA Purification system kit (Madison, WI) uses MagneSil™ paramagnetic particles (MPP) as affinity matrix and goes through a mobile solid phase extraction to obtain DNA. To establish the LOD, modification of the 200 mg protocol was needed. For 200 mg protocol, 250 µL food lysis B buffer, 500 µL Lysis buffer A and 5 µL RNase A was added to 150 mg of sample and incubated at the same condition. After incubation, 3.75ml precipitation solution was added to the sample and centrifuge at 4000rpm for 20 minutes. To create sample dilutions for LOD, digested sample was diluted (1:100) with lysis buffer and then serially diluted (1:10) to cover concentration range from 1,333 ppm to 1.3 ppb. The lysis buffer was a combination of lysis buffer B, lysis buffer A and precipitation solution (1:2:3, respectively). Afterward, 700 µL of sample dilution was used for recovery with 50 µl of MPP added to the mix. A 560 µL aliquot of isopropanol was added and incubated for 5 min at room temperature. Liquid waste was discarded after inserting the tube to a magnetic stand for 2 min. Afterward, 250 µL of Lysis B was added, mixed gently and liquid waste was discarded. A 1 mL aliquot of 70% ethanol was added to resuspend the MPP and liquid waste was discarded. The 70% ethanol wash was repeated 3 times and then incubated for 10 min at 65 ℃. After incubation,

100 µl of DNase-free water was added and heated to release DNA from MPP. Recovered DNA was then stored at -20 ℃. Furthermore, the magnetic beads were designed to interact with DNA in a suspension and thus the sample size can easily be adjusted. A larger sampling size was used to investigate if sample size could be a factor in increasing detection sensitivity. Therefore, 1g of sample was used and modification of the protocol was needed. For the 1 g sample, 1.25 mL food lysis B buffer, 2.5 ml Lysis buffer A and 25 µL RNase A was added during digestion. The MPP used remained the same but sample size after dilution and isopropanol increased to 3 mL and 24mL (respectively). Both lysis buffer B and 70% ethanol wash increased to 1.25mL and 5mL, respectively. The amount of DNase-free water remained the same and recovered DNA was then stored at -20 ℃.

3.6 EVALUATION OF BOTH DNA RECOVERY METHODS IN FOODS

3.6.1 Sample preparation, DNA extraction and Real-time PCR on Soy milk

Soy milk was obtained from the local grocery store and stored at 4 ℃. The selected soy milk, refrigerated Silk original contains roughly 2.5 % of protein, 1.5% total fat, and 3.3% total carbohydrates. To generate a LOD, 1mL of soy milk was first diluted with 99 mL of deionized water (1:100) and then serially diluted (1:10) to cover concentration range of 1000ppm to 100 ppb. DNA was then extracted using both Qiagen DNeasy Mericon Food Kit and Promega Wizard Magnetic DNA Purification system kit. For both kits, 200 mg protocol provided by the manufacture was followed and no modification was implemented. Recovered DNA was collected and stored at -20 ℃. DNA concentration and purity (A 260/280) were measure using Eppendorf BioPhotometer Plus with Hellma Traycell (Hauppauge, NY) for extracted DNA. Afterward, DNA amplification was performed using Bio-Rad CFX96 ™ Multiplex real-time PCR following the modified protocol used in previous section. A standard curve was constructed to assess the limit of detection (LOD). Percent PCR recovery was calculated based on the changes in actual Cq values and the expected Cq values

through a series of dilutions. For extracted DNA (soy and SPI), a serial dilution (1:10) with sterile nanopure water was carried out after DNA was recovered. However, for soy, SPI and soymilk, no dilution was performed.

3.6.2 Sample preparation, DNA extraction and real-time PCR on food sample

To determine the matrix effect, 4 different types of food sample (high carbohydrate: wheat flour batter, high fat: soy-free ranch dressing and high protein: 93:7 ground beef and water) was spiked with 10 % soy protein isolates (SPI). Store brand ground beef (93:7) and Hidden Valley original ranch dressing (soy-free) were obtained from the local grocery stores. The deep fried batter mix was prepared with 3:6 solid (wheat flour) to water ratio. In this case, instead of adding additional 10% of soy SPI, the amount needed was used to replace part of the wheat flour. Furthermore, a control sample (soy) was prepared by suspending 10% SPI in water. For each food sample, 20 g of SPI was added to 180 g of sample and homogenized using a food processor at low speed for 1 minute. Afterward, each food was divided into 2 equal portions and one of them was subjected to thermal treatment (water bath at 95ºC) for 1 hour.

After heat treatment, samples were cooled to room temperature and extracted using both DNA recovery systems. For each system, extractions were carried out using the 200 mg protocols provided by manufacture. Recovered DNA was collected and stored at -20 ℃. DNA concentration and purity (A 260/280) of extracted DNA were measure using Eppendorf BioPhotometer Plus with Hellma Traycell (Hauppauge, NY). Afterward, detection was determined using Bio-Rad CFX96 ™ Multiplex real-time PCR amplification following protocol used in detecting soy DNA in soy milk. All DNA extractions were repeated three times for each sample and PCR reactions were conducted in duplicates. To compare against ELISA method, Cq values were converted to concentration using regression equations generated by standard curves constructed for SPI using both DNA recovery methods.

3.7 COMPARING DNA DETECTION METHODS WITH ELISA METHOD

3.7.1 Sample preparation and ELISA test on soybean, soy protein isolates and soy milk

Approximately 5 grams of sample (Soybean, SPI, soy milk) was added to 125 mL of 10 mM buffered salt solution (PBS) and heated to 60 ℃ for 15 minutes in Thermo PRECISON reciprocal shaking bath (model 25; Waltham, MA) at 150 rpm. Samples were digested, filtered and serially diluted (1:10) with PBS buffer to cover concentration range of 4000 ppm to 4 ppb. Samples and controls (0, 10, 25, 50 and 100 ppm soy protein isolates) were subjected to the testing protocol provided by the manufacture. The reactions were read using Neogen Stat Fax 4700 microwell reader with a 650 nm filter. The optical densities (absorbance) of controls were used to construct the standard curve, and the sample absorbances were plotted against the curve to calculate the estimated concentration of soy protein (expressed as soy protein isolates). All protein extraction was conducted in duplicate.

3.7.2 Sample preparation and ELISA test on food samples

For this experiment, the same eight samples (4 different food samples and 2 different treatments) prepared for PCR methods were used. The ELISA protocol provided by manufacture was followed. After digestions, filtered sample was diluted (1:100) with deionized water and another dilution (1:10) was carried out to cover a sample concentration of 40 and 400 ppm. All protein extraction was conducted in triplicate. Percent recovery based on expected concentration was calculated.

3.7 Statistical Analysis

For primer verification and real-time PCR optimization, means and standard deviation were calculated. Statistical results presented for primer verification and

primer concentration for real-time PCR optimization and percent recovery for detecting soy in foods were determine by Two-Way Analysis of Variance (ANOVA) with a significant level of (p<0.05). In the meantime, Three-Way ANOVA with a significant level of (p<0.05) was used for optimization of annealing temperature and time. Statistical analysis was performed using SigmaPlot (Version 11.1, Systat Software, Inc., Chicago, IL).

PROC GLIMMIX of SAS (Version 9.2, Cary, NC) with a split-plot design was applied for real-time PCR and ELISA results for detecting soy in food. Treatments (heat or no-heat) were considered as the main plot while food matrices (meat, ranch, flour and soy) were use as subplot. All tests were done at the 0.05 level of significance. Since the data set for ELISA was substantially skewed, a Box-Cox transformation was used to find the best potentially nonlinear transformations of a dependent variable. Thus, a Log transformation was performed on concentration from ELISA method.

Chapter4

RESULTS AND DISCUSSION

4.1 EVALUATION PRIMER DESIGN

Previous research conducted by Tate and Dewitt (2008) developed a real-time PCR protocol to identify soy DNA targeting soy lectin 1 gene (le1). The protocol was developed with SYBR® Green I dye as florescent monitoring agent. Since SYBR® Green I dye binds to any double-stranded DNA presented, it is important to verify the specificity of the primer design. The Basic Local Alignment Search Tool (BLAST) screen conducted by Tate and Dewitt (2008) indicated that the primer sequence was specific to soy le 1. For further verification, green pea was chosen as a negative control to demonstrate the specificity of the primer design. Green pea and soybean are both legumes and have lectin gene in the root and seed (Hirsch 1999). Both lectin genes have been used as target gene in its respective source for real time PCR detection (Brezna and others 2006a; Espineira and others 2010; Meyer and others 1996; Soares and others 2010).

4.1.1 Specificity of primer design

As shown in table 4.1, no detection was found in the negative control (green pea) sample. The result indicated that no amplification occurred throughout the 40 cycles of real-time PCR and the primer set did not amplify the lectin gene in green pea. Gel electrophoresis was performed on both DNA template and amplified products to further confirm the result. Smeared DNA band was found in Figure 4.1 for both extracted soy and green peas. The smeared DNA indicated that DNA was successfully extracted from the samples. The diluted product from amplified product (for soy) indicate a strong band around 300bp (lane 4) while no visible band was observed in the diluted sample for pea (lane 3). Thus the result indicated that DNA was successfully extracted from both soy and green pea and amplified PCR product from soy was the only detectable product.

4.1.2 Versatility of primer design

4.1.2.1 DNA concentration and purity

The commercial forms of soy available for use in the food industry represent everything from mechanically processed soy flour to chemically processed soy protein. As a result, the primer designed for detection of soy needs to be robust enough to detect the presence of any form of soy used for food formulation. The DNA was extracted from different commercial forms of soy and concentration, purity and integrity were evaluated prior to real-time PCR testing. Concentration and purity of the extracted DNA are presented in table 4.2. The concentration of DNA recovered from soybean (207.1 ± 28.18 µg/µl) was significantly lower (p <0.05) than from all other tested products. The highest concentration of DNA was extracted from Ardex F (1650.9 ± 179.83 µg/µl) and Solae (both Soy Protein Isolates; 1505.6 ± 273.52 µg/µl; p<0.05). This was expected because DNA extraction is highly depending on particles size. Research have demonstrated that as particle size in a sample decreases, the DNA yield increases (Demeke and Jenkins 2010). Since soybean was frozen with liquid nitrogen and reduced to a powder formed using mortar and pestle, the particle size was higher than commercially refined products. As for DNA purity (A260/280), all soy samples were above 1.8 which is considered high DNA purity (Espineira and others 2010). Other than Prolia (2.10 ± 0.045), no significant differences were observed in purity of extracted DNA from soy products (p >0.05). High DNA purity is desirable for real-time PCR detection and it should be between 1.8 and 2.0. The results reported were slightly higher than 2.0. Normally, a low value in DNA purity can be caused by protein contamination or incomplete extraction. No negative correlation was linked to A260/A280 value slightly higher than 2.0. Since measurements were based on wavelength absorbance, slight shifts in wavelength can result in the higher values (Nanodrop 2007). Furthermore, integrity of the DNA was evaluated using gel electrophoresis. As shown in Figure 4.2 and 4.3, visible smeared bands (from lane 1-7) were observed for all extracted soy samples. During DNA

extraction, degradation of DNA can occur and thus producing DNA fragments of different sizes. Smeared bands in agarose gel indicate the spreading of fragmentized DNA within sample. Thus the integrity of the extracted DNA was affected by extraction. Also, the intensity of band indicates the concentration of the DNA at a specific size. Since no amplification was performed, the intensity would be low.

4.1.2.2 Real-time PCR results

The forward and reverse primer set was designed to detect the Soy lectin gene in soybean. A positive detection of the amplified product as measured by the Cq value indicates the ability of the primer to detect the presence of soy Also, for SYBR Green based amplicon detection, it is critical to also run a dissociation curve or melting curve following real-time amplification because SYBR Green is a nonspecific binding dye that binds to all double stranded DNA (Mackay and Landt 2004). A melt-curve analysis can be used to detect any undesirable double stranded DNA that includes primer dimers, contaminating DNA, and PCR product from misannealed primer (Kennedy 2011). Figure 4.4a represents the amplification curves produced for all products tested from one DNA extraction replication. Amplification was detected for each soy product. In addition, Figure 4.4b shows the resultant melt curve analyses from each amplified product. Since there is only one peak at a single temperature, the melt curve analysis demonstrates that no undesirable double stranded DNA was present in the amplified product. Successful amplification and melt-curve evaluation demonstrates the versatility and robustness of the designed primer set.

Amplification curves are interpreted by determining the cycle number (Cq) at which fluorescence begins to increase exponentially. Products with a low Cq value are interpreted as having more initial DNA template for the primer to target. The Cq values from all soy products tested are reported in Table 4.2. The highest Cq value ($p<0.05$) was reported for PFL (lecithin) and the lowest for Prolisse (SPI) which indicates there is more target DNA in that particular product (SPI). However, Cq values are very similar to each other and results generally demonstrate that the primer was specific enough to detect soy DNA from all products.

4.2 OPTIMIZATION OF REAL-TIME PCR PROTOCOL

4.2.1 Annealing temperature

Modification of the real-time PCR protocol previously developed by Tate and others (2009) was required due to an upgrade of the real-time PCR detection system. Optimization of the real time PCR protocol was carried out with DNA extracted from soybean following guidelines provided by BioRad for SsoAdvanced™ SYBR® Green Supermix.

Annealing temperature was optimized by evaluating three different primer concentrations. At low primer concentration (300nM), no amplification was observed for all temperatures investigated. When primer concentration was increased from 300 nM to 350nM, Cq values were reported for annealing temperature at and above 60.8ºC. Real-time reaction performed with high primer concentration (400nM) resulted in Cq values reported for all temperatures evaluated. At primer concentration of 350nM, the differences between annealing temperature were not significant (p>0.05). But at higher primer concentration (400nM), Cq values were highest at annealing temperature of 60ºC (26.42 ± 4.134) and lowest at annealing temperature of 64ºC (22.10 ± 0.183). The annealing temperature of 60 ºC was significantly different from almost all but annealing temperature of 61.6ºC (p<0.05). There was a primer concentration effect, however, no differences existed between primer concentrations of 350nM and 400nM when the temperature was above 60.8 ºC (p>0.05). As a result, the combined means of annealing time was reported for this segment of the data. For the combined means, the Cq result was highest for 62.6 ºC. However, Cq values need to be optimized for the lowest value. Melt curve analysis was conducted on all amplified products from the temperature and concentration optimization and no non-specific products were amplified (Figure 4.5). Since the real-time PCR protocol developed by Tate and Dewitt (2008) called for an annealing temperature 61 ºC, it was decided that this temperature continue to be used in the modified protocol.

4.2.2 Annealing time

The super mix utilized for the modified protocol was developed to anneal and extend amplified DNA in less time than the previous protocol. The recommended annealing time was between 10 and 30 seconds, as a result annealing time was evaluated for optimization at 10, 20, and 30 seconds at 3 different primer concentrations (Table 4.4). A Cq value was reported for all primer concentration levels evaluated and annealing times indicating successful amplification. The amplification graph for each annealing time is given in Figure 4.6. Overall, primer concentration has no significant effect within each annealing time ($p>0.05$). However, Cq values were significantly influenced by annealing time ($p<0.05$). Among the three, annealing time of 10 seconds has the highest average Cq value while annealing time of 20 seconds has the lowest average Cq value and standard deviation. As a result, annealing time of 20 seconds was chosen for the new real-time PCR protocol.

4.2.3 Primer Concentration

With the annealing conditions optimized, a broader range of primer concentrations (300-500 nM) were evaluated. All levels of primer concentration demonstrated strong signals for detecting soy DNA (Figure 4.7a) and no non-specific amplification was detected (Figure 4.7b). Furthermore, Cq values obtained using primer concentration 300-500nM ranged from 27.28 ± 0.29 (450nM) to 27.45 ± 0.29 (300nM). As indicated in Figure.4.8, no significant differences were observed for all real-time PCR results ($p>0.05$). Among the primer concentration, 350nM had the lowest standard deviation. The low standard deviation indicated that the real-time PCR reaction was the most consistent at primer concentration of 350nM. Thus, the final primer concentration of 350nM was chosen for the new real-time PCR protocol.

4.3 EVALUATION OF ELISA AND TWO DNA EXTRACTION METHODS (COLUMN-TYPE VS. BEAD-TYPE)

Studies have demonstrated that matrix interference and glycation can interfere with the detectability of commercial ELISA kits, and a more robust system is needed (L'Hocine and others 2007; Platteau and others 2011). Even so, ELISA is still the primary test for soybean allergen in the food industry. *Neogen Veratox*® (Lansing, MI) for soy allergen, quantitative sandwich-ELISA test, was chosen to compare against the developed DNA method to determine the LOD in soybean, SPI and soy milk.

4.3.1 ELISA

ELISA test was performed for soybean, SPI and soymilk sample. Standards provided by the manufacturer were used to generate a standard curve. The concentration results presented in Table 4.5 were calculated based on the standard curve for each product. Due to the limitation in equipment, detection can only be detected and reported from sample at the lower concentration range (0 to 100 ppm). The r^2 values (0.969- 0.978) suggested that the estimation based on standard curves were reliable. At the lowest concentration (4 ppm), soybean and SPI were overestimating the results while no detection was found for soymilk. At 40 ppm, results from soybean were closer to the expected concentration while higher in SPI and soymilk. Thus, the results suggested that the ELISA kit can overestimate the concentration of soy in samples. Similar observation was reported in an inter-laboratory validation study for peanut that focused on 5 different types of commercial ELISA kit (Poms and others 2005). The authors concluded that recoveries (or percent estimation of concentration) had a spread of 44-191% in concentration. Even though the ELISA kits were not designed for soybean, the results suggest that variations within ELISA kits can be high. Furthermore, another study conducted by L'Hocine and other (2007) reported that high variation within percent recovery occurred in solution spiked with soy protein. Two commercially available ELISA kit were tested

and percent recovery ranged from 86-948%. Even so, acceptance levels of protein recovery were restricted to 77% to 120%.

4.3.2 DNA extraction method comparison

Real-time PCR is highly specific, reproducible, sensitive and rapid in processing time. Taking into consideration the differences in type, composition and degree of processing, an efficient DNA extraction step would be crucial for a PCR detection system (Gryson and others 2004; Olexova and others 2004). Two different types of DNA extraction methods designed for food products were evaluated for their ability to recover DNA from the soy products. Both types of extraction methods have been suggested to produce high quality extracted DNA. In order to compare extraction method ability to recover DNA, soybean, SPI and soy milk were selected to represent high (99%), medium (40%), and low (2-3.5%) protein content soy products. In addition, SPI is an example of a highly processed product, while soybean is an example of an unprocessed sample. The SPI chosen, AMD Ardex F, had the highest DNA concentration and its average Cq value was comparable to soybean and can also be easily dispersed in water or mixed in food matrices. Soy milk represented not only a low protein product, but a product that was easily dispersed in water or mixed in food. Aside from evaluating DNA extraction, percent PCR efficiency and Limit of detection (LOD) was performed to determine the sensitivity of the developed real-time PCR protocol. Finally, a standard curve was constructed to check for linearity and relate concentration and Cq values.

4.3.2.1 DNA concentration and purity

Total DNA concentration and DNA purity of soybean and SPI were measured and presented in Table 4.6. Soymilk sample was diluted prior to extraction and thus the concentration was below the detection limit. Overall, total DNA extracted with column-DNA has high purity (A260/280) values (1.85 to 2.14) and similar concentration for both samples; 35.6 - 115.1µg/µl for soybean and 63.5- 171.7µg/µl

for SPI. On the other hand, total DNA extracted from magnetic beads-DNA was higher in concentration for soybean (444.9 -485.9µg/µl) but lower in DNA purity for both soybean (1.64-1.76 µg/µl) and SPI (1.76-2.05 µg/µl).

4.3.2.2 Amplification and Limit of Detection

Amplification plots were generated for each DNA extraction (Appendix 1.1 and 1.2). Each plot represents a 10-fold serial dilution of DNA extract recovered from Table 4.7 and 4.8. Table 4.7 demonstrates detection of soy DNA in soybean, SPI and soymilk at concentrations ranging from 2 to 2000 ppm using the column type DNA recovery method. Table 4.8 demonstrates amplification detection at concentrations ranging from 1.33 to 1330 ppm for samples extracted with the magnetic bead DNA recovery method. The subsequent melt-curve analysis (Appendix 1.3 and 1.4) indicated there were no non-specific amplified products. The real-time PCR developed was set to have 40 cycles of amplification. If the concentration of targeted DNA template was low, amplification might not occur within the designed amplification cycle. Thus, as indicated in Table 4.7 and 4.8, results that were beyond detection capability were stated as with Cq values >40. Cq values for column DNA method were detected samples with concentration ranging from 2 to 2000 ppm. However, at 2 ppm, only some of the extractions were detected. For magnetic beads, Cq values was detected for all but one replication in soybean. Since amplification did occurred on the same sample, the results could be due to an operator error. Cq values for SPI were detected for all sample with concentration ranging from 1.33 to 1333.3 ppm and only a few at 1.33 ppm. As for soymilk sample, Cq values were detected for concentration ranging 13.3 to 1333.3 ppm and only one detection from each extraction for concentration of 13.3 ppm. Overall, column method was consistent at higher concentration (more than 20 ppm) for all samples. However, the magnetic bead method was effective in soybean at low concentration (1.3ppm) but not in soymilk.

Limit of detection (LOD) is defined as the lowest concentration at which 95% of the positive samples are detected. Table 4.7 demonstrates that LOD for all 3 samples using the column-DNA recovery method is 20ppm. On the other hand, real time PCR results on magnetic-bead DNA recovery method (Table 4.8) indicated that LOD for soymilk and SPI is 133.3ppm and 13.3 ppm (respectively) while even lower for soybean (1.33ppm).

LOD have been reported to vary between food matrices. Known amount of soy protein have been added to food matrix to help determine the LOD. For example, through serial dilutions of known soy in food matrices, LOD for powdered peanut and whole milk was 100 ppm while LOD for powdered skim milk, wheat flour and mung bean was 10 ppm using CTAB extraction method and traditional PCR detection (Wang and others 2012). Another PCR detection reported by Yamakawa (2007) indicated that a similar LOD (10 ppm) was detected when wheat flour was spiked with soy flour and extracted using a silica-based extraction kit. Thus the results presented were either in agreement or lower with PCR detection methods designed to detect Soy DNA.

4.3.2.3 Percent PCR efficiency

In theory, a DNA product can be amplified and doubled in each cycle. However, primer–dimer and GC-rich regions of the template and the heating/cooling ratio of the PCR system may interfere with the efficiency of the PCR (Li and others 2005a). Hence, percent PCR efficiency is used to describe the ability of the amplification to double its product at each cycle. Non-detectable PCR results were omitted from the calculation. The percent PCR efficiency for extracted DNA from soybean (Table 4.9) were 100.2% for Column method and 98.3% for magnetic beads method with r^2 values of 0.987 and 0.991 (respectively). For SPI samples, the percent PCR efficiency were 105.6% for column method and 99.8% for magnetic beads method with r^2 values of 0.989 and 0.995 (respectively). Ideally, the percent PCR efficiency should fall between 90-110% and has correlation coefficients (r^2 values)

closer to 1. Thus the results suggested that the developed assay is precise for detecting Soy DNA in soybean and SPI. However, the results did not account for the efficiency of the actual DNA recovery method.

As listed in table 4.9, PCR efficiency for 200 mg SPI was within the ideal range (90-110%) with high r^2 values of 0.979 ± 0.007 (for column DNA) and 0.954 ± 0.067 for magnetic beads DNA method. The PCR efficiency for both extraction methods with 200 mg soybean were 101.7± 16.49% with r^2 values of 0.979 ± 0.007 (column-DNA) and 101.7 ± 16.49% with r^2 values of 0.954 ± 0.067 (magnetic beads-DNA). Even though the percent PCR efficiencies for soybean were not ideal, it was still acceptable and the r^2 values were mostly high. Soymilk samples showed high PCR efficiency greater than 105% for both column-DNA and magnetic beads-DNA method. The percent PCR efficiency for magnetic beads DNA method for 1 g sample was 98.85 ± 15.9% for soybean with r^2 value of 0.833 ± 0.101 and 103.4 ±13.4% with r^2 value of 0.906 ± 0.011 for SPI. Thus, increased in sampling size had no effect on PCR efficiency. In this part of the study, PCR efficiency results were based on DNA recovered from a serial dilution. Since the developed PCR assay is precise, the results can be used to describe the efficiency of the DNA recovery method. As the results suggested, the efficiency of both recovery methods can be affected by sample type. Overall, SPI sample had the lowest variations within extraction and is closest to a percent PCR efficiency of 100% for both DNA recovery methods.

4.3.2.4 Standard Curves

To further determine the relationship between real time PCR results and concentration for both DNA recovery methods, a standard curve for the combined data was constructed. Real-time PCR results, listed in Table 4.7 and 4.8 (soybean, SPI and soymilk), were plotted against the log of sample concentration. Results that were non-detectable were omitted. As shown in Figure 4.9, standard curve plots were generated from each sample type. Within each plot, results from both DNA recovery methods were plotted. For each set of results, a linear regression was plotted to construct the standard curve. An equation with r^2 value was generated from the regression. The

correlation coefficient (r^2 value) is the fraction of the variation that is shared by both Cq values and concentration. High r^2 value suggested that the variation is low and thus the equation generated can be used to estimate the unknown (Colton and Bower 2002).

Overall, the r^2 values for all samples extracted were above 0.8 for both extraction methods. Thus the high r^2 values suggested that the regression equations generated can be used to predict soy concentration based on Cq values. For the magnetic beads DNA recovery method, results for 1 g samples were presented in table 4.10 and plotted in Figure. 4.10. Standard curve was constructed for both samples and regression equation was generated. The r^2 values for soybean and SPI were both above 0.89. Such observation suggested that increasing the sampling size showed no significant changed in sampling variations.

Variations for real time PCR can occur at DNA extraction or during real-time PCR detection. Since sample dilution occurred at DNA extraction or recovery, a slight error could affect the downstream process. Furthermore, SsoAdvanced™ SYBR® Green Supermix required a final concentration of 50ng-5pg of DNA template in each reaction for amplification to occur. Thus if the concentration of extracted DNA is low, real-time PCR amplification will not performed. Also, sampling and pipetting error could contribute to the variation.

In summary, column DNA recovery method can be used to consistently detect DNA in soybean, SPI and soymilk down to 20 ppm while magnetic beads DNA method demonstrated a lower detection for soybean sample (1.33 ppm) but higher concentration detection for soymilk (133.3 ppm). Also, regression equations generated can be used to predict soy concentration with detected Cq values.

4.5 DETECTION OF SOY IN FOODS

Thus far, the evaluation of both ELISA and DNA methods have been on soy-only (except soy milk) samples and such evaluations do not represent the ability of methods to perform in the presence of a food matrix. ELISA kit methods have been

demonstrated to be negatively affected by both heat and food matrices (L'Hocine and others 2007; Platteau and others 2011). In addition, previous research suggests that different food matrices (such as ground beef, hotdog or breaded fish sticks) could also affect the efficiency of DNA recovery (Holzhauser and Vieths 1999). Therefore, evaluations of all recovery methods and the detection protocol are needed using food samples.

Since soy has been widely used as functional ingredient, it can easily be found in a processing plant and cross-contamination has been one of the major concerns. Also, further processing after addition of soy can pose a problem in detecting the allergenic compound. Each food matrix (protein, fat, carbohydrate, or water) was chosen based on its composition and possible interaction with soy products. For example, soy is often added to meat products as binders and thus ground beef was selected as a food system with high protein. Also, ranch dressing typically consist nearly 50% of fat and sometimes may contain soy lecithin as emulsifier or soy protein to enhance flavor. Thus a soy-free ranch dressing was chosen to represent food with high fat. In addition, soy flour can be added to deep fried meat/fish product batters. Results for soy-free Ranch (Ranch), ground beef (Meat), flour batter (Flour) and Water (Water) were compared and are explained below.

4.5.1 DNA and ELISA results on detected concentration

The detected SPI concentrations using ELISA and DNA (column and magnetic beads) methods in heated and non-heated food matrices were presented in Table 4.11. Real-time PCR results (appendix 1.8) were converted to concentration (ppm) using standard curves constructed for SPI sample in the previous section. ELISA results were calculated based on the regression equations generated for each food sample type using standards provided by the kit. Also, ELISA results were presented in log transformation means because the residual analysis of concentration means did not fit the normality.

Overall, treatment effect (heat or no heat) was significant for all food matrices detected with ELISA method ($p<0.05$). The log means of detected concentrations were higher for all non- heated matrices. For non-heated treatments, water and ranch matrices were significantly higher than other food matrices ($p<0.05$). The log means for heated food matrices were highest for water and lowest for meat matrix ($p<0.05$). For magnetic bead method, treatment effect was only significant in ranch matrix ($p<0.05$). For no-heat treatment, detected concentration was highest in water matrix and lowest in flour matrix. As heat was applied, detected concentration in water matrix remains the highest while lowest in ranch ($p<0.05$). Treatment effect was significant in column-DNA for ranch and water matrices ($p<0.05$). Within no-heat treatment, water and ranch matrices were similar ($p>0.05$), however, they were significantly higher from flour and meat matrix ($p<0.05$). As for heated matrices, there were no significant differences between ranch and flour matrices ($p>0.05$).

Percent recovery was calculated based on the expected SPI concentration (100,000 ppm) in each food matrix (table 4.12). Overall, percent recovery for water matrix was the highest for all detection methods. However, percent recovery was the lowest in meat matrix for ELISA detection and lowest in flour matrix for DNA detections. For ELISA method, the concentration for water matrix and flour were higher than expected and thus result in an overestimate of percent recovery. But as heat was applied to both food matrices, the percent recovery decreased significantly.

For ELISA method, Platteau and others (2011) reported that Neogen Veratox® for soy allergen (ELISA kit used in this study) can overestimate soy protein by as high as an average of 300% (depending on the food matrix). But with extensive period of heating (70ºC for 48 hours), the estimation will go down and eventually non-detectable. The observation suggested that heating can induced protein denaturation and thus underestimating the actual concentration. For this study, the temperature and time for heating was modeled on the conditions that a cooked product might encounter. Thus the heating temperature was much higher and the time was shorter.

Even so, for ELISA method, the heating effect was still significant in all food matrices.

The concentration results for DNA methods were based on real-time PCR results. Thus the changes in real-time PCR detection (Cq value) would results in different concentrations. As indicated in the DNA purity results for column-DNA method (appendix 1.10), DNA quality for water and ranch samples were higher in non-heated sample and thus resulted in lower Cq values and higher detected concentration. For magnetic beads DNA method, the DNA concentration was higher in non-heated water, flour and meat matrices. Higher DNA concentration can be associated with higher amount of target DNA template for amplification. Thus, higher DNA concentration can contribute to a lower Cq value and higher detected concentration. Also, different DNA extraction methods can influence real-time detection. Piknova (2006) stated that recovery rate using CTAB based liquid to liquid extraction would result in recovery rates of 3 to 54% and 8 to 66% on a chraotopic solid- phase extraction (SPE). Column DNA method used in this study is a combination of both extraction methods. Other than non-heated flour and heated ranch matrices, most of the matrices were within the expected range. Furthermore, Pinto and others (2007) stated that magnetic beads-DNA demonstrated a higher efficiency for vegetable matrices rich in polysaccharides and polyphenolics. However, DNeasy Tissue kit, a silica-column-based system (similar to column-DNA), demonstrated higher efficiency with complex and processed food matrices.

Overall, heat treatment (95ºC for 1 hour) can significantly reduce the ability of the ELISA method to detect soy in all food matrices. Depending on the food matrices, heat treatment can also improve the detection of soy of both DNA methods (magnetic beads and column). The detection recovery in ELISA method was the highest among non-treated samples while magnetic beads DNA method was the lowest. For column DNA method, heat treatment can reduce the percent recovery significantly for both water and ranch matrices. Similar trend was observed in magnetic beads-DNA method. For heated flour and meat, heat treatment can improve the detection;

however, the changes were not significant. As for food matrices, water matrix has the highest percent recovery while meat and flour matrix has the lowest percent recovery for ELISA and both DNA methods (respectively). Also, the concentration of DNA collected using magnetic beads method was limited by the amount of magnetic beads. If more magnetic beads were added, the results would be higher; however, the cost of the test would be higher too.

Table 4.1 : Verification of primer design with using real time PCR developed to detect soy DNA. Green pea (PEA) was used as negative control while soybean (SB) was used as positive control. Means ± standard deviation of means were presented.

Sample ID	Sample Type	Total DNA Concentration (µg/µl)	DNA Purity (A260/280)	Cq
PEA	Green Pea	344.5 ±10.34	2.13 ± 0.075	N.D.
SB	Soybean	207.1 ± 28.18	2.15 ± 0.034	24.19 ± 0.11

N.D. = no detection or detection beyond 40 cycles.

Table 4.2: Assessment of soy products using real-time PCR developed for detection of soy DNA. Means ± standard deviation of means were presented for total DNA, DNA purity and Cq values.

Sample ID	Sample Type	Total DNA Concentration (µg/µl)	DNA Purity (A260/280)	Cq
Arcon F	Concentrates	959.8^{b} ± 35.00	2.16^{a} ± 0.012	23.85^{b} ± 0.31
Prolia	Defatted Flour	1104.1^{b} ± 193.91	2.10^{b} ± 0.045	22.74^{ab} ± 0.53
PFL	Lecithin	854.8^{b} ± 90.82	2.17^{a} ±0.008	25.04^{bc} ± 0.40
SB	soybeans	207.1^{c} ± 28.18	2.15^{a} ± 0.034	22.98^{ab} ± 0.62
AFD	SPI	1650.9^{a} ± 179.83	2.15^{a} ± 0.005	23.47^{ab} ± 0.92
Prolisse	SPI	874.6^{b} ± 49.49	2.18^{a} ± 0.014	22.33^{a} ± 0.18
Solae	SPI	1505.6^{a} ± 273.52	2.15^{a} ± 0.022	23.41^{ab} ± 0.94

^{a,b,c} Means within a column with similar letter are not significantly different ($p > 0.05$).
All measurements were conducted in duplicated with n=3.
SPI= Soy protein isolates.
Prolia= Cargill Prolia ^{TM} 200/70)
Arcon F= ADM Arcon F
PLF= ADM Protein Fiber Lecithin
Solae= Solae ISP
AFD= ADM Ardex F Dispersible
Prolisse= Cargill Prolisse
SB= soybeans

Table 4.3 : Evaluation of annealing temperature and primer concentration for the new real time PCR protocol developed with SsoAdvanced™ SYBR® Green Supermix to detect Soy DNA in soybean. DNA extracted from soybean was used as template. Temperature gradient of 60-64 C° and primer concentration of 300nM, 350nM and 400nM were used to optimize the real-time PCR protocol. Means ± standard deviation of means were presented for all Cq values.

Temperature (C°)	**Cq Values**			
	300nM	**350nM**	**400nM**	**Mean***
60	N.D.	N.D.	26.42^{c} ± 4.134	N/A
60.3	N.D.	N.D.	23.32^{ab} ± 0.361	N/A
60.8	N.D.	22.95^{a} ± 0.168	22.96^{ab} ± 0.166	22.96^{b} ± 0.167
61.6	N.D.	23.54^{a} ± 0.507	25.39^{bc} ± 0.330	23.21^{b} ± 0.537
62.6	N.D.	23.58^{a} ± 0.778	25.16^{ab} ± 1.605	24.37^{a} ± 0.686
63.4	N.D.	22.37^{a} ± 0.282	22.61^{ab} ± 0.667	22.49^{b} ± 0.526
63.8	N.D.	22.50^{a} ± 0.214	22.21^{ab} ± 0.115	22.36^{b} ± 0.227
64	N.D.	22.40^{a} ± 0.123	22.10^{a} ± 0.183	22.25^{b} ± 0.215

N.D.= no detection or detection beyond 40 cycles.
Real-time PCR was conducted in duplicated with n=2.
Primer concentration effect was not significant (p>0.05).
^{a,b,c} Means within a column with similar letter are not significantly different (p > 0.05).
* Mean values are means values from 350nM and 400nM primer concentration at each temperature.

Table 4.4 : Evaluation of annealing time and primer concentration for optimizing real-time PCR protocol designed to detect soy DNA using Bio-Rad CFX-96 system. DNA extracted from soybean was used as DNA template. Means ± standard deviation of means were presented for all Cq values.

Primer concentration(nM)	**Cq values**		
	10 seconds*	**20 seconds***	**30 seconds***
300	31.45 ± 1.159	26.80 ± 0.327	28.62 ± 1.787
350	31.45 ± 1.823	27.00 ± 0.189	28.26 ± 1.639
400	29.43 ± 0.220	26.89 ± 0.250	28.03 ± 1.632
Mean	30.78 ± 1.401	26.89 ± 0.252	28.30 ± 1.537

*Means within column for 300, 350 and 400 nmol are not significantly different ($p > 0.05$).

[a,b,c] Mean values were significantly different ($p<0.05$).

Real-time PCR was conducted in duplicated with n=2.

SPI= soy protein isolates.

Table 4. 5: Assessment ELISA method in quantifying soy allergen for soybean, SPI soy protein isolates and soymilk.

Concentration (ppm, mg/L)	**soybean**		**SPI**		**soymilk**	
	Rep 1	Rep2	Rep 1	Rep2	Rep 1	Rep2
4	8.8	17.5	48.9	54.1	N.D.	N.D.
40	53.9	43.7	94.48	90.4	77.4	66.3
400	+	96.33	+	+	+	+
4000	+	+	+	+	+	+
40000	+	+	+	+	+	+
R^2	0.972	0.978	0.972	0.978	0.969	0.978

N.D.= no detection. + = protein concentration is high and out of range.
SPI= ADM Ardex F Dispersible soy protein isolates.
Extractions were conducted in duplicates.
Standard curve was generated using standard solutions provided by the kit.
All estimated results were based on the linear equations.

Table 4.6: Total DNA concentration and purity of extracted DNA prepared for real-time PCR detection of soy DNA.

Extraction type	Soybean		SPI	
	Total DNA Concentration (µg/µl)	**DNA Purity (A260/280)**	**Total DNA Concentration (µg/µl)**	**DNA Purity (A260/280)**
Column	35.6-115.1	2.02-2.13	63.4-171.7	1.85-2.14
Magnetic bead	444.9- 485.9	1.64-1.76	45.1-51.6	1.76-2.05

Measurements were conducted in duplicated with n=2.
SPI= ADM Ardex F Dispersible soy protein isolates.

Table 4.7: Real-time PCR results on column type DNA recovery method (Qiagen DNeasy mericon Food Kit) for soybean, soy protein isolates (SPI) and soymilk using the developed real-time PCR protocol. For each DNA extraction, 200mg of sample was used.

Concentration (ppm, mg/L)	Ext	Rep	Cq for soybean	Cq for SPI	Cq for Soymilk
2	1	1	36.6	36.34	>40
	1	2	36.55	>40	>40
	2	1	>40	36.36	>40
	2	2	>40	>40	35.97
	3	1	>40	35.56	37.19
	3	2	>40	36.54	>40
20	1	1	33.93	34.66	32.44
	1	2	32.54	34.04	32.25
	2	1	34.12	36.22	32.01
	2	2	32.92	34.7	32.55
	3	1	35.37	32.01	32.2
	3	2	36.91	32.49	32.3
200	1	1	28.72	30.89	27.91
	1	2	28.77	30.26	27.57
	2	1	28.81	31.48	27.95
	2	2	29.15	31.86	27.85
	3	1	33.07	28.34	27.6
	3	2	32.84	28.34	27.55
2000	1	1	25.42	27.11	27.81
	1	2	25.6	26.92	26.98
	2	1	26.13	27.71	27.33
	2	2	26.4	27.83	27.11
	3	1	28.92	26.5	27.33
	3	2	29.17	26.36	26.95

Data represent triplicate extractions and duplicate real-time PCR reactions.
> 40 = non-detectable within 40 cycles.
Rep= extraction replication.
Ext= extraction

Table 4.8: Real-time PCR results of magnetic beads type DNA recovery method (Promega Wizard Magnetic DNA Purification system kit) for soybean, soy protein isolates (SPI) and soymilk using the developed real-time PCR protocol. For each DNA extraction, 200mg of sample was used.

Concentration (ppm, mg/L)	Ext	Rep	Cq for soybean	Cq for SPI	Cq for Soymilk
1.333333	1	1	33.37	>40	>40
	1	2	33.36	>40	>40
	2	1	34.66	37.29	>40
	2	2	35.33	>40	>40
	3	1	34.84	>40	>40
	3	2	35.16	36.23	>40
13.33333	1	1	30.15	36.7	37.59
	1	2	30.25	36.12	>40
	2	1	30.96	35.78	37.43
	2	2	34.65	35.34	>40
	3	1	32.53	35.08	38.18
	3	2	31.29	35.1	>40
133.3333	1	1	26.79	32.01	34.17
	1	2	27.49	>40	35.36
	2	1	31.33	32.02	36.00
	2	2	>40	31.3	33.78
	3	1	28.85	31.89	35.16
	3	2	28.9	31.28	34.61
1333.333	1	1	24.22	27.81	33.02
	1	2	24.12	27.73	32.95
	2	1	27.04	28.17	31.7
	2	2	27.41	28.1	31.52
	3	1	25.51	27.04	31.12
	3	2	25.37	28.43	31.31

Data represent triplicate extractions and duplicate real-time PCR reactions.
> 40 = non-detectable within 40 cycles.
Rep= extraction replication.
Ext= extraction.

Table 4.9: PCR efficiency and correlation coefficients (R^2 values) from both extraction methods using the modified real time PCR protocol on soybean, SPI and soymilk. Verification was performed by diluted soybean and SPI extracted DNA. Bio Rad CFX manager software was used to generate both values. For each DNA extraction, 200 mg of sample was used. Means ± standard deviation of means were presented for all percent PCR efficiency and R^2.

Sample	**Sample size**	**Column**		**Magnetic beads**	
		% efficiency	**R^2**	**% efficiency**	**R^2**
Soybean DNA		100	0.987	98.3	0.991
SPI DNA		105.6	0.989	99.8	0.995
Soybean	200 mg	101.7 ± 16.49	0.979 ± 0.007	101.7 ± 16.49	0.954 ± 0.067
	1g	N/A	N/A	98.85 ± 15.9	0.833 ± 0.101
SPI	200 mg	102.8 ± 6.62	0.979 ± 0.003	101.73 ± 7.13	0.988 ± 0.011
	1g			103.4 ±13.4	0.906 ± 0.011
Soymilk	200 mg	106.3 ± 52.66	0.862 ± 0.214	110.00 ± 8.45	0.883 ± 0.025

Real-time PCR for soybean, SPI and soymilk was conducted in duplicated with n=3.
Real-time PCR for extracted DNA was conducted in triplicates.
SPI= ADM Ardex F Dispersible soy protein isolates

Table 4.10: Real-time PCR results of magnetic beads type DNA recovery method (Promega Wizard Magnetic DNA Purification system kit) for soybean, soy protein isolates (SPI) and soymilk using the developed real-time PCR protocol. For each DNA extraction, 1g of sample was used.

Concentration (ppm, mg/L)	Ext	Rep	CQ for Soy	CQ for SPI
1.33	1	1	>40	>40
	1	2	>40	>40
	2	1	>40	>40
	2	2	>40	>40
	3	1	>40	>40
	3	2	>40	>40
13.33	1	1	>40	36.45
	1	2	>40	36.16
	2	1	34.04	36.23
	2	2	34.01	37.33
	3	1	>40	33.06
	3	2	>40	35.81
133.33	1	1	35.63	33.59
	1	2	32.56	34.08
	2	1	30.9	33.59
	2	2	30.77	34.08
	3	1	32.22	32.8
	3	2	34.03	33.1
1333.33	1	1	33.1	29.3
	1	2	33.54	29.7
	2	1	28.11	31.55
	2	2	27.37	30.7
	3	1	32.37	31.13
	3	2	32.83	31.79

Data represent triplicate extractions and duplicate real-time PCR reactions.
> 40 = non-detectable within 40 cycles.
Rep= extraction replication.
Ext= extraction

Table 4.11: Evaluation of ELISA and DNA methods (column and magnetic beads) for heated and non-heated food samples. All food matrices were spiked with 100,000 ppm SPI (soy protein isolates). Real-time PCR results were converted to concentration (ppm) using standard curves constructed for SPI. ELISA results were presented in Log transformation means.

Extraction	Treat	Water	Flour	Meat	Ranch	S.E.M
Column	Con	68181.8^{a}*	2945.2^{c}	16620.0^{b}	63052.0^{a}*	2142.18
	Heat	35556.0^{a}*	6884.5^{b}	23847.0^{ab}	768.2^{c}*	2142.18
Magnetic bead	Con	35758.0^{a}	2803.9^{c}	12546.0^{ab}	26174.0^{a}*	1892.78
	Heat	21831.0^{a}	7291.9^{ab}	14053.0^{ab}	816.2^{c}*	1892.78
Transformed Means						
ELISA	Con	11.70^{a}*	11.17^{b}*	9.63^{c}*	11.66^{a}*	0.048
	Heat	9.51^{a}*	7.05^{b}*	6.33^{c}*	6.49^{d}*	0.048

S.E.M= Standard error of means for each sample within a row.
Treat=treatment.
Con= no heat.
Heat= heated at 95°C for 1 hour.
^{a,b,c,d} LS means appears in the same row with similar letter are not significantly different (p>0.05).
* LS means appears in the same column within each extraction type indicates that treatment effect was significant (p<0.05).

Table 4.12: Evaluation of percent recovery for ELISA and DNA methods (column and magnetic beads) for heated and control food matrices. Percent recovery calculation was based on the detected concentration and the initial spiked amount of SPI soy protein isolates (100,000 ppm). LS means (percent) were presented for all results.

Extraction	Treat	Water	Flour	Meat	Ranch	S.E.M
Magnetic beads	Con	35.75*	2.80	12.55	26.10*	1.60
	Heat	21.83*	7.29	14.05	0.82*	
Column	Con	68.18*	2.95	16.62	63.05*	1.789
	Heat	35.56*	6.89	23.85	0.78*	
ELISA	Con	120.8*	71.24*	15.22*	109.17*	11.76
	Heat	13.50*	1.15*	0.56*	0.66*	

S.E.M= Standard error of means for each sample within a row.
Treat=treatment.
Con= no heat.
Heat= heated at 95°C for 1 hour.
*LS means appears in the same column within each extraction type indicates that treatment effect was significant ($p<0.05$).

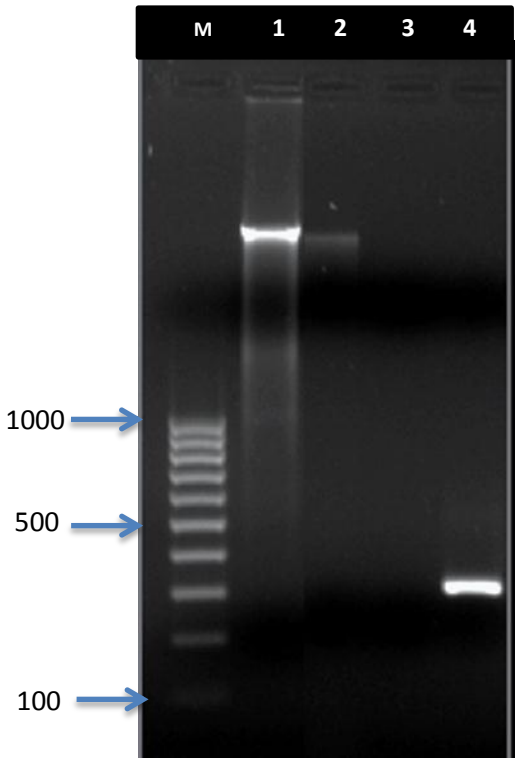


Figure 4.1: Specificity assessment for real-time PCR developed for detection of soy DNA sequence with green peas. Lane M: 100-1000 base pair (bp) Bio-rad DNA marker. Lane 1: Total DNA extracted from green pea. Lane 2: Total DNA extracted from soy. Lane 3: Amplified product from green pea DNA template, diluted 1:1000.Lane 4: Amplified product from soy DNA template, diluted 1:1000.

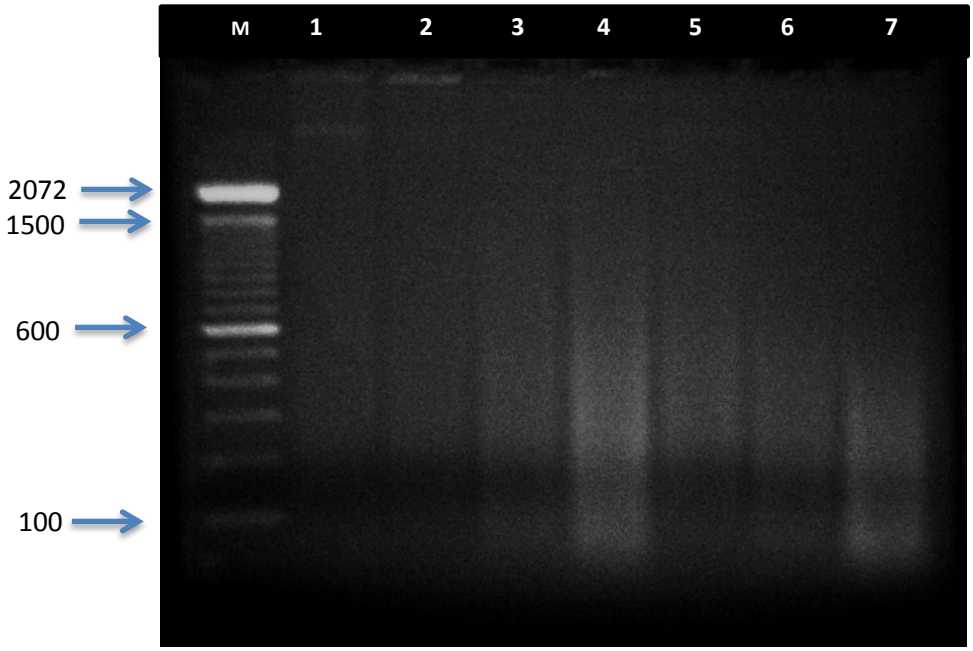


Figure 4.2: Evaluation of extracted DNA prepared for the versatility assessment of the real-time PCR detection of soy DNA. 100-1500 base pair (bp) Invitrogen DNA Ladder was used (M). DNA template used were extracted from (1) Soybean, (2) Prolia (defatted soy flour), (3)Arcon F(soy protein concentrates), (4) AFD (dispersible soy protein isolates), (5)Prolisse (soy protein isolates), (6) PFL (soy lecithin) and (7) Solae (soy protein isolates).

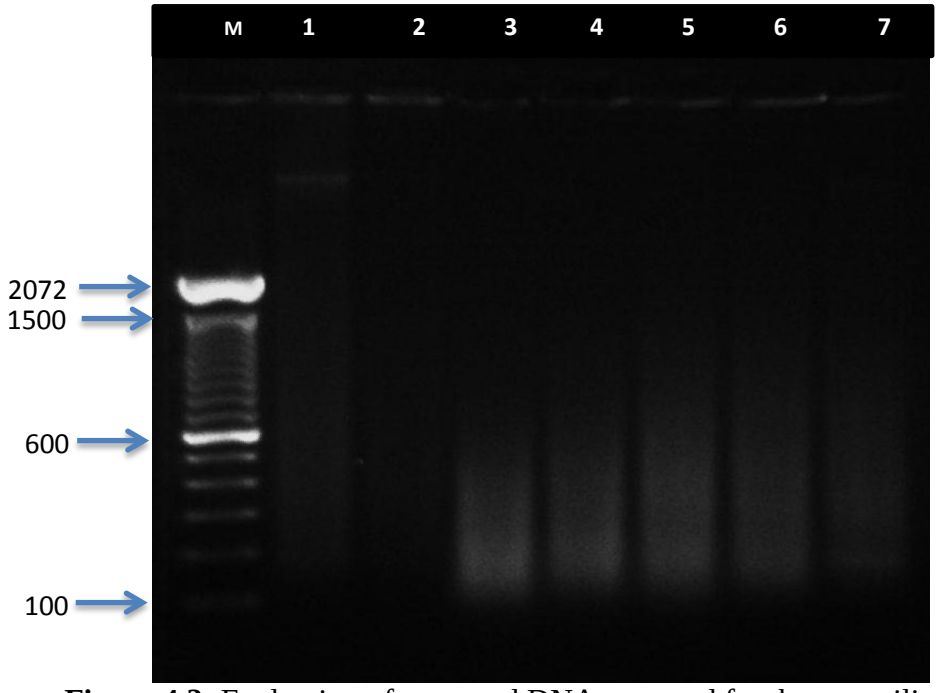


Figure 4.3: Evaluation of extracted DNA prepared for the versatility assessment of the real-time PCR detection of soy DNA. 100-1500 base pair (bp) Invitrogen DNA Ladder was used (M). DNA template used were extracted from (1) Soybean, (2) Prolia (defatted soy flour), (3) Solae (soy protein isolates), (4) PFL (soy lecithin), (5) SPI (dispersible soy protein isolates), (6) Arcon F (soy protein concentrates) and (7)Prolisse (soy protein isolates).

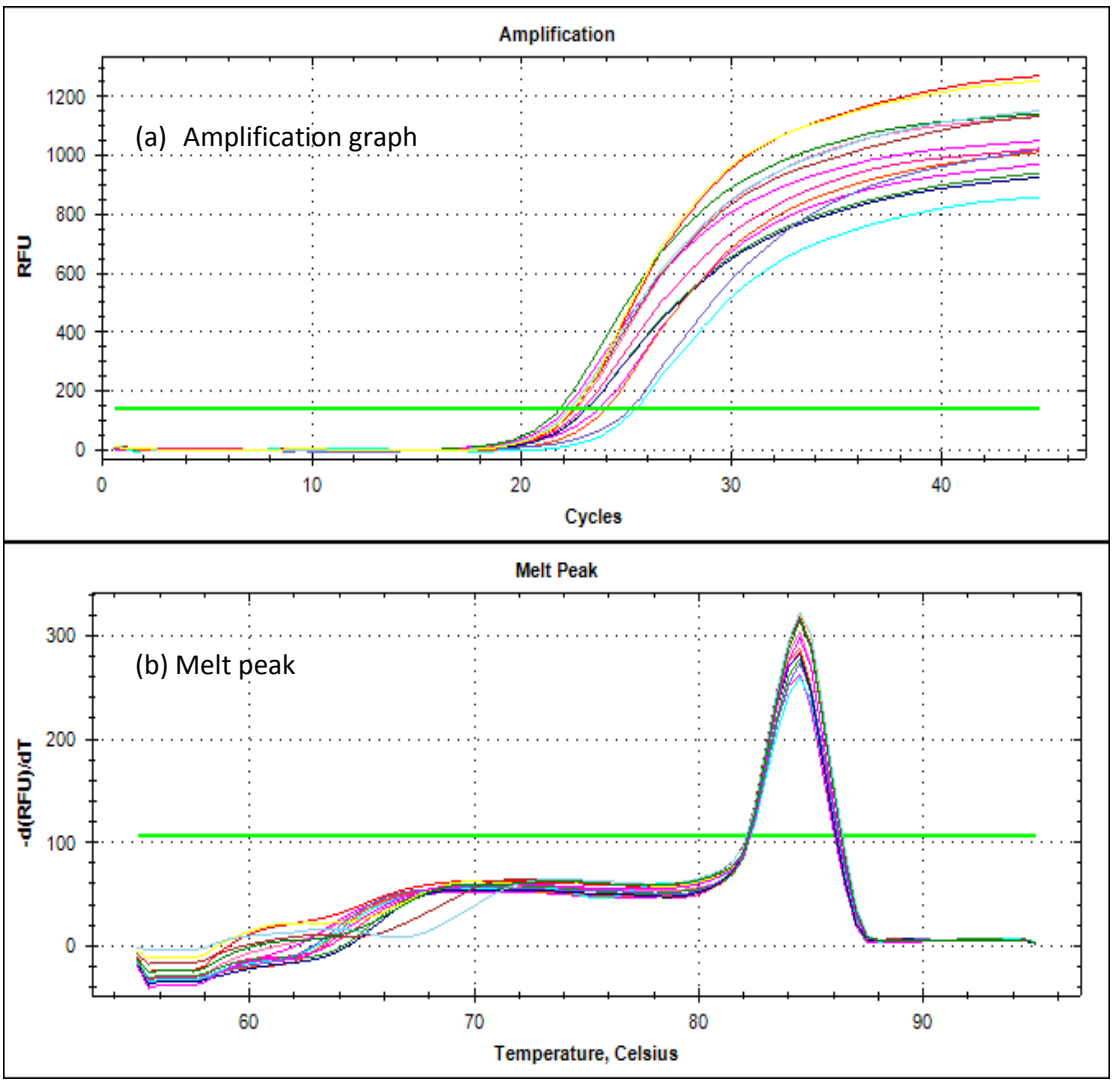


Figure 4.4: Real-time PCR results for evaluating the versatility of the primer design to detect soy DNA in various soy products. Melt-curve analysis confirmed that no non-specific products were amplified. RFU= relative florescent unit. –d(RFU/dT)= first negative derivative of fluorescence

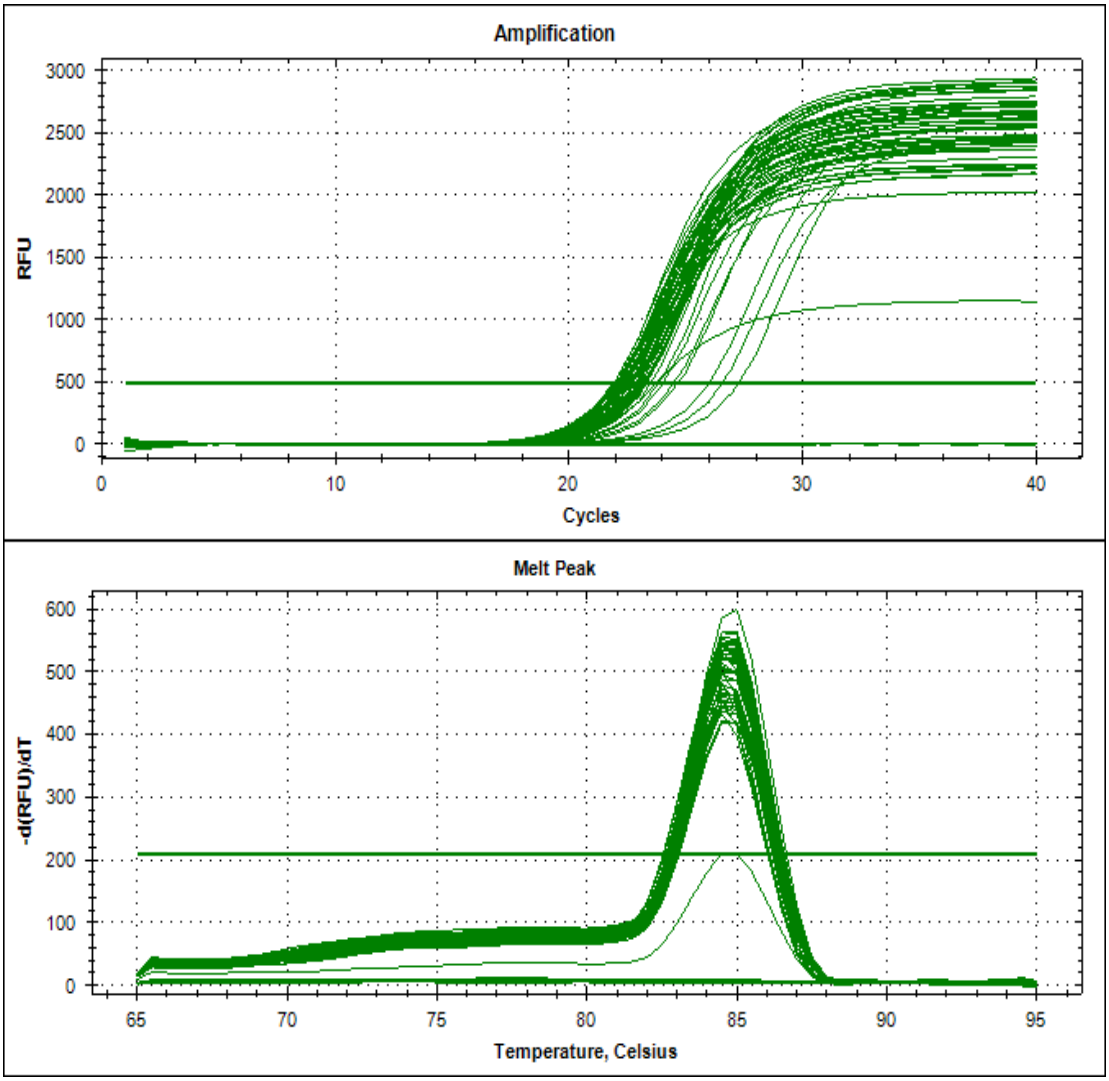


Figure 4.5: Real-time PCR results for evaluating annealing temperature using temperature gradient (60-64 C°) and primer concentration of 300nM, 350nM and 400nM. Melt-curve analysis confirmed that no non-specific products were amplified.

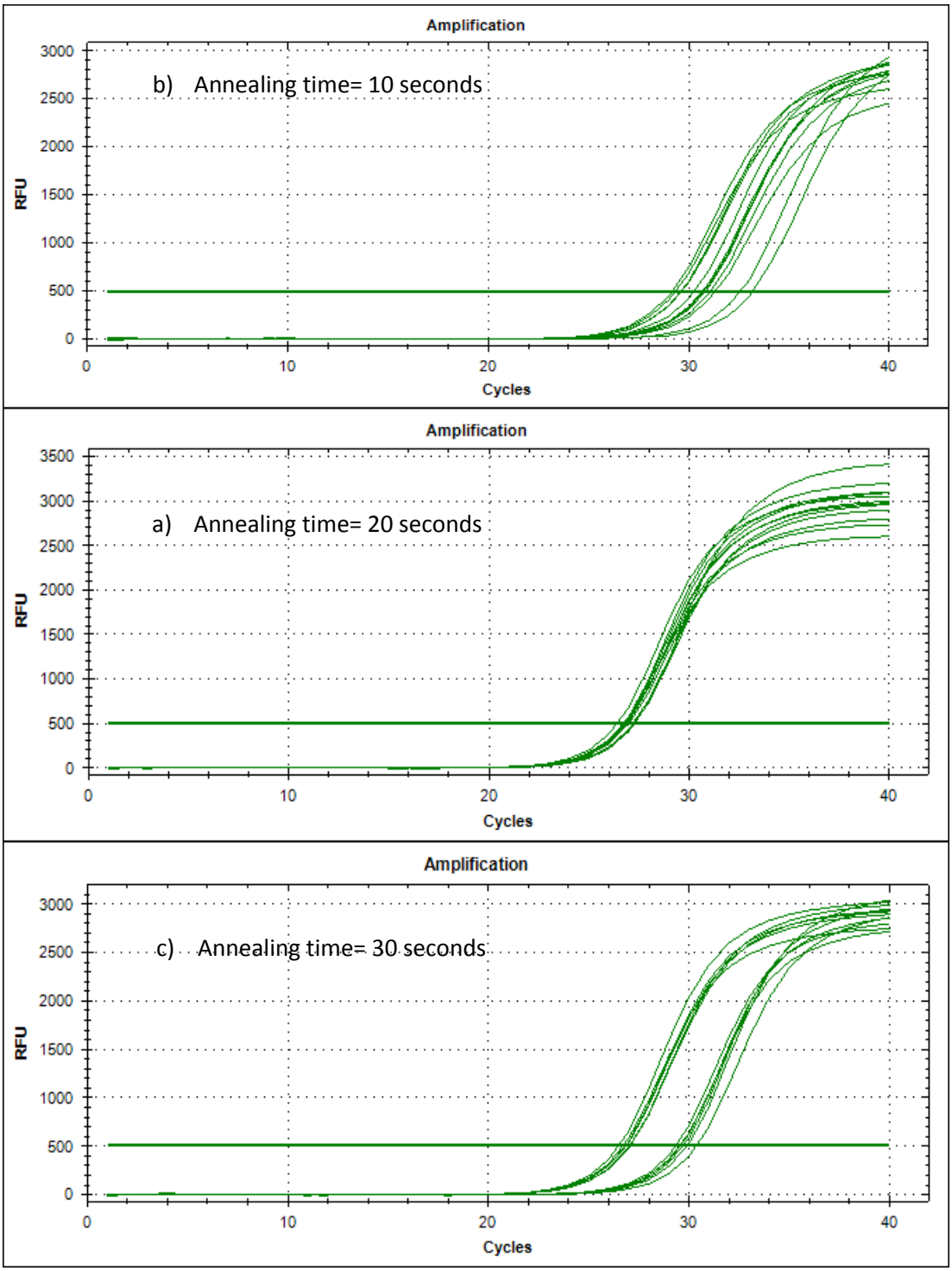


Figure 4.6: Real-time PCR results for evaluations of annealing time needed for optimizing real-time PCR protocol to detect soy DNA in soybean. Primer concentration of 300nM, 350nM and 400nM) were used with annealing time of (a) 10 seconds, (b) 20 seconds and (c)30 seconds.

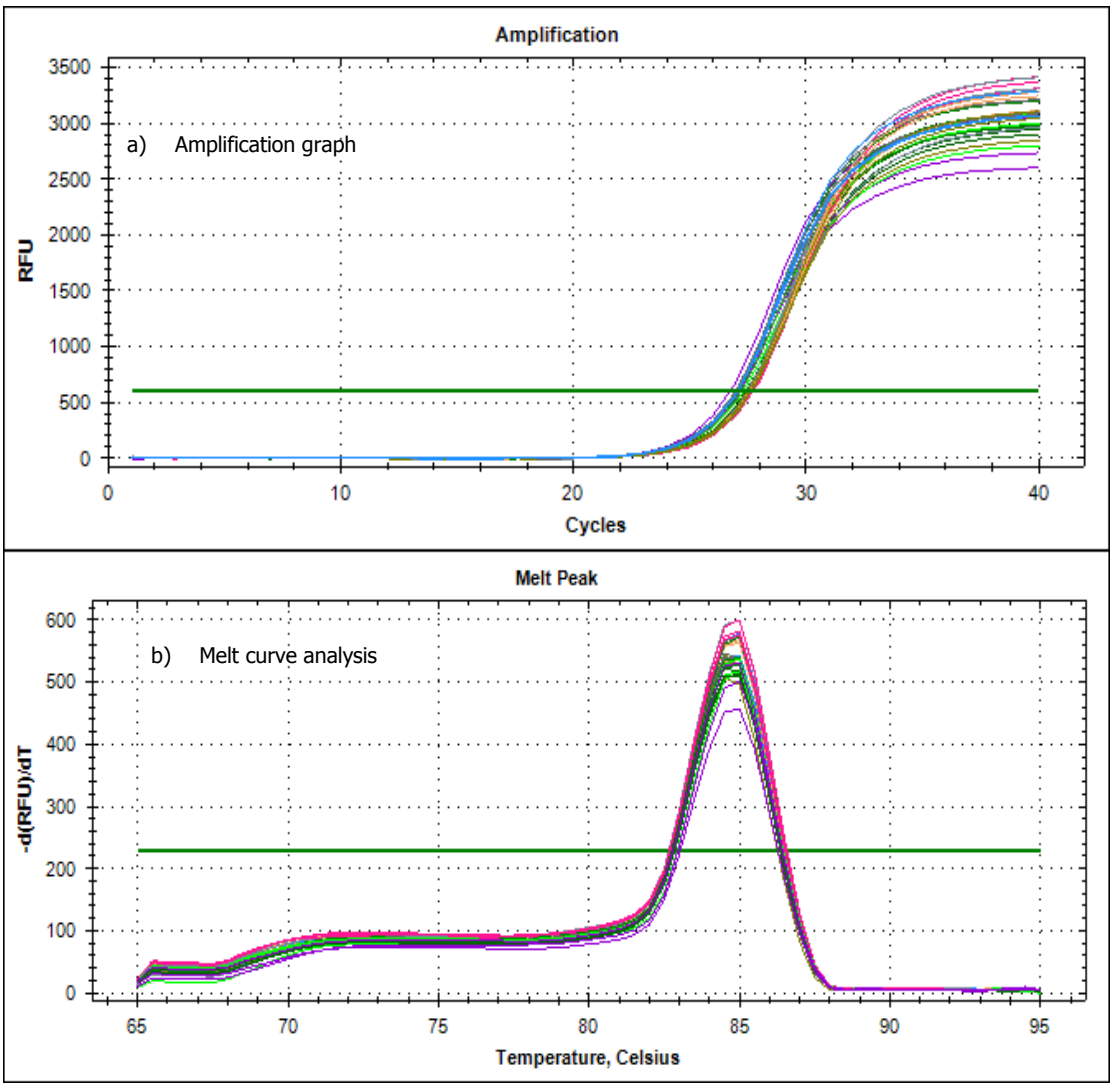


Figure 4.7: Real-time PCR results for evaluating primer concentration (300nM-500nM; increment of 50nM). Melt-curve analysis confirmed that no non-specific products were amplified.

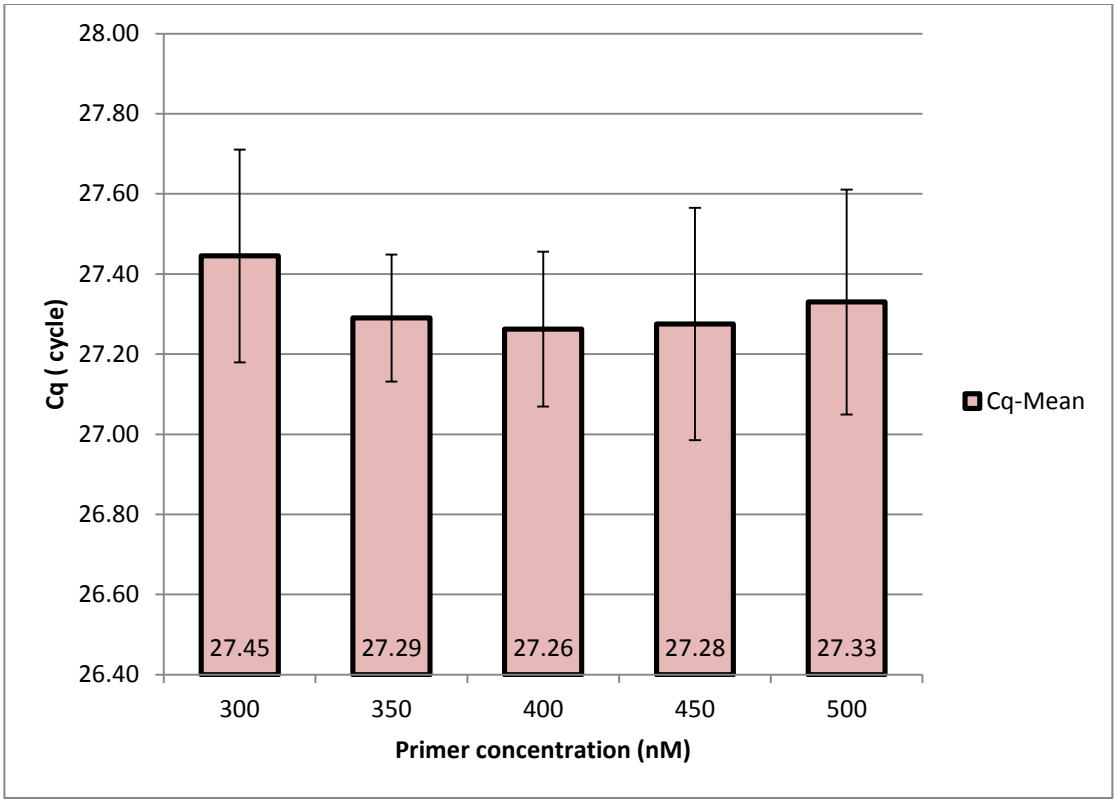


Figure 4.8: Assessment of primer concentration (300-500 nM; increment of 50nM) for optimizing real-time PCR protocol to detect soy DNA in soybean. All real-time PCR was conducted in duplicated with n=3. Real-time PCR results for all primer concentration were not significantly different (p>0.05)

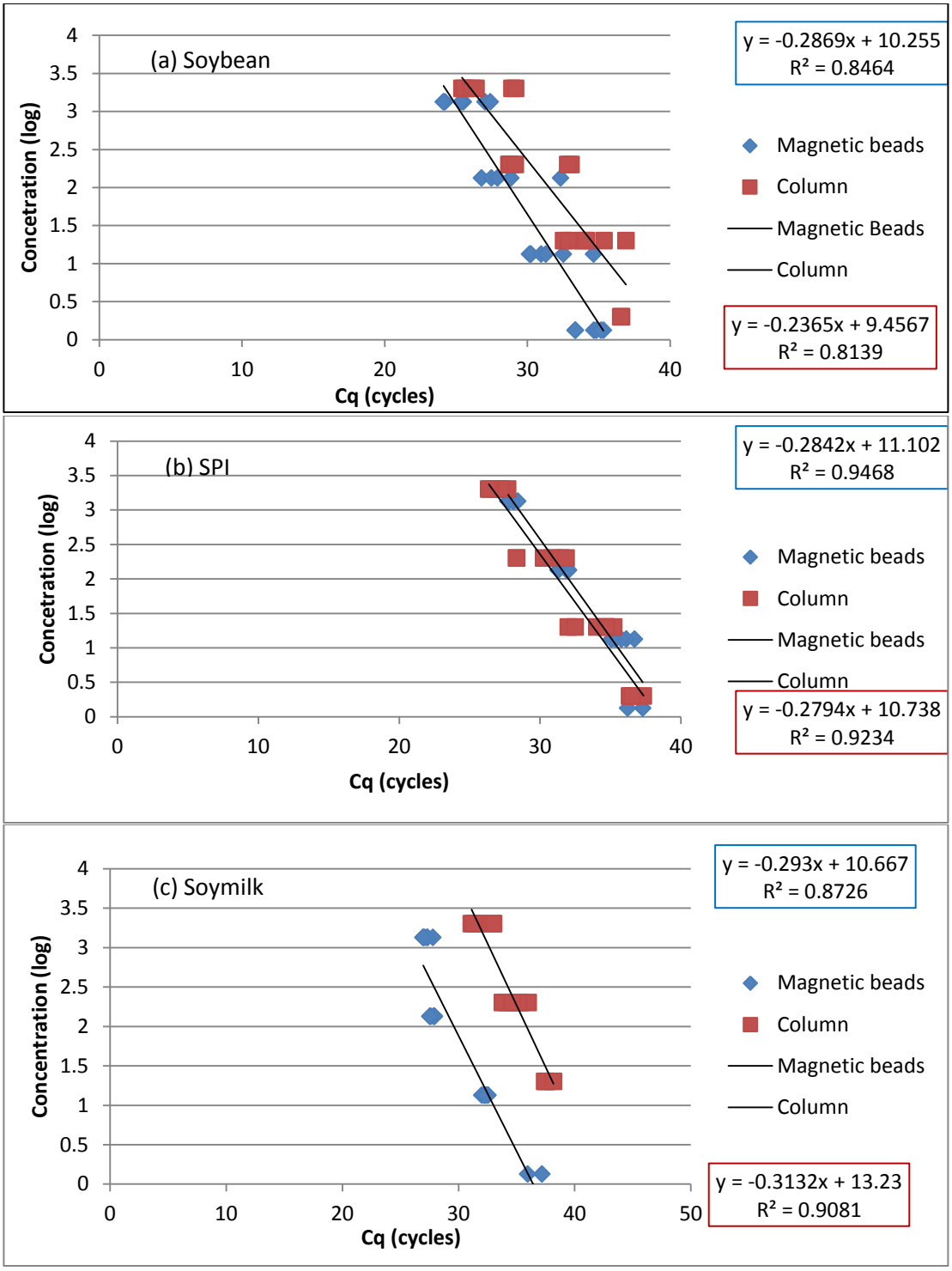


Figure 4.9: Linearity test on real-time PCR results for (a) soybean, (b) soy protein isolates (SPI) and (c) soymilk using both column and magnetic beads DNA recovery method. Results greater than 40 were replaced with 40. Top regression equation was constructed for magnetic beads and bottom for column.

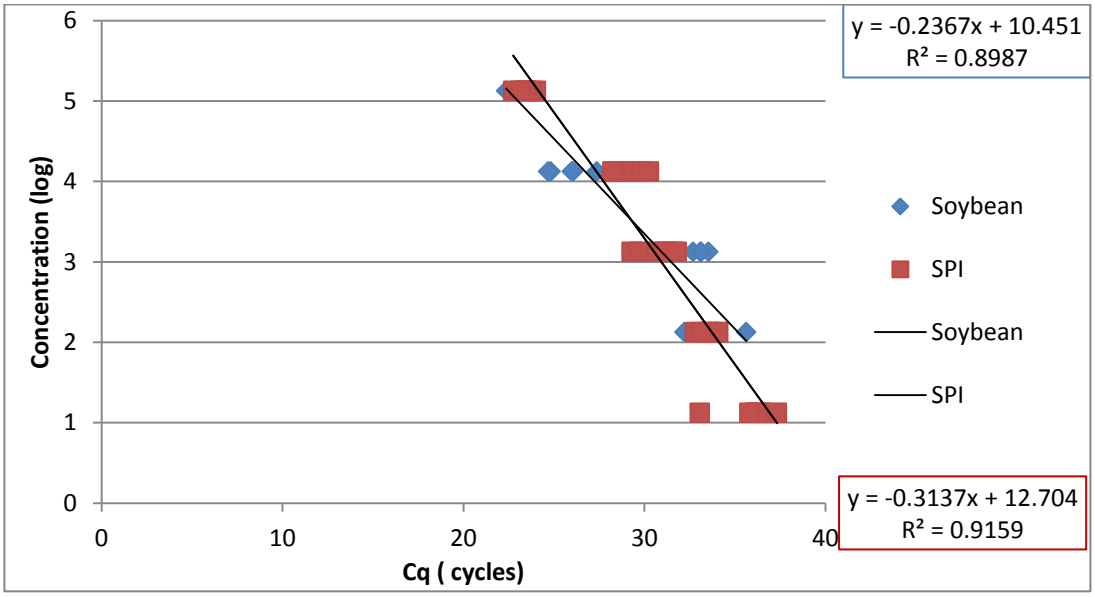


Figure 4.10: Linearity test on real-time PCR results of magnetic beads DNA recovery method for both soybean and SPI soy protein isolates. Results greater than 40 were replaced with 40. Top regression equation was constructed for soybean and bottom for SPI soy protein isolates.

Chapter 5

CONCLUSION

In conclusion, the primer designed exhibit the ability to react specifically and detect a wide range of processed soy products. The real-time PCR assay was optimized and verified for high PCR efficiency. Limit of detection for column DNA recovery method in soybean, SPI and soymilk can be as low as 20 ppm, however, for magnetic beads DNA method, the limit of detection was matrix dependent and demonstrated a lower detection limit for soybean sample (1.33 ppm) but higher detection limit for soymilk (133.3 ppm). The percent recovery for non-heated food matrices was higher in ELISA methods and lower in magnetic beads DNA method. For heated food matrices, soy recoveries for both DNA methods were higher than ELISA. Overall, heat treatment can significantly reduce the ability for ELISA method to detect soy in all food matrices. However, for DNA methods (column and magnetic beads), water and high fat (ranch) matrices were the only two that were significantly affected by thermal processing. As for matrix effect, water matrix has the highest percent recovery for all detection methods. However, flour matrix has the lowest percent recovery in non-heated sample for both DNA detection methods. As for heated matrices, high fat matrix has the lowest percent recovery in both DNA detection methods. Thus, other than flour matrix and heated ranch matrix, both DNA methods can be used to detect soy in heated and non-heated foods. Future studies are needed to optimize the amount of magnetic beads in magnetic beads DNA method for higher detections and lower the concentration of extracted DNA for higher soy recovery.

BIBLIOGRAPHY

Allan Bock S, Sampson HA, Atkins FM, Zeiger RS, Lehrer S, Sachs M, Bush RK, Metcafe DD. 1988. Double-blind, placebo-controlled food challenge (DBPCFC) as an office procedure: a manual. Journal of Allergy and Clinical Immunology 82(6):986-97.

Allmann M, Candrian U, Hofelein C, Luthy J. 1993. Polymerase chain reaction (PCR): a possible alternative to immunochemical methods assuring safety and quality of food Detection of wheat contamination in non-wheat food products. Zeitschrift fur Lebensmitteluntersuchung und-forschung A 196(3):248-51.

Arlorio M, Cereti E, Coisson JD, Travaglia F, Martelli A. 2007. Detection of hazelnut (*Corylus spp.*) in processed foods using real-time PCR. Food control 18(2):140-8.

Awazuhara H, Kawai H, Maruchi N. 2006. Major allergens in soybean and clinical significance of IgG4 antibodies investigated by IgE-and IgG4-immunoblotting with sera from soybean sensitive patients. Clinical & Experimental Allergy 27(3):325-32.

Ballmer-Weber BK, Vieths S. 2008. Soy allergy in perspective. Current opinion in allergy and clinical immunology 8(3):270.

Barac M, B., Stanojevic SP, Jovanovic ST, Pesic MB. 2004. Soy protein modification: A review. Acta Periodica Technologica 35:3-16.

Barnett D, Howden MEH. 1987. Lectins and the radioallergosorbent test. Journal of Allergy and Clinical Immunology 80(4):558-61.

Beardslee TA, Zeece MG, Sarath G, Markwell JP. 2000. Soybean glycinin G1 acidic chain shares IgE epitopes with peanut allergen Ara h 3. International archives of allergy and immunology 123(4):299-307.

Beyer K, Morrowa E, Li XM, Bardina L, Bannon GA, Burks A, Sampson HA. 2001. Effects of cooking methods on peanut allergenicity. Journal of Allergy and Clinical Immunology 107(6):1077-81.

Bindslev-Jensen C, Ballmer-Weber BK, Bengtsson U, Blanco C, Ebner C, Hourihane J, Knulst AC, Moneret? autrin DA, Nekam K, Niggemann B. 2004. Standardization of food challenges in patients with immediate reactions to foods? position paper from the European Academy of Allergology and Clinical Immunology. Allergy 59(7):690-7.

Bock SA. 1986. The natural history of adverse reactions to foods. Allergy and Asthma Proceedings: OceanSide Publications, Inc. p. 504-10.

Branum AM, Lukacs SL. 2009. Food allergy among children in the United States. Pediatrics 124(6):1549-55.

Brezna B, Hudecova L, Kuchta T. 2006a. A novel real-time polymerase chain reaction (PCR) method for the detection of walnuts in food. European Food Research and Technology 223(3):373-7.

Brezna B, Hudecova L, Kuchta T. 2006b. Detection of pea in food by real-time polymerase chain reaction (PCR). European Food Research and Technology 222(5):600-3.

Burks AW, Cockrell G, Connaughton C, Guin J, Allen W, Helm RM. 1994. Identification of peanut agglutinin and soybean trypsin inhibitor as minor legume allergens. International archives of allergy and immunology 105(2):143-9.

Careri M, Elviri L, Mangia A, Mucchino C. 2007. ICP-MS as a novel detection system for quantitative element-tagged immunoassay of hidden peanut allergens in foods. Analytical and bioanalytical chemistry 387(5):1851-4.

Chehade M, Mayer L. 2005. Oral tolerance and its relation to food hypersensitivities. Journal of Allergy and Clinical Immunology 115(1):3-12.

Chen Y, Wang Y, Ge Y, Xu B. 2005. Degradation of endogenous and exogenous genes of Roundup-Ready soybean during food processing. Journal of agricultural and food chemistry 53(26):10239-43.

Colton JA, Bower KM. 2002. Some misconceptions about r^2. International Society of Six Sigma Professionals, EXTRAOrdinary Sense 3(2):20-2.

Coombs R, Gell PGR. 1975. Classification of allergic reactions responsible for clinical hypersensitivity and disease. In: Gell PGH, Coombs R, Lachmann PJ, editors. Clinical Aspects of Immunology. 3rd ed. Oxford Blackwell Scientific publication p. chapter 25.

Cordle CT. 2004. Soy protein allergy: incidence and relative severity. The Journal of Nutrition 134(5):1213S-9S.

Datukishvili N, Gabriadze I, Kutateladze T, Karseladze M, Vishnepolsky B. 2010. Comparative evaluation of DNA extraction methods for food crops. International journal of food science & technology 45(6):1316-20.

Daul CB, Slattery M, Reese G, Lehrer SB. 1994. Identification of the Major Brown Shrimp(*Penaeus aztecus*) Allergen as the Muscle Protein Tropomyosin. International archives of allergy and immunology 105(1):49-55.

Demeke T, Jenkins GR. 2010. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. Analytical and bioanalytical chemistry 396(6):1977-90.

Desroches S, Mauger JF, Ausman LM, Lichtenstein AH, Lamarche B. 2004. Soy protein favorably affects LDL size independently of isoflavones in hypercholesterolemic men and women. The Journal of Nutrition 134(3):574-9.

Dieffenbach CW, Lowe TM, Dveksler GS. 1993. General concepts for PCR primer design. Genome Research 3(3):S30-S7.

Dorak T. 2006. Real time PCR: Taylor & Francis.

Dragan AI, Pavlovic R, McGivney JB, Casas-Finet JR, Bishop ES, Strouse RJ, Schenerman MA, Geddes CD. 2012. SYBR Green I: Fluorescence Properties and Interaction with DNA. Journal of fluorescence:1-11.

Espineira M, Herrero B, Vieites JM, Santaclara FJ. 2010. Validation of end-point and real-time PCR methods for the rapid detection of soy allergen in processed products. Food Additives and Contaminants 27(4):426-32.

Foti N, Onori R, Donnarumma E, De Santis B, Miraglia M. 2006. Real-time PCR multiplex method for the quantification of Roundup Ready soybean in raw material and processed food. European Food Research and Technology 222(1):209-16.

Gatti M, Ferretti C. 2009. Soy Allergen Detection. Molecular Biological and Immunological Techniques and Applications for Food Chemists:335.

Gryson N, Messens K, Dewettinck K. 2004. Evaluation and optimisation of five different extraction methods for soy DNA in chocolate and biscuits. Extraction of DNA as a first step in GMO analysis. Journal of the Science of Food and Agriculture 84(11):1357-63.

Hefle SL, Nordlee JA, Taylor SL. 1996. Allergenic foods. Critical Reviews in Food Science & Nutrition 36(S1):69-89.

Heid CA, Stevens J, Livak KJ, Williams PM. 1996. Real time quantitative PCR. Genome Research 6(10):986-94.

Helm R, Cockrell G, Herman E, Burks A, Sampson HA, Bannon G. 1998. Cellular and molecular characterization of a major soybean allergen. International archives of allergy and immunology 117(1):29-37.

Herman L, Block JD, Viane R. 2003. Detection of hazelnut DNA traces in chocolate by PCR. International journal of food science & technology 38(6):633-40.

Hird H, Lloyd J, Goodier R, Brown J, Reece P. 2003. Detection of peanut using real-time polymerase chain reaction. European Food Research and Technology 217(3):265-8.

Hirsch AM. 1999. Role of lectins (and rhizobial exopolysaccharides) in legume nodulation. Current opinion in plant biology 2(4):320-6.

Holland PM, Abramson RD, Watson R, Gelfand DH. 1991. Detection of specific polymerase chain reaction product by utilizing the 5'----3'exonuclease activity of Thermus aquaticus DNA polymerase. Proceedings of the National Academy of Sciences 88(16):7276-80.

Holzhauser T, Vieths S. 1999. Indirect competitive ELISA for determination of traces of peanut (Arachis hypogaea L.) protein in complex food matrices. Journal of agricultural and food chemistry 47(2):603-11.

Holzhauser T, Wackermann O, Ballmer-Weber BK, Bindslev-Jensen C, Scibilia J, Perono-Garoffo L, Utsumi S, Poulsen LK, Vieths S. 2009. Soybean (Glycine max) allergy in Europe: Gly m 5 (beta-conglycinin) and Gly m 6 (glycinin) are potential diagnostic markers for severe allergic reactions to soy. Journal of Allergy and Clinical Immunology 123(2):452-8. e4.

Hong KW. 2007. Genomic DNA sequence of mackerel parvalbumin and a PCR test for rapid detection of allergenic mackerel ingredients in food. Food Science and Biotechnology 16(1):67-70.

Hrckova M, Rusnakova M, Zemanovic J. 2002. Enzymatic hydrolysis of defatted soy flour by three different proteases and their effect on the functional properties of resulting protein hydrolysates. Czech journal of food sciences 20(1):7-14.

Jones SM, Pons L, Roberts JL, Scurlock AM, Perry TT, Kulis M, Shreffler WG, Steele P, Henry KA, Adair M. 2009. Clinical efficacy and immune regulation

with peanut oral immunotherapy. Journal of Allergy and Clinical Immunology 124(2):292-300. e97.
Kagan RS. 2003. Food allergy: an overview. Environmental health perspectives 111(2):223.
Kennedy S. 2011. PCR troubleshooting and optimization: The essential guide: Horizon Scientific Press.
Kinsella JE. 1979. Functional properties of soy proteins. Journal of the American Oil Chemists' Society 56(3):242-58.
Kirsch S, Fourdrilis S, Dobson R, Scippo ML, Maghuin-Rogister G, De Pauw E. 2009. Quantitative methods for food allergens: a review. Analytical and bioanalytical chemistry 395(1):57-67.
Kuipers BJH, van Koningsveld GA, Alting AC, Driehuis F, Gruppen H, Voragen AGJ. 2005. Enzymatic hydrolysis as a means of expanding the cold gelation conditions of soy proteins. Journal of agricultural and food chemistry 53(4):1031-8.
L'Hocine L, Boye JI, Munyana C. 2007. Detection and Quantification of Soy Allergens in Food: Study of Two Commercial Enzyme-linked Immunosorbent Assays. Journal of food science 72(3):C145-C53.
Lee K, Brennand CP. 2005. Physico-chemical, textural and sensory properties of a fried cookie system containing soy protein isolate. International journal of food science & technology 40(5):501-8.
Lehrer SB, Ayuso R, Reese G. 2006. Current understanding of food allergens. Annals of the New York Academy of Sciences 964(1):69-85.
Li M, Lin YC, Wu CC, Liu HS. 2005a. Enhancing the efficiency of a PCR using gold nanoparticles. Nucleic acids research 33(21):e184-e.
Li S, Zhang HQ, Tony Jin Z, Hsieh F. 2005b. Textural modification of soya bean/corn extrudates as affected by moisture content, screw speed and soya bean concentration. International journal of food science & technology 40(7):731-41.
Lovati MR, Manzoni C, Gianazza E, Arnoldi A, Kurowska E, Carroll KK, Sirtori CR. 2000. Soy protein peptides regulate cholesterol homeostasis in Hep G2 cells. The Journal of Nutrition 130(10):2543-9.
Mackay J, Landt O. 2004. Real-time PCR fluorescent chemistries. METHODS IN MOLECULAR BIOLOGY-CLIFTON THEN TOTOWA- 353:237.
Maleki SJ, Viquez O, Jacks T, Dodo H, Champagne ET, Chung SY, Landry SJ. 2003. The major peanut allergen, Ara h 2, functions as a trypsin inhibitor, and roasting enhances this function. Journal of Allergy and Clinical Immunology 112(1):190-5.
Maltas E, Dageri N, Vural HC, Yildiz S. 2011. Biochemical and molecular analysis of soybean seed from Turkey. Journal of Medicinal Plants Research 5(9):1575-81.
Meyer R, Chardonnens F, Hubner P, Luthy J. 1996. Polymerase chain reaction (PCR) in the quality and safety assurance of food: detection of soya in processed meat

products. Zeitschrift fur Lebensmitteluntersuchung und-forschung A 203(4):339-44.
Moroz LA, Yang WH. 1980. Kunitz soybean trypsin inhibitor. A specific allergen in food anaphylaxis. New England Journal of Medicine 302(20):1126-8.
260/280 and 260/230 Ratios: NanoDrop® ND-1000 and ND-8000 8-Sample Spectrophotometers technical Bulletin Wilmington; 2007 [Accessed 2012.
Niestijl Jansen JJ, Kardinaal AFM, Huijbers G, Vlieg-Boerstra BJ, Martens BPM, Ockhuizen T. 1994. Prevalence of food allergy and intolerance in the adult Dutch population. Journal of Allergy and Clinical Immunology 93(2):446-56.
Nowak-Wegrzyn A, Sampson HA. 2006. Adverse reactions to foods. Med Clin North Am. 2006 Jan;90(1):97-127.
Ogawa T, Bando N, Tsuji H, Nishikawa K, Kitamura K. 1995. Alpha-subunit of beta-conglycinin, an allergenic protein recognized by IgE antibodies of soybean-sensitive patients with atopic dermatitis. Bioscience, biotechnology, and biochemistry 59(5):831.
Olexova L, Dovicovicova L, Kuchta T. 2004. Comparison of three types of methods for the isolation of DNA from flours, biscuits and instant paps. European Food Research and Technology 218(4):390-3.
Ortolani C, Ispano M, Scibilia J. 2004. Introducing chemists to food allergy. Allergy 56(s67):5-8.
Panyam D, Kilara A. 1996. Enhancing the functionality of food proteins by enzymatic modification. Trends in food science & technology 7(4):120-5.
Piknova L, Kuchta T. 2006. Concentration dependence of recovery rates of an exogenous short linear DNA fragment from soya flour. European Food Research and Technology 223(1):143-5.
Pinto AD, Forte VT, Guastadisegni MC, Martino C, Schena FP, Tantillo G. 2007. A comparison of DNA extraction methods for food analysis. Food control 18(1):76-80.
Platteau C, Cucu T, De Meulenaer B, Devreese B, De Loose M, Taverniers I. 2011. Effect of protein glycation in the presence or absence of wheat proteins on detection of soybean proteins by commercial ELISA. Food Additives and Contaminants 28(2):127-35.
Poms RE, Agazzi ME, Bau A, Brohee M, Capelletti C, Norgaard JV, Anklam E. 2005. Inter-laboratory validation study of five commercial ELISA test kits for the determination of peanut proteins in biscuits and dark chocolate. Food Additives and Contaminants 22(2):104-12.
Poms RE, Klein CL, Anklam E. 2004. Methods for allergen analysis in food: a review. Food Additives and Contaminants 21(1):1-31.
Querci M, Jermini M, Eede GVD. 2006. The analysis of food samples for the presence of genetically modified organisms. World Health Organization-Joint Research Centre, European Commission Directorate General, Luxembourg.
Rhee KC. 1994. Functionality of soy proteins. Protein functionality in food systems:311-24.
Riaz MN. 1999. Soybeans as functional foods. Cereal foods world 44:88-93.

Roberts G, Lack G. 2005. Diagnosing peanut allergy with skin prick and specific IgE testing. Journal of Allergy and Clinical Immunology 115(6):1291-6.

Sampson HA. 1999a. Food allergy. Part 1: immunopathogenesis and clinical disorders. Journal of Allergy and Clinical Immunology 103(5):717-28.

Sampson HA. 1999b. Food allergy. Part 2: diagnosis and management. Journal of Allergy and Clinical Immunology 103(6):981-9.

Sampson HA. 2003. 9. Food allergy. J Allergy Clin Immunol 111(2 Suppl):S540-S7.

Sampson HA. 2004. Update on food allergy. Journal of Allergy and Clinical Immunology 113(5):805-19.

Scheibe B, Weiss W, Rueff F, Przybilla B, Gorg A. 2001. Detection of trace amounts of hidden allergens: hazelnut and almond proteins in chocolate. Journal of Chromatography B: Biomedical Sciences and Applications 756(1):229-37.

Schubert-Ullrich P, Rudolf J, Ansari P, Galler B, Fuhrer M, Molinelli A, Baumgartner S. 2009. Commercialized rapid immunoanalytical tests for determination of allergenic food proteins: an overview. Analytical and bioanalytical chemistry 395(1):69-81.

Shanti KN, Martin BM, Nagpal S, Metcalfe DD, Rao PV. 1993. Identification of tropomyosin as the major shrimp allergen and characterization of its IgE-binding epitopes. The Journal of Immunology 151(10):5354-63.

Sicherer SH, Sampson HA. 2007. Peanut allergy: emerging concepts and approaches for an apparent epidemic. Journal of Allergy and Clinical Immunology 120(3):491-503.

Sicherer SH, Sampson HA. 2009. Food allergy: recent advances in pathophysiology and treatment. Annual review of medicine 60:261-77.

Sicherer SH, Sampson HA, Burks AW. 2001. Peanut and soy allergy: a clinical and therapeutic dilemma. Allergy 55(6):515-21.

Siegel M, Schnur K, Boernsen B, Pietsch K, Waiblinger HU. 2012. First ring-trial validation of real-time PCR methods for the quantification of allergenic food ingredients. European Food Research and Technology:1-12.

Soares S, Mafra I, Amaral JS, Oliveira MBPP. 2010. A PCR assay to detect trace amounts of soybean in meat sausages. International journal of food science & technology 45(12):2581-8.

Stephan O, Vieths S. 2004. Development of a real-time PCR and a sandwich ELISA for detection of potentially allergenic trace amounts of peanut (Arachis hypogaea) in processed foods. Journal of agricultural and food chemistry 52(12):3754-60.

Studer E, Rhyner C, Luthy J, Hubner P. 1998. Quantitative competitive PCR for the detection of genetically modified soybean and maize. Zeitschrift fur Lebensmitteluntersuchung und-forschung A 207(3):207-13.

Tate FT, Dewitt C. 2008. Evaluation of a method to detect soybean DNA in soy protein and flour products using real-time PCR. Presented at 2008 Institute of Food Technologist annual meeting and Food Expo. New Orleans, LA.

Taylor SL, Gendel SM, Houben GF, Julien E. 2009. The Key Events Dose-Response Framework: A foundation for examining variability in elicitation thresholds for food allergens. Critical reviews in food science and nutrition 49(8):729-39.

Taylor SL, Hefle SL, Bindslev-Jensen C, Atkins FM, Andre C, Bruijnzeel-Koomen C, Burks AW, Bush RK, Ebisawa M, Eigenmann PA, Host A, Hourihane JO, Isolauri E, Hill DJ, Knulst A, Lack G, Sampson HA, Moneret-Vautrin DA, Rance F, Vadas PA, Yunginger JW, Zeiger RS, Salminen JW, Madsen C, Abbott P. 2004. A consensus protocol for the determination of the threshold doses for allergenic foods: how much is too much? Clin Exp Allergy. 2004 May;34(5):689-95.

Taylor SL, Hefle SL, Bindslev-Jensen C, Bock SA, Burks AW, Jr., Christie L, Hill DJ, Host A, Hourihane JO, Lack G, Metcalfe DD, Moneret-Vautrin DA, Vadas PA, Rance F, Skrypec DJ, Trautman TA, Yman IM, Zeiger RS. 2002. Factors affecting the determination of threshold doses for allergenic foods: how much is too much? J Allergy Clin Immunol. 2002 Jan;109(1):24-30.

Taylor SL, Hourihane JOB. 2008. Food allergen thresholds of reactivity. In: Metcalfe DD, Sampson HA, Simon RA, editors. Food Allergy – Adverse Reactions to Foods and Food Additives. 4th ed. Malden Blackwell Publishing. p. 82-9.

Taylor SL, Lemanske RF, Bush RK, Busse WW. 1987. Food allergens: structure and immunologic properties. Annals of allergy 59(5 Pt 2):93-9.

Torres YW, DeSalle R, Giribet G. 2002. Michele K. Nishiguchi, Phaedra Doukakis, Mary Egan, David Kizirian, Aloysius Phillips, Lorenzo Prendini, Howard C. Rosenbaum, Elizabeth. Techniques in Molecular Systematics and Evolution 5:249.

Soybeans and oil crops 2012 [Accessed 2012 10/18/12] Available from: http://www.ers.usda.gov/topics/crops/soybeans-oil-crops/trade.aspx#US.

Van der Velden VHJ, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, Van Dongen JJM. 2003. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. Leukemia 17(6):1013-34.

van Hengel AJ. 2007. Food allergen detection methods and the challenge to protect food-allergic consumers. Analytical and bioanalytical chemistry 389(1):111-8.

Van Hoef AMA, Kok EJ, Bouw E, Kuiper HA, Keijer J. 1998. Development and application of a selective detection method for genetically modified soy and soy? erived products. Food Additives & Contaminants 15(7):767-74.

Van Wijk F, Knippels L. 2007. Initiating mechanisms of food allergy: Oral tolerance versus allergic sensitization. Biomedicine & pharmacotherapy 61(1):8-20.

Wang BP, Li ZX, Pawar R, Wang XF, Lin H. 2012. PCR-Based Detection of Traces of Potentially Allergenic Soybean (Glycine max) in Food Matrices. Advanced Materials Research 343:618-24.

Weiner HL. 2000. Oral tolerance, an active immunologic process mediated by multiple mechanisms. Journal of Clinical Investigation 106(8):935-8.

Wolfe KA, Breadmore MC, Ferrance JP, Power ME, Conroy JF, Norris PM, Landers JP. 2002. Toward a microchip? ased solid? hase extraction method for isolation of nucleic acids. Electrophoresis 23(5):727-33.

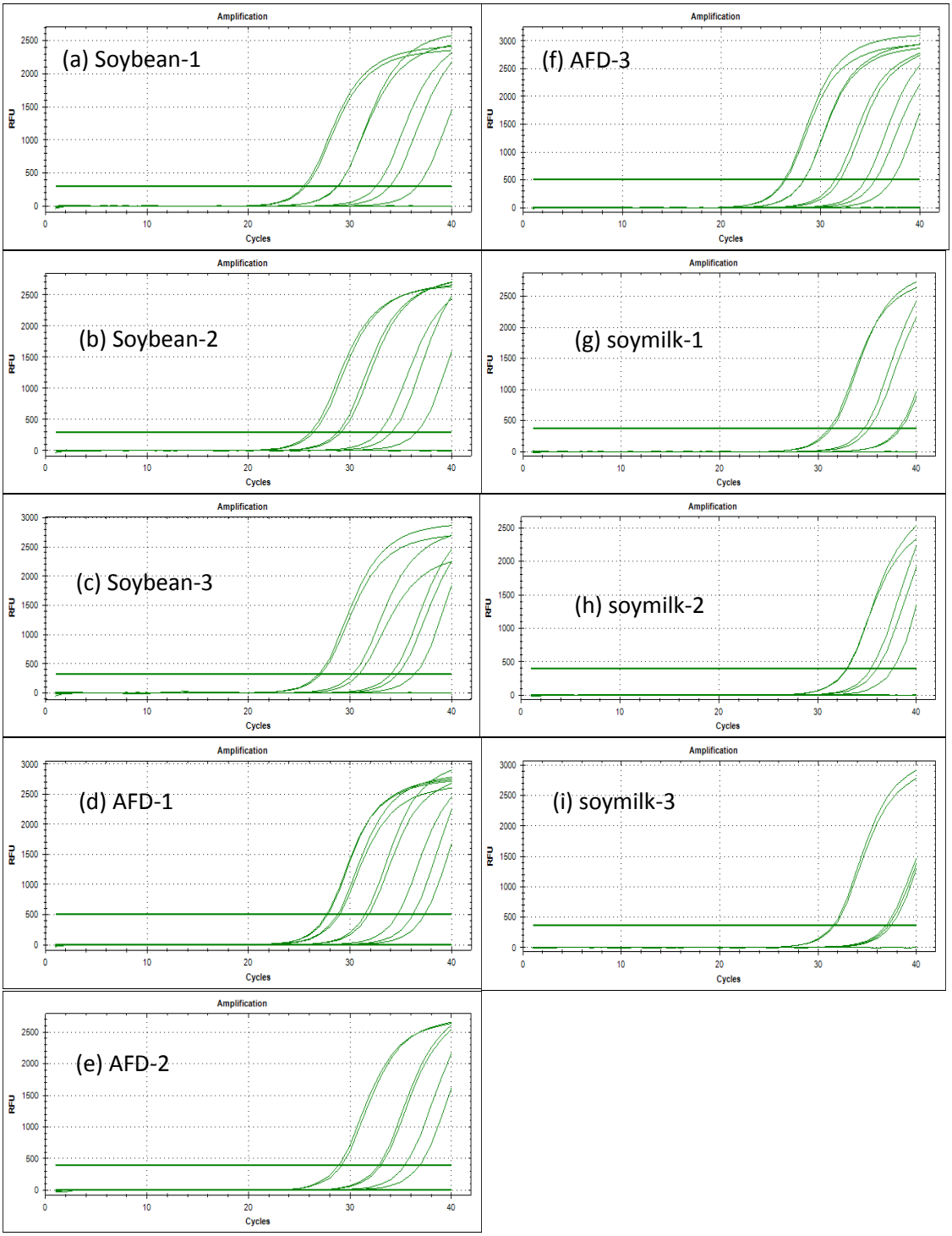
Wood RA. 2003. The natural history of food allergy. Pediatrics. 2003 Jun;111(6 Pt 3):1631-7.

Yamakawa H, Akiyama H, Endo Y, Miyatake K, Sakata K, Sakai S, Moriyama T, Urisu A, Maitani T. 2007. Specific detection of soybean residues in processed foods by the polymerase chain reaction. Bioscience, biotechnology, and biochemistry 71(1):269-72.

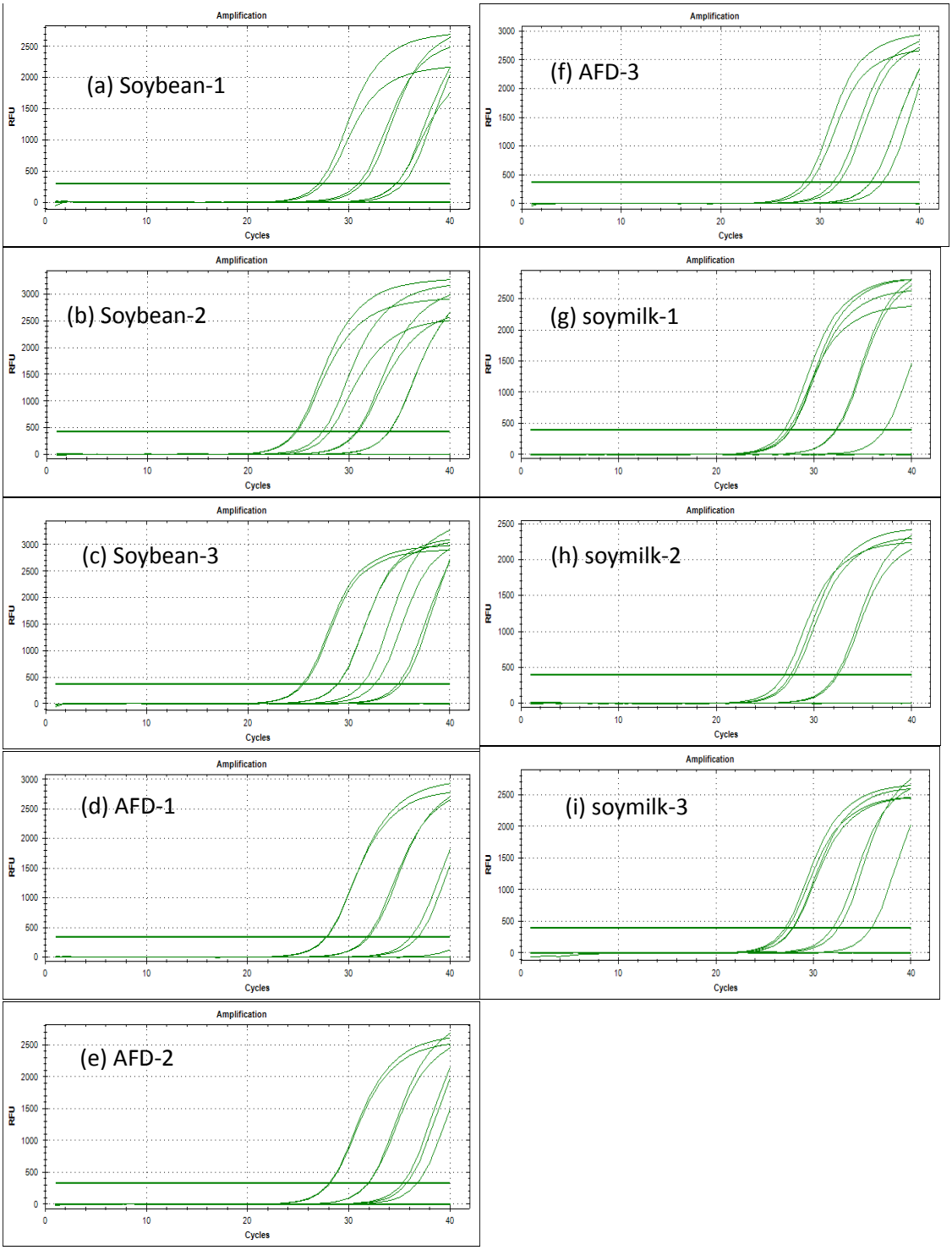
Yang DY, Eng B, Waye JS, Dudar JC, Saunders SR. 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. American journal of physical anthropology 105(4):539-43.

Yang WW, de Mejia EG, Zheng H, Lee Y. 2011. Soybean Allergens: Presence, Detection and Methods for Mitigation.

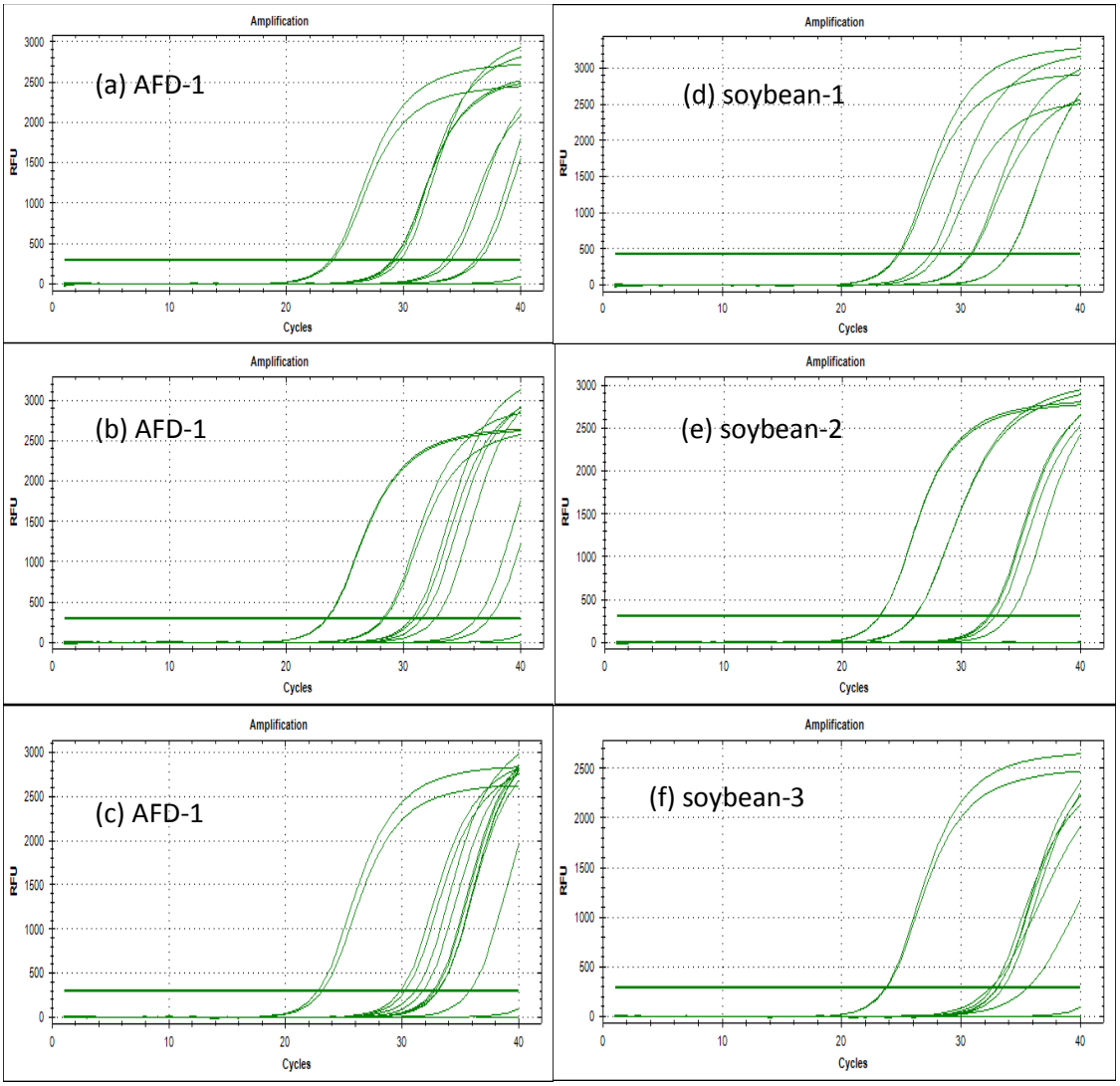
Appendix



A1.1: Real time PCR results for Qiagen DNA recovery methods. All samples (soybean, SPI soy protein isolates and soymilk) were extracted (n=3) and duplicate real-time PCR reaction.

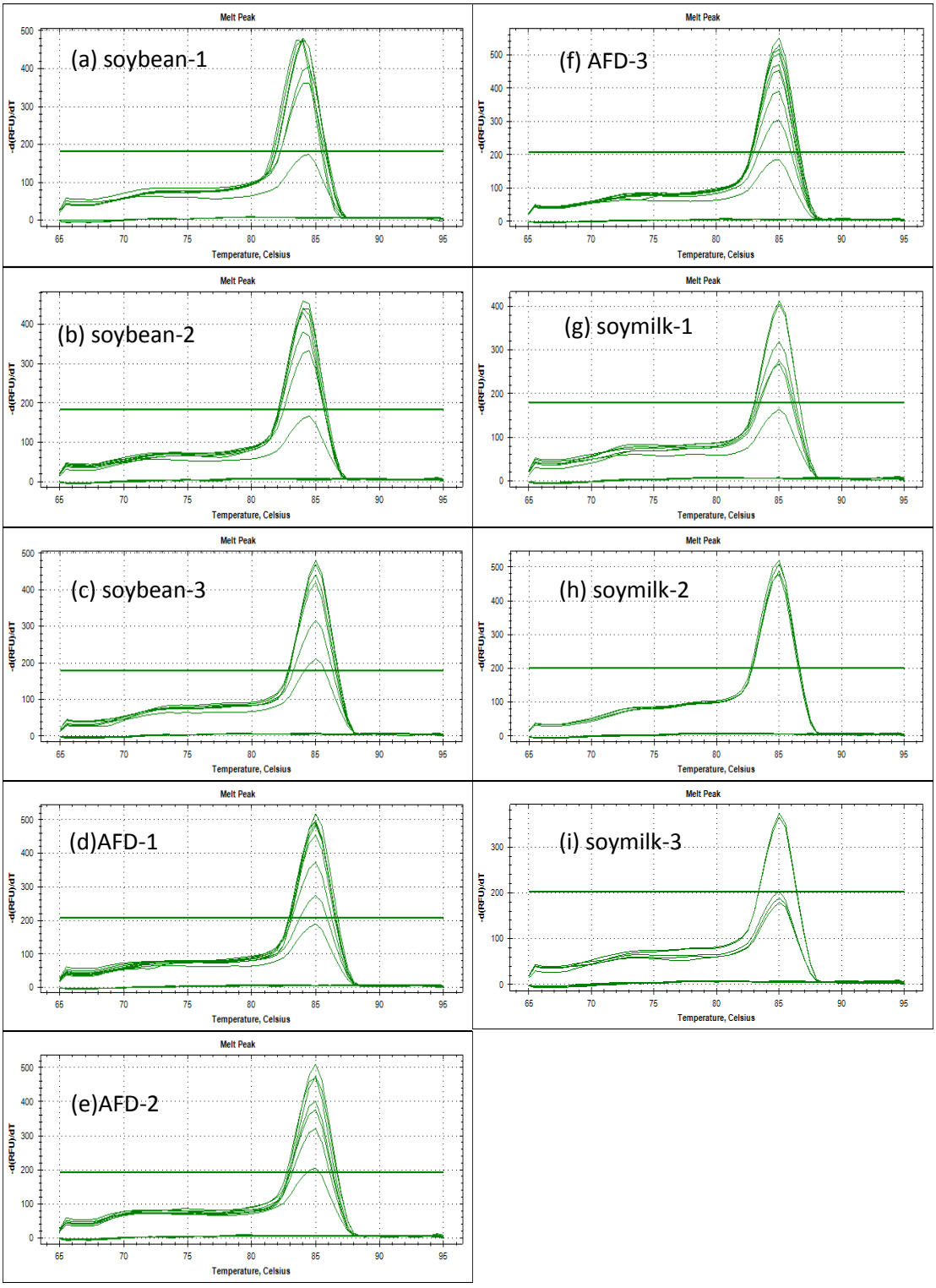


A1.2: Real time PCR results for Promega DNA recovery methods. All samples (soybean, SPI soy protein isolates and soymilk) were extracted (n=3) and duplicate real-time PCR reaction.

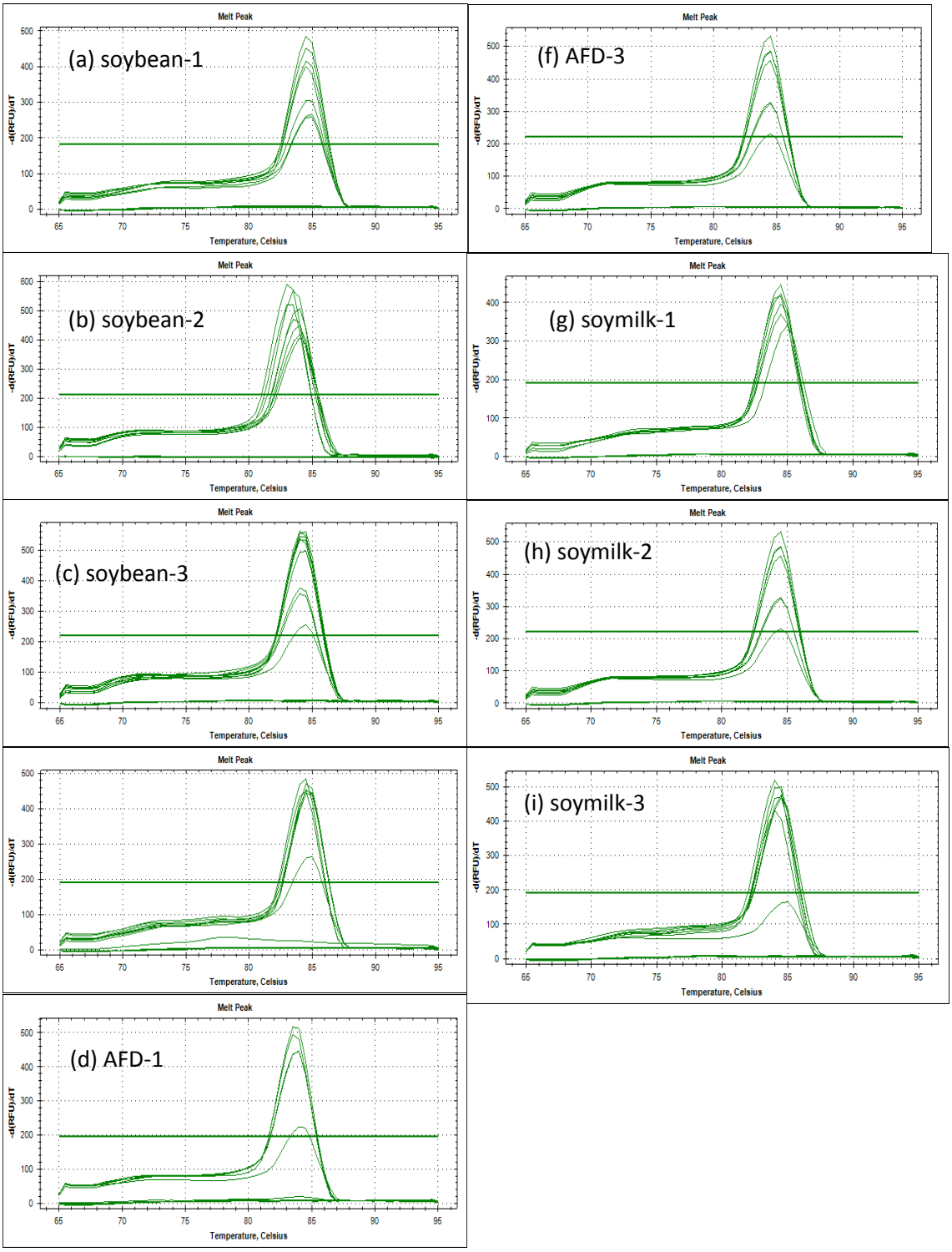


A1.3: Real time PCR results for magnetic beads DNA recovery methods for 1 g sample. All samples (soybean, SPI soy protein isolates and soymilk) were extracted (n=3) and duplicate real-time PCR reaction

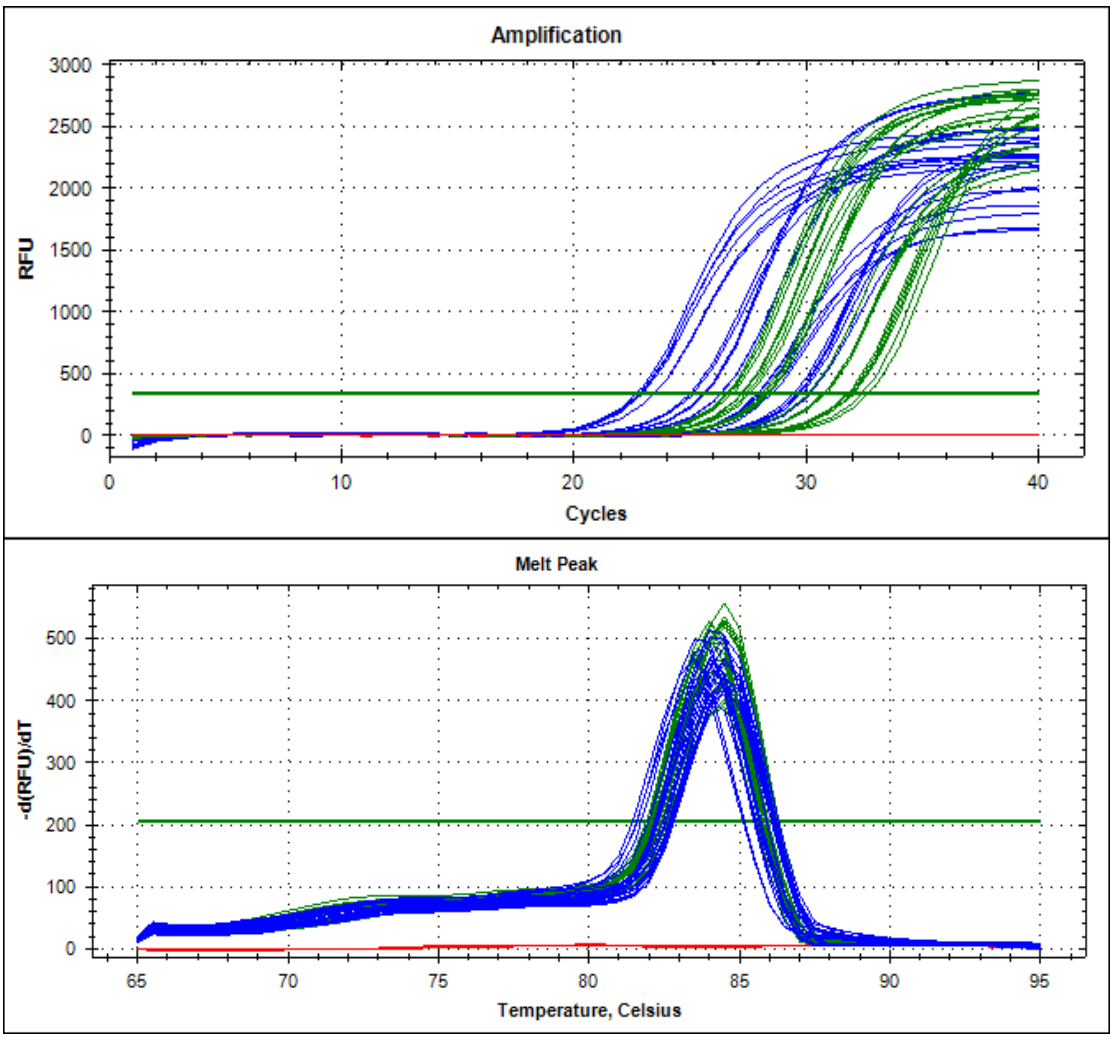
.



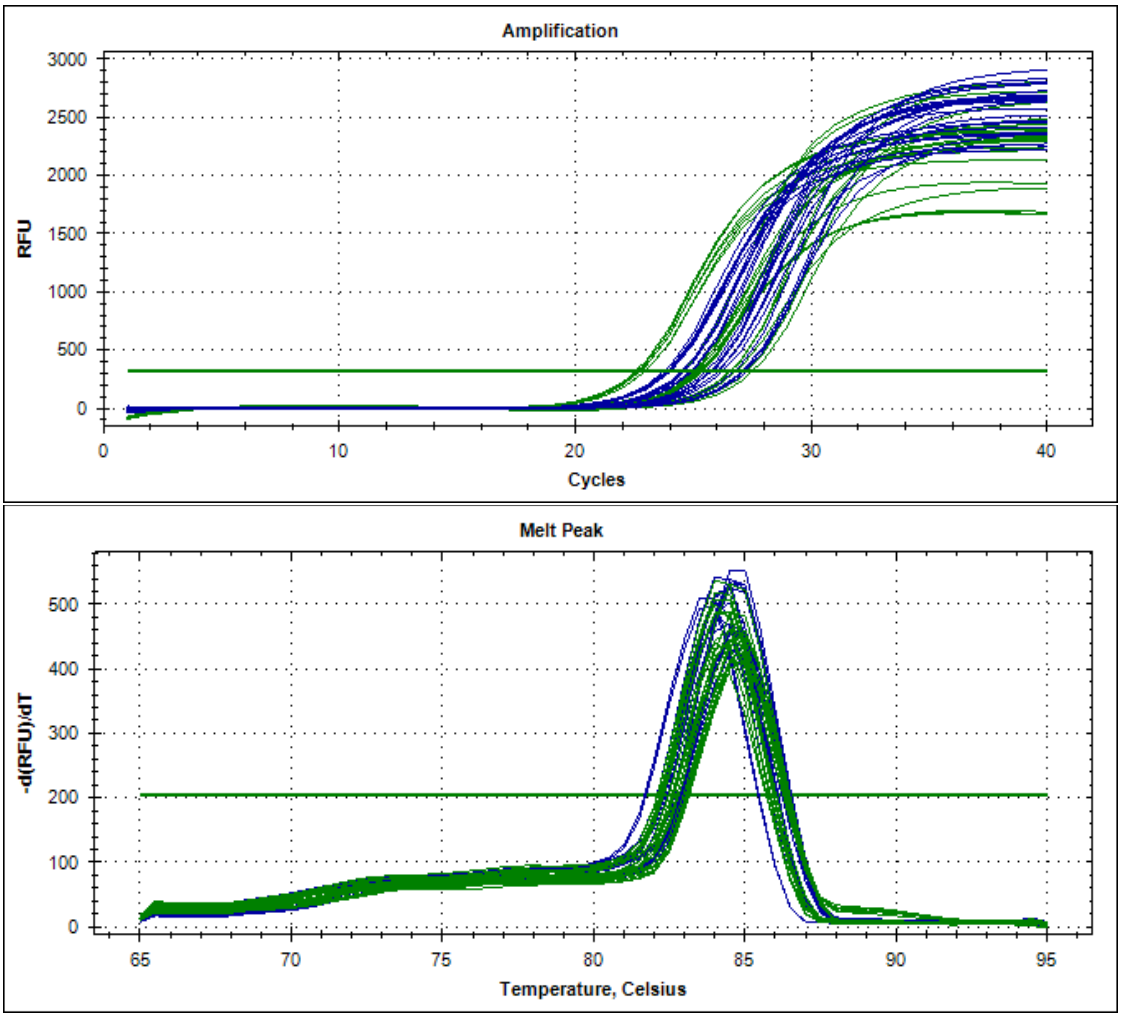
A1.4: Melting peaks for column DNA recovery methods. All samples (soybean, SPI soy protein isolates and soymilk) were extracted (n=3) and duplicate real-time PCR reaction.



A1.5: Melting peaks for magnetic beads DNA recovery methods. All samples (soybean, SPI soy protein isolates and soymilk) were extracted (n=3) and duplicate real-time PCR reaction.



A.1.6 Amplification graph and melting peaks for (heated and non-heated) flour and ranch matrices using both DNA recovery methods.

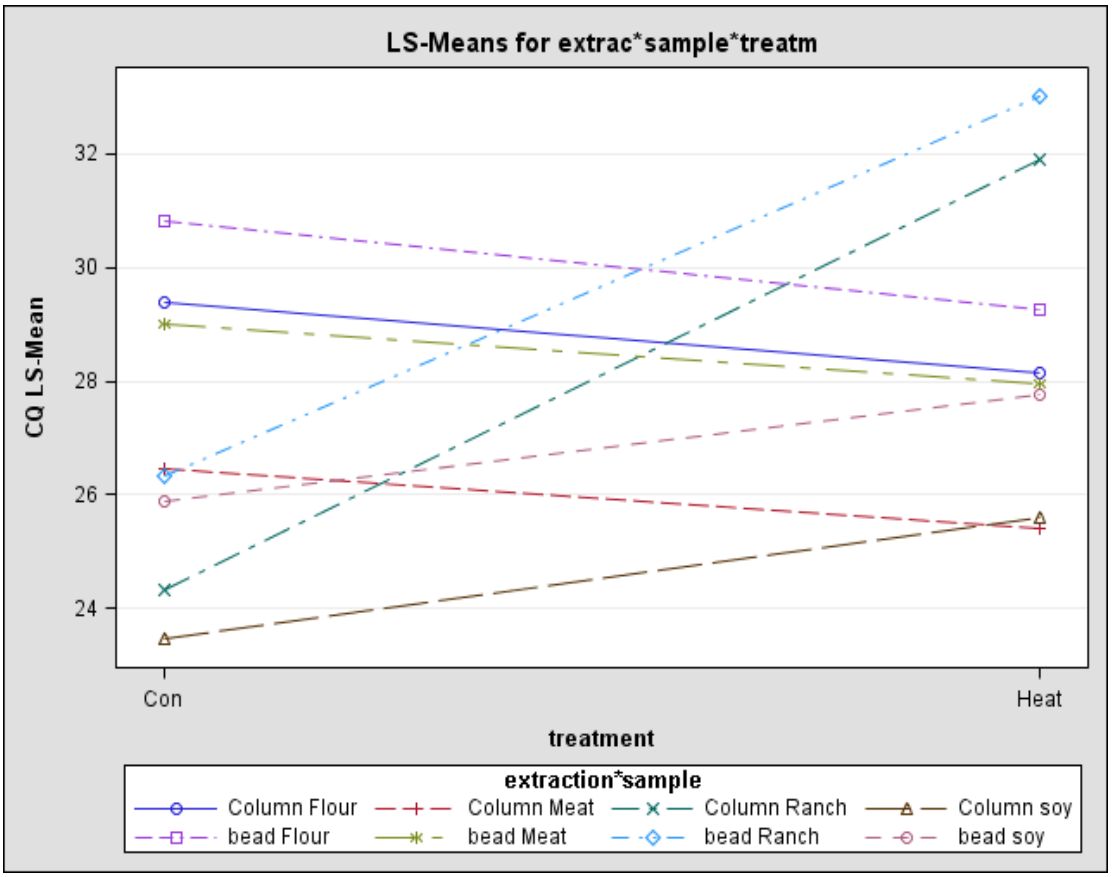


A.1.7 Amplification graph and melting peaks for (heated and non-heated) water and meat matrices using both DNA recovery methods.

A.1.8: Evaluation of column and magnetic beads DNA recovery methods for heated and control food samples. All food samples were spiked with 10% SPI. Cq values were adjusted to account for the dilutions required during DNA extraction. LS means were presented for all Cq values.

Ext Type	treatment	Soy	flour	meat	ranch	SE
Column	Con	23.46	29.38	26.47	24.32	0.12
	Heat	25.59	28.13	25.42	31.90	0.12
Magnetic bead	Con	25.90	30.82	29.00	26.33	0.12
	Heat	27.76	29.25	27.97	33.01	0.12

Ext= extraction.
All LS means were significantly different ($p<0.05$)



A.1.9: LS means plot using real-time PCR results (Cq values) to showcase the interaction between DNA recovery methods (column vs magnetic beads) and treatments (heat vs no heat (Con)) in four different food matrix (meat, ranch, flour and soy (water)).

A.1.10: Total DNA concentration and purity of heated and control food samples extracted with column and magnetic beads DNA recovery methods. All food samples were spiked with 10% SPI. Soy = control sample. LS means were presented for all DNA concentration and purity.

Sample	Treat	Total DNA concentration (µg/µl)		DNA Purity(A260/280)	
		Column	Magnetic beads	Column	Magnetic beads
Flour	Con	187	203.7[b]	1.5*	1.16
Flour	Heated	207.3[a]	363[ab]	1.11[c]*	1.14[c]
Meat	Con	56.7	88.7	1.58	1.1
Meat	Heated	74.3	28.7	1.66	1.08
Ranch	Con	19.9[a]	362.6[ab]	1.92*	1.22
Ranch	Heated	26.4[a]	169.3[ab]	1.23[c]*	1.22[c]
Soy	Con	38	96	1.42	1.05
Soy	Heated	38	45	1.55	1.04
SEM		37.55	37.55	0.059	0.059

Ext= extraction.
Treat= treatment.
[a] LS means appears in the same row (column vs. magnetic beads) for total DNA concentration are significantly different ($p<0.05$).
[b] LS means appears in the same column (Heated vs. con) within the same food sample for total DNA concentration are significantly different ($p<0.05$).
[c]LS means appears in the same row (column vs. magnetic beads) for DNA purity are not significantly different ($p>0.05$).
*LS means appears in the same column (Heated vs. con) within the same food sample for DNA purity are significantly different ($p<0.05$).