universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Development of a real-time PCR assay for potentially allergenic soy and its validation as part of a multiplex assay for soy, celery and white mustard"

verfasst von / submitted by

Leonie Lester, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 659
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Lebensmittelchemie / Master's degree programme Food Chemistry
Betreut von / Supervisor:	Univ.-Prof. Mag. Dr. Margit Cichna-Markl

Preface

The work of the underlying master's thesis was a cooperation between the University of Vienna and the Austrian Agency for Health and Food Safety (AGES). It was carried out from January 2021 to July 2021 at the Department of Molecular Biology and Microbiology, Institute for Food Safety, AGES, under the supervision of Univ.-Prof. Mag. Dr. Margit Cichna-Markl, Department of Analytical Chemistry, University of Vienna, and Ing. Verena Peterseil and Walter Mayer, Department of Molecular Biology and Microbiology, Institute for Food Safety, AGES.

List of contents

1 Abstract

Food allergy cases have been on the rise over the last decades, particularly in Western countries. Moreover, children seem to be even more affected than adults are, making this type of allergy not only an economic and safety burden but an emotional one, too. The most efficient management strategy of food allergy is dietary exclusion of the offending food, requiring first and foremost reliable food labelling and analytical methods to check the label's compliance. The main purpose of the master's thesis was to develop a qualitative singleplex real-time PCR assay for the detection of soy (*Glycine max*) and its validation as part of a qualitative triplex real-time PCR assay for the simultaneous detection of soy, celery (*Apium graveolens*) and white mustard (*Sinapis alba*) in food. The primers and the TaqMan probe, labelled with Cy5 as reporter dye and modified with a minor groove binder, of the singleplex PCR assay target a specific, short sequence within the chloroplast genome of soy. Six primer/probe systems were tested. The concentrations of primers and TaqMan probes were optimized. 86 species, both closely and not closely related to soy, were investigated for cross-reactivity. The limit of detection (LOD) was found to be 0.016 mg soy/kg food with a calculated C_t cut-off value of 39.3. The LODs and calculated C_t cut-off values for soy, celery and white mustard were found to be 0.06 mg soy/kg food and 38.5, 0.125 mg celery/kg food and 38.9, and 0.125 mg white mustard/kg food and 38.8, respectively.

2 Introduction

2.1 Food allergy

2.1.1 Definition

Correctly defining the term *food allergy* had caused a great deal of confusion among the population, not only limited to the average citizen but also physicians struggled. A mix of frequent occurrence of adverse reactions to food and different interpretations created the need for standardization (Bruijnzeel-Koomen et al., 1995; Untersmayr & Jensen-Jarolim, 2006; Johansson et al., 2004).

In 1995 the European Academy of Allergy and Clinical Immunology first published a classification system for adverse reactions to food exclusively based on mechanisms to standardize and also simplify diagnostics (Bruijnzeel-Koomen et al., 1995). As knowledge had been growing, several revisions of the original classification were made with the update of the Academy's position paper in 2003 (Johansson et al., 2004) being the last one. The most recent version of the classification system of adverse food reactions is displayed in Figure 1.

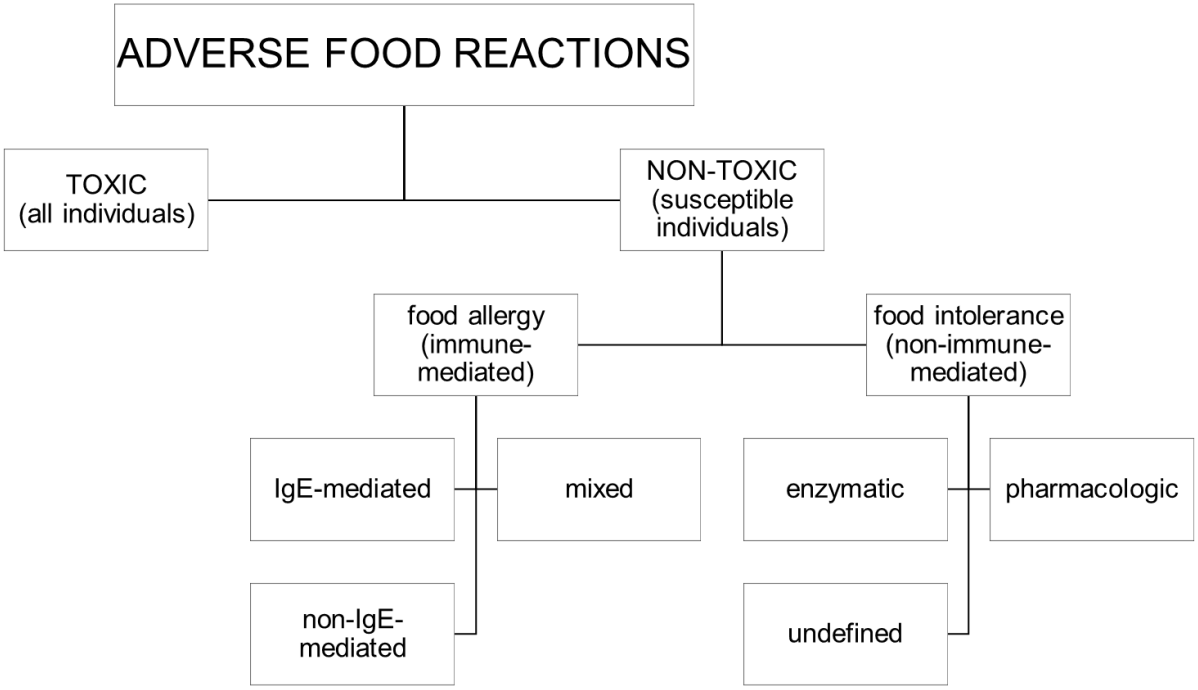

Figure 1 Classification system of adverse food reactions. Adapted from Asero et al., 2007 and Sampson, 2004.

The generic term *adverse food reactions* discriminates between toxic and non-toxic reactions. *Toxic reactions* describe reactions that are the result of an individual's exposure to a toxic food or component at an eliciting dose (ED), for example (e.g.) food poisoning, whereas *non-toxic*

reactions only occur in individuals susceptible to certain food or components. It is further differentiated between *immune-mediated reactions* – food allergy – and *non-immune-mediated reactions* that comprise *enzymatic, pharmacologic and undefined reactions* (Johansson et al., 2004; Johansson et al., 2001). To give an example, lactose intolerance is a classic non-immune mediated reaction; the ingested lactose as a food component, e.g. in milk, cannot be digested properly due to the lack of the enzyme lactase (enzymatic reaction) (Paige, 2005). The update of the European Academy of Allergy and Clinical Immunology position paper in 2003 proposed using the term *non-allergic food sensitivity* for non-immune-mediated reactions to avoid misinterpretation of the alternative description *food intolerance* (Johansson et al., 2004). Still, those two terms – food allergy and food intolerance – are often incorrectly used interchangeably (Boyce et al., 2010).

The most recent definition of *food allergy* is "*an adverse health effect arising from a specific immune response that occurs reproducibly on exposure to a given food*" (Boyce et al., 2010). A food allergy can be non-immunoglobulin E (non-IgE)-mediated (cell-mediated) and IgE-mediated, the latter being well-investigated. When both IgE- and non-IgE-mediated/cell-mediated reactions are involved, it should be referred to as *mixed* pathophysiology (Sampson, 2004). In 2010, the National Institute of Allergy and Infectious Diseases (NIAID)-Sponsored Expert Panel additionally suggested the introduction of a fourth subcategory termed *cell-mediated* describing the clinical disorder allergic contact dermatitis. What is more, according to the expert panel, celiac disease should be categorized as non-IgE mediated (Boyce et al., 2010).

2.1.2 IgE-mediated immune response

All kinds of antigens are able to trigger immune responses, leading to different hypersensitivity reactions. Non-infectious environmental antigens, as ingested food antigens, can induce an IgE antibody response. The outcome is an *immediate* hypersensitivity reaction (type I out of four types) shortly after ingestion (Abbas et al., 2016). The response can be divided into two phases: induction phase (sensitization) and triggering phase (provocation). The triggering phase consists of an acute- and a late-phase reaction (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014; Abbas & Lichtman, 2016).

After food ingestion, IgE antibodies are excessively produced by the immune system as a response to this particular food antigen, representing the first phase, *sensitization*. IgE antibodies can bind to both high-affinity FcεI and low-affinity FcεII receptors (Abbas et al., 2016). FcεI receptors are found on mast cells and basophils, whereas FcεII receptors are found on macrophages, monocytes, lymphocytes as well as eosinophils and platelets

(Sampson & Burks, 1996; Burks & Ballmer-Weber, 2006). The second phase, *provocation*, is triggered by repeated exposure to the same antigen (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). The specific, cell-bound IgE molecules bind the antigen and become cross-linked generating signals from the associated receptors (Fcε receptors). This signaling leads to mast cell granule release, enzymatic generation of leukotrienes and prostaglandins, and synthesis of cytokines. Biogenic amines such as histamine, released from the granules and the secreted prostaglandins, cause acute vascular changes, which results in increased blood vessel permeability (Abbas et al., 2016). This condition also promotes smooth muscle contraction and mucus secretion. Acute effects usually occur in a matter of a few minutes (min) after ingestion of the trigger food (Burks & Ballmer-Weber, 2006; Burks et al., 2012). Contrary to this acute phase of the immediate hypersensitivity reaction, there is also the late-phase reaction representing an inflammatory response. In the late phase, leukocytes are recruited to the site of mast cell degranulation with symptoms appearing within a few hours (Abbas et al., 2016).

2.1.3 Clinical symptoms

Due to an existing food allergy, immune responses can lead to clinically definable symptoms in an individual already sensitized to a certain food antigen. Provocation and associated clinical symptoms, if developed, only occur with sensitization to this certain food antigen as the prior step, making those two phases – sensitization and provocation – a requirement for clinical manifestation (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

Generally speaking, manifestations of clinical symptoms are related to different organ systems. An overview of the most common ones including associated clinical symptoms is given in Table 1.

Table 1 Common clinical symptoms of food allergy. Adapted from EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014 and Burks et al., 2012.

organ system	clinical symptom
eyes	conjunctivitis
respiratory tract	asthma
	rhinitis
	cough
	stridor
gastrointestinal tract	oral allergy syndrome (OAS, pollen-associated food allergy syndrome)
	nausea/vomiting
	gastro-oesophageal reflux disease
	diarrhea
	enteropathies
	abdominal pain
	constipation
	infantile colic
	failure to thrive
skin	atopic dermatitis
	pruritus
	angioedema
	urticaria
	erythema
systemic	anaphylaxis

Allergic reactions can manifest mildly to severely, with food anaphylaxis as an example of the most severe but at the same time one of the rarest reactions (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). The symptoms' onset after food ingestion and duration can also vary based on a multitude of factors; for example, it can vary individually and based on pathology (IgE-mediated, mixed or cell-mediated) as well as the symptom itself. Likewise, the severity of the allergic reaction is affected by multiple factors, such as the amount of ingested food and food processing, e.g. raw or cooked, and the absorption rate. The latter can be increased when the trigger food was eaten on an empty stomach (Burks et al., 2012).

The OAS, also known as pollen-associated food allergy syndrome, is one of the most common clinical features of food allergy (Sampson, 2005; Burks et al., 2012). In particular frequently occurring in adults, it is an IgE-mediated immediate-type immune reaction causing contact urticaria within a few minutes after food intake. This type of urticaria comprises lips, oral mucosa, as well as pharynx (Amlot et al., 1987; Ortolani et al., 1988). The skin is an organ system regularly associated with IgE-mediated allergic reactions to food, in which the course is rather mild.

Among clinical symptoms associated with food allergy, anaphylaxis is the most serious one (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Anaphylaxis is by definition a severe, systemic or generalized immediate hypersensitivity (IgE-mediated) reaction characterized by shock and airway obstruction resulting from mast cell degranulation in many tissue sites, typically after exposure to an antigen, which is injected or ingested as food (Abbas et al., 2016). When not treated properly and/or fast enough, this condition can be fatal. Various organ systems can be involved; this includes the skin, the mouth, the pharynx, the respiratory tract as well as the cardiovascular system, either separately or combined. Symptoms may occur within a few minutes after food contact. However, a reaction only seconds after exposure is not unusual either (Yunginger, 1988; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). What is more, even traces of the food allergen can be sufficient to trigger an anaphylactic reaction (Morisset et al., 2003). This reaction is characterized by (the combination of) symptoms like itching, erythema generalized, urticaria, nausea, vomiting, dyspnea, dizziness, palpitations, fainting or even collapsing. Adolescents and young adults with both allergy to peanuts or tree nuts and asthma are vulnerable groups (Bock et al., 2001). In addition, there are certain augmentation factors, such as alcohol, acetylsalicylic acid, non-steroidal anti-inflammatory drugs, menstruation, illness and, particularly, exercise, that in combination with ingesting certain food can trigger food-dependent anaphylaxis, while without one of these stimuli the anaphylactic reaction would not occur. Concerning food-dependent exercise-induced anaphylaxis, symptoms develop during the session when the offending food was taken in a few hours before exercising (Feldweg, 2017; Sicherer & Sampson, 2018; Niggemann & Beyer, 2014).

Even though those common clinical features of food allergy are mostly regarded as well-investigated, other adverse reactions to food may mimic the symptoms complicating diagnosis. Also, determining which food or food component elicited the allergic reaction can be very challenging, which in addition complicates allergy management (Sicherer & Sampson, 2014, 2018). The mainstay of allergy management is dietary exclusion of the offending food, requiring, besides knowledge of this particular trigger, first and foremost reliable food labelling

and analytical methods (Muraro et al., 2014; Holzhauser et al., 2020; Sicherer & Sampson, 2018).

2.1.4 Epidemiology

Allergy prevalence is difficult to pinpoint as multiple factors need to be brought into the equation. Sicherer et al. identified such factors as the definition of allergy, methodology, study population and geographical variation, dietary exposure as well as age, race, ethnicity, all of which interplaying in a complex network (Sicherer, 2011; Sicherer & Sampson, 2018).

Any food containing proteins is able to trigger an allergic reaction. Yet, only a small number of foods – the so-called “big eight” – account for the vast majority of food allergies. These foods comprise peanut, tree nuts, fish, crustacean shellfish, milk, egg, wheat and soy (Boyce et al., 2010; Sicherer & Sampson, 2018). In 2010 estimates showed that more than 1-2 % but less than 10 % of the world’s population suffers from food allergy, unknowing whether the prevalence was raising at that point (Chafen et al., 2010). Follow-up research concluded that almost 5 % of adults and 8 % of children are affected by food allergy (Sicherer & Sampson, 2014). In the same paper, the conclusion was drawn that the occurrence of food allergy had been raising but the reason for the raise was and still is not fully understood (Sicherer & Sampson, 2014, 2018). Young children, if allergic, tend to develop a much more severe immune reaction to certain foods. On the other hand, allergies occurring during (early) childhood generally outgrow more often than allergies developed as an adult (Sampson, 2004). Data may well point up that mild forms of food allergy are more frequent (Asero et al., 2007). Also noteworthy is that data on food allergy based on the patient’s self-evaluation is often overrated (Rona et al., 2007; Nwaru et al., 2014).

Within the framework of the so-called EuroPrevall project, Rona et al. drew on an extensive range of sources to assess the prevalence of food allergy in the European Union (EU). Data from food challenges showed prevalence rates for fish and shellfish near 0 %, milk 0 % to 3 % and egg 0 % to 1.7 % (Rona et al., 2007).

Nwaru et al. found the prevalence of food-challenge-defined allergy to cow’s milk, egg, wheat, soy, peanut, tree nuts, fish and shellfish to be 0.6 %, 0.2 %, 0.1 %, 0.3 %, 0.2 %, 0.5 %, 0.1 %, and 0.1 %, respectively, among Europeans. Milk and egg allergy were more frequent in younger children, whereas allergy to peanut, tree nuts, fish and shellfish were more frequent in older children. Additionally, there appeared to be a geographic difference in allergy occurrence: except for soy and peanut allergy, food allergy was more widespread in Northern Europe regions (Nwaru et al., 2014).

The population-based HealthNuts study (Melbourne, Australia) revealed the highest rates (11 %) of food allergy compared to every other country worldwide at that point (Osborne et al., 2011). In a follow-up analysis of the same cohort, the children's allergy occurrence was reassessed three years later (age one/age four); a decrease was observed but the total number was still high (Peters et al., 2017). In 2018, another Australian population-based study (SchoolNuts study) aimed to pinpoint allergy prevalence, obtaining a total of 4.5 %. The most common trigger food was peanut (2.7 %) and tree nut (2.3 %) (Sasaki et al., 2018).

A recent study (2019) did research on food allergy prevalence in South Africa; the authors also aimed to discriminate between the prevalence in rural and urban regions. The results of the overall population-based prevalence were similar to those from Western countries. The differentiation between rural and urban regions within South Africa, however, varied immensely. For urban regions, a much higher allergy prevalence was found (Botha et al., 2019).

In a nutshell, food allergy cases have been on the rise over the last decades. The Western population/urban regions seem to be more affected. There is a more frequent occurrence in children compared to adults (Chafen et al., 2010; Sicherer & Sampson, 2018; Boyce et al., 2010).

2.2 Food allergens

2.2.1 Background knowledge

To date, approximately 400 food allergens have been characterized. A summary of all these allergens including isoforms is given on a regularly updated official website (WHO/IUIS Allergen Nomenclature Sub-Committee, Allergen Nomenclature, 2021). The first successfully established systematic nomenclature of allergens, the International Union of Immunological Societies (IUIS) Allergen Nomenclature, originated in the 1980s. This original nomenclature forms the basis of the revised version that is in use at present: three letters represent the genus, the single letter following stands for the species name according to the Linnaean taxonomic system and the number in the end names the chronological order, in which the allergen was identified (Chapman et al., 2007; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014; King et al., 1994), e.g. soy (*Glycine max*) is listed as Gly m 1, (Gly m 2, *no food allergen*), Gly m 3, and so forth. Depending on the frequency (more or less than 50 %) subjects in a subject group exhibit allergen-specific IgE-binding to the allergen in question in a defined test system, food allergens are classified as *major* allergen (> 50 %) or *minor* (< 50 %) allergen. However, this terminology only refers to specific antibody binding, which does not automatically mean clinical symptoms occur.

Food allergens are proteins or glycoproteins that are foreign to the body (King et al., 1994; Chapman, 1988). It is almost always the case that the allergenicity of a food is the result of multiple allergenic proteins acting together. The protein's epitopes that bind IgE antibodies give the molecule its allergenic potential. Important to note is that the allergenicity of a certain protein cannot be precisely predicted, even though properties might be shared. Only with the inclusion of clinical as well as immunological information, the complex interplay of a protein's allergenicity can be described (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

Properties commonly shared by food allergens are structural properties and their stability against human digestion and food processing (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

2.2.2 Structural properties

Plant food allergens can be grouped into protein families and superfamilies depending on both their three-dimensional structure and their biological function (Breiteneder & Radauer, 2004). Though, they only belong to a fraction of the thousands of protein families known (Jenkins et al., 2005; Pfam, 2021). Lots of those allergens can be classified into the prolamin superfamily or the cupin superfamily. In addition, many plant food allergens are homologous to so-called pathogenesis-related proteins (PRs) (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). PRs consist of 14 protein families, which are part of the plant defense system (van Loon & van Strien, 1999). Besides those three main (super)families, there are other allergenic structural and metabolic proteins, e.g. profilins (structural proteins) or flavin adenine dinucleotide-dependent oxidases, just to name a few.

The prolamin superfamily comprises the largest subset of plant food allergens. A prominent example are 2S seed storage albumins. They are important for the development of seeds. Major allergens, in e.g. mustard seeds, are categorized into this subgroup. Another example are non-specific lipid transfer proteins frequently occurring in fruits from the *Rosaceae* family (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014), e.g. apple and peach. They are able to elicit severe allergic symptoms (van Ree, 2002).

Proteins of the cupin superfamily comprise the major globulin storage proteins causing the majority of allergies against nuts and legumes. They can be distinguished based on the number of cupin domains within the protein. Based on their sedimentation coefficient globulins can be further discriminated into the globulins vicilins and the legumins. The best-studied vicilin is the major peanut allergen Ara h 1 accounting for the majority of anaphylaxis cases with fatal outcome (Burks et al., 1991; Breiteneder & Radauer, 2004).

Food allergens originating from animals can be grouped into three families: calcium-modulated ("EF-hand") proteins, tropomyosins and caseins. In comparison to plant food allergens, they are less abundant (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

Furthermore, two types of epitopes are discriminated based on their structure: linear epitopes and conformational epitopes (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Linear epitopes consist of a continuous sequence of amino acids residues, whereas conformational epitopes consist of amino acid residues that are not in a sequence. Yet, these residues come within close proximity as a result of protein folding (Abbas & Lichtman, 2016). Denaturation mostly provokes modification or disintegration of conformational epitopes. Linear epitopes, on the contrary, remain unaltered (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Heat, pressure, radiation or ultrasound mainly have an impact on conformational epitopes as these physical processes change both the secondary and tertiary structure of proteins. Linear epitopes, if altered at all, may be affected by biochemical food processing, e.g. fermentation or enzymatic hydrolysis. New allergens, so-called neo-allergens, can be the result of Maillard reactions between food allergens and other compounds present in the food (Rahaman et al., 2016).

2.2.3 Effects of food processing on allergenicity

In general, food processing can modulate the allergenic features of food proteins. A processed food may exhibit reduced or enhanced allergenicity; some allergenic proteins, though, are not affected at all by processing (Rahaman et al., 2016). Microbial fermentation and enzymatic or acid hydrolysis, in particular coupled with heat treatment, can decrease the integrity of allergenic proteins so as to allergic reactions are not triggered anymore in an allergic subject. Pressure treatment may have similar effects (Verhoeckx et al., 2015).

2.2.4 Cross-reactivity

Cross-reactivity occurs when IgE antibodies bind not only one antigen but several. The more similar the 3D structure of proteins, the more likely cross-reactivity occurs. Additionally, the homology of the amino acid's sequence as well as the presence of common epitopes among proteins play a role. Both linear and conformational epitopes can be cross-reactive (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

Highly cross-reacting allergen groups are, for instance, profilins and lipid transfer proteins (LTPs) (Bonds et al., 2008; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). If sensitization, as the first step of an IgE-mediated immune response, happens due to

inhalation of allergens (pollen), food allergy may manifest itself as the so-called pollen-food allergy syndrome. Hence, with an existing birch pollen allergy, the likelihood to develop an allergy to apple, hazelnut, carrot and celery increases. Other prominent examples are cross-reactions between latex and fruits, dust mite and shrimp tropomyosin, and mold and spinach (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

2.2.5 European legislation of food allergen labelling

To enable allergic individuals to successfully avoid food allergens, full information about food components is required. In the EU, it is set down in food law that information on ingredients or processing aids potentially causing allergies or intolerances must be provided by the producer. This legislation applies to certain substances or (derived) products listed in *Annex II* of the Regulation (EU) No 1169/2011 of the European Parliament and the Council. If the substance or product is still present in the end product including altered forms, it must be properly labelled in any case. Allergen labelling shall be integrated into the list of ingredients and emphasis shall be added on the potential allergen within the text by using another font or style. As of December 2014, allergen labelling has been mandatory for non-prepacked food too. For non-prepacked food, which does not have a list of ingredients, allergen information must be given with the help of the term “contains [name of the substance or products]”.

Annex II includes the following 14 substances or (derived) products potentially causing allergies or intolerances:

"1. Cereals containing gluten, namely: wheat (such as spelt and khorasan wheat), rye, barley, oats or their hybridised strains, and products thereof, except:
(a) wheat based glucose syrups including dextrose ([1]);
(b) wheat based maltodextrins ([1]);
(c) glucose syrups based on barley;
(d) cereals used for making alcoholic distillates including ethyl alcohol of agricultural origin;
2. Crustaceans and products thereof;
3. Eggs and products thereof;
4. Fish and products thereof, except:
(a) fish gelatine used as carrier for vitamin or carotenoid preparations;
(b) fish gelatine or Isinglass used as fining agent in beer and wine;
5. Peanuts and products thereof;
***6. Soybeans and products thereof,** except:*
(a) fully refined soybean oil and fat ([1]);
(b) natural mixed tocopherols (E306), natural D-alpha tocopherol, natural D-alpha tocopherol acetate, and natural D-alpha tocopherol succinate from soybean sources;
(c) vegetable oils derived phytosterols and phytosterol esters from soybean sources;
(d) plant stanol ester produced from vegetable oil sterols from soybean sources;
7. Milk and products thereof (including lactose), except:
(a) whey used for making alcoholic distillates including ethyl alcohol of agricultural origin;
(b) lactitol;
8. Nuts, namely: almonds (Amygdalus communis L.), hazelnuts (Corylus avellana), walnuts (Juglans regia), cashews (Anacardium occidentale), pecan nuts (Carya illinoinensis (Wangenh.) K. Koch), Brazil nuts (Bertholletia excelsa), pistachio nuts (Pistacia vera), macadamia or Queensland nuts (Macadamia ternifolia), and products thereof, except for nuts used for making alcoholic distillates including ethyl alcohol of agricultural origin;
9. Celery and products thereof;
10. Mustard and products thereof;
11. Sesame seeds and products thereof;
12. Sulphur dioxide and sulphites at concentrations of more than 10 mg/[kilogram] (kg) or 10 mg/litre [L] in terms of the total SO_2 which are to be calculated for products as proposed ready for consumption or as instructions of the manufacturers;
13. Lupin and products thereof;
14. Molluscs and products thereof.

([1]) And the products thereof, in so far as the process that they have undergone is not likely to increase the level of allergenicity assessed by the Authority for the relevant product from which they originated" (European Parlament and Council, 2011).

Furthermore, the law lays down options to give voluntary food information in case of "*possible and unintentional presence in food of substances or products causing allergies or intolerances*" (European Parlament and Council, 2011), so-called *cross-contamination*. This kind of (inadvertent) inclusion of food allergens is a ubiquitous problem in the food industry (Pádua et al., 2016). Foodstuff may be cross-contaminated during storage and shipping, manufacturing or also from carryover caused by insufficient cleaning of processing equipment (Hefle et al., 1996; Deibel et al., 1997; Huggett & Hischenhuber, 1998). This voluntary regulation tempts producers to label all sorts of their products with "May contain …" or "May contain traces of …" allergen statements (so-called "precautionary allergen labelling", PAL) being *on the safe side*, even though there might not be any allergenic substance in it. Precautionary labelling does, however, the direct opposite: instead of protecting vulnerable groups, it is rather a protection to producers themselves and restricts allergic individuals even more on food choices, further contributing to a decreased life-quality of affected people (Bundesinstitut für Risikobewertung, 2020; Lieberman & Sicherer, 2011). The other extreme to very limited food choices is that those foods with PAL are yet consumed, posing the risk of eliciting an allergic reaction. Allergic reactions occurring due to accidental exposure are a threat to public health (Blom et al., 2018).

One strategy to tackle this issue is to introduce regulatory thresholds for allergenic foods. According to the Concise Oxford English Dictionary (9th edition), in the discipline of Physiology, the term *threshold* is defined as "*a limit below which a stimulus causes no reaction*" (Fowler et al., 1995; Crevel et al., 2008). By this definition, the *threshold* would have to correspond with the lowest observed adverse effect level (LOAEL); it represents the amount of allergenic food triggering mild clinical symptoms in highly sensitive allergic subjects (Taylor et al., 2002). Importantly, this term is always associated with the second of the two phases involved in an allergic reaction, the elicitation phase (Crevel et al., 2008). Taylor et al., however, reported a few years later that in experimental settings (double-blind, placebo-controlled, food challenges, DBPCFCs) *individual* thresholds lie within the interval of the no observed adverse effect level (NOAEL) and LOAEL (Taylor et al., 2009). Though, the individual degree of sensitivity to allergens may differ enormously. A good example are individuals allergic to peanut. They typically react to minute amounts (down to 2 mg) (Taylor et al., 2002); but also soy, especially soy protein, in trace amounts can trigger reactions (Koppelman et al., 2004). Since risk assessment for the general population including all allergic individuals requires a for this group adapted dose, the *minimum eliciting dose* (MED) was introduced. The MED can be described as "*a threshold for a defined proportion of the allergic population" (Crevel et al., 2008)*.

The ultimate goal is to define legally binding, population-based amounts of allergenic foods (reference doses). By exceeding these reference doses in the product, labelling of the corresponding allergenic food would be mandatory, superseding labelling of potential traces of allergenic foods as a precaution. The most recent proposal for reference doses of allergenic foods is VITAL (Voluntary Incidental Trace Allergen Labelling) Scientific Expert Panel (VSEP) 3.0, a work done by Allergen Bureau (Allergen Bureau, 2019) and scientifically evaluated by the German Institute for Risk Assessment (Bundesinstitut für Risikobewertung, 2020). It includes proposed amounts of total protein for each of the 14 substances or (derived) products triggering an allergic reaction after ingestion in allergic subjects with a certain probability. These proposed amounts are based on the total protein content of the corresponding allergenic foods. In VITAL 3.0, some of the reference doses of the allergenic foods already discussed in the former version, VITAL 2.0, were adapted; amongst others the reference dose of soy, which was lowered. A reference dose for celery was determined for the first time (Bundesinstitut für Risikobewertung, 2020). The minimal eliciting dose 01, ED_{01}, is a newly introduced numeric value of Vital 3.0. ED_{01} is defined as the "*dose of the total allergen protein that is predicted to produce objective symptoms in 1 % of the allergic population*" (Allergen Bureau, VITAL Science, 2021). Alternatively in case of insufficient data, the minimal eliciting dose 05, ED_{05}, is used, which describes the same as ED_{01} but refers to 5 % of the allergic population. ED_{01} and ED_{05} of soy, celery and mustard are given in the following subchapters.

2.2.5.1 Soy

Cultivated soy (cultivated soybean, *Glycine max*) is an edible legume of Asian origin, which is taxonomically grouped into the family of *Fabaceae* (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). In addition, the large *Fabaceae* family comprises other agriculturally relevant plants (plant seeds), such as the common bean, lentil, chickpea, pea and peanut (Christenhusz & Byng, 2016; Graham & Vance, 2003).

In Europe, soy consumption has been rising in recent years, whereas in Asia and the USA soy and soy products are already an integral part of everyday nutrition (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Among European countries, in 2015-2019 Hungary was the leading agricultural producer of soybean, followed by Romania and Croatia. In this timeframe, agricultural soybean production was rising in Austria; in 2019 production reached its peak (76,341 [current thousand US$]) (Food and Agriculture Organization of the United Nations (FAO), Value of Agricultural Production, 2021). Soy products are particularly popular among vegetarians, vegans and individuals that want to restrain meat, e.g. for ethical reasons. Examples are soy flour, soy oil, soy milk and soy-based drinks, soy cheese, soy mayonnaise and soy flakes (Friedman & Brandon, 2001; Jayachandran & Xu, 2019; Kumar et

al., 2017; Rizzo & Baroni, 2018; Messina & Messina, 2010; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Also, fermented soybean products are gaining in importance (Jayachandran & Xu, 2019). Fermented soybean products include miso, soy sauce and tofu, to name but a few (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

Soy is one of the most important legumes, not only in terms of agricultural relevance but also of nutritional and technical functionality (Frias et al., 2008). It is the most commonly planted genetically modified crop worldwide (Saski et al., 2005). Soybeans score with their high protein content (38.2 %) (Souci et al., 2016). The seeds contain about 20 % oil, which typically serves as cooking oil (Friedman & Brandon, 2001; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Soy protein cannot be regarded as ideal, though, as it lacks the essential amino acid methionine (Friedman & Brandon, 2001). Nevertheless, on PDCAAS (protein digestibility corrected amino acid score), a parameter to demonstrate protein value in human nutrition, soy protein scored a value equally high to beef (Schaafsma, 2000). Soy has the highest content of protein among food crops (Koppelman et al., 2004). With improving protein isolation methods, the use of soy protein isolates as well as concentrates in foodstuff has been extended. In food industry, soy is now used as texturizer and emulsifier. Since it is a cheap protein source, soy protein is also frequently used as fillers, for instance, in meat preparations, potentially unlabeled. Moreover, it is becoming more and more attractive as a non-animal source in surrogate products, such as meat substitutes (Kumar et al., 2017; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014; Costa et al., 2017). Infant formulas based on soy are used as an alternative to formulas based on cow's milk to feed infants suffering from cow's milk allergy (Klemola et al., 2002; Businco et al., 1998). In feed industry, soy is a highly relevant source of protein, too (Frias et al., 2008). What comes with various advantages for industry and the general population, is an increasing health threat for individuals allergic to soy (Herian et al., 1990; Vidal et al., 1997; Herman et al., 2003).

Nwaru et al. systemically reviewed food allergy prevalence in Europe among all age groups between 2000 and 2012. Self-reported lifetime and food-challenge-defined soy allergy were shown to be 1.5 % (95 %, CI 1.2-1.8) and 0.3 % (95 %, CI 0.1-0.4), respectively. Specific IgE-positivity was with 3.2 % (95 %, CI 2.7-3.6) comparably high. Estimates of soy allergy prevalence seemed to be higher in younger children than in subjects of older age (Nwaru et al., 2014). In addition, Katz et al. did research on the prevalence of IgE-mediated soy allergy. Studies between 1909-2013 were systemically reviewed, resulting in a weighted prevalence of soy allergy of 0.27 % for the general population and 2.7 % for allergic children (Katz et al., 2014). Soy allergy is common in children (0.4 %) (Sicherer & Sampson, 2006). A survey conducted between 1993-1996 in Sweden included 45 young children severely reacting to

tree nuts, peanuts and/or soy (61 cases in total); four were fatal. The four youngsters who died from soy anaphylaxis (asthma) were highly allergic to peanut, too (Foucard & Malmheden Yman, 1999). Further research corroborates the assumption that children already allergic to peanut are more prone to have a soy allergy too (Savage & Johns, 2015). This data suggests that there may be cross-reactivity between soy and peanut allergens. Furthermore, pollen-food allergy syndrome is associated with soy allergy as birch pollen cross-react with soy allergens (Ballmer-Weber & Vieths, 2008). Soy allergy has, however, a high rate of resolution in the course of childhood. By the age of six, about 45 % of children have outgrown the allergy and the resolution appears to continue into adolescence (Savage et al., 2016).

Soy allergy may clinically manifest itself as, e.g. enterocolitis; however, also more severe symptoms, such as food protein-induced enterocolitis syndrome (FPIES), can occur. Anaphylaxis is rare with soy allergy but possible (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

The globulins (storage proteins) b-conglycinin (Gly m 5), glycinin (Gly m 6) and trypsin inhibitors represent more than 85 % of total soy protein and are the most relevant regarding allergenicity among Europeans. They are a main target for protein-based detection (Koppelman et al., 2004; García et al., 1997; Holzhauser et al., 2009). As of yet, there have been found 44 IgE-binding allergens in soybean (*Glycine max* and its wild progenitor *Glycine soja*), 14 of which are also by WHO/IUIS designated allergens (University of Nebraska–Lincoln, Food Allergy Research and Resource Program (FARRP) AllergenOnline.org database, 2021). Only a few of those allergens account for approximately 90 % of allergic reactions to soy (Frias et al., 2008). One of the suspected major soy allergens is Gly m Bd 30K. Gly m Bd 60K and Gly m Bd 28K may also be dominant among soy seed allergens (Ogawa et al., 2000). However, due to a lack of information concerning clinical reactivity of these IgE-binding proteins, it is not yet clear how clinically relevant they are (Verhoeckx et al., 2015).

In the majority of cases, soy allergic individuals do not severely react to both wholly refined soybean oil and fat because the levels of soy allergens after refinement are assumed to be below EDs (Verhoeckx et al., 2015; European Food Safety Authority, 2007). The degradation of allergenic proteins in soybean due to microbial fermentation has been an object of research. Investigations of Kobayashi indicated that soy sauce has hypoallergenic features. In the course of fermentation, soy protein is hydrolyzed into smaller peptides; each of the allergenic proteins contained in raw soy is cleaved and, therefore, no longer present in the fermented product (Kobayashi, 2005). Hefle et al., however, proved the opposite; they claimed soy sauce can retain some of its allergenic potential, even though the product underwent the process of fermentation (Hefle et al., 2005). Correspondingly, Frias et al. found only a decrease in

immunoreactivity of soybean foods after fermentation (Frias et al., 2008). Since the majority of experiments on soy and soy products allergenicity is based on antibody-based assays rather than high-quality investigations *in vivo* or sera from soy allergic donors, results should be treated with caution (University of Portsmouth, 2013; Verhoeckx et al., 2015). Research indicates that soy allergenicity may either decrease or remain unchanged due to food processing (Verhoeckx et al., 2015).

In VITAL 3.0, based on an increased amount of data compared to the former recommendations (VITAL 2.0), the ED_{01} for soy (i.e. soy drink, soy flour) was lowered to 0.5 mg protein (VITAL 2.0: ED_{01} = 1 mg protein) as well as ED_{05} was set to 10 mg protein (Bundesinstitut für Risikobewertung, 2020).

2.2.5.2 Celery

Celery (*Apium graveolens*) is taxonomically grouped into the family of *Apiaceae*. In human nutrition, celery roots, celery stalks and leaf celery are typically used, either raw (salad) or cooked/processed (an ingredient in soups, sauces, meat dishes, as celery powder and extract) (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). Celery allergy is the most common food allergy among adults in Central and Western European countries (M. Etesamifar, 1998; Ballmer-Weber et al., 2002; Ballmer-Weber et al., 2000). The severity of allergic reactions ranges from mild to severe, i.e. life-threatening anaphylaxis (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). In VITAL 3.0, ED_{01} and ED_{05} are 0.05 mg and 1.3 mg protein, respectively (Bundesinstitut für Risikobewertung, 2020).

2.2.5.3 Mustard

Mustard plants are food crops, which are taxonomically grouped into the family of *Brassicaceae*. The seeds of three mustard species, one of which is the so-called white or yellow mustard (*Sinapis alba*), are typically used in human nutrition (Rancé, 2003; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). *Mustard* (i.e. whole, ground and/or processed seeds) is frequently part of meat containing food preparations, such as hamburgers or spices (curry), salads and sauces (mayonnaise, ketchup) (Rancé et al., 2000; EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014). André et al. did research on the change in frequency of sensitization to potential (hidden) food allergens. In a population of 580 subjects living in France, sensitization to certain food including soy, mustard and celery was shown to be increasing in the course of 9 years of analysis. This increase may be linked to increased consumption of those foods (André et al., 1994). However, the prevalence in the general population is still uncertain (EFSA Panel on Dietetic Products, Nutrition and Allergies

(NDA), 2014). Clinical manifestations of mustard allergy include mild forms but also severe, systemic reactions (anaphylaxis) are possible (Figueroa et al., 2005). The reference doses for mustard as part of the VITAL 3.0 recommendations on allergen thresholds are for ED_{01} a value of 0.05 mg protein and for ED_{05} a value of 0.4 mg protein; compared to VITAL 2.0, it remained unchanged (Bundesinstitut für Risikobewertung, 2020).

2.2.6 Chloroplast genome of soy

The chloroplast genome of soy (*Glycine max*) is entirely sequenced. It consists of 152,218 base pairs (bp). This includes a pair of inverted repeats (25,574 bp) of identical sequence, which are divided by a small single-copy region (17,895 bp) and a large single-copy region (83,175 bp). Furthermore, the chloroplast genome has 130 genes in total, 60 % coding regions and 40 % non-coding regions are included and the content of the ratio of guanine to cytosine (G:C) and adenine to thymine (A:T) is 34 % and 66 %, respectively (Saski et al., 2005).

2.3 Analytical methods for the detection and quantification of food allergens

Analytical methods for the detection and quantification of food allergens have appreciably improved in recent years (Xu et al., 2021). Largely, they comprise DNA- and protein-based approaches.

Krska et al. defined requirements for analytical methods for the detection of allergenic proteins. They have to be sensitive, being able to detect minute amounts of the corresponding protein. As a general rule, the LOD must lie between 1 and 100 mg allergenic protein/kg food. The methods must be specific and allow for detection in all kinds of food matrices. The equipment must be easily applicable and fast in use (Krska et al., 2004). Analytical methods also require adequate reliability and reproducibility (Demeke & Jenkins, 2010). DNA-based methods tendentially exhibit lower detection limits than protein-based approaches. Protein-based methods' LODs lie in the range of high mg/kg to g/kg, whereas DNA-based methods' lie in the range of low mg/kg (Poms et al., 2004b; Holzhauser et al., 2020).

For routine analysis of allergenic foods, such as soy, celery and white mustard, PCR and enzyme-linked immunosorbent assay (ELISA) are commonly employed (Scharf et al., 2013).

With PCR, specific DNA stretches can be amplified and, thus, detected. Almost each of the 14 foods or (derived) products potentially causing allergies or intolerances are specifically detectable, except milk, egg, gluten as well as sulphites (Holzhauser, 2018; Holzhauser & Röder, 2015). The PCR method is explained in detail in subchapter 4.4.

Options for PCR targets have become more and more diverse, being no longer limited to genes encoding allergenic proteins. Thanks to advances in sequencing, also other specific DNA sequences located in mitochondria or chloroplast genomes now constitute potential targets (Holzhauser & Röder, 2015). So that PCR assays are specific, sensitive and efficient, target regions in the genome that are both adequately variable for differentiation of taxonomically related species and high in copy number are required (Caldwell, 2017). Comparing mitochondrial and chloroplast DNA, chloroplast DNA is structurally more variable (Daniell et al., 2016; Puente-Lelievre & Eischeid, 2018). Also, there are multiple copies of non-nuclear DNA in each of the many chloroplasts present in a plant cell (Alberts et al., 2002). The potential of chloroplast DNA in PCR assays has been repeatedly underlined (Mayer et al., 2019; Ladenburger et al., 2018).

ELISA as an immunoanalytical detection method enables the analysis of major allergens or groups of potentially allergenic proteins or mixtures of proteins of the trigger food *(antigens)* by antibody binding followed by colorimetric reaction. It is a comparatively simple method using low-cost equipment. Antibodies are specific and sensitive. Results are obtained within a few hours and data can simply be analyzed. An antigen-antibody complex and subsequently the antigen itself is quantified with the help of a standard curve generated with standards. There are two different formats of ELISA, competitive ELISA and sandwich ELISA, the latter being the most often applied type. The principle of the sandwich ELISA is based on two antibodies ultimately forming a “sandwich”. One antibody, the capturing antibody, is immobilized, e.g. on the bottom of a 96-well plate. The antigen – the allergenic protein of interest present in the sample – is captured by this first antibody; the second antibody, however, also binds to the antigen already complexed with the first antibody. This second antibody is labelled with an enzyme, which in the further course of the analysis reacts with substrate added so that a colored product is generated. The spectrophotometrically measured absorbance of the product is directly proportional to the concentration of the antigen (analyte, for example, an allergenic protein). There is a multitude of ELISA kits available on the market (Morinaga Institute of Biological Science, Food Allergen ELISA Kit Ⅱ, 2021; Elisa Systems, Elisa Kits, 2021; Eurofins, Allergen ELISA Kits, 2021). Those kits may differ, though, in antibody quality. What is more, they may utilize different target proteins, extraction and calibration procedures, all of which leading to highly varying results among brands and even batches. Food matrix effects, cross-reactivity, biological changes of the original sample material, inadequate protein extraction and a lack of multi-residue analysis are serious disadvantages of ELISA (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014; Senyuva et al., 2019; Poms et al., 2004a; Yeung, 2006).

As a consequence, some foodstuff is difficult to analyze with ELISA. Methods operating on the DNA level constitute a specific and sensitive alternative. Concerning the analysis of, for example, processed foods, DNA-based analytical methods are evidently superior to protein-based approaches as DNA is more stable than proteins. Evidence is given by a proficiency testing investigation for the detection of soy. In this investigation, Scharf et al. reported that performances of ELISA qualitatively detecting soy suffered. With ELISA a high number of false-negative results were obtained. With PCR, on the contrary, the qualitative results were reliable (Scharf et al., 2013; Poms et al., 2004a). There are conflicting findings on DNA degradation due to high temperatures over a long period during food processing as well as its effect on PCR performance (Iniesto et al., 2013; Costa et al., 2013). Earlier literature stresses food processing might differently influence the DNA of the allergenic food and allergenic proteins in respect to their integrity, possibly leading to a false negative and, therefore, incorrect safety evaluation (Allmann et al., 1993). Poms et al. stated that DNA is fairly stable against temperatures below 120 degree centigrade (°C); low pH (< 4), however, results in DNA truncation as well as do shearing forces and enzymatic degradation. Pressure does not seem to negatively affect DNA analysis (Poms et al., 2004a). Gryson did similar research on the effect of food processing on plant DNA degradation, showing amplification of degraded DNA can still be achievable (Gryson, 2010). In the case of highly processed foods, literature now supports the idea to design primers that flank a short sequence (amplicon < 200 bp), enhancing the chance to detect the target sequence even if the DNA strand has already been partly fragmented (Holzhauser, 2018). Nevertheless, food processing and its significant impact on both the analyte and its extractability needs to be taken into consideration (Bauer et al., 2004) as it will directly affect the accuracy of the assay (Platteau et al., 2011). Heat treatment has most certainly the greatest impact thereon (Senyuva et al., 2019). A false-negative result may also be due to dilution of the target to absence (Holzhauser & Röder, 2015). Taken together, PCR not only shows considerable advantages over antibody-based ELISA but the design, development and optimization of methods is easier (Holzhauser & Röder, 2015).

Important to note is that DNA is a surrogate target molecule of the allergenic compound. As part of data interpretation, quantified DNA needs to be extrapolated for proteins of the allergenic food. This kind of extrapolation is, however, also necessary for some antibody-based methods intended for food allergen detection (Poms et al., 2004a; Holzhauser & Röder, 2015). In brief outline, the principle of PCR as a detection method for food allergens is based on the assumption that positive results mean the presence of both the allergenic food and allergenic protein. Thanks to their good correlation, it is not considered an actual limitation (Holzhauser & Röder, 2015). Yet, research has been attending itself to the matter. Provided

calibration, the DNA amount (i.e. total allergenic food) can be converted into the relative amount of allergenic component (i.e. protein) using conversion factors (Holzhauser & Röder, 2015). For real-time PCR assays expressing results in an absolute manner, e.g. in pg, pg/milliliter (mL) or copy number, this conversion is not possible due to the lack of such factors for DNA. The influence of the matrix, though, is not considered with this approach (Köppel et al., 2010). Food is a complex matrix, in which inhibitory components commonly occur. Such inhibitors negatively affect PCR efficiency (Platteau et al., 2011). If they are not removed during DNA extraction, they can pose a significant problem to analysis. In the worst case, there is no signal at all owing to these inhibitory effects on amplification, which leads to an incorrect evaluation of the food sample (Poms et al., 2004a). With certified reference materials for allergen analysis still lacking, such limitations have to be tolerated. Certified reference materials would be required for each food allergen of analytical interest as well as for the matrix of the analyzed food itself as, in the end, in food analysis analytical performance is strongly dependent on the matrix (Costa et al., 2017). Another approach is to apply internal standards; a known amount of allergen is added to the sample (Platteau et al., 2011). There are several methodologies for the quantification of food allergens, all of which trying to minimize the influence of food matrix effects on amplification efficiency (Holzhauser, 2018).

Lots of commercial and research methods for the detection of soy(bean) DNA have been developed. Mostly, single-copy genes serve as target, such as the lectin gene (Holzhauser et al., 2020). Table 2 gives an extract of research methods developed in the last five years (2017-2021) with their most relevant features. The publications are listed chronologically with those for detection of soy only listed first, followed by those describing methods for the detection of other allergenic foods too. Table 3 lists relevant research multiplex PCR methods for the detection of soy, celery and/or white mustard with their most relevant features.

Table 2 Extract of research PCR methods for the detection of soy with their most relevant features, developed between 2017-2021.

allergenic food *(reference)*	method	type of detection	target	specificity [*number of species tested (number of cross-reactivities)]*	calibration	matrix applicability	LOD [mg soybean/kg]
soybean *(Costa et al., 2017)*	real-time PCR	TaqMan probe *(FAM-BHQ1)*	lectin gene nuclear 18S rRNA gene	-	model mixtures of pork meat spiked with known amounts of soybean	processed meat products	*9.8 pg soybean DNA*
soybean *(Espiñeira & Santaclara, 2017)*	real-time PCR	TaqMan probe *(6-FAM-TAMRA)*	lectin gene	*[Espiñeira & Santaclara only provide a protocol, method optimization and validation must be done separately]*			
soybean *(Mayer et al., 2019)*	ddPCR	TaqMan probe	chloroplast DNA NAD(P)H-quinone oxidoreductase subunit H (ndhH) gene	72 (*practically none*)	SureFood® QUANTARD Allergen 40	meat products flour milk fatty creams	0.16
soybean peanut *(Ladenburger et al., 2018)*	competitive real-time PCR	fluorescent hydrolysis probe *(assumingly TaqMan)*	mitochondrial DNA bait8	69 (13)	spiked skimmed milk powder rice cookies hollandaise sauce powder sausage	skimmed milk powder rice cookie hollandaise sauce powder sausage	1
soybean wheat *plants* *(Shin et al., 2021)*	PCR	agarose gel electrophoresis	Gly m Bd 30K gene	22	diluted target DNA (water, pea DNA)	processed food products	1000

Table 3 Extract of research multiplex PCR methods for the detection of soy, celery and/or white mustard

allergenic food *(reference)*	method	type of detection	target	specificity [*number of species tested (number of cross-reactivities)]*	calibration	matrix applicability	LOD [mg soybean/kg]
hexaplex *cashew* *peanut* *hazelnut* **soy** **celery** **mustard** *(Köppel et al., 2012)*	real-time PCR	TaqMan probe *(ROX-BHQ2)* *(Joe-BHQ1)* *(DY681-BHQ-2; DY681-BHQ-2)*	lectin gene Le1 mannitol dehydrogenase (mdh) gene SinA1 protein	41 (≤ 10; 2+*Brassicaceae*)	spiked rice cookies boiled sausages	all kinds of food samples	5 (rice cookies) 32 (boiled sausages)
triplex **celery** **white mustard** black/brown mustard *(Palle-Reisch et al., 2015)*	real-time PCR	TaqMan probe *(Cy5-BHQ2)* *(FAM-TAMRA)*	mRNA NADPH-dependent mannose 6-phosphate receptor (M6PR) gene MADS D protein	75 (7)	spiked raw, brewed model sausages	processed food	celery: 50 white mustard: 50
duplex **celery** **white** **mustard** *(Fuchs et al., 2013)*	real-time PCR	fluorescent *hydrolysis* probe *(assumingly TaqMan)*	mRNA NADPH-dependent M6PR gene MADS D protein	64	diluted target DNA spiked raw, brewed model sausages	raw, processed food	celery: 50 white mustard: 10
tetraplex **soy** **celery** **white mustard** brown mustard *(Luber et al., 2015)*	real-time PCR standard-addition method	fluorescent *hydrolysis* probe *(assumingly TaqMan)* *(ATTO 425-DDQI)* *(FAM-BBQ)* *(HEX-BBQ)*	lectin gene mdh gene mRNA MADS D protein	-	dilution of standard materials (*unk*) sausage lysate mixtures	commercial food products	*only LOQ*

unk unknown

In view of the recently updated VITAL 3.0 reference doses for allergens in foods, food analysis may be challenged in the near future. If such legislation came into force, several analytical methods would require some adaption regarding sensitivity. Related to soy detection, an adaption may be required to meet the demands of detecting 0.5 mg soybean protein (revised ED_{01} value) in a serving size up to 500 g. PCR protocols can easily be tailored to meet analytical needs (McPherson & Møller, 2006a). So can the LOD and limit of quantification (LOQ) be decreased ten- to 100-fold by taking advantage of the high number of DNA copies chloroplast DNA offers (Bauer et al., 2011; Demmel et al., 2008; Hirao et al., 2009; Holzhauser & Röder, 2015). A very recent research article gathered information on, amongst others, analytical methods for soy detection and quantification (ELISA, PCR, MS) in order to examine whether the current methodology can verify VITAL 2.0(/3.0) reference doses (Holzhauser et al., 2020). Literature published within the timeframe from the mid-1990s to February 2018 was considered for inclusion. They concluded that commercial ELISA and PCR methods were able to detect soy allergens at VITAL 2.0 reference doses for all serving sizes. However, only one research method, a multiplex DNA micro assay based on a digital versatile disc (Tortajada-Genaro et al., 2012), achieved adequate sensitivity to verify the soy protein reference dose of 1 mg protein in a large serving size (500 g) in each investigated matrix (Holzhauser et al., 2020).

3 Aims

Food allergy cases have been on the rise over the last decades, particularly in Western countries. Moreover, children seem to be even more affected than adults are, making this type of allergy not only an economic and safety burden but an emotional one, too. Even traces of a food allergen can be sufficient to trigger clinical symptoms in sensitized subjects, with anaphylaxis being the most severe and potentially life-threatening. The mainstay of managing food allergy is dietary exclusion of the offending food, requiring first and foremost reliable food labelling and analytical methods. With the introduction of the Regulation (EU) No 1169/2011 of the European Parliament and the Council, crucial steps were taken towards the protection of allergic individuals and more are likely to come. In view of the recently updated VITAL 3.0 reference doses for allergens in foods, food analysis may be challenged in the near future. If such legislation came into force, several analytical methods would require some adaption regarding sensitivity. However, too many analytical methods fall short on meeting possible future demands for the detection of allergenic foods, such as soy, celery and white mustard.

The main purpose of the master's thesis was to develop a qualitative singleplex real-time PCR assay for the detection of soy (*Glycine max*) and its validation as part of a qualitative triplex real-time PCR assay for the simultaneous detection of soy, celery (*Apium graveolens*) and white mustard (*Sinapis alba*) in food. The first step was to develop the qualitative singleplex real-time PCR assay for the detection of soy, followed by the second step to combine this novel singleplex assay with the already developed and successfully validated qualitative duplex real-time PCR assay for the simultaneous detection of celery and white mustard to form a triplex system. By targeting a specific sequence in the chloroplast genome of soy, advantage of the high number of DNA copies was taken in order to significantly lower detection limits of the two real-time PCR assays. Aiming the singleplex PCR assay to be run on the same 96-well reaction plate with other real-time PCR assays, the annealing temperature had to be 60 °C as a prerequisite. The three allergenic foods, soy, celery and white mustard, are likely to be present at the same time in food. Thus, this triplex real-time PCR assay is purposed to save both time and resources.

4 Theoretical part

4.1 Structure of DNA

The genetic material of eukaryotes – animals, plants, fungi and protists – is stored deep inside eukaryotic cells. There in the cell's organelles, intertwined with a multitude of other components, the nitrogen-containing nucleobases are located. They are the fundamental units of each DNA molecule. According to the sequence of nucleobases, genetic information is encoded.

Nucleobases are grouped into *purine* nucleobases (G and A) and *pyrimidine* nucleobases (C and T). In the molecule, the nucleobase is chemically linked to a sugar residue, which can be either oxygenated or deoxygenated at position 2. Oxygenated, it is a ribose (later ribonucleic acid, *RNA*) and deoxygenated, it is a deoxyribose (later deoxyribonucleic acid, *DNA*). Together, the nucleobase and the sugar make up a *nucleoside*; plus an additionally associated phosphate group, it is a *nucleotide* (Alberts et al., 2002).

Nucleobases pair according to the Watson and Crick base pairing rules: only G binds to C (G:C) and A binds to T (A:T) (McPherson & Møller, 2006a). Stability is given due to hydrogen bonds; between C:G there are three of those bonds, between A:T only two. The latter is, thus, less stable, whereas the link between C:G is more (Alberts et al., 2002). When nucleotide monomers congregate, they form polynucleotides, such as the well-known nucleic acids. They are held together by phosphodiester bonds. These phosphodiester bonds alongside hydrogen bonds allow the ultimate formation of double-stranded DNA, featuring an alternating sugar-phosphate backbone. Every single strand runs in a different direction, meaning they run antiparallel to each other. DNA synthesis is always directed from the 5′-end to the 3'-end. DNA double strands are coiled by nature resulting in the development of grooves within the DNA helix. During the incorporation of nucleotides into the DNA chain, pyrophosphate is cleaved off (Alberts et al., 2002).

DNA is stored in several organelles of the cell. In the beginning, researchers assumed genetic information is only kept in the nucleus of eukaryotes (nuclear DNA). DNA is, though, found in chloroplasts, too (chloroplast DNA) (Sugiura, 2005), opening up a broader repertoire for analysis (Daniell et al., 2016). DNA is not only synthesized in eukaryotic organisms constantly but can also be amplified *in vitro* by PCR (Mullis et al., 1986).

4.2 DNA extraction

The first steps towards DNA analysis are taking a representative subsample of the food under investigation, homogenizing as well as extracting DNA from this subsample to be analyzed (Waiblinger, 2010; Holzhauser & Röder, 2015). In order to properly continue with downstream methods of analysis, there are some requirements put on the DNA extract. A DNA extract with the following features is considered high-quality:

High degree of purity: Any kind of unwanted component, e.g. proteins, salts, other nucleic acids and nucleases, have to be removed. Depending on the source material, different strategies for obtaining high-purity DNA are applied.

High yield: The different steps of preparation before the extraction itself have a great impact on DNA yield. Generally speaking, the more DNA yielded from the source material, the better (Demeke & Jenkins, 2010; Mülhardt, 2013).

High integrity: Some downstream methods require DNA of high integrity. If food is highly processed, the likelihood of DNA still being of sufficient integrity is low (Popping & Diaz-Amigo, 2010).

At large, there are six steps within a standard protocol for DNA extraction:

1. Homogenization
2. Cell lysis: the cellular structure is disrupted
3. Digestion with proteinases
4. Organic extraction: DNA is dissolved from cell debris and other non-soluble components
5. Alcohol precipitation: removes other components aside from DNA, e.g. proteins.
6. Dissolution of DNA precipitate (Moore & Dowhan, 2012)

Some matrices contain substances possibly inhibiting downstream methods, like PCR. Concerning food, matrices tend to be complex since they comprise a variety of different components. For instance, plants are typically rich in polysaccharides, which are inhibitory. Polyphenols, pectin and xylan also contained in plants have the same effect. Milk and milk-based products contain calcium ions and certain proteases, which inhibit the reaction in an equal manner. Further examples of food matrices containing polymerase chain reaction (PCR) inhibitors are berries (phenols and polysaccharides) and seafood (glycogen and polysaccharides) (Popping & Diaz-Amigo, 2010; Demeke & Jenkins, 2010; Schrader et al., 2012).

4.2.1 Hexadecyltrimethylammonium bromide (CTAB) method

In 1980, Murray and Thompson succeeded in adapting a method using CTAB in order to extract DNA from plants (Murray & Thompson, 1980). The CTAB method has been widely used for plant DNA extraction as it is applicable to various plant species and tissue types, e.g. whole seedlings, leaves or grains (Moore & Dowhan, 2012). It can be applied to food and feed samples too, even when they are processed (International Organization for Standardization, 2005; Demeke & Jenkins, 2010). Nucleic acids are effectively complexed by the ionic detergent CTAB. The advantages of this method are its simplicity and the rather short duration of the protocol. Unwanted compounds like polysaccharides, phenolic compounds or other enzyme inhibiting substances frequently present in plants can be rapidly removed so that genomic DNA of high purity is prepared. For those DNA preparations, milligrams (mg) of plant tissue is sufficient. On the other hand, the sample size can also be increased up to grams (g). The condition of the starting material is very variable too (lyophilized, dehydrated, fresh or frozen) (Moore & Dowhan, 2012).

The underlying principle is that the detergent CTAB forms an insoluble complex with nucleic acids once the initial NaCl concentration is reduced. Especially polysaccharides are problematic contaminants when aiming for pure plant DNA extracts. However, with this method polysaccharides, phenolic compounds or other enzyme inhibiting substances do not form complexes; they remain dissolved in the supernatant, whereas the complex precipitates. This complex of interest, CTAB-nucleic acids, is exclusively soluble in a solution with a high salt concentration. For that reason, the detergent itself is removed by increasing the concentration of NaCl in the solution. CTAB is more soluble in alcohol than DNA is. After adding ethanol for the precipitation of nucleic acids, the pellet is washed several times, which removes the remaining CTAB. Shear forces and nuclease activity affect the DNA yield. It is, therefore, crucial to add CTAB buffer to the material before grinding so that nucleases do not degrade nucleic acids (Moore & Dowhan, 2012).

Alternatively to the original, manual CTAB-based method, a DNA extraction machine, e.g. Maxwell® 16 Instrument, can be utilized for DNA extraction. The DNA extraction machine Maxwell® 16 Instrument, for instance, makes use of paramagnetic particles for sample purification, particularly effective for removal of PCR inhibitors. Those particles provide a mobile solid phase, which the target is captured on and again eluted from after several washing steps (Promega Corporation, Maxwell® 16 Instrument Operating Manual, 2007–2015; Schrader et al., 2012).

4.3 Quantification of total DNA by spectroscopy

DNA yield can be determined by photometry (optical density) and the use of fluorescent DNA-binding dyes. The most common optical method to assess DNA yield and purity is the measurement of absorbance. With the help of a spectrophotometer, the approximate DNA concentration can be easily determined within a short time (Parkes & Saunders, 2007).

The principle is based on the nucleic acids' specific absorption of ultraviolet light at certain wavelengths (Gallagher, 2011). DNA molecules have their absorption maximum at a wavelength of 260 nanometers (nm) (A_{260}). Put simply, the more light is absorbed at this wavelength, the higher the DNA yield. For calculation of the concentration of double-stranded DNA, a concentration factor of 50 micrograms (µg)/mL, which is derived from the Beer-Lambert Law, has to be considered, additionally to the reading of A_{260}.

It is possible to evaluate DNA purity by determining the absorbance ratio of 260 nm and 280 nm. A A_{260}/A_{280} ratio of 1.7-2.0 is considered high-purity DNA. The lower the ratio, the higher the degree of contamination. It is important to note that this is only a rough estimation of DNA purity. Readings below 1.7 do not rule out the functionality in downstream assays, which are, by far, the best test of DNA extract quality.

The ratio of the measured absorbance at wavelength 260 and 230 [nm] may be helpful to gain further insight into DNA purity. Organic substances and chaotropic salts absorb at 230 nm; here a A_{260}/A_{230} ratio higher than 1.5 is considered high-purity DNA.

Nevertheless, limitations of absorption spectroscopy have to be kept in mind in order to make proper conclusions. Aromatic amino acids absorb UV light at 280 nm and RNA, for instance, greatly absorbs at 260 nm too, leading to a possible overestimation of DNA yield (Gallagher, 2011; Parkes & Saunders, 2007).

4.4 Polymerase chain reaction

The polymerase chain reaction (PCR) is a one-of-a-kind method of analysis. With no other comparable method, it is possible to amplify the analyte (target) *in vitro*. ELISA as an alternative analytical approach targets proteins. PCR, however, targets the DNA, allowing the specific identification of species. *Decoding* this into the analysis of food allergens: the qualitative detection of the DNA sequence of interest, e.g. of a certain allergenic species, means this certain allergen is present in the food sample at or above a certain limit. This limit, the limit of detection, can be very low thanks to target amplification, contrary to the sensitivity of the method being very high (Poms et al., 2004a; McPherson & Møller, 2006a). High-throughput screening is achievable, too. Unsurprisingly, PCR, mostly applied as real-time

PCR, has been increasing in popularity in the sector of food allergen analysis (Druml et al., 2015).

In 1983, Kary Mullis had the idea to invent the method of PCR. For this contribution to the scientific community, he was awarded the Nobel Prize in Chemistry a decade later (Broll, 2010).

4.4.1 Principle

The PCR's reaction mix contains several essential components, which will now be described in detail.

Nucleic acid template: original nucleic acid fragment that contains the target sequence to be amplified

In PCR, templates of even a very complex nature, for instance, genomic DNA or cDNA, can be amplified (McPherson & Møller, 2006b). Both applying too little and too much template leads to a suboptimal reaction (Broll, 2010). For genomic DNA, a maximum of 100 ng per PCR tube and well is recommended. Exceeding this limit may bring about negative effects on amplification, such as non-specific products (Müller & Prange, 2015). There are also more conservative recommendations, which say 10 ng/microliter (µL) as final DNA concentration (equaling 50 nanograms [ng] DNA per PCR tube and well) should not be exceeded. Generally, the input of about 10^4 copies of the target DNA should yield a signal at cycle 25 to 30 (Broll, 2010). High-quality DNA extracts yield the best results in downstream assays. DNA extract purity is especially important when aiming for amplification of long sequences (> 1000 bp). There are scenarios, in which relevant levels of contamination may suppress or fail the amplification (McPherson & Møller, 2006b). For example, starting material containing oil and fats might not give any amplification as those substances inhibit DNA polymerase activity (Popping & Diaz-Amigo, 2010).

Deoxynucleotide triphosphates (dNTPs): essential components for sequence elongation

Forward and reverse primer: sequence-specific oligonucleotides

The two primers, the forward and the reverse primer, define the sequence to be amplified (Mullis & Faloona, 1987). The DNA sequence of the analyte has to be known, at least partly, to allow for primer design. Primers have to be complementary to the sequence of interest. In

addition, some other rules should be taken into consideration in the course of design. Both primers ideally have the same melting temperature (T_m) or at least a similar one so that they anneal simultaneously to the target sequence at a given annealing temperature (T_a) (McPherson & Møller, 2006b; Broll, 2010). At T_m, 50 % of the DNA is single-stranded; the other 50 % is still double-stranded (Mülhardt, 2013). The simplest calculation of T_m is based on the G:C content of the primer, though is limited to short sequences (20 nucleotides) (Mülhardt, 2013). The T_a of the individual amplification reaction is calculated from the (mean) T_m of both primers and ideally lies 5-10 °C below (Müller & Prange, 2015; Mülhardt, 2013). Primers should be between 18-30 nucleotides long and the G:C to A:T content should be balanced. Literature recommendations on G:C content is 40-60 % (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014; Broll, 2010; McPherson & Møller, 2006c). With a higher G:C content, primer binding is much more stable due to stronger hydrogen bonding whereas a high A:T content leads to a reduced T_m. G:C-rich sequences at the 3'-end, therefore, promote annealing as well as PCR efficiency. Preferably, the last nucleobase of the 3'-end is a G or C, known as "GC clamp" (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014; Thermo Fisher Scientific (Life Technologies), PCR Primer Design Tips, 2019). Repetitive sequences should be avoided since they come with an increased likelihood of primer *"slipping"* (McPherson & Møller, 2006b). Also, it is important that the primers' 3'-end is not internally complementary. It should not feature degenerate nucleotides, so-called "wobble"-structures, and it should be steered clear of three or more Gs or Cs in this region, all of which to avoid mispriming. Primers should not form any secondary structures, such as hairpins or dimers. Primer dimers may be generated due to internal complementarity, in particular at the 3'-end. Equally produced within the elongation phase, either on the primer itself (self-priming) or on the second primer, they can govern the whole reaction (McPherson & Møller, 2006b, 2006d; Broll, 2010; Müller & Prange, 2015).

Even though primers are applied in vast excess, the concentration of both forward and reverse primer has a great impact on the amplification reaction. Optimization of the primer concentration can help improve the outcome. A primer concentration of up to 2 mM is applicable. Above this limit, mispriming and the generation of non-specific products become more likely. In case of artifact formation, the concentration of primers taking part in the amplification reaction is insufficient. Purified primers, e.g. by HPLC, are preferable (McPherson & Møller, 2006b, 2006d; Broll, 2010; Müller & Prange, 2015).

DNA polymerase: serves as the catalyst of DNA synthesis in the 5′→3′ direction

In addition, these enzymes may feature 3′→5′ exonuclease activity. The fidelity and the efficiency of DNA synthesis are two key factors associated with DNA polymerase and its influence on PCR. The efficiency is a product of synthesis rate and processivity; the latter provides insight into the enzyme's affinity to the target sequence. At the very beginning of PCR application, Klenow fragments of DNA polymerase I from *E. coli* were used as such catalysts. This early strategy was not only more laborious and expensive but also amplification of non-target sequences was enhanced due to the required low temperature in the phase of elongation (37 °C). The thermostable *Taq* DNA polymerase superseded its precursor soon. This enzyme is extracted from the thermophilic bacterium *Thermus aquaticus (Taq)*. *Taq* DNA polymerase has its optimum temperature at around 72-75 °C. At these high temperatures, the specificity of the generation of non-target products is minimized.

As a "good general-purpose enzyme" Taq DNA polymerase is commonly applied in PCR runs in routine laboratories. It has high efficiency and specificity. One drawback is, however, that the enzyme is already active during the PCR setup. This might lead to the formation of non-specific products (McPherson & Møller, 2006b). So-called "hot-start" DNA polymerases may require heat activation (Kubista et al., 2006). Before subjection to high temperatures ("hot start", e.g. 95 °C for 15 minutes), these special DNA polymerases are inactive so as to avoid the formation of primer dimers due to mispriming. Since the likelihood of the generation of so-called PCR artifacts is therefore kept to a minimum, high PCR specificity is assured (Qiagen, QuantiTect® Multiplex PCR Handbook, 2011).

A buffer containing magnesium chloride ($MgCl_2$) or magnesium sulfate ($MgSO_4$): magnesium functions as a metabolic cofactor of the enzyme DNA polymerase

Magnesium is another determining factor for reaction specificity and efficiency. Also, the cation's concentration is linked with the T_m of the dsDNA. The concentration optimum that supports DNA polymerase activity is around 1.2-1.3 mM free Mg^{2+}. It is dependent on the concentration of the other components in the reaction (template DNA concentration, chelating agents, such as EDTA, dNTP concentration, proteins). For instance, the free Mg^{2+} concentration correlates with the dNTP concentration as dNTPs bind the cation. Like a too low concentration of Mg^{2+}, a too high concentration of Mg^{2+} may also prevent a successful PCR. The *Taq* DNA polymerase is more prone to errors under these conditions (McPherson & Møller, 2006b; Broll, 2010; Müller & Prange, 2015). QuantiTect Multiplex PCR Buffer, for example, a product of Qiagen, has a $MgCl_2$ concentration of 11 mM. Under these buffered

conditions, negative effects of simultaneous amplification reactions on each other in PCR multiplexing are compensated (Qiagen, QuantiTect® Multiplex PCR Handbook, 2011).

PCR premixes including buffer, dNTPs and *Taq* DNA polymerase can be acquired commercially. This is especially advantageous for high-throughput screening, e.g. in routine analysis.

Regarding necessary equipment, the minimum requirement is a thermal cycler (thermocycler), a heating block that can be programmed (McPherson & Møller, 2006b). The temperature program goes on, once more in every cycle.

Each cycle comprises three distinct steps varying in temperature:

1. Denaturation

   Due to high temperatures (normally around 94 °C), hydrogen bonds between G:C and A:T are cleaved. Instead of one double strand, two complementary single strands are now present. Only those single strands can function as the template in the following step (McPherson & Møller, 2006d). Complete denaturation of the double strand is crucial for efficient annealing (Mülhardt, 2013; Müller & Prange, 2015). Only partially separated double strands reanneal as soon as the temperature is lowered again, preventing correct priming (Kubista et al., 2006). For most efficient bond cleavage, often the highest possible temperature the DNA polymerase can tolerate is chosen (95 °C). This temperature is held for 1-5 minutes at the beginning (cycle 1); following cycles can be run with a denaturation step of a few seconds. If G:C is abundant in the template sequence, the time of denaturation may be raised. However, as these repeated high temperatures lead to degradation of the other reaction components too, this step should be kept as short as possible (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014; Mülhardt, 2013; Müller & Prange, 2015).

2. Annealing

   In this step, the temperature needs to be rapidly lowered so as to enable the hybridization of both primers. One factor greatly contributing to assay specificity is the T_a. The T_a can be chosen between 40-72 °C but, ultimately, it depends on the assay design. Higher T_a comes with higher assay specificity. However, with a T_a too high, annealing cannot take place at all (McPherson & Møller, 2006d; Müller & Prange, 2015). When a small template sequence is to be amplified, annealing may be combined with the last step, elongation;

60 °C can then be used at both steps (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014).

3. Elongation

   In the last step, the enzyme, DNA polymerase, plays a key role. It catalyzes the complementation of deoxynucleotides to the template from the 3'-end to 5'-end direction of both corresponding single strands, ultimately resulting in two new duplex DNA strands. (McPherson & Møller, 2006d; Coen, 2012). The chosen elongation temperature has to be matched with the length of the target sequence and the processivity of the DNA polymerase. As a general rule: 30-60 seconds (s) for the elongation of 1000 bp by using *Taq* DNA polymerase. After the last cycle, there may be added another step for the extension of partially elongated strands, lasting 5-15 minutes at 72 °C (Mülhardt, 2013; Müller & Prange, 2015). Increased temperatures during the elongation step are additionally advantageous as unwantedly formed secondary structures, which may inhibit extension, are degraded (Kubista et al., 2006).

In the subsequent cycle, the just generated two double strands act as additional templates. These copies are much shorter than the original template but also still longer than the actual amplicon. After the second cycle, two single-stranded DNA sequences are first generated that match the intended amplicon size. Then, these fragments of defined length are repeatedly amplified. This is why, in theory, the number of double strands doubles both exponentially and indefinitely. Assuming PCR runs 100 % efficiently, a single genomic DNA template theoretically yields almost 1.050.000 copies of double-stranded targets after 20 cycles. Perfect efficiency is in PCR, however, rather an illusion than reality. Thus, more cycles need to be added to the first 20 ones for compensation, enabling sensitive detection (normally 25-40 cycles in total) (McPherson & Møller, 2006d; Holzhauser & Röder, 2015).

The kinetics of PCR take course in three phases as depicted in Figure 2.

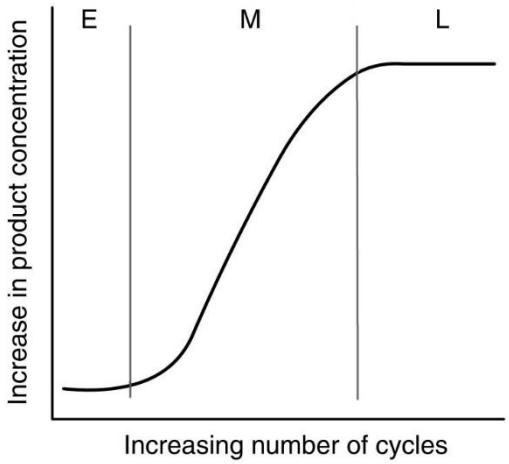

Figure 2 Kinetics of accumulation of the target product over the course of PCR. E: early cycles. M: mid cycles. L: late cycles. Adapted from McPherson & Møller, 2006d.

The first phase, E, stands for the early cycles. Primers (and probes) bind to complementary sequences within the template strand. Since the amplification reaction has just started, the amount of amplicon is low yet. The mid cycles, M, is the most relevant phase for analysis. Here the amplicon is exponentially amplified: after primers acting as initiation sites for DNA synthesis, the polymerase takes them up on their job and starts to complement the template with dNTPs. Exclusively during this phase, correct qualitative and quantitative determination of the analyte is given. The goal is to maximize amplification efficiency (high increase in product concentration). The late cycles, L, signalize amplification coming to an end as a consequence of substrate loss and a simultaneous increase of DNA product, pyrophosphate, as a byproduct of the incorporation of nucleotides into the DNA chain, and monophosphate nucleotides. The presence of these products, especially pyrophosphate, is accompanied by activity loss of the thermostable DNA polymerase. As soon as more DNA product than DNA polymerase is present, there is a loss of exponential amplification; not each strand can be used as a template anymore. Additional reasons are other suboptimal working conditions, e.g. inhibition of the reaction. Product sequences are generally longer than primers, which allow them to anneal to complementary product strands at higher temperatures. Hence, in the plateau phase, the chance of the generation of non-specific DNA products increases. Before this plateau phase begins, the reaction is ideally stopped (McPherson & Møller, 2006d; Mülhardt, 2013).

4.4.2 Variants of PCR

The conventional PCR as described above was supplemented with a more advanced technique, the real-time PCR. Real-time PCR allows real-time monitoring of target amplification based on fluorescence detection throughout the whole process and is even quicker than the conventional variant (Coen, 2012; McPherson & Møller, 2006c). A detailed elaboration on real-time PCR is given in the following subchapters.

One for this work relevant key advantage of real-time PCR is the possibility to multiplex. PCR multiplexing enables specific amplification and detection of multiple analytes in the very same reaction, in contrast to singleplex PCR assays (Mülhardt, 2013).

4.4.2.1 Real-time PCR

The main difference between real-time and conventional PCR is that in real-time PCR the accumulation of DNA product and associated fluorescence is continuously detected, i.e. at each cycle, by a fluorescence detector. Conventional PCR, in contrast, has to make use of post-PCR detection methods, such as agarose gel electrophoresis. As a PCR technique with endpoint detection, quantification is not possible because the endpoint is not proportional to the initially input DNA amount. The variant of real-time PCR, developed by Higuchi et al. in 1992, has become an indispensable tool for the detection as well as quantification of RNA and DNA (McPherson & Møller, 2006a; Higuchi et al., 1992; Holzhauser & Röder, 2015).

Thanks to advances in chemistry and instrumentation, there have been emerging quite a few methods for fluorescence detection in real-time PCR. Those methods are based on either intercalating dyes or probes displaying fluorescence properties. In general, the fluorescence signal is directly proportional to the number of amplicons generated (McPherson & Møller, 2006c). Intercalating fluorescence dyes bind to double-stranded DNA in a non-sequence specific manner. One frequently used dye is Sybr Green, an asymmetric cyanine dye (Mülhardt, 2013; Kubista et al., 2006). A different sequence-specific approach is the usage of an oligonucleotide hydrolysis probe, most often a TaqMan probe (McPherson & Møller, 2006c; Kubista et al., 2006).

TaqMan probes are widely used in techniques of molecular biology, such as in real-time PCR. What started with application in the medical field, is now being applied in multiple sectors, e.g. analysis of genetically modified organisms (GMO) and food allergens (Fu et al., 2020; Lee et al., 1993; Terry et al., 2002a).

Similarly to primers, probes are synthetic sequences of DNA of short length. The feature that makes the application of probes so attractive is their contribution to the specificity of the assay. Primer dimers do not lead to an increase of the fluorescence signal, making the assay more

selective, too. Melting point analysis is no longer necessary. Nevertheless, if present, primer dimers do affect PCR efficiency and dynamics. (McPherson & Møller, 2006c).

Figure 3 illustrates TaqMan probe interactions with the target.

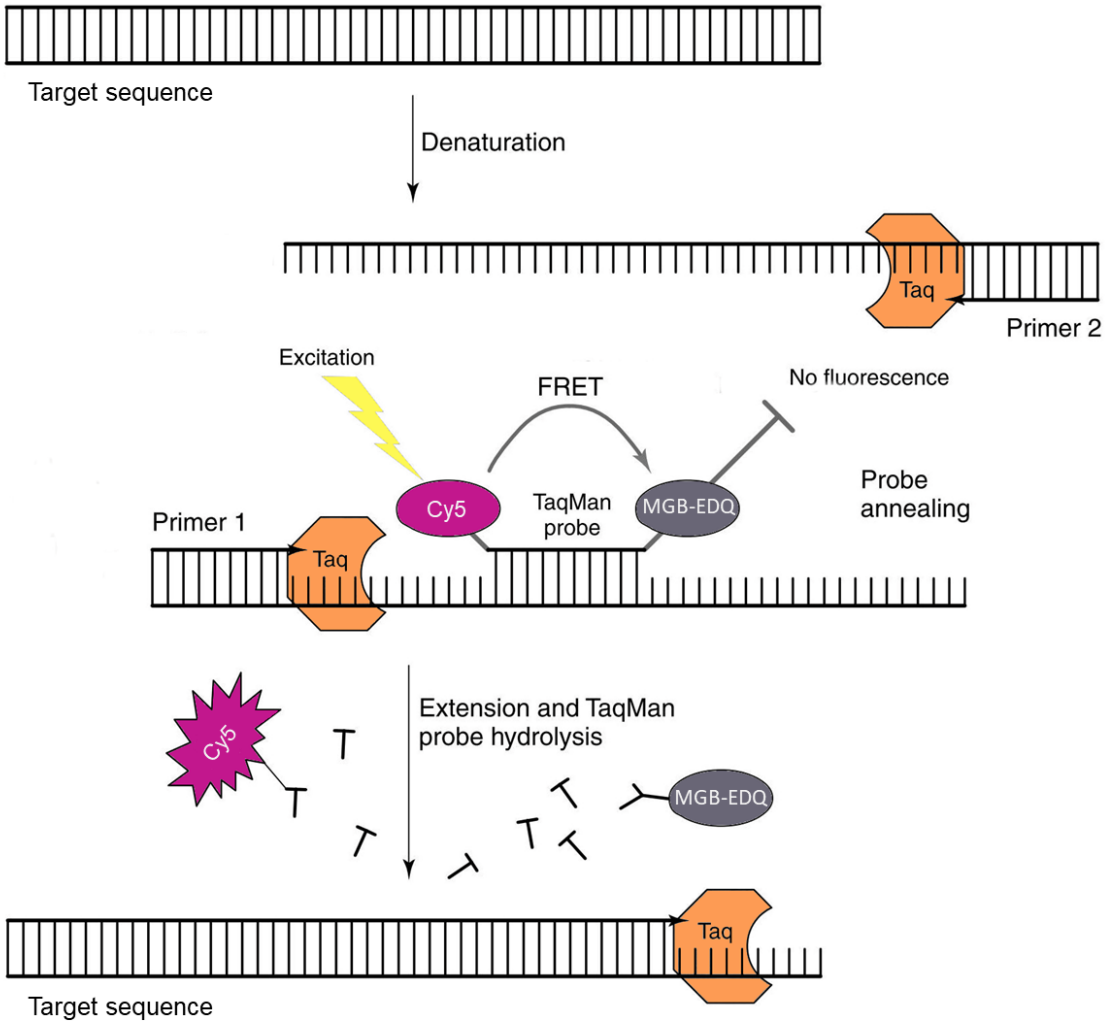

Figure 3 Schematic diagram showing the principle of real-time PCR using the TaqMan approach. Cy5: cyanine-5 (fluorescent reporter). EDQ-MGB: Eclipse™ Dark Quencher-minor groove binder (non-fluorescent quencher). FRET: fluorescence resonance energy transfer. Adapted from McPherson & Møller, 2006c.

Fluorescent probes are designed to hybridize to the target DNA, additionally to both hybridizing primers. To be more precise, the TaqMan probes are supposed to bind in between the forward and the reverse primer. Such probes are labelled with two covalently bound molecules: a fluorescent dye, also called *reporter*, at their 5'-end and a *quencher* at their 3'-end. Examples for the reporter are (6-)FAM (6-carboxyfluorescein), VIC (2-chloro-7-phenyl-1,4-dichloro-6-carboxyfluorescein), HEX (6-carboxy-2',4,4',5',7,7'-hexachlorofluorescein) or Cy5 (cyanine-5); the quencher can be, for instance, a non-fluorescent Eclipse™ Dark

Quencher (EDQ) coupled with a minor groove binder (MGB) (as exemplified in Figure 3). On molecular level, fluorescence resonance energy transfer (FRET) occurs. Emission of the fluorescent reporter can be fully suppressed by the quencher. This is because the quencher has higher emission energy than the reporter. Consequently, as long as the TaqMan probe is intact, meaning the quencher is in close proximity to the reporter, the fluorescent signal is *quenched*. It does not matter at this point whether the probe is free in solution or already hybridized. The TaqMan assay exploits the *Taq* DNA polymerase's exonuclease activity (5'→3'). As the amplification reaction goes on, primers and probe anneal. DNA *Taq* polymerase starts to complement the sequence. However, when the *Taq* polymerase reaches the 5'-end of the TaqMan probe, it cleaves one deoxynucleotide after another because the probe's oligonucleotides are recognized as DNA by the enzyme. In the further course of degradation, the reporter is cleaved (here in Figure 3 Cy5); now the quencher and reporter of the probe are not in close proximity anymore. FRET is lost between the two, so fluorescence is not quenched anymore. Eventually, probe displacement leads to the actual optical detection of a fluorescent signal at a defined wavelength. Since in real-time PCR fluorescence is recorded during each cycle, the level of probe displacement is proportional to the amount of PCR product (McPherson & Møller, 2006c; Mülhardt, 2013; Müller & Prange, 2015).

Such TaqMan probes are still high in price. Also, more sensitive approaches for amplicon detection already exist, such as the innovative method of molecular beacons. In contrast to TaqMan probes, they have a relatively low background fluorescence (McPherson & Møller, 2006c).

In food allergen analysis, one essential requirement is that food samples can be analyzed with a selective and specific assay without suffering from cross-reactivity for closely related species (Druml et al., 2015). Homologs among those species often pose problems. They exhibit similar sequences as the target sequence. Hence, non-specific binding is more likely to occur. In order to tackle this common problem, special care has to be taken during primer and probe design (Popping & Diaz-Amigo, 2010). As a basic principle, TaqMan probes should be longer than primers (between 18-30 oligonucleotides); a length of 20 oligonucleotides is considered ideal. The T_m of a TaqMan probe should be approximately 10°C higher than for the primers, which makes hybridization to the target sequence in the elongation step also possible. Most importantly, G at the 5′-end should be avoided since this results in quenching of the fluorescent signal even after cleavage of the probe. More than three consecutive Gs should be avoided too (McPherson & Møller, 2006c; Müller & Prange, 2015).

Primer specificity is impacted mostly by a mismatch between the template sequence and the homolog sequence at the 3'-end of the primer sequence. In TaqMan probes mismatches also within the probes' sequence have a significant impact on specificity. This is even more distinct

in TaqMan MGB (minor groove binding) probes, which are shorter in length. Therefore, TaqMan MGB probes show lower T_m, which means that T_a does not need to be so high either to enable successful cleavage. Structurally, they exhibit a minor groove binding molecule at their 3'-end, which enhances the affinity for DNA. By binding to the target, a minor groove is formed within the DNA molecule. TaqMan MGB probes are recommended when aiming for specific assays (McPherson & Møller, 2006c; Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014; Kubista et al., 2006).

4.4.2.2 Multiplex real-time PCR

Multiplex real-time PCR is a powerful variant of the original singleplex real-time PCR assay. On the one hand, multiplexing saves time and money and the required amount of sample DNA and reagents are lower. On the other hand, sample throughput is increased. However, these benefits come with increased assay optimization efforts to analyze every single sequence of interest adequately. In a multiplex real-time PCR experiment, the compatibility of several primers and probes must be ensured, as well as similar T_as. Another limiting factor is the number of channels the fluorescence detector of the thermocycler is equipped with. For accurate detection of different fluorescent dyes, overlapping of emission spectra must be avoided. Multiplex real-time PCR can be used for both qualitative and quantitative analysis. The duplex assay is a common type of multiplex assays (Mülhardt, 2013; Henegariu et al., 1997). However, not surprisingly, the trend is going to higher-order multiplexes (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014).

In the course of developing a multiplex assay, multiplex PCR saturation can pose a challenge. Multiplex PCR saturation is caused by saturation of the DNA polymerase due to excessive amplification of one of the target sequences that is abundant. Normally, this problem can be overcome by primer limitation of the corresponding primer(/probe) system. Yet, limited concentrations need to be sufficient to result in exponential amplification for (qualitative) data analysis (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014).

Homologies among multiplexed assays are also challenging. These homologies come with the likelihood of undesired primer interactions since multiple components come into contact with each other. Generally speaking, when there is a significant difference in C_t values between the singleplex and the multiplex assay, the corresponding assay may be excluded from the multiplex PCR, at least for quantification (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014).

4.4.3 Data analysis

In real-time PCR, the detector system records fluorescence after each cycle. Thereby, the fluorescence signal is plotted against the cycle number, which results in the amplification plot (McPherson & Møller, 2006c) as depicted in Figure 4.

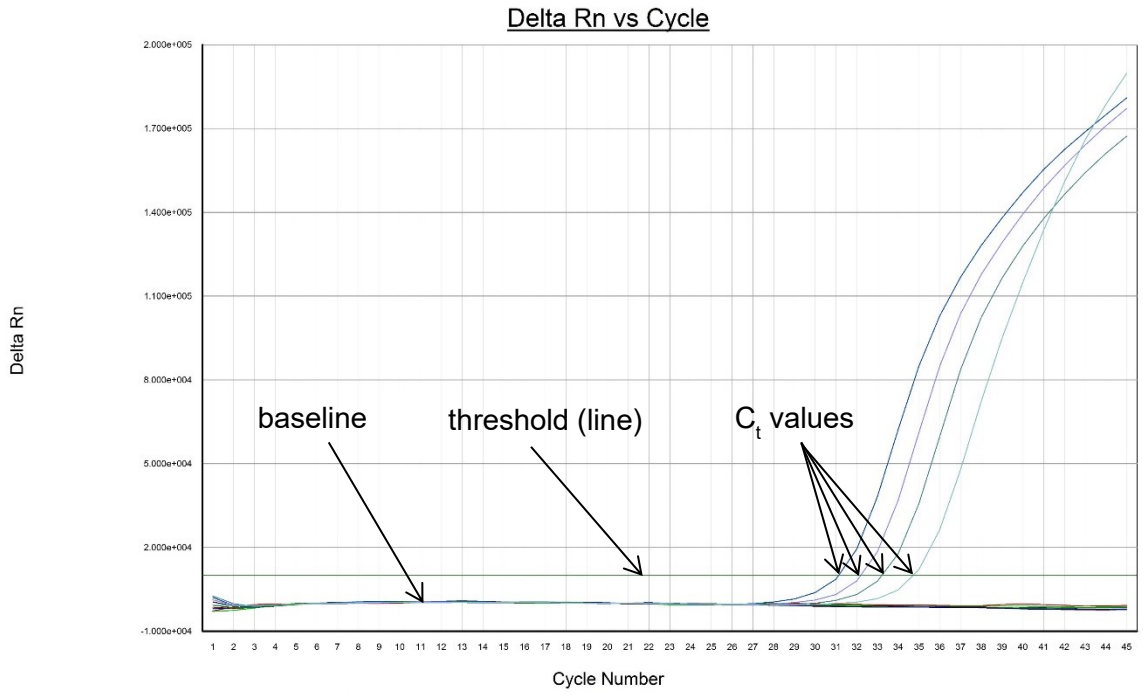

Figure 4 Amplification plot and its components, baseline, threshold (line) and C_t value(s).

The *baseline* reflects low-level signaling within the first three to 15 cycles where the fluorescence signal changes negligibly. Everything below the threshold line is considered as so-called *noise* or *background* from the detector, which is recorded at the very beginning of the amplification reaction. When the baseline is manually set, care should be taken since it influences both C_t determination and reaction efficiency: only background fluorescence needs to be excluded. In order to enable direct comparison between PCR runs, the baseline, i.e. the start and the end cycle, should be equally set. Manual adaptions may be improving results (Adams, 2007; Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014).

At the beginning of fluorescence measurement, there is no increase yet as the signal intensity has not exceeded the LOD of the fluorescence detector of the thermocycler (see Figure 4). In the further course of the reaction, the signal emerges from the baseline and crosses a certain line. This *threshold line* or *threshold* is to be set within the exponential phase (mid cycles) of the amplification curve (McPherson & Møller, 2006d; Adams, 2007). The value resulting, the value of the threshold cycle (C_t value), is considered most important for data analysis of real-

time PCR. It is inversely related to the logarithm of the initial DNA copy number (Adams, 2007). For example, 1:2 serial dilutions of analyte should yield C_t value differences of one cycle, provided that the amplification reaction operates at 100 % (see *amplification efficiency*). Every fluorescence signal recorded but not exceeding the threshold is considered as a result *not determined*, thus, a negative one (Thermo Fisher Scientific (Life Technologies), Real-time PCR handbook, 2014). If comparing two or more PCR runs with one another, threshold lines have to be set at the same level.

The *amplification efficiency* is another very important parameter in real-time PCR, particularly for quantification. With this parameter, inhibitory effects on the enzyme *Taq* DNA polymerase and consequently on PCR can be demonstrated (Adams, 2007).

5 Experimental part

All chemicals used during the laboratory work within the scope of the underlying master's thesis were of analytical or molecular biological grade. A lot of the solutions described in this chapter were provided by the AGES. A summary of chemicals as well as required amounts for the preparation of those solutions is given in subchapter 12.1. In the following, the term *soy* refers to *soybean* and vice versa and the term *tube* refers to *safe-lock tube* (Eppendorf, Germany).

Soy contamination posed one of the main issues during the practical work of the underlying thesis. Even though special care was taken to avoid any kind of soy cross-contamination/contamination, it sporadically cropped up as a laborious problem. At the AGES, seed/GMO analysis are part of everyday laboratory work, and so are samples that are highly likely to contain soy or are soy(-based). Since the practical work of the underlying thesis and these kinds of work could not be fully separated due to shared laboratory spaces, this might be an explanation.

5.1 Reference material

As reference material, SureFood® QUANTARD Allergen 40 (R-Biopharm, Germany) was used. It consists of corn flour additionally containing the substances or (derived) products that potentially cause allergies or intolerances, according to Annex II of the Regulation (EU) No 1169/2011 (with exception of lactose, molluscs and sulphur dioxide, see subchapter 2.2.5), each at a concentration of 40 mg/kg.

5.2 Plant material

Plant material used in experiments was taken from stocks available at the AGES or kindly provided by Walter Mayer or other colleagues. Some plant material was purchased at a local market in Vienna. In case there was already extracted genomic DNA available in AGES stocks, it was utilized.

5.3 Plant material preparation

From the very start of the laboratory work, strenuous efforts had been made to remove surficial contamination of plant material. Whenever possible, a piece of plant material was cut out from the inside so as to circumvent potential input. Otherwise, the material was thoroughly washed with distilled water multiple times. Particular care was taken to avoid any cross-contamination.

Depending on the type of plant material, it was dried or prepared as acetone dry powder for the subsequent lysis. This additional preparation step was conducted for plant material rich in fat, e.g. sesame and millet. Sage, dill and ginger were also prepared this way. The following will describe the procedure.

5.3.1 Preparation of acetone dry powder of plant material

Approximately 5 g material was reduced to very small pieces by cutting with a clean knife and/or grinding and put into a 50 mL falcon tube (VWR, USA). 35 mL acetone, roughly measured on the falcon tube's integrated marks, was added to the very small pieces of plant material. The tube was then placed into the fridge (Liebherr, Germany) overnight for soaking. Subsequently, the material and acetone were mixed employing a homogenizer (Polytron PT 3000, Kinematica, Switzerland). In the case of preparing several types of plant material at once, the homogenizer was thoroughly cleaned after finishing up with one type by use of a small cleaning brush, soap and distilled water in order to avoid any cross-contamination. The next step was centrifuging at room temperature (RT) (e.g. 24 °C), 1025 rpm (590 rcf) for 5 minutes. The supernatant was removed, another 35 mL of acetone was added into the tube and the mixture was vortexed (Vortexmixer, Scientific Industries, USA) for approximately 10 seconds. This sequence of adding acetone, centrifugation, discarding the supernatant and vortexing was repeated 3-4 times. The follow-up step was carried out identically except for using petroleum benzine (Merck, Germany) instead of acetone. The supernatant was removed. To evaporate the remaining petroleum benzine (boiling range 40-60 °C) in the material, the tube was left open for several hours.

5.4 Lysis

If the plant material was prepared with acetone as elaborated in subchapter 5.3.1, 200-250 mg of acetone dry powder were weighed into a 50 mL tube. The actual weighed-out quantity correlated with the amount of acetone dry powder yielded. For some preparations, e.g. dill, even less than 100 mg was obtained. If this preparation step of the source material was not required, about 1 g of material was weighed out in a 50 mL tube.

To each 50 mL tube of all kinds of plant material, whether it was pretreated or not, 10-20 mL of CTAB extraction buffer (see subchapter 12.1.1) was added. Whenever the upstream step of material preparation did not include the generation of acetone dry powder of the component of interest (which already included reduction to small pieces), further reduction to small pieces of the material in CTAB extraction buffer was carried out at all times in order to yield the highest possible amount of DNA; for some plant material, the homogenizer was used. After adding 80

µL proteinase K (Sigma-Aldrich, USA), the 50 mL tube was placed onto a spinning device situated in the middle of an incubator (Incubator Hybaid, MWG-Biotech, Germany; Incubator Unihood 750, Uniequip, Germany) where it was incubated at 50 °C for 15 hours/overnight whilst rotating.

5.5 Genomic DNA extraction – hexadecyltrimethylammonium bromide (CTAB)-based method

Here, too, particular care was taken to avoid any (cross-)contamination. As a general rule, reagents were added into blank tubes, i.e. solutions containing sample DNA were always added last. If not feasible, an aliquot from the reagent was taken and worked with. Tubes were kept closed when not in use. Concerning DNA extraction for the development of the triplex real-time PCR assay, particularly of celery and white mustard, a blank only consisting of CTAB extraction buffer and proteinase K was additionally extracted and analyzed as negative extraction control.

5.5.1 CTAB-based method by DNA extraction machine Maxwell® 16 Instrument

After digestion of the plant material, the 50 mL tube was centrifuged (5415 R, Eppendorf, Germany; 5810 R, Eppendorf) at RT, 3,900 rpm (2,245 rcf) for 10 minutes. 1 mL of the supernatant was transferred into a 2 mL tube already containing 600 µL chloroform/isoamyl alcohol (24:1, v/v; Sigma-Aldrich/ Merck). The steps using the vortex mixer, which are repeated multiple times throughout this protocol and the one of the manual CTAB-based method (see subchapter 5.5.2), are important for yielding DNA in sufficient amounts. Thus, vortexing was carried out for at least 1 minute. The 50 mL tube was then centrifuged at RT at 13,200 rpm (16,363 rcf). 300 µL lysis buffer (Promega, USA) and 5 µL RNase as part of the Maxwell® RSC PureFood GMO and Authentication Kit (Promega) was added into another 2 mL tube as well as 300 µL of supernatant. The tube was placed in a thermomixer (Eppendorf) at 65 °C for 15 minutes. Preparing for the following extraction with the help of the extraction machine Maxwell® 16 Instrument (AS2000) (Promega), the program *Blood* (RUN → DNA → Blood) in the corresponding software Maxwell RSC, version 3.0 (Promega) was selected. The standard procedure is given in the operating manual of the instrument (Promega Corporation, Maxwell® 16 Instrument Operating Manual, 2007–2015). For details see the technical manual of the corresponding kit (Promega Corporation, Maxwell® RSC PureFood GMO and Authentication Kit Technical Manual, 2016–2020). The next steps were carried out as described in those two manuals. Quantification of the total DNA yield was conducted as elaborated in subchapter 5.6.

5.5.2 Manual CTAB-based method

The subsequent protocol is an adapted version of the original paper "*Rapid isolation of high molecular weight plant DNA*" by Murray and Thompson described in 1980 (Murray & Thompson, 1980). In the AGES, this adapted protocol is applied to routine analysis.

800 µL chloroform/isoamyl alcohol (24:1, v/v) was added into a 2 mL tube as well as 5 µL of bovine serum albumin (20 mg/mL) (BSA; Roche, Switzerland). Then, 800 µL of supernatant of the lyophilized plant material that had been prepared prior to this step (see subchapter 5.4) was added into the tube. The tube was vortexed for at least 30 seconds and centrifuged at 13,200 rpm (16,363 rcf) for 10 minutes. Meanwhile, 1320 µL of precipitation solution was transferred into another tube. From the just centrifuged tube, 660 µL supernatant was transferred to the prepared tube containing precipitation solution; the mixture was vortexed thoroughly. Afterwards, it was incubated at RT for no less than 1 hour, followed by another centrifugation at 13,200 rpm (16,363 rcf) for 10 minutes. The supernatant was discarded and 500 µL of a mixture consisting of 450 µL sodium chloride solution (1.2 M), 50 µL 10x RNase buffer, and 5 µL RNase was added, which had been prepared in advance. The mixture was then incubated in a thermomixer at 56 °C for 15 minutes. 500 µL phenol:chloroform:isoamyl alcohol (25:24:1, saturated with 10 mM Tris, pH 8.0, 1 mM EDTA; Sigma-Aldrich) was added; the mixture was vortexed for at least 30 seconds. Afterwards, 400 µL isopropanol (Merck) was transferred to a 1.5 mL tube into which 2 µL of GlycoBlue Coprecipitant (Thermo Fisher Scientific, USA) was added in addition. 400 µL of the top (aqueous) phase of the supernatant was recovered and transferred to the prepared solution of alcohol and coprecipitant. Special care was taken during this process so as to remove nothing but the aqueous phase; the organic phase or a phase in the middle that may have been developed was left in the tube. The tube containing alcohol, coprecipitant and 400 µL of the aqueous phase was vortexed and placed into the freezer (Liebherr) for 1 hour minimum. After incubation, it was centrifuged at 13,200 rpm (16,363 rcf) for 10 minutes. The supernatant was discarded and the precipitate was washed with 500 µL of 70 % ethanol (v/v; VWR). After vortexing for at least 30 seconds and centrifuging, the supernatant was carefully removed from a well-visible blue precipitate. The tube was left open to evaporate the remaining alcohol. After several hours (often overnight) and checking for left alcohol, 100 µL of elution buffer (Promega) was added and the DNA pellet was dissolved in a thermomixer at 54 °C for 15 minutes. After cooling down of the solution, quantification of the total DNA yield was conducted as elaborated in subchapter 5.6.

5.6 Quantification of total DNA by spectroscopy

After genomic DNA extraction, total DNA was quantified by spectroscopy employing the UV/VIS spectrophotometer QIAxpert (Qiagen, Germany). The setting *A260 dsDNA* was selected for measurement, which determined the underlying calculation automatically done by the device (multiplication with factor 50; for further explanation, see subchapter 4.3 and Qiagen, QIAxpert® User Manual, 2019. Every time a measurement of DNA extracts was carried out, a blank only consisting of elution buffer was additionally measured on the same cartridge (QIAxpert Slide-40; Qiagen). Elution buffer was used for DNA elution as the last step of both DNA extraction by machine and manually. 2 µL of each solution to be measured was added onto the cartridge.

If the DNA extract was expected to be stored for only a short period, it was placed into the fridge directly after total DNA quantification. In case of longer storage, it was put into the freezer until analysis.

5.7 Real-time PCR

This and the following subchapters are divided into the different steps of developing and validating the underlying assay/s. In this subchapter, 5.7 Real-time PCR, the principal workflow of the real-time PCR is elaborated, followed by further subchapters in detail describing the work that was done during that particular step (see subchapters 5.8 and 5.9).

Real-time PCR assays were run with an Applied Biosystems® 7500 Real Time PCR System using the associated software Applied Biosystems® 7500 System SDS (Sequence Detection System) Software, version 1.4.0 (Thermo Fisher Scientific). The term *water*, subsequently, always refers to bidistilled/sterile ultra-pure water for PCR (H_2O_{dd}), either obtained in-house by the Synergy® Water Purification System (Milli Q Water; Millipore, USA) or purchased RNase-free water (Qiagen).

For PCR experiments, the commercially available QuantiTect® Multiplex PCR NoROX Master Mix (Qiagen) was used. This ready-to-use master mix consists of a dNTP mix containing dATP, dCTP, dGTP, and dTTP/dUTP of ultrapure quality, QuantiTect Multiplex PCR Buffer and HotStarTaq DNA Polymerase.

The required amounts of diluent and DNA extract were mixed in a tube and vortexed for at least 30 seconds. To minimize any contamination risk stemming from general handling, the tube containing DNA solution was immediately spun down. Also, for the sake of keeping the risk of contamination down to the lowest level possible, there are two separate, adjacent laboratories in the AGES. Only in one of these two, working with DNA is allowed. After thoroughly mixing the solutions of primers, probes and the master mix, the mix as part of the

PCR reaction mixture was prepared in a PCR workstation (LTF Labortechnik, Germany) of the laboratory in which working with DNA is not allowed. For this reason, water, master mix and the forward and reverse primer as well as the probe was combined into a tube in the required amounts (see Table 5 and Table 10). After vortexing and spinning down, 20 µL each of the just prepared mix was transferred to a well of a 96-well plate (Sarstedt, Austria; Thermo Fisher Scientific). The 96-well plate was taken into a different PCR workstation of the other laboratory before adding DNA to the plate. Following protocol, 5 µL of DNA solution was added into the well. Therefore, the total volume in one well was 25 µL. After the plate being sealed either by foil or by 8-cap strips and quickly spun down, the run was started under the following conditions that were given in the instructions of the master mix in use.

The temperature program of the thermocycler remained unaltered during the course of every experiment (see Table 4). In addition, the data obtained were analyzed as described in subchapter 4.4.3.

Table 4 Applied temperature program for the real-time PCR assays.

step	T [°C]	time [min]	repeats
1. denaturation including enzyme activation	95 °C	15	1
2. denaturation	94	1	45
elongation	60 °C	1	

In every single experiment, a no-template control (NTC) as well as a positive control were included, thereby tracking any contamination on the spot and checking if successful amplification occurred, respectively. As for the NTCs, both water used for preparing the PCR reaction mixture and water used for preparing DNA solutions were always tested when running a PCR. Positive controls of soy were applied at a concentration of 0.05 µg/mL, which was obtained by dilution of soy DNA extract with water. Each NTC, positive control and sample was tested in each run at a minimum of two replicates.

For qualitative data analysis, the threshold was identified using the logarithmic view of the amplification plot and the end cycle as part of the baseline setting was adjusted as described in subchapter 4.4.3. In case different runs needed to be compared with each other, the thresholds had to be set at the same level. The start cycle was never changed (start cycle = 3); it was automatically set by the corresponding software. For end cycle adjustment, four cycles were subtracted from the C_t value first recorded, following the literature recommendation of "*at least two cycles*" (McPherson & Møller, 2006c) as well as providing a

way of standardization, which also supported comparison between different runs. Obtained C_t values were compared with positive controls and amplification curves were evaluated (using the logarithmic and linear view).

5.8 Development and validation of the singleplex real-time PCR assay for the detection of soy

5.8.1 Primer/probe design

Sequence data for the design of forward and reverse primers and TaqMan probes were derived from NCBI GenBank database. The software CLC Genomics Workbench, version 10.1.1, (Qiagen) and Primer Express Software for Real-Time PCR, version 3.0.1, (Thermo Fisher Scientific) with setting *TaqMan MGB quantification* selected, was used.

The alignment soy-curcuma was intensively screened for sections containing sequences potentially specific for soy. When such a section was found, up to 100 bp were marked and pasted into Primer Express Software for Real-Time PCR so as to quickly locate potential primers and probes within this section by software. This procedure was only successful in very few cases (e.g. system 8, Table 18). Either there was no primer and/or probe found at all or by the software suggested primers and probes were not specific for soy, i.e. there were binding sites in the chloroplast genome of curcuma too.

The alternative approach was to design manually. For manual design, potential forward and reverse primers were checked with the Primer Probe Test Tool of Primer Express Software for T_m (calculated from the recommended master mix of the corresponding company, Thermo Fisher Scientific) and G:C content; the difference in T_m was taken into consideration too. For probes, sequences featuring G at the 5'-end were avoided. Potential probe sequences were also checked with the help of the Primer Probe Test Tool, including an *in silico* compatibility check with corresponding primers. Potential primer/probe systems were then checked for binding sites in the other alignment (soy-common bean-peanut-carob-macadamia-chickpea), excluding any foreseeable cross-reactions with these species. Nevertheless, the options were quite limited; finding a sequence specific for soy was set as the most important goal. Parameters provided by the software for primer as well as probe evaluation were acknowledged but strict compliance was subordinated to specificity.

Eight different primer/probe systems were designed and six different primer/probe systems, diverse in the number of forward and reverse primers and TaqMan probes, were tested in PCR experiments (see Table 18). One primer pair and corresponding TaqMan probe and one reverse primer were designed in the course of the underlying master's thesis, the others were designed by my colleague Walter Mayer (AGES). TaqMan probes were modified with Cy5 as

reporter dye at their 5'-end, allowing detection of soy DNA in the red detector channel of the real-time thermocycler (excitation at 625 nm, emission at 670 nm), and a minor groove binder-Eclipse™ Dark Quencher (MGB-EDQ) at their 3'-end. All primers and TaqMan probes were ordered from Eurogentec, BE, each of which with a stock concentration of 100 µM dissolved in TE buffer.

Example of a pipetting scheme of the singleplex real-time PCR assay (Table 5)

Each preparation of the reaction mixture (1) and, thus, of the total PCR reaction mixture (2) of the following experiments testing different primer/probe systems, also in varying combinations, specificity and different primer and probe concentrations was based on the exact same calculatory principle. Only the total number of reactions (3) as well as the concentrations of the working solutions were changed. Sometimes it was useful to adjust the aliquoted stock solution to a lower concentration so that a bigger volume could be added. Principally, it was avoided to handle volumes below 1 µL.

Table 5 Example of a pipetting scheme of the singleplex real-time PCR assay (total number of reactions = 1).

(1) reaction mixture [µL]	20					
sample DNA [µL]	5					
component	primer/probe	c stock solution [µM]	c working solution [µM]	one PCR reaction mixture [µL]	**(3)** total number of reactions	**(2) total PCR reaction mixture [µL]**
QuantiTect® Multiplex PCR NoROX Master Mix		2x	1x	12.5	*x (1)*	**12.5**
soy system	*x*	100	0.2	0.05		**0.05**
	x	100	0.2	0.05		**0.05**
	x	100	0.1	0.025		**0.025**
total water				7.20		**7.20**

5.8.2 Specificity tests

5.8.2.1 Specificity tests with species closely related to soy

DNA extracts of nine species closely related to soy and/or relevant were diluted with water to prepare DNA solutions with a concentration of 5 µg/mL, which were tested for cross-reactivity under non-optimized conditions (FW/RV/PR [µM] 0.5/0.5/0.2) in at least two replicates each. Tests were carried out with different combinations of forward and reverse primers and probes of systems 1, 2, 3, 5, 7, and 8 (FW 1b/RV 1b/PR 1b, FW 2a/RV 2a/PR 2, FW 3b/RV 3b/PR 3b, FW 5b/RV 5c/PR 5b, FW 7a/RV 7a/PR 7a, FW 8/RV 8/PR 8). A list of the species closely

related to soy and/or relevant is given in subchapter 6.2.2.1, Table 19. Thresholds were set at the same level (10,000) for comparison.

5.8.2.2 Specificity tests with species not closely related to soy

With the selected primer/probe system FW 7e/RV 7c/PR 7b at working concentrations of 0.2, 0.2, and 0.1 µM for forward and reverse primer and probe, respectively, follow-up specificity tests with species not closely related to soy were carried out. DNA extracts of species not closely related to soy were diluted with water to prepare DNA solutions with a concentration of 5 µg/mL, analyzed in at least two replicates each. The DNA solutions of these species were tested for cross-reactivity, in addition to species closely related to soy and/or relevant that had been tested (see 5.8.2.1). Thresholds were set at the same level (10,000) for comparison.

In total, 86 species comprising species not and closely related to soy and/or relevant were investigated. These species include almond, anise, apple, atyidae (“shrimp”), beetroot, Brazil nut, broccoli, buckwheat, caraway, cardamon, carob, carrot, cashew, cauliflower, celery, cherry, chicken/turkey/beef/pork, chickpea, chili, chive, cilantro, common bean, cress, cucumber, cumin, curcuma, dill, fennel, garlic, ginger, hazelnut, hazelnut, herring, horseradish, house cricket, laurel, leek, lentil, lesser mealworm (“buffalo worm”), lovage, lupin, macadamia, maize, marjoram, yellow mealworm, migratory locust, mustard, black, mustard, brown, mustard, white, oat, onion, oregano, paprika, parsley, parsnip, pea, peach, peanut, pecan, pepper, pimento, pistachio nut, potato, quince, radish, radish, rape, rice, rosemary, rye, sage, sesame, sorghum, squash, squid, strawberry, summer savory, tarragon, thyme, tomato, turnip, walnut, wheat and white cabbage. Another list of all species along with the obtained mean C_t values in experiments is given in subchapter 6.2.2.2, Table 21.

Agarose gel electrophoresis

200 mL of 10x TBE buffer (Bio-Rad, USA) was diluted with water to 2 L of 1x TBE buffer. For the required 2 % (weight per weight [w/w]) agarose gel, 1.6 g agarose powder (Sigma-Aldrich) was weighed out and dissolved in 80 mL of 1x TBE buffer. The mixture was heated until boiling in the microwave; each 20-30 seconds the flask containing the mixture was carefully stirred manually so that complete dissolution was guaranteed. Subsequently, 8 µL of GelRed (Merck) was added, well-mixed and after a brief cool-down period to 60 C° the gel was poured into the provided equipment, already prepared with the comb correctly placed, in which it coagulated within 30-45 minutes.

As soon as polymerization of the agarose gel was finished (more solid, white-colored gel), it was transferred to the electrophoresis chamber. Then, a volume of 1x TBE buffer sufficient to cover the gel with approximately 2 cm of the solution was also added; the comb was removed. To 25 µL of PCR product, 5 µL of loading buffer (Biozym, Germany) was added. 10 µL each of the solution containing PCR product and loading buffer was transferred into the wells, and so was 2 times 10 µL of DNA ladder (25 bp) into the two external wells of the gel. Electrophoresis was run at 190 V for approximately 90 minutes. Immediately after stopping the run, the gel was analyzed under UV light with a transilluminator (Bio-Rad) using Quantity One Basic Software, version 4.6.6 (Bio-Rad). A picture was taken for protocolling.

5.8.3 Optimization by primer/probe titration

5.8.3.1 Primer titration

The primer titration was carried out with forward and reverse primers of the primer/probe systems FW 1b/RV 1b/PR 1b and FW 7a/RV 7a/PR 7a. The standard primer concentration of 0.5 µM was varied for both systems. In Table 6, the tested concentrations are listed. The standard probe concentration of 0.2 µM was kept constant for both systems. DNA extracts of three species closely related to soy and/or relevant, curcuma, common bean and carob, as well as of soy as positive control were diluted with water to prepare DNA solutions with concentrations of 5 and 0.05 µg/mL, respectively. The DNA solutions were analyzed in at least two replicates each.

Table 6 Varying primer concentrations tested with the primer/probe systems FW 1b/RV 1b/PR 1b and FW 7a/RV 7a/PR 7a.

c forward primer [µM]	c reverse primer [µM]
0.2	0.5
0.2	0.9
0.5	0.2
0.5	0.9
0.9	0.2
0.9	0.5

5.8.3.2 Primer/probe titration

The probe titration was carried out with the probe of the primer/probe system FW 7a/RV 7a/PR 7a. The standard probe concentration of 0.2 µM was varied. Equimolar concentrations of primers and varying concentrations of the probe were tested. In Table 7, the tested concentrations are listed. DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was diluted with water to prepare a DNA solution corresponding to a concentration of 1 mg/kg. The DNA solution was analyzed in at least two replicates each.

Table 7 Varying probe concentrations tested with the primer/probe system FW 7a/RV 7a/PR 7a.

c forward primer [µM]	c reverse primer [µM]	c probe [µM]
0.1	0.1	0.2
0.2	0.2	0.2
0.2	0.2	0.1
0.2	0.2	0.15
0.35	0.35	0.2
0.35	0.35	0.15
0.35	0.35	0.1
0.5	0.5	0.1

5.8.4 Summary of the singleplex real-time PCR assay

The optimized PCR assay was performed with the primer/probe system FW 7e/RV 7c/PR 7b. Reactions were carried out in a total reaction volume of 25 µL, consisting of 12.5 µL QuantiTect® Multiplex PCR NoROX Master Mix, 0.2 µM forward primer, 0.2 µM reverse primer as well as 0.1 µM probe, 5 µL DNA solution and water. The following temperature program was used: 1x 15 min at 95 °C; 45x 1 min at 94 °C, 1 min at 60 °C (see subchapter 5.7).

5.8.5 Determination of limit of detection (LOD) and C_t cut-off value

The validation parameters LOD and robustness were determined based on the European Network of GMO Laboratories (ENGL) guidelines (European Commission. Joint Research Centre., Verification of analytical methods for GMO testing when implementing interlaboratory validated methods :guidance document from the European Network of GMO Laboratories, version 2, 2017).

The CTAB-based method by the DNA extraction machine Maxwell® 16 Instrument (see subchapter 5.5.1) was used to extract DNA of a mean weighed-out quantity of 0.298 g SureFood® QUANTARD Allergen 40 (40 mg/kg) three times for three independent PCR runs, which were carried out on three different days (PCR run 1-3, see Table 26). The DNA extract was diluted with non-target DNA (herring sperm DNA solution [Böhringer Mannheim, Germany], 100 µg DNA/mL) to prepare DNA solutions corresponding to concentrations of 1, 0.5, 0.25, 0.125, 0.06, 0.03, 0.016, 0.008, 0.004, 0.002 and 0.001 mg/kg (serial dilution, 1:2). The DNA solutions corresponding to concentrations of 0.06-0.004 mg/kg (PCR run 1 and 2) and 0.03-0.004 mg/kg (PCR run 3) were analyzed in twelve replicates each.

The determination of the LOD was based on the lowest concentration that led to an increase of the fluorescence signal within 45 cycles (i.e. positive result). The C_t cut-off value was calculated from this concentration, too. First, the mean C_t value was calculated from the C_t values obtained at the lowest concentration at which all (twelve out of twelve) replicate measurements were positive. Second, the standard deviation was also calculated and its value was added two times to the mean C_t value, yielding the *C_t cut-off value*. All C_t values obtained at the lowest concentration were checked to see if each of them was below the C_t cut-off value. If this was the case, the corresponding concentration was determined as the LOD. The LOD of the singleplex real-time PCR assay was, therefore, defined as the lowest concentration that led to an increase of the fluorescence signal below the C_t cut-off value in all replicate measurements.

The robustness was examined by running the exact same PCR assay on a duplicate thermocycler of the model 7500 Real Time PCR System of Applied Biosystems®, again using the associated software Applied Biosystems® 7500 System SDS Software.

Thresholds were set at the same level (10,000) for comparison.

5.9 Development and validation of the triplex real-time PCR assay for the simultaneous detection of soy, celery and white mustard

Table 8 shows the initial concentrations of primers and probes including reporter dye for the three targets of the triplex real-time PCR assay. Further information on primers and probes of the duplex assay for celery and white mustard is given in Table 9.

Table 8 Initial concentrations of the forward and reverse primers and probes and reporter dye of the triplex real-time PCR assay.

target	initial concentration [µM]			probe reporter dye (5'-end)
	forward primer	reverse primer	probe	
soy	0.2	0.2	0.1	Cy5
celery	0.4	0.4	0.1	FAM
white mustard	0.4	0.4	0.1	HEX

Table 9 Information on the primers and probes of the duplex assay for celery and white mustard and the working concentrations.

primer/probe	sequence 5'→3'	T_m [°C]*	length [bp]	amplicon length [bp]	working concentrations [µM]
Sel FW4	AGG TCA TTT CTA TAC TAT CAT TTC TAT AC	51.0	29	89	0.4
Sel RV2	GAA TTT TCC TCC TTT TCC TTT TCT	56.5	24		0.4
Sel PR 0	**FAM**-AGG GAG AGT TTC ACT AAC-**MGB**	68.0	18		0.1
AB669975 FW1	AGC AGG GTA TTT TTG AGT TTT CGT	58.4	24	99	0.4
AB669975 RV1	CTC CTC TTT GGT TTC CAT CAT ATT TAT	58.1	27		0.4
AB669975 PR1	**HEX**-CCT TTA GGA AGA ATC CT-**MGB**	68.0	17		0.1

*according to T_m calculator (Thermo Fisher Scientific)

Since the probe for white mustard was also in use otherwise at the routine laboratory of the AGES, it was labelled with the fluorescent dye HEX at the 5'-end (enabling applicability in digital droplet PCR). HEX, however, could not be selected as detection channel on the used thermocycler. For this reason, the detection channel VIC was selected instead. As HEX and VIC have almost identical excitation and emission spectra, an exception to use them interchangeably could be made (HEX: excitation 535 nm, emission 556 nm; VIC: excitation

538 nm, emission 554 nm) (AAT Bioquest, AAT Bioquest's interactive Spectrum Viewer, 2021).

Example of a pipetting scheme of the triplex real-time PCR assay (Table 10)

Each preparation of the reaction mixture (1) and, thus, of the total PCR reaction mixture (2) of the following experiments testing different primer/probe systems, also in varying combinations, specificity and different primer and probe concentrations was based on the exact same calculatory principle. Only the total number of reactions (3) as well as the concentrations of the working solutions were changed (for details on general handling see Table 5).

Table 10 Example of a pipetting scheme of the triplex real-time PCR assay (total number of reactions = 1).

(1) reaction mixture [µL]		20				
sample DNA [µL]		5				
component	primer/probe	c stock solution [µM]	c working solution [µM]	one PCR reaction mixture [µL]	**(3)** total number of reactions	**(2) total PCR reaction mixture [µL]**
QuantiTect® Multiplex PCR NoROX Master Mix		2x	1x	12.5		**12.5**
soy system	*So60_FW 7e*	100	0.2	0.05		**0.05**
	So60_RV 7c	100	0.2	0.05		**0.05**
	So60_PR 7b	100	0.1	0.025		**0.025**
celery system	*Sel FW 4*	100	0.4	0.10		**0.10**
	Sel RV 2	100	0.4	0.10	*x (1)*	**0.10**
	Sel PR 0	100	0.1	0.025		**0.025**
white mustard system	*AB669975 FW 1*	100	0.4	0.10		**0.10**
	AB669975 RV 1	100	0.4	0.10		**0.10**
	AB669975 PR 1	100	0.1	0.025		**0.025**
total water				6.70		**6.70**

DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was diluted with water to prepare DNA solutions corresponding to concentrations of 1, 0.5, 0.25, 0.125, 0.06, 0.03 and 0.016 mg target/kg (target = soy/celery/white mustard; serial dilution, 1:2). The DNA solutions corresponding to concentrations of 1-0.125 mg target/kg were analyzed in at least twelve replicates each.

5.9.1 Optimization by primer/probe titration

DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was diluted with water as well as high amounts of target DNA to prepare DNA solutions corresponding to concentrations of 1, 0.5, 0.25, 0.125, 0.06, 0.03 and 0.016 mg target/kg (target DNA for soy: celery, white

mustard; target DNA for celery: soy, white mustard; target DNA for white mustard: soy, celery; serial dilution, 1:2). The DNA solutions were analyzed in at least two replicates each.

5.9.1.1 Optimization of white mustard PCR assay

A forward and reverse primer concentration of 0.8 µM was applied, which was tested with a continuously increasing probe concentration (see Table 11). The different combinations for white mustard were investigated in presence of high amounts of celery (20 µg DNA/mL) as background DNA. The primer and probe concentrations of the PCR assays for soy and celery were kept constant.

Table 11 Varying concentrations of the probe for white mustard in presence of high amounts of celery (20 µg DNA/mL) as background DNA for the optimization of the triplex real-time PCR assay.

primer/probe system white mustard			
primer/probe # optimization attempt	c FW [µM]	c RV [µM]	c PR [µM]
#1	0.8	0.8	0.25
#2	0.8	0.8	0.3
#3	0.8	0.8	0.35
#4	0.8	0.8	0.4

5.9.1.2 Optimization of soy PCR assay

Three different combinations of lower concentrations of forward and reverse primer and probe for soy (see Table 12) were investigated in presence of high amounts of soy (20 µg DNA/mL) as background DNA. The primer and probe concentrations of the PCR assays for celery and white mustard were kept constant.

Table 12 Varying concentrations of the forward and reverse primer and probe for soy in presence of high amounts of soy (20 µg DNA/mL) as background DNA for the optimization of the triplex real-time PCR assay.

primer/probe system soy			
primer/probe # optimization attempt	c FW [µM]	c RV [µM]	c PR [µM]
#1	0.1	0.1	0.05
#2	0.05	0.05	0.025
#3	0.025	0.025	0.0125

The first combination with forward primer, reverse primer and probe concentrations of 0.1, 0.1 and 0.05 µM, respectively, (see Table 12) was also tested applying high amounts of DNA of the other two targets of the triplex assay, celery and white mustard, (20 µg celery DNA/mL, 5 µg white mustard DNA/mL) as background DNA.

In addition, both primer and probe concentrations for soy were varied, as summarized in Table 13. Four different combinations were investigated in presence of high amounts of celery (20 µg DNA/mL) as background DNA. For the white mustard PCR assay, optimized concentrations (FW/RV/PR [µM] 0.8/0.8/0.3, see subchapter 5.9.1.1) were applied. The primer and probe concentrations of the PCR assay for celery were kept constant.

Table 13 Varying concentrations of the forward and reverse primer and probe for soy in presence of high amounts of celery (20 µg DNA/mL) as background DNA for the optimization of the triplex real-time PCR assay.

primer/probe system soy			
primer/probe # optimization attempt	c FW [µM]	c RV [µM]	c PR [µM]
#1	0.1	0.1	0.1
#2	0.1	0.1	0.15
#3	0.15	0.15	0.1
#4	0.2	0.2	0.1

The probe concentration for soy was continuously increased (see Table 14). The forward and reverse primer concentration of 0.1 µM was kept constant.

Table 14 Varying concentrations of the probe for soy in presence of high amounts of celery (20 µg DNA/mL) as background DNA for the optimization of the triplex real-time PCR assay.

primer/probe system soy			
primer/probe # optimization attempt	c FW [µM]	c RV [µM]	c PR [µM]
#1	0.1	0.1	0.2
#2	0.1	0.1	0.25
#3	0.1	0.1	0.3
#4	0.1	0.1	0.35

The first combination with forward primer, reverse primer and probe concentrations of 0.1, 0.1 and 0.2 µM, respectively, (see Table 14), was also tested applying high amounts of DNA of the other two targets of the triplex assay, soy and white mustard, (20 µg soy DNA/mL, 20 µg

white mustard DNA/mL) as background DNA, as well as water as diluent. The primer and probe concentrations of the PCR assays for celery and white mustard were kept constant.

5.9.2 Summary of the triplex real-time PCR assay

Every PCR tube was filled with 12.5 µL QuantiTect® Multiplex PCR NoROX Master Mix, 0.2 µM forward primer, 0.2 µM reverse primer as well as 0.1 µM probe for soy, 0.4 µM forward primer, 0.4 µM reverse primer as well as 0.1 µM probe for celery and 0.4 µM forward primer, 0.4 µM reverse primer as well as 0.1 µM probe for white mustard. Furthermore, 5 µL DNA solution and water were contained, resulting in a total reaction volume of 25 µL. The following temperature program was used: 1x 15 min at 95 °C; 45x 1 min at 94 °C, 1 min at 60 °C (see subchapter 5.7).

5.9.3 Inhibition control (IC)

DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was diluted with very high amounts of target DNA (100 µg target DNA/mL; target DNA for soy: celery, white mustard; target DNA for celery: soy, white mustard; target DNA for white mustard: soy, celery) to prepare DNA solutions corresponding to concentrations of 1, 0.5, 0.25, 0.125, 0.06, 0.03 and 0.016 mg target/kg (serial dilution, 1:2). The DNA solutions corresponding to concentrations of 0.25-0.016 mg target/kg were spiked with a soy/celery/white mustard DNA solution corresponding to a concentration of 0.25 mg target/kg, prepared from DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg). A positive/spike control, containing QuantiTect® Multiplex PCR NoROX Master Mix, the forward and reverse primers as well as the probes for soy, celery and white mustard (see subchapter 5.9.2) and water, was spiked, too, with the same soy/celery/white mustard DNA solution corresponding to a concentration of 0.25 mg target/kg. This positive/spike control was analyzed together with the equally spiked prepared DNA solutions (0.25-0.016 mg target/kg), in at least two replicates each.

5.9.4 Determination of limits of detection (LODs) and C_t cut-off values

The validation parameter LOD was determined based on the ENGL guidelines (European Commission. Joint Research Centre., Verification of analytical methods for GMO testing when implementing interlaboratory validated methods :guidance document from the European Network of GMO Laboratories, version 2, 2017). DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was diluted with high amounts of target DNA (5 or 20 µg target DNA/mL; target DNA for soy: celery, white mustard; target DNA for celery: soy, white mustard;

target DNA for white mustard: soy, celery) to prepare DNA solutions corresponding to concentrations of 1, 0.5, 0.25, 0.125, 0.06, 0.03 and 0.016 mg target/kg (serial dilution, 1:2). The DNA solutions corresponding to concentrations of 0.25-0.016 mg target/kg (PCR run 1A, B), 0.125-0.016 mg target/kg (PCR run 2A, B) and 0.25-0.016 or 1-0.016 mg target/kg (PCR run 3A, B) were analyzed in twelve or sixteen replicates each (see PCR runs 1A-3B in Table 32). Thresholds were set at the same level for comparison (soy: 10,000; celery: 50,000; white mustard: 10,000).

The determination of the LOD was based on the preliminary assumption that the C_t cut-off value is 40. Consequently, an increase of the fluorescence signal at or above cycle 40 (C_t value equal or greater 40) was interpreted as a negative result and an increase of the fluorescence signal below cycle 40 (C_t value less 40) was interpreted as a positive result. According to this assumption, the mean C_t value of the C_t values, obtained at the lowest concentration at which all (twelve out of twelve or sixteen out of sixteen) replicate measurements were positive, was calculated. The standard deviation was also calculated and its value was added two times to the mean C_t value, yielding the adjusted *C_t cut-off value* (instead of the preliminarily assumed C_t cut-off value of 40). All C_t values obtained at the lowest concentration were checked to see if each of them was below the adjusted C_t cut-off value. If this was the case, the corresponding concentration was determined as the LOD. For each target (soy, celery, white mustard), two corresponding concentrations and two C_t cut-off values resulted. As the final LODs of the triplex real-time PCR assay, the higher concentration and the lower C_t cut-off value was used.

6 Results and discussion

The main purpose of the master's thesis was to develop a qualitative singleplex real-time PCR assay for the detection of soy (*Glycine max*) and its validation as part of a qualitative triplex real-time PCR assay for the simultaneous detection of soy, celery (*Apium graveolens*) and white mustard (*Sinapis alba*) in food. Aiming the singleplex PCR assay to be run on the same 96-well reaction plate with other real-time PCR assays, the annealing temperature had to be 60 °C as a prerequisite.

In the first step, the qualitative singleplex real-time PCR assay for the detection of soy was developed. In the second step, this novel singleplex assay was combined with the already developed and successfully validated qualitative duplex real-time PCR assay for the simultaneous detection of celery and white mustard to form a triplex system.

In the following, the term *soy* refers to *soybean* and vice versa.

6.1 Genomic DNA extraction by CTAB-based method and quantification of total DNA by spectroscopy

After genomic DNA extraction employing either the CTAB-based method by the DNA extraction machine Maxwell® 16 Instrument or manually (see subchapter 5.5), total DNA was spectrophotometrically quantified (see subchapter 5.6).

Table 15 and Table 16 give the results of the quantification of total DNA by spectroscopy for different species and celery and white mustard.

Table 15 Results of the quantification of total DNA by spectroscopy after the DNA extraction with CTAB-based methods for different species. The mean concentration and standard deviation (S) were calculated from two DNA extractions.

sample	mean concentration (S, RSD %) [µg/mL]	type of CTAB-based method
carob	36.3*	manually
chickpea	453.4 (244.7, 54.0 %)	manually
common bean	144.3*	manually
curcuma	4.2*	manually
dill	126.3*	Maxwell
ginger	37.1*	Maxwell
lentil	854.6 (41.5, 4.9 %)	manually
millet	63.7*	Maxwell
pea	284.0 (41.9, 14.8 %)	manually
pepper	46.7 (8.5, 18.1 %)	Maxwell
rosemary	73.2*	Maxwell
sage	37.0*	Maxwell
sesame	210.7*	Maxwell

*one DNA extraction

DNA was extracted from plant material of the species listed in Table 15 because (trace) soy contamination of material priorly used and/or their DNA extracts was suspected based on preliminary experiments by Walter Mayer (for details on suspicion of soy contamination see subchapter 6.2.2.2). Obtained DNA extract concentrations and the standard deviations for different species varied; for instance, the standard deviation for chickpea was high (244.7 µg/mL, RSD 54.0 %). With the DNA extraction machine, the standard deviation tended to be lower (see pepper in Table 15). This is probably due to the automatized DNA extraction, in contrast to the manual extraction, which is known to be less reproducible. However, there is not enough data to discuss this tendency further. The DNA concentration of the curcuma

extract was with 4.2 µg/mL very low, presumably because the curcuma root was used for the extraction procedure due to the unavailability of other source material.

The DNA extraction machine was utilized for DNA extraction of the reference material SureFood® QUANTARD Allergen 40 (40 mg/kg) because the same protocol is applied in routine analysis of the AGES, too. A mean DNA concentration of 139 µg/mL with a standard deviation of 12 µg/mL and a relative standard deviation (RSD) of 8.6 % was obtained from triplicates.

Table 16 Results of the quantification of total DNA by spectroscopy after the DNA extraction with CTAB-based methods for celery and white mustard.

sample	mean concentration (S, RSD %) [µg/mL]	concentration range [µg/mL]
celery	89.9 (16.5, 18.4 %)[1]	78.2-101.6
type of CTAB-based method	$A_{260/280}$	$A_{260/230}$
manually	1.97	2.53
Maxwell	2.18	1.63
mustard, white	192.3 (200.4, 104.2 %)[2]	19.6-469.1
type of CTAB-based method	mean $A_{260/280}$	mean $A_{260/230}$
manually	2.06[3]	2.44[3]
Maxwell	2.01[4]	2.98[4]

[1] two DNA extractions
[2] five DNA extractions
[3] mean absorbance ratio calculated from two extractions (manually)
[4] mean absorbance ratio calculated from three extractions (Maxwell)

Celery and white mustard DNA had to be extracted more often in the framework of developing the triplex real-time PCR assay (see Table 16). Overall, the smaller the plant material was reduced before lysis, the higher DNA yield. This was especially noticeable for white mustard, whose DNA extracts showed DNA concentrations in the range from 19.6 to 469.1 µg/mL. At this stage, the most important goal was, however, to obtain DNA of the corresponding plants free from soy. The less laboratory equipment was used, the less contamination of soy could be brought in. For instance, some plant material of white mustard was reduced to very small pieces just by cutting with a clean knife; no homogenizer was used. This negatively affected DNA yield but led to soy-free DNA extracts of white mustard. As additionally evident from

Table 16 when comparing readings of $A_{260/280}$ and $A_{260/230}$ for both celery and white mustard, the manual extraction appears to yield DNA of slightly higher purity. For celery, the reading of both absorbance ratios for the manual extraction is within the ranges that are considered high-purity DNA (see subchapter 4.3) whereas the reading of $A_{260/280}$ for the extraction by machine is not. For white mustard, both extraction types yielded equally pure DNA. With respect to the well-known limitations of absorption spectroscopy, readings including their ratios as indicators for DNA extract purity as well as quality were taken into account but were of lower priority. In routine laboratories, such as the AGES is, time and cost management is crucial; most preferable are simple, fast and reproducible DNA extraction methods (Demeke & Jenkins, 2010; Terry et al., 2002b). Additional purification steps are only carried out in case the amplification reaction was not successful using crude sample preparations and/or automatized DNA extraction instruments (e.g. Maxwell® 16 Instrument). As already stressed in subchapter 4.3, readings of $A_{260/280}$ above 2.0 do not rule out target amplification in PCR assays since it is only a rough estimation of the purity of DNA extracts. During the practical work of the underlying thesis, DNA purity did not pose any problems with the amplification of the target molecule.

6.2 Development and validation of the singleplex real-time PCR assay for the detection of soy

6.2.1 Primer/probe design

Research was done on species closely related to soy and/or being relevant for human nutrition. These species included common bean, peanut, carob, macadamia, chickpea, and curcuma. The complete sequences of the chloroplast genome of soybean (*Glycine max*, accession no.: NC_007942, 152,218 bp) as well as the chloroplast genomes of the species closely related to soy and/or relevant (see Table 17) were downloaded from NCBI GenBank database; using CLC Genomics Workbench, the complete sequence of the chloroplast genome of soy was aligned with that of curcuma, on the one hand, and with those of common bean, peanut, carob, macadamia and chickpea, on the other hand.

Table 17 Species closely related to soy and/or being relevant for human nutrition and their scientific name, accession number and sequence length of their chloroplast genome.

common name	family	scientific name	accession no.	length [bp]
curcuma	*Zingiberaceae*	*Curcuma flaviflora*	NC_028729	160,478
common bean	*Fabaceae*	*Phaseolus vulgaris*	EU_196765	150,284
peanut	*Fabaceae*	*Arachis hypogaea*	KX257487	156,391
carob	*Fabaceae*	*Ceratonia siliqua*	KJ468096	156,367
macadamia	*Proteaceae*	*Macadamia ternifolia*	KF862711	159,714
chickpea	*Fabaceae*	*Cicer arietinum*	NC_011163	125,319

Table 18 lists forward and reverse primers and TaqMan probes tested in PCR experiments, which target different DNA sequences of the chloroplast genome of soy. The tested primer/probe systems (no. 1, 2, 3, 5, 7, 8; six systems in total) are diverse in the number of forward and reverse primers and TaqMan probes. The light grey shaded ones were designed in the course of the underlying master's thesis (primer/probe set 1c, reverse primer 5d). The others were designed by my colleague Walter Mayer (AGES) (see subchapter 5.8.1).

Table 18 Tested forward and reverse primers and TaqMan probes with their sequence and T_m. The six tested primer/probe systems (no. 1, 2, 3, 5, 7, 8) are diverse in the number of forward and reverse primers and TaqMan probes. FW: forward primer. RV: reverse primer. PR: probe.

primer/probe system no.	primer/probe	sequence 5'→3'	T_m [°C]
1	So60_FW 1b	CGT GGG TGT ATA TAT CCA ATC	54.6*
	So60_FW 1c	AGA TTG AAT TGT CTA ATA AAA TAT CAG AC	53.4
	So60_RV 1b	TCA ATT CGA CTT CAA CAT TTT G	54.9*
	So60_RV 1c	GTA TAA CAT AGT AGA TAA AGA GGT GGT ATA CC	54.8
	So60_PR 1b	**Cy5**-AAT TGT CTA ATA AAA TAT CAG AC-**MGB**	67.0
	So60_PR 1c	**Cy5**-CAA AAT GTT GAA GTC GA-**MGB**	64.0
2	So60_FW 2a	CAA CAC ATC CGG AAA CAT C	54.2*
	So60_RV 2a	ATC TTA TTA TTC CTA CCT GTT AGT AAC	53.8*
	So60_PR 2	**Cy5**-CGG AAG TTT TAA GTA AAT G-**MGB**	68.0
3	So60_FW 3b	AAG ACA GAT TCA TTT GCA TAC TTA	54.1*
	So60_RV 3b	GCA AGC TAT CTT CCG ATT ATT T	54.9*
	So60_PR 3b	**Cy5**-TTC GTA ATG TCT CCC G-**MGB**	65.0
5	So60_FW 5b	TCC CTT ATC TTG ATA AAG ATT TTG A	54.5*
	So60_RV 5b	TTT TCT TAT TTG TAT TTA TCC CTT TTA CCT	57.0
	So60_RV 5c	CCT TTT ACC TTT ATT TGG GAA TAA G	54.9*
	So60_RV 5d	GTA TTT ATC CCT TTT ACC TTT ATT TGG	55.4
	So60_PR 5b	**Cy5**-CCT ATA TTG GTA AAG ATT-**MGB**	65.0
7	So60_FW 7a	CCT GCA ATT CAT TTT TTT CC	54.1*
	So60_ FW 7e	TTC CAT CCC TGC AAT TCA TTT	58.0
	So60_RV 7a	TCG TAA ACG TAG AAA AAA ATT C	54.3*
	So60_RV 7c	CAA AAG ACC CGG TAG ATT TTG TAC	57.0
	So60_PR 7	**Cy5**-CTA TGG ATA GAG TAT TTT C-**MGB**	67.0
	So60_PR 7a	**Cy5**-TGG ATA GAG TAT TTT C-**MGB**	65.0
	So60_PR 7b	**Cy5**-CCT CTA CTA TGG ATA GAG TA-**MGB**	65.0
8	So60_FW 8	GGG ATG GAA ATG AAG GAA TGT C	58.2
	So60_RV 8	TGA ACC TAC AAA ACC CTT CAA ATT G	59.3
	So60_PR 8	**Cy5**-CCT GGA TTG AAT CAG-**MGB**	66.0

*according to T_m calculator (Thermo Fisher Scientific)
Cy5: cyanine-5 (fluorescent reporter)
MGB: minor groove binder (non-fluorescent quencher)

With the MGB modification, cross-talk between adjacent channels is reduced and a higher T_m could be reached even though the sequence is comparatively short in length. Care was taken to design systems that generate amplicons short in length for efficient amplification, as recommended, particularly for the analysis of processed food.

6.2.2 Specificity tests

6.2.2.1 Specificity tests with species closely related to soy

DNA solutions (5 µg/mL) of only a few species, all of which closely related to soy and/or being relevant for human nutrition, were tested under non-optimized conditions with different combinations of primers and probes of the six systems, 1, 2, 3, 5, 7 and 8, (see subchapter 5.8.2.1 and Table 18). In Table 19, a list of the preliminarily tested species closely related to soy and/or being relevant for human nutrition is given.

Table 19 Species closely related to soy and/or being relevant for human nutrition and their scientific name.

common name	family	scientific name
carob	*Fabaceae*	*Ceratonia siliqua*
chickpea	*Fabaceae*	*Cicer arietinum*
common bean	*Fabaceae*	*Phaseolus vulgaris*
curcuma	*Zingiberaceae*	*Curcuma flaviflora*
lentil	*Fabaceae*	*Lens culinaris*
lupin	*Fabaceae*	*Lupinus arboreus*
navy bean *(white)*	*Fabaceae*	variety of *Phaseolus vulgaris*
pea	*Fabaceae*	*Pisum sativum*
peanut	*Fabaceae*	*Arachis hypogaea*

Tests with species closely related to soy and/or being relevant for human nutrition were carried out not only to investigate cross-reactivity but also to preselect primer/probe systems.

The smallest differences between the mean C_t value for the soy positive control and the mean C_t values of the cross-reacting species (ΔC_t values) were obtained with the primer/probe system 8. Only with system 8, a cross-reaction resulting in a lower mean C_t value than the mean C_t value for the soy positive control was obtained. Figure 5 shows the amplification curves obtained with this system.

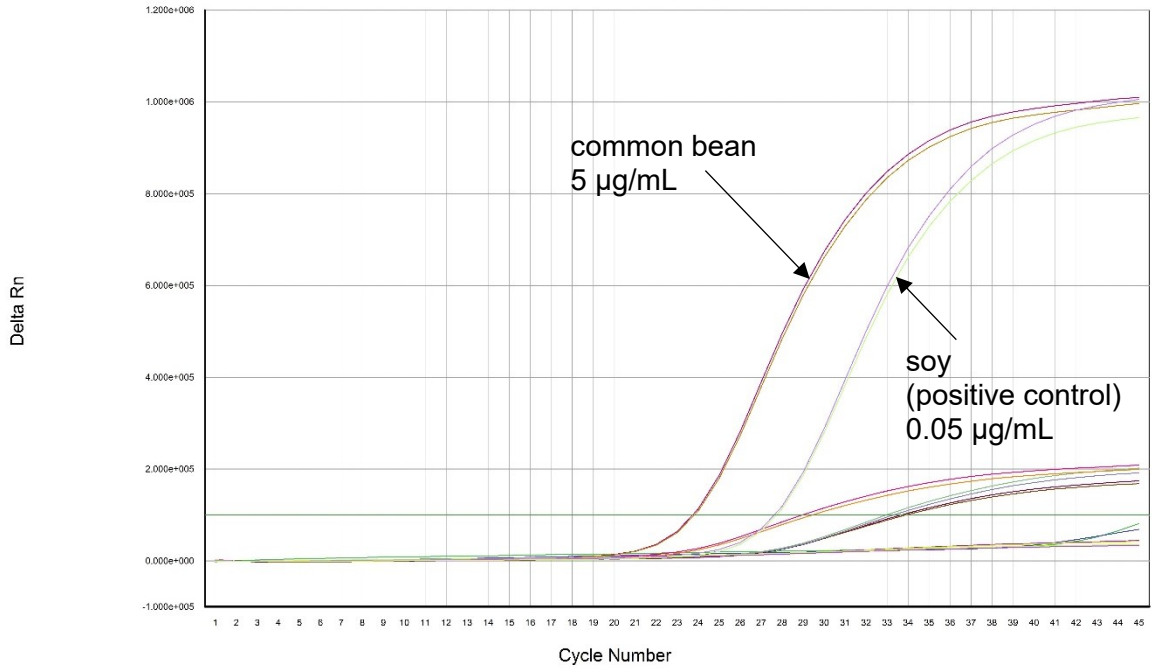

Figure 5 Cross-reaction with common bean (DNA solution: 5 µg/mL)) obtained with the primer/probe system 8 (FW/RV/PR [µM] 0.5/0.5/0.2).

Two out of two replicates of common bean with a concentration of 5 µg/mL showed an increase of the fluorescence signal (see Figure 5). The mean C_t value for the soy positive control was 27.67 and the mean C_t value for common bean (5 µg/mL) was 23.78, resulting in a ΔC_t value of 3.89. Primer/probe system 8 was the only one tested that was not manually designed and checked for binding sites in the other alignment (soy-common bean-peanut-carob-macadamia-chickpea). Primers and probe were suggested by Primer Express Software (see subchapters 5.8.1 and 6.2.1). Suggestions by Primer Express Software come along with the advantage of low probability of the formation of primer dimers and other secondary structures. The disadvantage that this system was not specific for soy, however, outweighed this advantage and led to the exclusion of system 8.

Subsequent to further exclusion of primer/probe systems 2, 3 and 5 due to cross-reactivity, the number of primer/probe systems was ultimately narrowed down to the best performing two, system 1 and system 7. With those two primer/probe systems, no cross-reactivity with

four of the nine tested species closely related to soy and/or relevant (navy bean, chickpea, lentil, pea) was observed, as evident from Figure 6.

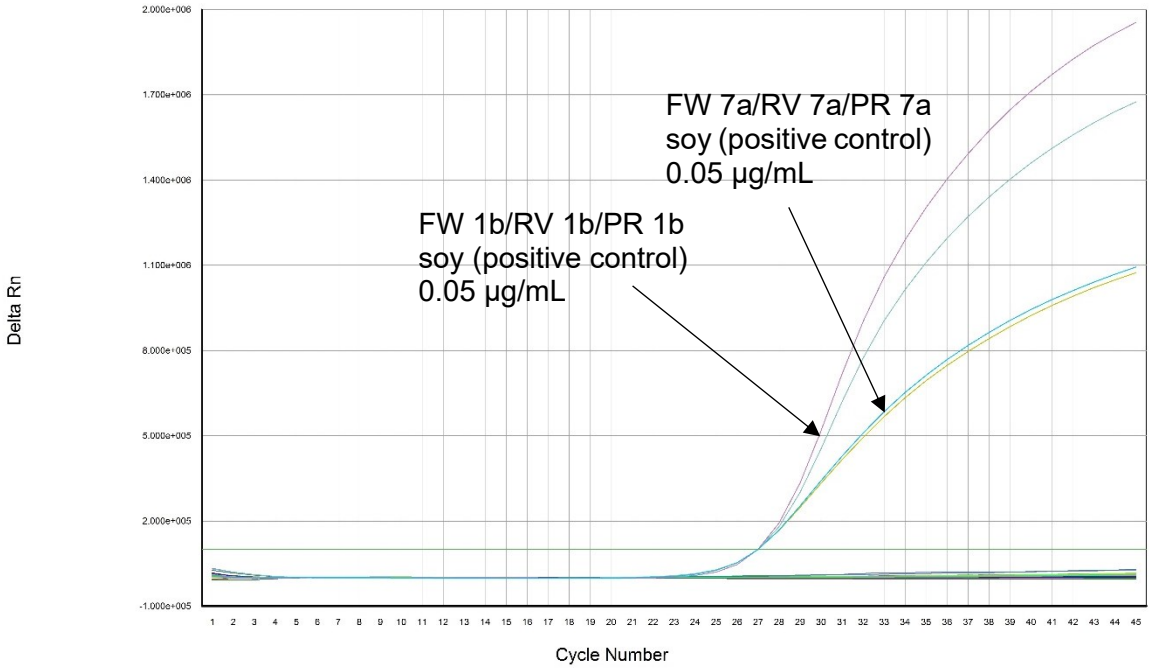

Figure 6 Amplification plot for the specificity test with species closely related to soy and/or being relevant for human nutrition (navy bean, chickpea, lentil, pea) with the primer/probe system 1 (FW 1b/RV 1b/PR 1b) and 7 (FW 7a/RV 7a/PR 7a) (FW/RV/PR [µM] 0.5/0.5/0.2).

Due to these results, primer/probe system 1 (FW 1b/RV 1b/PR 1b) and primer/probe system 7 (FW 7a/RV 7a/PR 7a) were used in optimization experiments (see subchapter 6.2.3).

Optimization experiments showed that primer/probe system 7 performs better than primer/probe system 1 as a higher ΔC_t value between soy and cross-reacting species was achieved (see subchapter 6.2.3). For this reason, two further combinations of forward and reverse primers and probes of primer/probe system 7 were tested with all species closely related to soy and/or being relevant for human nutrition (see Table 19) at optimized concentrations (FW/RV/PR [µM] 0.2/0.2/0.1, see subchapter 6.2.3.2). Table 20 presents the results obtained with the two combinations of forward and reverse primers and probes of primer/probe system 7, FW 7a/RV 7a/PR 7 and FW 7e/RV 7c/PR 7b. The combination of system 7 FW 7a/RV 7a/PR 7a was also tested with all species closely related to soy and/or being relevant for human nutrition but not included into Table 20. Comparison of mean C_t and ΔRn values was not possible because these specificity tests were only carried out with the concentrations FW/RV/PR [µM] 0.35/0.35/0.1.

Table 20 Results obtained with the primers and probes of system 7 testing all species closely related to soy and/or being relevant for human nutrition (DNA solution: 5 µg/mL) at optimized concentrations of 0.2, 0.2 and 0.1 µM for forward and reverse primer and probe, respectively, and the ΔRn values of the soy positive control (DNA solution: 0.05 µg/mL). The mean C_t value was calculated from two PCR replicates.

primer/probe system *[FW/RV/PR]* c [µM]	7a/7a/7 0.2/0.2/0.1		7e/7c/7b 0.2/0.2/0.1	
	mean C_t value	ΔC_t	mean C_t value	ΔC_t
carob	38.08/-	12.65	40.14/-	13.90
chickpea	39.58/-	14.16	-	-
common bean	-	-	-	-
curcuma	38.73[2]	13.31	37.82/-[2]	11.58
lentil	-	-	41.13/-[1]	14.89
lupin	-	-	-	-
navy bean (white)	-	-	-	-
pea	-	-	-	-
peanut	39.33/-	13.91	-	-
soy positive control (0.05 µg/mL)	25.42		26.24	
	ΔRn			
	177,519		155,000	

- negative PCR result: no increase of the fluorescence signal within 45 cycles
[1] 1 out of 4 replicates positive
[2] DNA concentration < LOD of the QIAxpert instrument

Comparing the two combinations of primers and probes of primer/probe system 7 shown in Table 20, the system FW 7e/RV 7c/PR 7b yielded generally higher ΔC_t values between soy and cross-reacting species. Thus, with this system, less cross-reactivity was observed. Figure 7 shows the corresponding amplification plot for primer/probe system FW 7e/RV 7c/PR 7b.

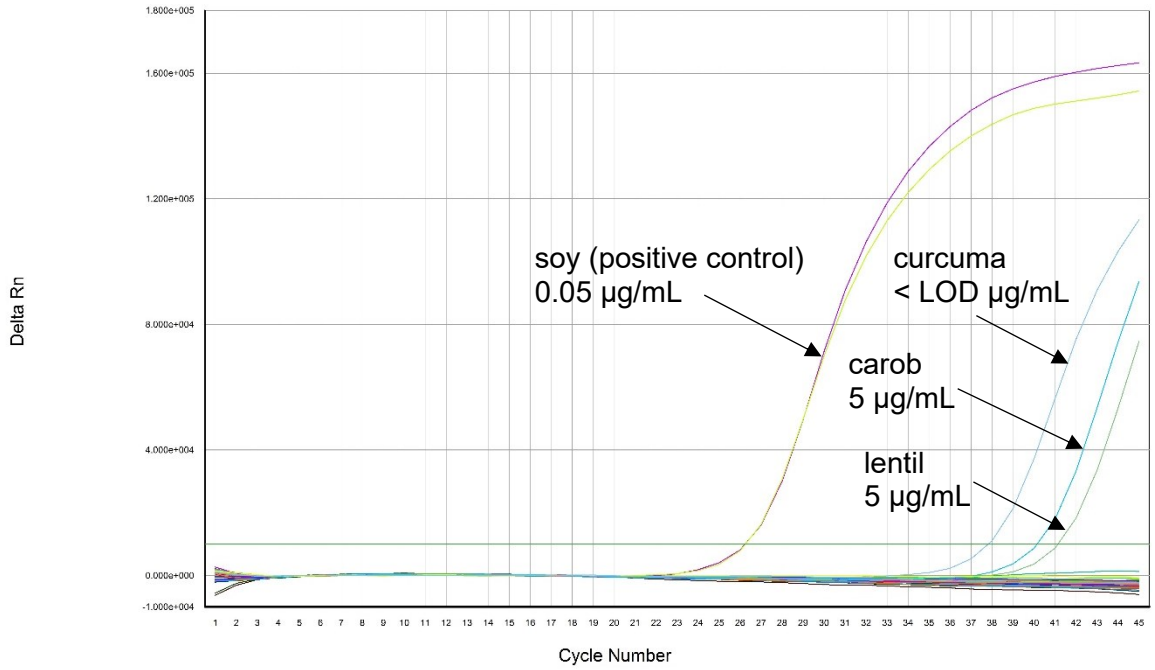

Figure 7 Amplification plot for the specificity test with species closely related to soy and/or being relevant for human nutrition at optimized concentrations of 0.2, 0.2 and 0.1 µM for forward and reverse primer and probe, respectively, with the primer/probe system FW 7e/RV 7c/PR 7b.

From Table 20 and Figure 7 it can be seen that only one out of two replicates of curcuma led to an increase of the fluorescence signal with the primer/probe system FW 7e/RV 7c/PR 7b. The spectrophotometric measurement of the curcuma DNA extract prior to PCR analysis showed that its concentration was below the LOD of the UV/VIS spectrophotometer QIAxpert. The increase of the fluorescence signal of only one out of two replicates of curcuma is explained by either soy contamination of minute amounts (for details on this issue see subchapter 6.2.2.2) or cross-reactivity with curcuma. However, the limited amount of data given does not allow further discussion. The ΔC_t value between the mean C_t value of the soy positive control (26.24) and this one positive result of curcuma (37.82) makes 11.58, which is great enough to correctly identify soy and therefore avoid false-positive results. For carob, only one out of two replicates led to an increase of the fluorescence signal (C_t value: 40.14) and for lentil, only one out of four replicates led to an increase of the fluorescence signal (C_t value: 41.13). Both C_t values were above 39.3, meaning their increase could be neglected in

further consequence of the introduction of a so-called *C_t cut-off value* (see subchapters 5.8.5 and 6.2.5).

For the selection of the most suitable primer/probe system, the obtained mean C_t and ΔRn values, both of the soy positive control (see Table 20), were also taken into consideration. Yet for the final selection, the specificity of the primer/probe system was higher prioritized. Due to this optimal outcome of the specificity tests with species closely related to soy and/or being relevant for human nutrition achieved with primer/probe system FW 7e/RV 7c/PR 7b (see Table 20), no further optimization experiments by primer/probe titration were carried out with this system.

The primer/probe system FW 7e/RV 7c/PR 7b at concentrations of 0.2, 0.2 and 0.1 µM for forward and reverse primer and probe, respectively, was selected as the most suitable system for the specific detection of soy. Information on the selected system FW 7e/RV 7c/PR 7b is summarized in subchapters 5.8.4 and 6.2.4.

6.2.2.2 Specificity tests with species not closely related to soy

With the selected primer/probe system FW 7e/RV 7c/PR 7b at working concentrations of 0.2, 0.2 and 0.1 µM for forward and reverse primer and probe, respectively, follow-up tests investigated a multitude of species not closely related to soy for cross-reactivity, in addition to species closely related to soy and/or being relevant for human nutrition that had been investigated (see subchapters 6.2.2.1 and 5.8.2.2). In total, DNA solutions (5 µg/mL) of 86 species relevant to food production were tested, including species that are often used in spices or as ingredients in processed food. In the majority of cases, no increase of the fluorescence signal was observed, indicating that the primer/probe system did not cross-react with these species. A summary of all tested species including results is given in Table 21.

Table 21 Results of the specificity tests with the primer/probe system FW 7e/RV 7c/PR 7b at optimized concentrations of 0.2, 0.2 and 0.1 µM for forward and reverse primer and probe, respectively, (DNA solutions: 5 µg/mL). The mean C_t value was calculated from two PCR replicates.

common name	scientific name	increase of fluorescence signal	mean C_t value
almond	*Prunus dulcis*	-	-
anise	*Pimpinella anisum*	-[1]	39.84
apple	*Malus*	-	-
atyidae (“shrimp”)	*Caridina*	-	-
beetroot	*Beta vulgaris*	+	36.96
Brazil nut	*Bertholletia excelsa*	-	-
broccoli	*Brassica oleracea*	-	-
buckwheat	*Fagopyrum esculentum*	-	-
caraway	*Carum carvi*	-[1,2]	41.09
cardamon	*Elettaria cardamomum*	-[1]	39.87
carob	*Ceratonia siliqua*	±[1]	40.14/-
carrot	*Daucus carota*	-	-
cashew	*Anacardium occidentale*	-	-
cauliflower	*Brassica oleracea var. botrytis*	-	-
celery	*Apium graveolens*	-	-
cherry	*Prunus avium*	±[1]	40.03/-
chicken/turkey/ beef/pork		-[1]	40.46
chickpea	*Cicer arietinum*	-	-

(continued on next page)

Table 21 (continued)

common name	scientific name	increase of fluorescence signal	mean C_t value
chili	*Capsicum*	-	-
chive	*Allium schoenoprasum*	-[1]	40.05
cilantro	*Coriandrum sativum*	-	-
common bean	*Phaseolus vulgaris*	-	-
cress	*Lepidium sativum*	-	-
cucumber	*Cucumis sativus*	-	-
cumin	*Cuminum cyminum*	-[1]	39.38
curcuma	*Curcuma flaviflora*	±[4]	37.82/-
dill	*Anethum graveolens*	-	-
fennel	*Foeniculum vulgare*	-	-
garlic	*Allium sativum*	-	-
ginger	*Zingiber officinale*	-[1]	40.73
hazelnut	*Corylus avellana*	-	-
hazelnut	*Corylus avellana*		-
herring	*Clupea harengus*	-	-
horseradish	*Armoracia rusticana*	-	-
house cricket	*Acheta domesticus*	+	33.26
laurel	*Laurus nobilis*	±[1]	40.46/-
leek	*Allium porrum*	-	-
lentil	*Lens culinaris*	-	-
lesser mealworm (“buffalo worm”)	*Alphitobius diaperinus*	+	34.87
lovage	*Levisticum officinale*	-	-
lupin	*Lupinus*	-	-
macadamia	*Macadamia ternifolia*	-	-
maize	*Zea mays*	-	-
marjoram	*Origanum majorana*	-	-
yellow mealworm	*Tenebrio molitor*	+	35.83
migratory locust	*Locusta migratoria*	-	-
mustard, black	*Brassica nigra*	-	-
mustard, brown	*Brassica juncea*	±[1]	40.41/-
mustard, white	*Sinapis alba*	-	-
oat	*Avena sativa*	±[1]	41.10/-

(continued on next page)

Table 21 (continued)

common name	scientific name	increase of fluorescence signal	mean C_t value
onion	*Allium cepa*	-	-
oregano	*Origanum vulgare*	-	-
paprika	*Capsicum annum*	-	-
parsley	*Petroselinum crispum*	-	-
parsnip	*Pastinaca sativa*	-	-
pea	*Pisum sativum*	-	-
peach	*Prunus persica*	-[1]	39.71
peanut	*Arachis hypogaea*	-	-
pecan	*Carya illinoinensis*	-	-
pepper	*Piper nigrum*	±[1]	40.66/-[3]
pimento	*Pimenta dioica*	-	-
pistachio nut	*Pistacia vera*	-[1]	40.34
potato	*Solanum tuberosum*	-	-
quince	*Cydonia oblonga*	-	-
radish	*Raphanus sativus var. sativus*	+	38.04
radish	*Raphanus sativus*	±[1]	40.55/-
rape	*Brassica napus*	-	-
rice	*Oryza sativa*	-[1]	39.81
rosemary	*Rosmarinus officinalis*	±[1]	40.74/-
rye	*Secale cereale*	-	-
sage	*Salvia officinalis*	±[1]	39.60/-
sesame	*Sesamum indicum*	±[1]	41.09/-
sorghum	*Sorghum bicolor*	-	-
squash	*Cucurbita*	-	-
squid	*Coleoidea/Dibranchiata*	-	-
strawberry	*Fragaria × ananassa*	+	37.26
summer savory	*Satureja hortensis*	-	-
tarragon	*Artemisia dracunculus*	-	-
thyme	*Thymus vulgaris*	-	-
tomato	*Solanum lycopersicum*	-	-
turnip	*Brassica rapa subsp. rapa subvar. esculenta*	±[1]	39.85/-

(continued on next page)

Table 21 (continued)

common name	scientific name	increase of fluorescence signal	mean C_t value
walnut	*Juglans regia*	-	-
wheat	*Triticum aestivum*	-	-
white cabbage	*Brassica oleracea var. capitata f. alba*	-	-

- negative PCR result: no increase of the fluorescence signal within 45 cycles
± negative/positive PCR result
+ positive PCR result: mean C_t value < 39.3
[1] negative PCR result: mean C_t value ≥ C_t cut-off value of 39.3
[2] DNA concentration: 0.05 µg/mL
[3] mean C_t value calculated from six C_t values
[4] DNA concentration < LOD of the QIAxpert instrument

For a few species (anise, mean C_t value: 39.84; caraway, mean C_t value: 41.09; cardamon, mean C_t value: 39.87; carob, mean C_t value: 40.14/-; cherry, mean C_t value: 40.03/-; chicken/turkey/beef/pork, mean C_t value: 40.46; chive, mean C_t value: 40.05; cumin, mean C_t value: 39.38; ginger, mean C_t value: 40.73; laurel, mean C_t value: 40.46/-; mustard, brown, mean C_t value: 40.41/-; oat, mean C_t value: 41.10/-; peach, mean C_t value: 39.71; pepper, mean C_t value: 40.66/-; pistachio nut, mean C_t value: 40.34; radish, mean C_t value: 40.55/-; rice, mean C_t value: 39.81; rosemary, mean C_t value: 40.74/-; sage, mean C_t value: 39.60/-; sesame, mean C_t value: 41.09/-; turnip, mean C_t value: 39.85/-), there was an increase of the fluorescence signal observed but with C_t values higher than 39.3. Since this C_t value was introduced as C_t cut-off value as part of the determination of the LOD, all C_t values equal to or higher were regarded as negative PCR result (see subchapters 5.8.5 and 6.2.5). After exclusion of these obtained C_t values due to the C_t cut-off value of 39.3, there were seven species left showing positive PCR results with lower C_t values. Apart from curcuma (mean C_t value: 37.82), which has already been discussed (see subchapter 6.2.2.1), these species included lesser mealworm (“buffalo worm”) (mean C_t value: 34.87), strawberry (mean C_t value: 37.26), house cricket (mean C_t value: 33.26), yellow mealworm (mean C_t value: 35.83), radish (*Raphanus sativus var. sativus*, mean C_t value: 38.04) and beetroot (mean C_t value: 36.96).

Figure 8 presents corresponding amplification plots (A lesser mealworm (“buffalo worm”) and other edible insects, B curcuma, C pistachio, D radish and beetroot).

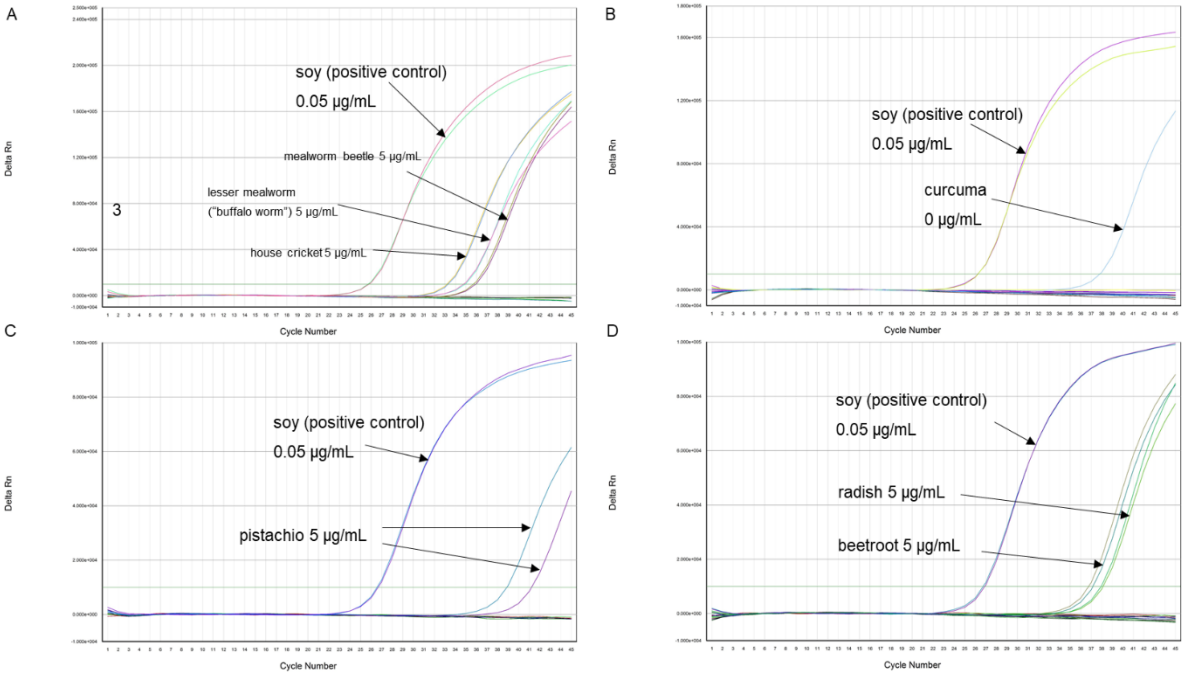

Figure 8 Amplification plots for the specificity tests with A lesser mealworm (“buffalo worm”) and other edible insects, B curcuma, C pistachio nut, D radish and beetroot implying cross-reactivity with the primer/probe system FW 7e/RV 7c/PR 7b.

Since these positive PCR results caused some doubt about cross-reactivity for species not closely related to soy and simultaneously raised questions about potential soy contamination of source material, an agarose gel electrophoresis was carried out to analyze the obtained PCR products. The tested eatable insects (lesser mealworm [“buffalo worm”], house cricket, yellow mealworm) were not included; soy contamination of source material is highly likely since soy meal, for instance, is commonly used for bedding and food of those animals.

Agarose gel electrophoresis

An agarose gel electrophoresis was carried out with the amplicons that were generated with the primer/probe system FW 7e/RV 7c/PR 7b, testing some of the species not closely related to soy that caused some doubt about specificity (applied in PCR assay: DNA solution 5 µg/mL) (see subchapter 5.8.2.2). These primers and the probe were designed to exclusively amplify a certain sequence of 133 bp within the chloroplast genome of soy. In all cases except one, a DNA band with the exact same length as the amplicon to generate with the primer/probe

system FW 7e/RV 7c/PR 7b (133 bp) was visible on the agarose gel, as evident from Figure 9.

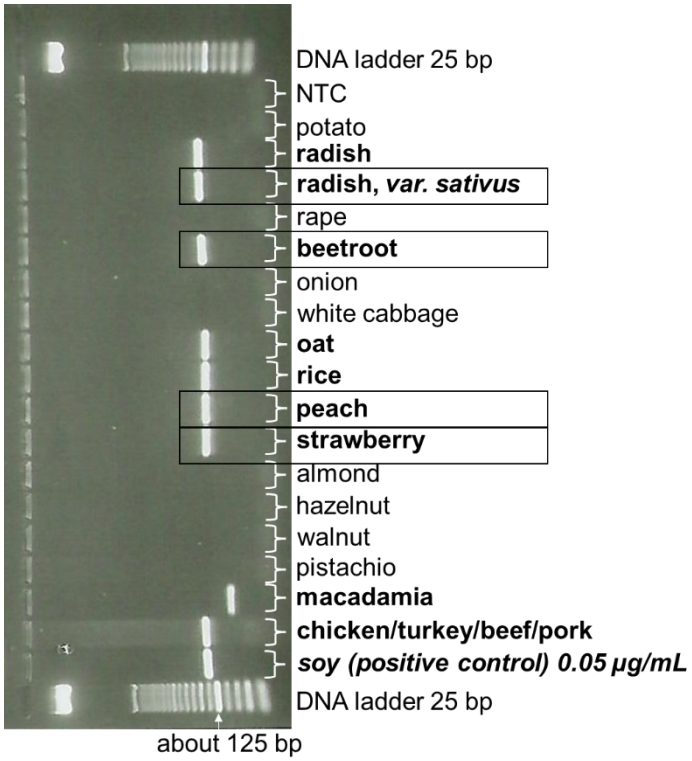

Figure 9 Example of an agarose gel and obtained DNA bands with the amplicons that were generated with the primer/probe system FW 7e/RV 7c/PR 7b, analyzing species not closely related to soy. Positive results of species (visible DNA bands) are highlighted by bold text.

Agarose gel electrophoresis was performed twice, on two different days using two different primer/probe systems (additionally to FW 7e/RV 7c/PR 7b with system 1 [FW 1b/RV 1b/PR 1], solely for verification of identity of the PCR product). The soy amplicon generated with system 1 is 116 bp long. In the experiment using primer/probe system 1, the same DNA solution (5 µg/mL) of radish (*var. sativus*), beetroot, peach and strawberry (highlighted in Figure 9) was used. The generated PCR products were equally applied on an agarose gel (see subchapter 5.8.2.2). The same results were obtained: thick DNA bands of those species suspected to be contaminated with traces of soy (see Figure 10).

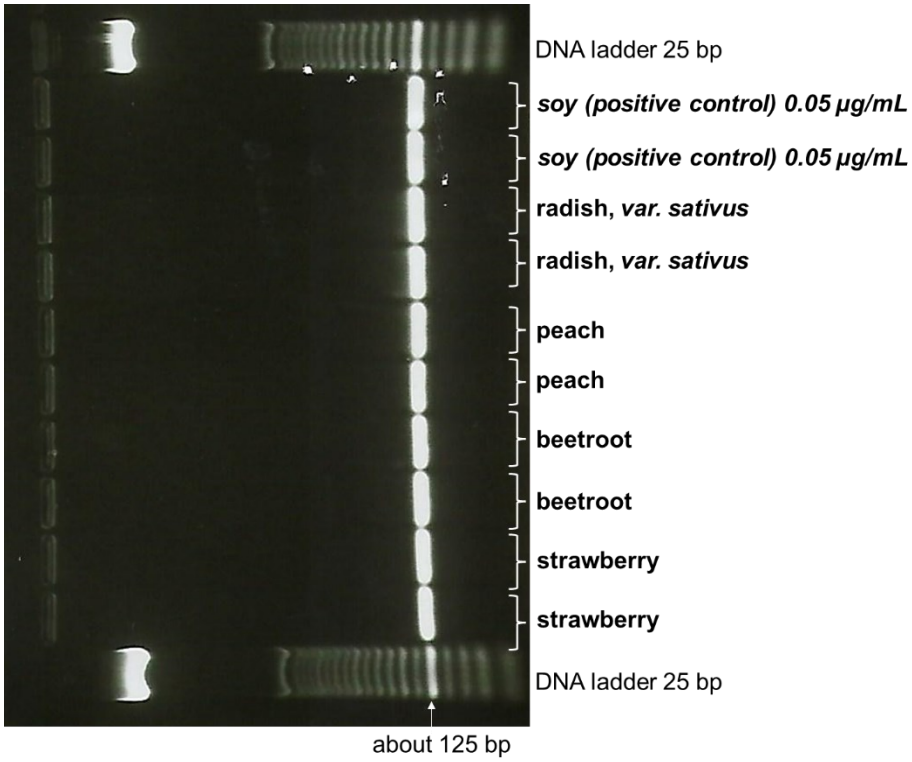

Figure 10 Example of an agarose gel and obtained DNA bands with the amplicons that were generated with the primer/probe system 1 (FW 1b/RV 1b/PR 1), analyzing species not closely related to soy in duplicates. Positive results of species (visible DNA bands) are highlighted by bold text.

Macadamia, the only outliner among the results of the agarose gel electrophoresis using primer/probe system FW 7e/RV 7c/PR 7b (see Figure 9), gave a thick DNA band of about 75 bp. The thickness of the band indicates high abundance of DNA molecules. The DNA molecules are smaller in size than the expected amplicon of soy (about 133 bp) as they migrated more quickly. It is not completely clear for what reason this band was obtained. In real-time PCR, no increase of the fluorescence signal was observed. It may be the result of other contamination with some DNA product. Non-specific priming and extension generating non-specific products or secondary structures, e.g. primer dimers, are possible too. The DNA molecules of a DNA band, which would result from primer dimers, would also migrate more quickly than the DNA molecules of target band because of their small size.

Chicken/turkey/beef/pork gave a DNA band with exactly the same length as the expected amplicon of soy (about 133 bp) (see Figure 9). The smear above and behind the visible DNA band could be attributed to protein contamination of the DNA extract due to insufficient protein removal during DNA extraction. Since this DNA extract was taken from AGES stocks, there is not enough data for further discussion.

Figure 8 also shows the amplification plot for the specificity test with pistachio nut (Figure 8C) of one PCR run; a few days after this run, however, this sample was tested again resulting in

a negative/positive PCR result. The PCR product of the second PCR run, leading to the negative PCR result, was investigated by agarose gel electrophoresis, which explains why the mean C_t value in Table 21 is 40.34 but the gel (Figure 9) lacks a band for pistachio nut.

Generally speaking, consistently positive results in real-time PCR coupled with a visible, thick DNA band in agarose gel electrophoresis with exactly the same length as the amplicon of the analyte supposed to be generated provides fairly compelling evidence of contamination. Nevertheless, an *in silico* test was carried out with the seven plant species (radish, radish [*var. sativus*], beetroot, oat, rice, peach, strawberry) using Basic Local Alignment Search Tool (BLAST, (Nucleotide BLAST: Search nucleotide databases using a nucleotide query, 2021)) in addition. None of the seven species was found to be 100 % identical to any of the binding sites of the three oligonucleotides (forward, reverse primer, probe) at once. The suspicion that some plant samples were contaminated with traces of soy was hereby confirmed.

Soy contamination, in general, posed one of the main issues during the practical work of the underlying thesis. Even though special care was taken to avoid any kind of soy cross-contamination/contamination, e.g. avoidance of working with different DNA extracts in parallel, repeated extraction of DNA from new source material, it sporadically cropped up as a laborious problem. At the AGES, seed/GMO analyses are part of everyday laboratory work, and so are samples that are highly likely to contain soy or are soy(-based). Sources of DNA contamination can be chemical solutions and water but also tubes, pipettes including tips, homogenizers, and even aerosols (Müller & Prange, 2015). Since the practical work of the underlying thesis and these kinds of work could not be fully separated due to shared laboratory spaces, this might be an explanation.

As already argued for curcuma, if there was cross-reactivity for these few species, the ΔC_t value between the mean C_t value of the soy positive control obtained (26.24) and the mean C_t value of the corresponding species is great enough to correctly identify soy and therefore avoid false-positive results. The ΔC_t value between the soy positive control (DNA solution: 0.05 µg/mL, mean C_t value: 26.24) and house cricket (DNA solution: 5 µg/mL), showing the lowest mean C_t value (33.26) of these six species (see Figure 8A), is 7.02. Furthermore, it should be considered that species, such as curcuma or beetroot, are typically present as spices in food containing soy, thus, only in traces, which makes the probability of falsely interpreting results (false positive) even rarer.

Similar research did also cover a multitude of species in the course of investigating cross-reactivity. Ladenburger et al. investigated 69 species by competitive real-time PCR targeting bait8 within mitochondrial DNA of soybean. Thirteen species showed some cross-reactivity (Ladenburger et al., 2018). Mayer et al. tested 72 species by digital droplet PCR targeting the

ndhH gene within chloroplast DNA of soybean; after exclusion of some of the obtained C_t values with the help of an introduced C_t cut-off value, such as it was done in the underlying thesis, they did not find any tested species cross-reacting (Mayer et al., 2019). This data suggests that the designed primer/probe system is suitable for the specific detection of soy.

6.2.3 Optimization by primer/probe titration

The aim of the optimization by primer/probe titration was to find the optimal forward and reverse primer and probe concentrations with which the C_t values for soy was the lowest, constantly applying 5 µL soy DNA solution of a constant concentration. Lower C_t values for the target are beneficial as a lower amount of sample can be used for the same level of signal. At the same time, assay specificity should be optimized so that ΔC_t values between soy and cross-reacting species were the highest.

6.2.3.1 Primer titration

The primer titration was carried out with primers of the primer/probe systems 1 (FW 1b/RV 1b/PR 1b) and 7 (FW 7a/RV 7a/PR 7a). The standard primer concentrations were varied between 0.2-0.9 µM for both forward and reverse primer and both systems. The standard probe concentration of 0.2 µM was kept constant for both systems (see subchapter 5.8.3.1 and Table 6).

Results of the primer titration with the primer/probe systems 1 and 7 testing species closely related to soy and/or being relevant for human nutrition (DNA solution: 5 µg/mL) are given in Table 22 and Table 23, respectively.

Table 22 Results of the primer titration with the primer/probe system 1 (FW 1b/RV 1b/PR 1b) testing species closely related to soy and/or being relevant for human nutrition (DNA solution: 5 µg/mL, soy positive control: 0.05 µg/mL). The mean C_t value was calculated from two PCR replicates.

FW/RV [µM]	0.2/0.5	ΔC_t	0.2/0.9	ΔC_t	0.5/0.2	ΔC_t	0.5/0.9	ΔC_t	0.9/0.2	ΔC_t	**0.9/0.5**	ΔC_t
	mean C_t value											
carob	39.75	13.48	-	-	-	-	38.90/-	12.87	39.41	13.02	**40.16/-**	**13.89**
common bean	-	-	-	-	39.10/-	12.89	-	-	39.61/-	13.22	**-**	**-**
curcuma[1]	37.89	11.62	38.55	12.48	38.50/-	12.29	37.22/-	11.19	40.38/-	14.00	**39.48/-**	**13.21**
soy	26.27		26.08		26.21		**26.03**		26.39		26.27	

- negative PCR result: no increase of the fluorescence signal within 45 cycles
[1] DNA concentration: 4.2 µg/mL

Table 23 Results of the primer titration with the primer/probe system 7 (FW 7a/RV 7a/PR 7a) testing species closely related to soy and/or being relevant for human nutrition (DNA solution: 5 µg/mL, soy positive control: 0.05 µg/mL). The mean C_t value was calculated from two PCR replicates.

FW/RV [µM]	0.2/0.5	ΔC_t	**0.2/0.9**	ΔC_t	0.5/0.2	ΔC_t	0.5/0.9	ΔC_t	0.9/0.2	ΔC_t	0.9/0.5	ΔC_t
	mean C_t value											
carob	40.00/-	14.15	**-**	**-**	-	-	38.83/-	12.83	41.02/-	14.31	-	-
common bean	-	-	**-**	**-**	39.11/-	12.76	41.04/-	15.04	-	-	-	-
curcuma[1]	38.23/-	12.38	**39.88/-**	**14.19**	39.27	12.92	39.33/-	13.33	36.99/-	10.28	38.30/-	11.88
soy	25.85		**25.70**		26.35		26.00		26.71		26.42	

- negative PCR result: no increase of the fluorescence signal within 45 cycles
[1] DNA concentration: 4.2 µg/mL

In these two optimization experiments, cross-reactivity was most important for evaluation of results. Similar cross-reactivity was found for both primer/probe systems at varied concentrations (0.2-0.9 µM for forward and reverse primer).

For system 1, the forward and reverse primer concentrations 0.9/0.5 µM yielded the highest ΔC_t value between soy and cross-reacting species (curcuma: 39.48, ΔC_t value: 13.21). The lowest mean C_t value for soy was 26.03, however, obtained with different concentrations (FW/RV [µM] 0.5/0.9). The mean C_t value for soy with 0.9/0.5 µM was 26.27. For system 7,

the forward and reverse primer concentrations 0.2/0.9 µM yielded the highest ΔC_t value between soy and cross-reacting species (curcuma: 39.88, ΔC_t value: 14.19). The lowest mean C_t value for soy, also obtained with 0.2/0.9 µM, was 25.70. Comparing the two primer/probe systems tested, the higher ΔC_t value between soy and cross-reacting species was obtained with system 7.

6.2.3.2 Primer/probe titration

Since the higher ΔC_t value between soy and cross-reacting species was obtained with primer/probe system 7 with the best performing forward and reverse primer concentrations (see subchapter 6.2.3.1), the primer/probe titration was carried out with this system. The aim was to further optimize primer and probe concentrations to achieve a low C_t value for soy. The standard probe concentration was varied between 0.1-0.2 µM. In connection with the primer titration (see subchapters 5.8.3.1 and 6.2.3.1), concentrations were varied between 0.1-0.5 µM for both forward and reverse primer. These lower equimolar primer concentrations were tested along with varying probe concentrations (see subchapter 5.8.3.2 and Table 7). Results of the titration with the primer/probe system 7 testing the soy positive control (0.05 µg/mL) are given in Table 24.

Table 24 Results of the titration with the primer/probe system 7 (FW 7a/RV 7a/PR 7a) testing the soy positive control (DNA solution: 0.05 µg/mL). The mean C_t value was calculated from two PCR replicates.

FW/RV/PR [µM]	0.1/0.1/0.2	**0.2/0.2/0.1**	**0.2/0.2/0.15**	0.2/0.2/0.2	0.35/0.35/0.1	0.35/0.35/0.15	0.35/0.35/0.2	0.5/0.5/0.1
	mean C_t value							
soy	31.61	**31.04**	**30.04**	30.18	30.72	30.27	29.97	31.00

The forward and reverse primer and probe concentration 0.2/0.2/0.15 µM yielded the lowest mean C_t value for soy (30.04). Comparing the concentrations 0.2/0.2/0.1 µM, 0.35/0.35/0.1 µM and 0.5/0.5/0.1 µM, the mean C_t values for soy at first slightly decreased by equimolarly increasing the forward and reverse primer concentrations (ΔC_t value: 0.32). Further increase, however, led to almost the same value from which the optimization was started (FW/RV/PR [µM] 0.2/0.2/0.1 mean C_t values: 31.04, 0.5/0.5/0.1: 31.00). With increasing probe concentration (0.1-0.2 µM) together with equimolar primer concentrations, the C_t value for soy marginally decreased (highest ΔC_t value between FW/RV/PR [µM] 0.2/0.2/0.1 and 0.2/0.2/0.15: 1.00). The equimolar forward and reverse primer concentrations 0.2/0.2 µM, in general, showed low mean C_t values for soy. Even though the concentrations 0.2/0.2/0.15 µM

yielded the lowest mean C_t value for soy, it was decided to further work with 0.2/0.2/0.1 µM, which yielded a neglectable higher mean C_t value for soy (31.04). The reason for this decision was the potential advantage of primer and/or probe limitation, considering the aim of the underlying thesis to develop a multiplex real-time PCR assay including the real-time PCR assay for soy.

6.2.4 Selected primer/probe system for the singleplex real-time PCR assay

The primer/probe system FW 7e/RV 7c/PR 7b at optimized concentrations of 0.2, 0.2 and 0.1 µM for forward and reverse primer and probe, respectively, was selected as the most suitable system for the specific detection of soy. Table 25 provides information on the system and its final concentrations (working concentrations). The primer/probe system FW 7e/RV 7c/PR 7b generates an amplicon of 133 bp. Figure 11 shows the corresponding chloroplast-DNA alignment of soybean and closely related and/or being relevant for human nutrition species with highlighted regions for both primers and the probe.

Table 25 Information on the selected primer/probe system FW 7e/RV 7c/PR 7b and the working concentrations.

primer/probe	sequence 5'→3'	T_m [°C]	region*	length [bp]	amplicon length [bp]	working concentrations [µM]
So60_ FW 7e	TTC CAT CCC TGC AAT TCA TTT	58.0	68939-68959	21	133	0.2
So60_RV 7c	CAA AAG ACC CGG TAG ATT TTG TAC	57.0	69048-69071	24		0.2
So60_PR 7b	**Cy5**-CCT CTA CTA TGG ATA GAG TA-**MGB**	65.0	68964-68983	20		0.1

*within the chloroplast genome of soy

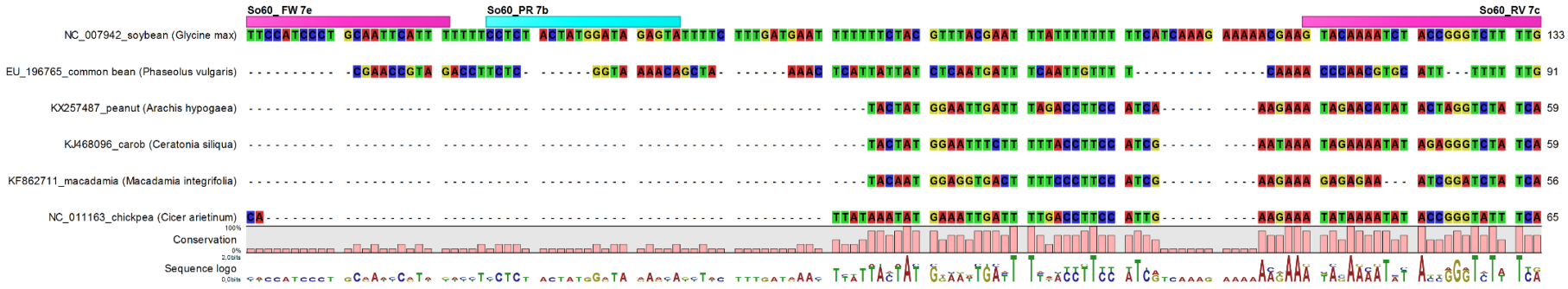

Figure 11 Chloroplast-DNA alignment of soybean and closely related and/or relevant species with highlighted regions for forward primer (So60_ FW 7e), reverse primer (So60_RV 7c) and probe (So60_PR 7b).

6.2.5 Determination of limit of detection (LOD) and Ct cut-off value

The LOD of the singleplex assay was determined by analyzing DNA solutions corresponding to target DNA concentrations of 0.06-0.004 mg/kg (PCR run 1 and 2) and 0.03-0.004 mg/kg (PCR run 3) in twelve replicates each. The DNA solutions were prepared from DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) and serially diluted with non-target DNA (herring sperm DNA solution, 100 µg DNA/mL) (see subchapter 5.8.5).

The determination of the LOD was based on the lowest concentration that led to an increase of the fluorescence signal within 45 cycles (i.e. positive result). The C_t cut-off value was calculated from this concentration, too. First, the mean C_t value was calculated from the C_t values obtained at the lowest concentration at which all (twelve out of twelve) replicate measurements were positive. Second, the standard deviation was also calculated and its value was added two times to the mean C_t value, yielding the C_t cut-off value. All C_t values obtained at the lowest concentration were checked to see if each of them was below the C_t cut-off value. If this was the case, the corresponding concentration was determined as the LOD. The LOD of the singleplex real-time PCR assay was, therefore, defined as the lowest concentration that led to an increase of the fluorescence signal below the C_t cut-off value in all replicate measurements. Following this strategy, three individual C_t cut-off values from the three PCR runs 1-3 were obtained (PCR run 1: 38.49, PCR run 2: 39.58, PCR run 3: 39.80). The results including the three individual C_t cut-off values are shown in Table 26.

Table 26 Results obtained as part of the determination of the LOD and C_t cut-off value of the singleplex real-time PCR assay. PCR run 1-3 were carried out independently on three different days. The mean C_t value was calculated from twelve PCR replicates.

PCR run 1				PCR run 2				PCR run 3			
c [mg/kg]	mean C_t value	S	C_t cut-off value	c [mg/kg]	mean C_t value	S	C_t cut-off value	c [mg/kg]	mean C_t value	S	C_t cut-off value
0.06	35.94	0.35		0.06	36.15	0.42		0.06	n.a.	n.a.	
0.03	36.75	0.54		0.03	37.17	0.51		0.03	37.13	0.59	
0.016	37.56	0.46	**38.49**	**0.016**	38.11	0.73	**39.58**	**0.016**	38.16	0.82	**39.80**
0.008	39.20/-	0.94		0.008	38.86/-	0.98		0.008	38.57/-	0.88	
0.004	38.82/-	0.88		0.004	39.75/-	0.92		0.004	39.30/-	0.70	

n.a. not analyzed
- negative PCR result: no increase of the fluorescence signal within 45 cycles

For the singleplex real-time PCR assay for the specific detection of the allergenic food soy, a LOD of 0.016 mg soy/kg food was achieved. A C_t value of 39.3 was calculated as the C_t cut-off value, which is the mean value of the three calculated individual C_t cut-off values from the three PCR runs 1-3.

In all cases, at concentrations lower than 0.016 mg soy/kg (0.008 and 0.004 mg/kg), some of the replicates did not show an increase of the fluorescence signal, i.e. the concentration was below the detection limit. In PCR run 1, the concentrations 0.008 and 0.004 mg/kg only led to an increase of the fluorescence signal in ten out of twelve and nine out of twelve replicates, respectively. In PCR run 2, the concentrations 0.008 and 0.004 mg/kg only led to an increase of the fluorescence signal in nine out of twelve and eight out of twelve replicates, respectively. In PCR run 3, the concentrations 0.008 and 0.004 mg/kg only led to an increase of the fluorescence signal in eleven out of twelve and ten out of twelve replicates, respectively.

Since the aim of the master's thesis was to develop a qualitative assay, the amplification efficiency was not calculated. Yet, linearity between the mean C_t values of the different concentrations (0.06-0.004 mg soy/kg) was assessed. In PCR run 1 (see Table 26), good linearity was achieved between the mean C_t values of the concentration 0.06 mg soy/kg and the LOD (0.016 mg/kg) (R^2 = 0.9992). There was also a linear relationship between the mean C_t values of the concentrations 0.06-0.004 mg/kg (R^2 = 0.8936). In PCR run 2, very good linearity was achieved between the mean C_t values of the concentration 0.06 mg/kg and the LOD (0.016 mg/kg) (R^2 = 1). There was also a linear relationship between the mean C_t values of the concentrations 0.06-0.004 mg/kg (R^2 of 0.9963). In the PCR run 3, there was a linear

relationship between the mean C_t values of the concentrations 0.03-0.004 mg/kg (R^2 = 0.9658).

The robustness was examined by running the same PCR assay (PCR run 3 in Table 26) on a duplicate thermocycler of the model 7500 Real Time PCR System of Applied Biosystems®. There is barely any difference between the mean C_t values of those three runs (PCR run 1-3), indicating that the singleplex real-time PCR assay in question is robust to switching between Applied Biosystems® 7500 Real Time PCR System instruments.

Since the LOD of the singleplex real-time PCR assay was only investigated with DNA solutions prepared from DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) and serially diluted with non-target DNA, direct comparison with other real-time PCR assays investigating detection limits in model food (see Table 2) may not be advisable. In some cases, direct comparison with other soy PCR assays was not possible at all, as the LOD was given in copies per microliters or grams (Mayer et al., 2019; Costa et al., 2017). Ladenburger et al. recently developed a competitive real-time PCR assay for the detection and quantification of soy and peanut using a fluorescent hydrolysis probe. By targeting mitochondrial DNA, soy could be detected in different food matrices (sausage, rice cookie, hollandaise sauce powder, skimmed milk powder) with a LOD of 1 mg/kg (Ladenburger et al., 2018). Differences in the LOD between TaqMan real-time PCR assays targeting soy-specific genes may be due to differences in measuring DNA concentration (Costa et al., 2017). What is more, reference materials used for validation, such as SureFood® QUANTARD Allergen 40, are not certified, which complicates comparison, too (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

The LOD determined for the singleplex real-time PCR assay, 0.016 mg soy/kg food, strongly confirms other researcher's statements of the means of lowering detection limits by taking advantage of the high number of DNA copies chloroplast DNA offers (Bauer et al., 2011; Mayer et al., 2019). For instance, Zhang et al. developed a real-time PCR assay targeting a soy-specific gene encoding for Gly m Bd 28K. They obtained a LOD of 50 mg soy/kg in different plant food powders as matrix (Zhang et al., 2018). Other real-time PCR assays, targeting the soy lectin gene, had LODs of 25 mg soy/kg in pork meat (Soares et al., 2014) and 5 mg soy/kg in rice cookie as well as 32 mg soy/kg in boiled sausage as matrix (Köppel et al., 2012).

The determined LOD of the developed singleplex real-time PCR assay is with 0.016 mg soy/kg food very low. The requirement of adequate sensitivity for such detection methods may be becoming even more important in light of the recently updated VITAL 3.0 reference doses for allergens in foods. It should be noted, though, that by targeting sequences in the chloroplast

but also in the mitochondrial genome, a calibrator is required in case of quantification. This is because the number of copies significantly differs among different parts of plants (Caldwell, 2017; Alberts et al., 2002; Mayer et al., 2019). Mayer et al., however, did research on the degree of deviation associated with varying copy number of chloroplast DNA in the soy plant by repeatedly analyzing undiluted native soy material. The experiments were carried out with a primer/probe system for the specific detection of soy by targeting chloroplast DNA in digital droplet PCR. They reported the deviation did not exceed 30 % (Mayer et al., 2019).

The LODs in commercial food spiked with soy were not tested in this work. With the intention of working at conditions similar to those when analyzing real food with its matrix constituents, high amounts of non-target DNA as background were applied in experiments. This procedure in the course of assay development is also strongly advised by a very recent paper discussing gaps, needs and recommendations for assay development and validation (Holzhauser et al., 2020).

6.3 Development and validation of the triplex real-time PCR assay for the simultaneous detection of soy, celery and white mustard

For the purpose of investigating the combinability of all primers and probes of the PCR assays, DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was serially diluted (1:2) with water to prepare DNA solutions corresponding to concentrations of 1-0.016 mg target/kg (target = soy/celery/white mustard). The DNA solutions corresponding to concentrations of 1-0.125 mg target/kg were analyzed (see subchapter 5.9). As evident from both Figure 12 and Figure 13, the singleplex (soy detection, Cy5) and the duplex (celery detection, FAM; white mustard detection, VIC) PCR assay were combinable and suitable for multiplexing in practice.

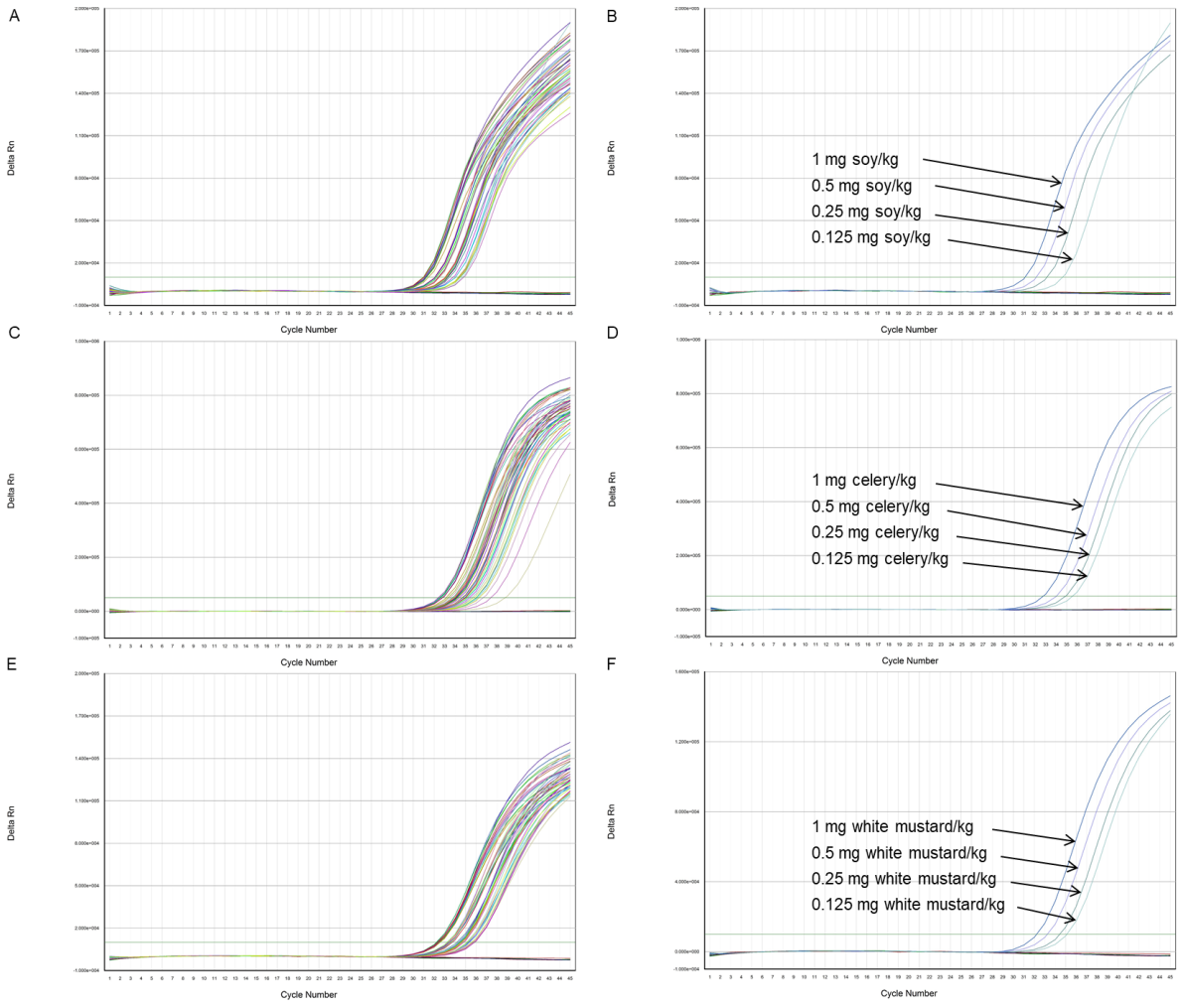

Figure 12 Amplification plots for the initial experiments with the triplex real-time PCR assay. In A, C and E, the amplification curves illustrate the performance of a single assay as part of the triplex assay with all twelve replicates of the concentrations 1, 0.5, 0.25 and 0.125 mg target/kg (diluent: water). In B, D and F, the same is illustrated but with only one replicate of each concentration. A and B show the soy PCR assay (Cy5), C and D show the celery PCR assay (FAM), E and F show the white mustard PCR assay (VIC).

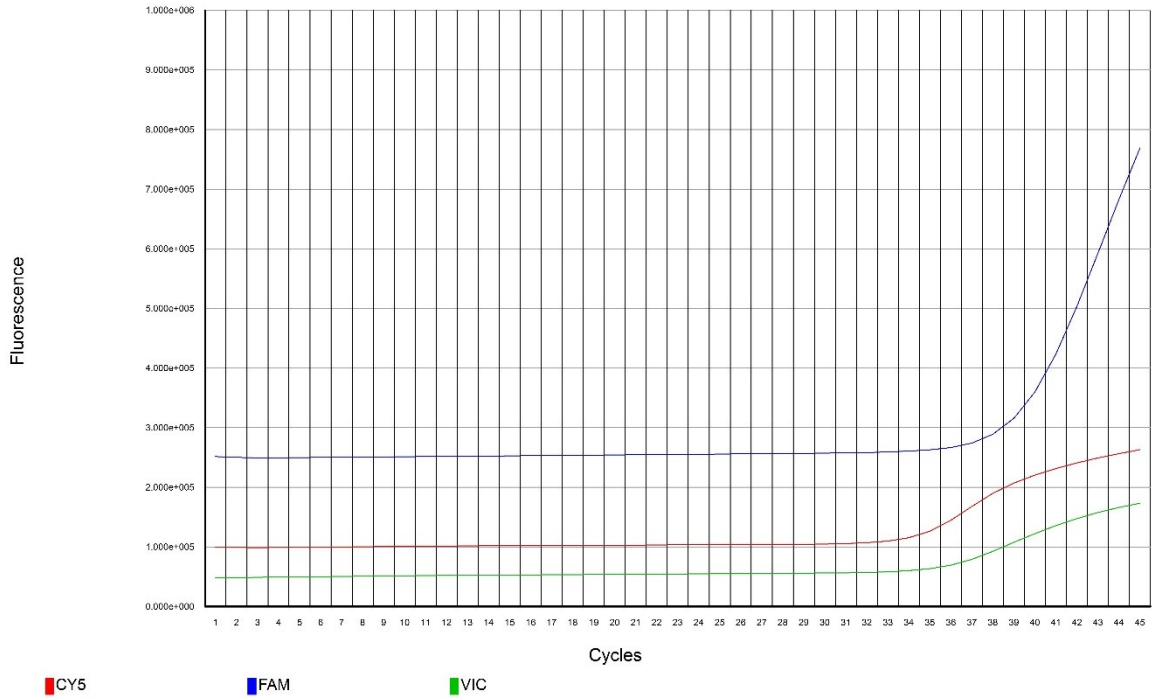

Figure 13 Component view for the triplex real-time PCR assay for the detection of soy (Cy5, red), celery (FAM, blue) and white mustard (VIC, green). The DNA solution corresponds to 0.125 mg target (soy, celery, white mustard)/kg (diluent: water).

Figure 12 includes amplification plots for the initial experiments with the triplex real-time PCR assay in which analyzed DNA solutions were prepared with water. In the first amplification plot, the amplification curves illustrate the performance of a single assay as part of the triplex assay with all twelve replicates (1-0.125 mg target/kg). In the second, the same is illustrated but with only one replicate of each concentration. In the soy PCR assay (Figure 12A, B), the celery PCR assay (Figure 12C, D) as well as the white mustard PCR assay (Figure 12E, F), there was a linear relationship between the mean C_t values of the concentrations 1, 0.5, 0.25 and 0.125 mg target/kg. Including the C_t values of all twelve replicates, a R^2 of 0.999, 0.9863 and 1 for soy, celery and white mustard, respectively, was achieved.

Since the aim of the master's thesis was to develop a qualitative assay, the amplification efficiency was not calculated. With the help of Figure 13, however, the increase of the fluorescence signal of Cy5 (soy), FAM (celery) and VIC (white mustard) could be directly compared. The comparison demonstrated that the increase of the fluorescence signal of Cy5 (soy) and of VIC (white mustard) were not as high as that of FAM (celery). The amplification of soy and white mustard seemed to be less efficient. In an attempt to find the optimal forward and reverse primer and probe concentrations, primer/probe titration of the soy and the white mustard PCR assay were carried out.

6.3.1 Optimization by primer/probe titration

DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) was serially diluted (1:2) with water as well as high amounts of target DNA to prepare DNA solutions corresponding to concentrations of 1-0.016 mg target/kg. In case of soy, celery and white mustard, high amounts of target DNA of celery as well as white mustard, and soy as well as white mustard, and soy as well as celery, respectively, were used. The DNA solutions were analyzed (see subchapter 5.9.1).

6.3.1.1 Optimization of white mustard PCR assay

A forward and reverse primer concentration of 0.8 µM was applied, which was tested with a continuously increasing probe concentration (0.25, 0.3, 0.35, 0.4 µM) to achieve efficient amplification. The different combinations for white mustard were investigated in presence of high amounts of celery (20 µg DNA/mL, equal to 100 ng DNA per tube/well) as background DNA (see subchapters 5.9.1, 5.9.1.1 and Table 11). The primer and probe concentrations of the soy and celery PCR assay were kept constant (FW/RV/PR [µM] soy: 0.1/0.1/0.05, see subchapter 6.3.1.2; celery: 0.4/0.4/0.1, see Table 8). The amplification curves obtained with the tested combinations (#1-4) are shown in Figure 14, corresponding ΔRn values are shown in Table 27.

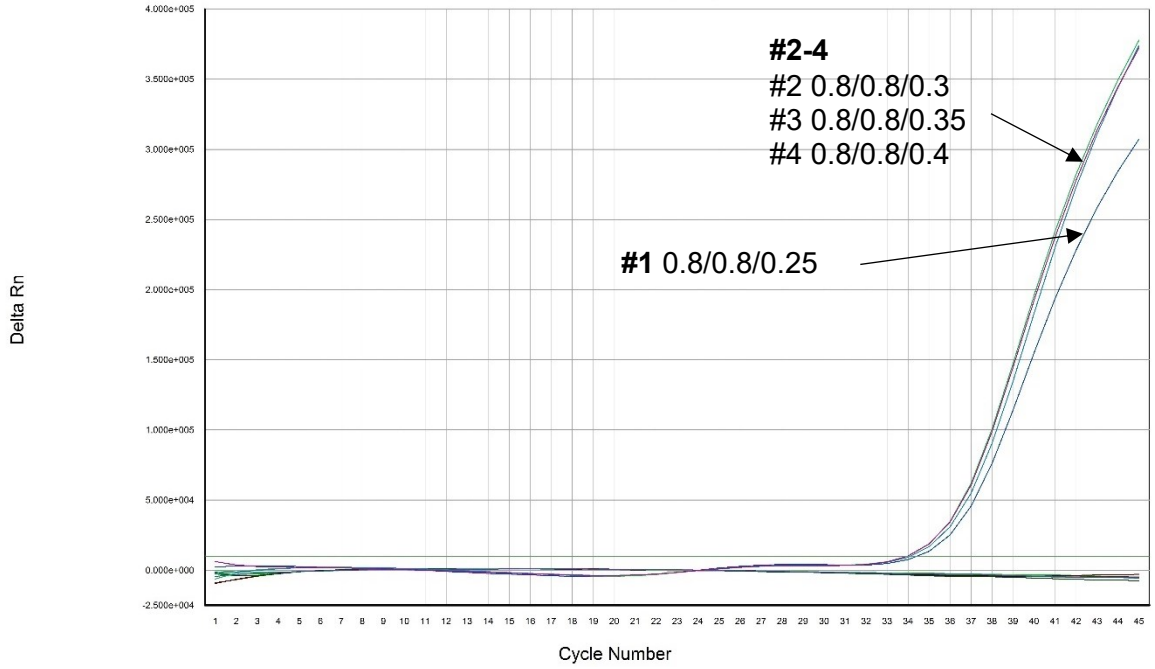

Figure 14 Amplification plot showing representative amplification curves (0.125 mg white mustard/kg) obtained with the tested combinations (#1-4, FW/RV/PR [µM]) of the white mustard PCR assay in presence of high amounts of celery (20 µg DNA/mL) as background DNA.

Table 27 ∆Rn values obtained with the tested combinations (#1-4, FW/RV/PR [µM]) of the white mustard PCR assay in presence of high amounts of celery (20 µg DNA/mL) as background DNA. The concentration was 0.125 mg white mustard/kg.

primer/probe system white mustard	
# optimization attempt [*FW/RV/PR* µM]	∆Rn
#1 0.8/0.8/0.25	306,269
#2 0.8/0.8/0.3	373,036
#3 0.8/0.8/0.35	373,036
#4 0.8/0.8/0.4	373,036

With the second combination (#2 FW/RV/PR [µM] 0.8/0.8/0.3, see Table 27), a high ∆Rn value was obtained (∆Rn value: 373,036), compared to the first combination 0.8/0.8/0.25 µM. The concentrations 0.8, 0.8 and 0.3 µM for forward and reverse primer and probe, respectively, were selected since higher probe concentrations did not result in higher ∆Rn values.

6.3.1.2 Optimization of soy PCR assay

Preliminary experiments with the newly developed singleplex PCR assay for soy combined with the celery/white mustard duplex PCR assay showed that the target soy DNA (Cy5) may amplify more efficiently than the target white mustard (VIC) (see Figure 13). In order to achieve more efficient amplification with the white mustard PCR assay, the primer as well as the probe concentrations for soy were lowered. Three different combinations of lower concentrations (FW/RV/PR [µM] 0.1/0.1/0.05, 0.05/0.05/0.025, 0.025/0.025/0.0125) were investigated in presence of high amounts of soy (20 µg DNA/mL, equal to 100 ng DNA per tube/well) as background DNA (see subchapters 5.9.1, 5.9.1.2 and Table 12). It should be tested if soy DNA does amplify at all with lower primer and probe concentrations. If so, the lowest primer and probe concentrations at which sufficient amplification is given should be found. The primer and probe concentrations of the celery and white mustard PCR assay were kept constant (see Table 8). The amplification curves obtained with the tested combinations (#1-3) are shown in Figure 15, corresponding ∆Rn values are shown in Table 28.

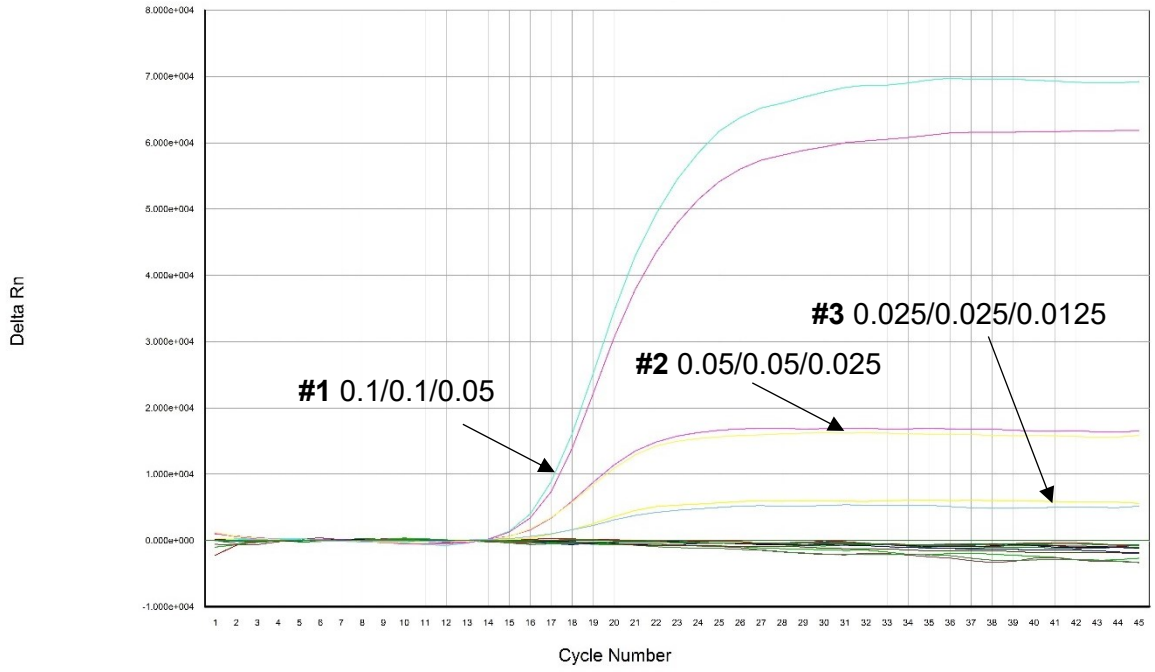

Figure 15 Amplification plot showing representative amplification curves (0.125 mg soy/kg) obtained with the tested combinations (#1-3, FW/RV/PR [µM]) of the soy PCR assay in presence of high amounts of soy (20 µg DNA/mL) as background DNA.

Table 28 ΔRn values obtained with the tested combinations (#1-3, FW/RV/PR [µM]) of the soy PCR assay in presence of high amounts of soy (20 µg DNA/mL) as background DNA. The concentration was 0.125 mg soy/kg.

primer/probe system soy	
# optimization attempt [*FW/RV/PR* µM]	ΔRn
#1 0.1/0.1/0.05	65,453
#2 0.05/0.05/0.025	16,511
#3 0.025/0.025/0.0125	5,363

With the first combination (#1 FW/RV/PR [µM] 0.1/0.1/0.05, see Table 28), the highest ΔRn value was obtained (ΔRn value: 65,453). Sufficient amplification was still achieved, in contrast to the other two combinations (#2, #3).

The concentrations 0.1, 0.1 and 0.05 µM for forward and reverse primer and probe, respectively, were also tested applying high amounts of DNA of the other two targets of the triplex assay, celery and white mustard, (20 µg celery DNA/mL, 5 µg white mustard DNA/mL) as background DNA. The concentration of 5 µg white mustard DNA/mL was chosen due to contamination problems with soy DNA in the DNA extract of white mustard. This concentration

(5 µg/mL) was 100-fold higher than the soy positive control (0.05 µg/mL) applied in the singleplex assay and, therefore, still considered a sufficient amount of background DNA.

In another attempt to further optimize the soy PCR assay, both primer and probe concentrations were varied (FW/RV/PR [µM] 0.1/0.1/0.1, 0.1/0.1/0.15, 0.15/0.15/0.1, 0.2/0.2/0.1). Four different combinations were investigated in presence of high amounts of celery (20 µg DNA/mL) as background DNA (see subchapters 5.9.1, 5.9.1.2 and Table 13). For the white mustard PCR assay, optimized concentrations were applied (see subchapter 6.3.1.1). The primer and probe concentrations of the celery PCR assay were kept constant (see Table 8). The corresponding ∆Rn values obtained with the tested combinations are shown in Table 29.

Table 29 ∆Rn values obtained with the tested combinations (#1-4, FW/RV/PR [µM]) of the soy PCR assay in presence of high amounts of celery (20 µg DNA/mL) as background DNA. The concentration was 0.125 mg soy/kg.

primer/probe system soy	
# optimization attempt [*FW/RV/PR* µM]	∆Rn
#1 0.1/0.1/0.1	85,511
#2 0.1/0.1/0.15	122,779
#3 0.15/0.15/0.1	104,653
#4 0.2/0.2/0.1	88,565

With the second combination (#2 FW/RV/PR [µM] 0.1/0.1/0.15, see Table 29*),* the highest ∆Rn value was obtained (∆Rn value: 122,779).

The aim of additional experiments was to test the effects of further increased concentrations of the probe for soy on amplification. The probe concentration for soy was varied (0.2, 0.3, 0.25, 0.35 µM). The forward and reverse primer concentration of 0.1 µM was kept constant (see subchapters 5.9.1, 5.9.1.2 and Table 14). For the white mustard PCR assay, optimized concentrations were applied (see subchapter 6.3.1.1). The primer and probe concentrations of the celery PCR assay were kept constant (see Table 8). The amplification curves obtained with the tested combinations (#1-4) varying in probe concentration are shown in Figure 16, mean C_t values and the ∆Rn values are given in Table 30.

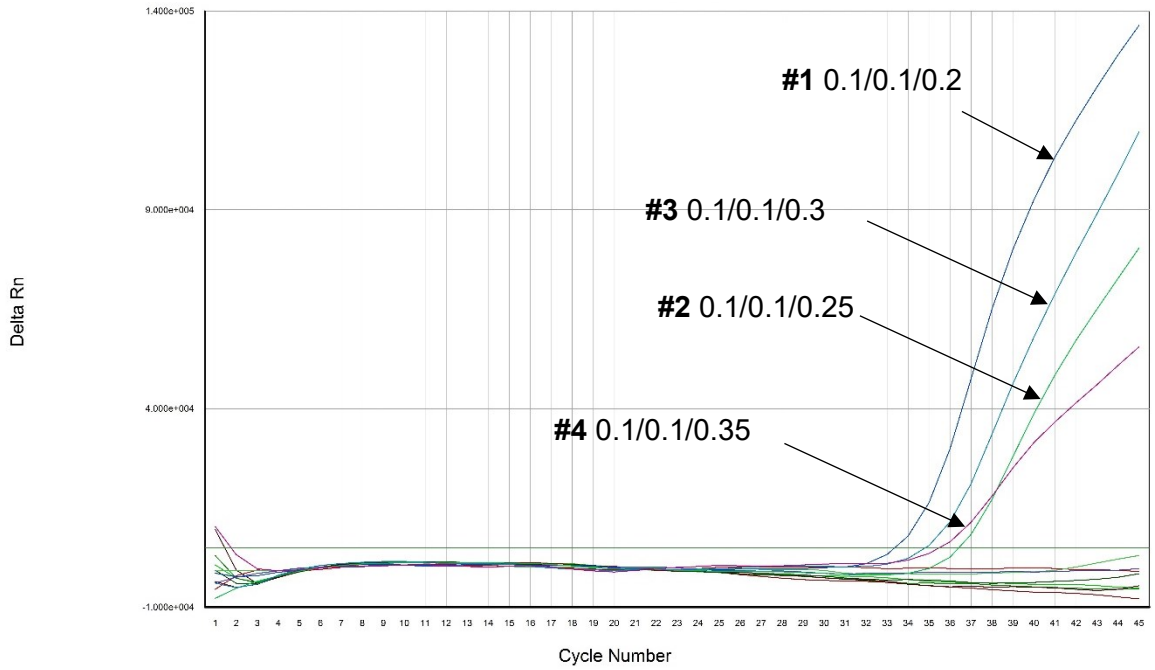

Figure 16 Amplification plot showing representative amplification curves (0.125 mg soy/kg) obtained with the tested combinations (#1-4, FW/RV/PR [µM]) of the soy PCR assay in presence of high amounts of celery (20 µg DNA/mL) as background DNA.

Table 30 ΔRn values obtained with the tested combinations (#1-4, FW/RV/PR [µM]) of the soy PCR assay in presence of high amounts of celery (20 µg DNA/mL) as background DNA. The concentration was 0.125 mg soy/kg and the mean C_t value was calculated from two PCR replicates.

primer/probe system soy		
# optimization attempt [*FW/RV/PR* µM]	ΔRn	mean C_t value
#1 0.1/0.1/0.2	135,921	33.38
#2 0.1/0.1/0.25	80,408	36.42
#3 0.1/0.1/0.3	80,408	34.84
#4 0.1/0.1/0.35	55,030	35.50

With the first combination (#1 FW/RV/PR [µM] 0.1/0.1/0.2, see Table 30*)*, the highest ΔRn value (ΔRn value: 135,921) and the lowest mean C_t value were obtained. The concentrations 0.1, 0.1 and 0.2 µM for forward and reverse primer and probe, respectively, also resulted in a higher ΔRn value than the combination with 0.1, 0.1 and 0.15 µM, respectively (ΔRn value: 122,779, see Table 29).

The concentrations 0.1, 0.1 and 0.2 µM for forward and reverse primer and probe, respectively, were also tested applying high amounts of DNA of the other two targets of the triplex assay, soy and white mustard, (20 µg soy DNA/mL, 20 µg white mustard DNA/mL) as background DNA, as well as water as diluent (see subchapters 5.9.1 and 5.9.1.2). Full inhibition of the amplification of soy DNA in presence of high amounts of white mustard (20 µg DNA/mL) as background DNA was repeatedly observed in experiments with optimized concentrations (soy: FW/RV/PR [µM] 0.1/0.1/0.2, celery: see Table 8, white mustard: FW/RV/PR [µM] 0.8/0.8/0.3). A representative amplification plot is given in Figure 17.

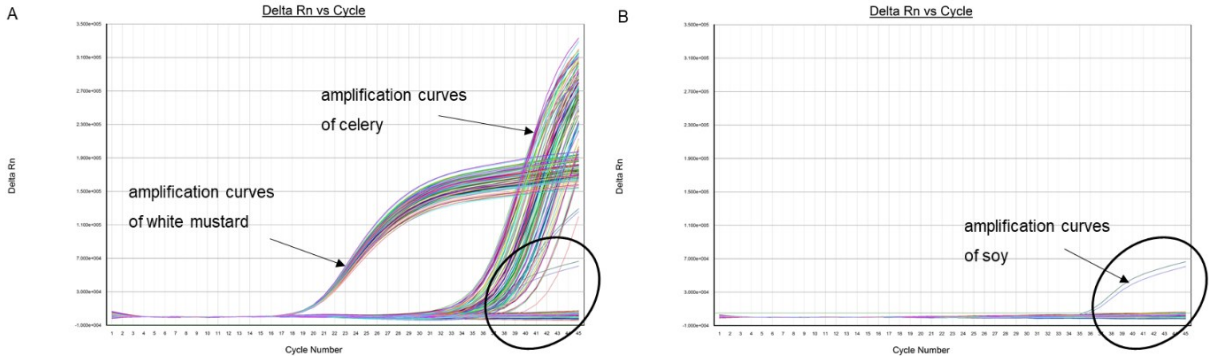

Figure 17 Representative amplification plots for an experiment with optimized primer and probe concentrations resulting in full inhibition of the amplification of soy in presence of high amounts of white mustard (20 µg DNA/mL) as background DNA. A shows the amplification curves obtained with the soy, celery and white mustard PCR assay with all twelve replicates of the concentrations 1, 0.5, 0.25, 0.125, 0.06, 0.03 and 0.016 mg target/kg of the triplex real-time PCR assay. B only shows the amplification curves obtained with the soy PCR assay.

In Figure 17, the two amplification plots originate from the same PCR run but in B only the amplification curves obtained with the soy PCR assay are shown. As evident from this plot (Figure 17B), the amplification reaction of soy was fully inhibited in presence of high amounts of white mustard. The celery PCR assay did not appear to be negatively affected (Figure 17A).

Due to this result, the conclusion was drawn that the initial forward and reverse primer and probe concentrations of the soy and white mustard PCR assay led to the best results after all (0.2 µM forward primer, 0.2 µM reverse primer, 0.1 µM probe for soy; 0.4 µM forward primer, 0.4 µM reverse primer, 0.1 µM probe for white mustard; see Table 8). With the initial concentrations of primers and probes for soy, celery and white mustard, the triplex assay performed best at each of the tested conditions applying high amounts of target DNA (5 or 20 µg soy/celery/white mustard DNA/mL) as background. By applying these different kinds of background DNA, the common scenario in a multiplex reaction that targets are not present in the same amount could be simulated. Using the initial concentrations of primers and probes, the triplex PCR assay was suitable for multiplexing in practice and a sufficiently high increase

of the fluorescence signal for qualitative analysis was obtained. Requirements for analytical methods for the detection of allergenic food were also met (high sensitivity/low LOD).

The selected initial concentrations in the triplex PCR assay are 0.2 µM forward primer, 0.2 µM reverse primer and 0.1 µM probe for soy, 0.4 µM forward primer, 0.4 µM reverse primer and 0.1 µM probe for celery and 0.4 µM forward primer, 0.4 µM reverse primer and 0.1 µM probe for white mustard. These selected initial concentrations and further information on the triplex real-time PCR assay are summarized in Table 8 and subchapter 5.9.2, respectively.

6.3.2 Inhibition control (IC)

Experiments were performed to investigate potential inhibition of the amplification reaction due to DNA of one of the three targets, soy, celery and white mustard, being present in very high amounts (100 µg target DNA/mL) as background. DNA solutions corresponding to concentrations of 0.25-0.016 mg target/kg, which were prepared from DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) and serially diluted with very high amounts of target DNA (soy/celery/white mustard DNA solution, 100 µg target DNA/mL), were spiked with a soy/celery/white mustard DNA solution corresponding to a concentration of 0.25 mg target/kg (see subchapter 5.9.3). A positive/spike control (containing master mix, primers, probes [see subchapter 5.9.2], water) was spiked, too, with the same soy/celery/white mustard DNA solution corresponding to a concentration of 0.25 mg target/kg. This positive/spike control was analyzed together with the equally spiked prepared solutions (0.25-0.016 mg target/kg), to compare and evaluate the difference between the positive/spike control and the spiked DNA solutions. A ΔC_t value of 1 or less was considered a neglectable inhibitory effect whereas above 1 C_t value was considered inhibitory. An inhibitory effect on the amplification reactions was observed, particularly in the celery and white mustard PCR assay in the case of soy being present in a concentration of 100 µg DNA/mL as background. Figure 18 shows the amplification curves obtained in such an inhibition experiment and Table 31 displays the results obtained in experiments (PCR run 1-3) with IC in presence of very high amounts (100 µg DNA/mL) of the three targets.

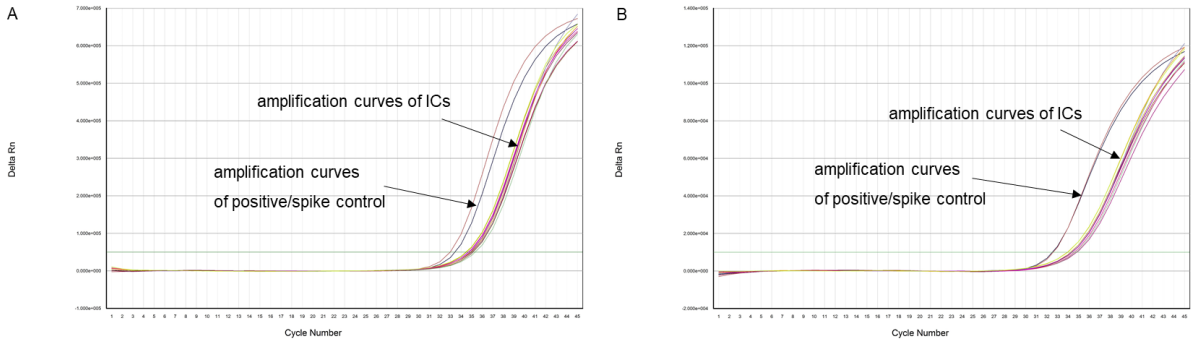

Figure 18 Representative amplification plots for an experiment with ICs in presence of very high amounts of soy (100 µg DNA/mL) as background DNA. A shows the amplification curves obtained with the celery PCR assay. B shows the amplification curves obtained with the white mustard PCR assay. Both A and B show amplification curves obtained with the positive/spike control (spike level: 0.25 mg target/kg) and ICs (0.25, 0.125, 0.06, 0.03 and 0.016 mg target/kg, spike level: 0.25 mg target/kg).

Table 31 Results of experiments with ICs in presence of very high amounts of target DNA (100 µg target DNA/mL) of soy, celery or white mustard (target DNA for soy: celery, white mustard; target DNA for celery: soy, white mustard; target DNA for white mustard: soy, celery). The spike level was 0.25 mg target/kg and the mean C_t value was calculated from two PCR replicates.

PCR run 1				
background DNA soy 100 µg DNA/mL	celery (FAM)		white mustard (VIC)	
concentration [mg target/kg]	positive/spike control mean C_t value *(spike level 0.25)*	IC mean C_t value (ΔC_t)	positive/spike control mean C_t value *(spike level 0.25)*	IC mean C_t value (ΔC_t)
0.25	33.20	34.59 (+ 1.38)	32.55	34.16 (+ 1.61)
0.125		34.82 (+ 1.62)		34.26 (+ 1.71)
0.06		35.20 (+ 1.99)		34.52 (+ 1.97)
0.03		34.77 (+ 1.57)		34.53 (+ 1.98)
0.016		35.01 (+ 1.81)		34.36 (+ 1.81)
PCR run 2				
background DNA celery 100 µg DNA/mL	soy (Cy5)		white mustard (VIC)	
concentration [mg target/kg]	positive/spike control mean C_t value *(spike level 0.25)*	IC mean C_t value (ΔC_t)	positive/spike control mean C_t value *(spike level 0.25)*	IC mean C_t value (ΔC_t)
0.25	30.70	31.93 (+ 1.24)	32.49	32.13 (- 0.36)
0.125		31.99 (+ 1.29)		32.61 (+ 0.12)
0.06		32.19 (+ 1.49)		32.74 (+ 0.25)
0.03		32.20 (+ 1.50)		33.06 (+ 0.57)
0.016		32.36 (+ 1.67)		33.28 (+ 0.79)
PCR run 3				
background DNA white mustard 100 µg DNA/mL	soy (Cy5)		celery (FAM)	
concentration [mg target/kg]	positive/spike control mean C_t value *(spike level 0.25)*	IC mean C_t value (ΔC_t)	positive/spike control mean C_t value *(spike level 0.25)*	IC mean C_t value (ΔC_t)
0.25	31.01	31.33 (+ 0.3)	33.09	32.62 (+ 0.47)
0.125		31.53 (+ 0.5)		32.88 (+ 0.21)
0.06		31.73 (+ 0.7)		32.65 (+ 0.44)
0.03		31.45 (+ 0.4)		32.62 (+ 0.47)
0.016		32.00 (+ 1.0)		32.80 (+ 0.29)

Inspection of Figure 18 as well as Table 31 (PCR run 1, background of 100 µg soy DNA/mL) indicates that the difference between the mean C_t values of the positive/spike control (spike level 0.25 µg target DNA/mL) and the mean C_t values of the ICs at the concentrations 0.25-0.016 mg target/kg clearly exceed a ΔC_t value of 1; in a few cases, it even reached a difference of almost 2 C_t values (at concentration 0.06 mg target/kg: celery: 35.20 (+ 1.99, compared to positive/spike control 33.20); white mustard: 34.52 (+ 1.97, compared to positive/spike control 32.55)). An inhibitory effect, linked to soy DNA being present in a concentration of 100 µg/mL as background, on the amplification reaction in the celery and the white mustard PCR assay is assumed.

In the case of celery DNA being present in a concentration of 100 µg/mL as background in the soy PCR assay, there was a difference between the mean C_t values of the positive/spike control (spike level 0.25 mg target/kg) and the mean C_t values of the ICs at the concentrations 0.25-0.016 mg target/kg exceeding 1 C_t value (see Table 31, PCR run 2, left column)) too, however, less pronounced than for soy as background. Inspecting the results for the white mustard PCR assay (see Table 31, PCR run 2, right column), no such inhibitory effect was observed within this experiment series (difference ≤ 1 C_t value). No inhibitory effect was observed for white mustard DNA being present in a concentration of 100 µg/mL as background in the soy and the celery PCR assay either (see Table 31, PCR run 3).

This data suggests that the applicability of the triplex real-time PCR assay may be limited to certain foods to avoid false-negative results. Experiments were, however, carried out with very high amounts of target DNA. The highest applied diluent concentration of 100 µg target DNA/mL equals 500 ng target DNA per tube/well. For genomic DNA, a maximum of 100 ng per PCR tube or well is recommended (Müller & Prange, 2015).

6.3.3 Determination of limits of detection (LODs) and C_t cut-off values

The LODs of the triplex assay were determined by analyzing DNA solutions corresponding to target DNA concentrations of 0.25-0.016 mg target/kg (PCR run 1A, B), 0.125-0.016 mg target/kg (PCR run 2A, B) and 0.25-0.016 or 1-0.016 mg target/kg (PCR run 3A, B) in twelve replicates each. The DNA solutions were prepared from DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) and serially diluted with high amounts of target DNA (soy/celery/white mustard DNA solution, 5 or 20 µg target DNA/mL) (see subchapter 5.9.4).

The determination of the LOD was based on the preliminary assumption that the C_t cut-off value is 40. Consequently, an increase of the fluorescence signal at or above cycle 40 (C_t value equal or greater 40) was interpreted as a negative result and an increase of the fluorescence signal below cycle 40 (C_t value less 40) was interpreted as a positive result.

According to this assumption, the mean C_t value of the C_t values, obtained at the lowest concentration at which all (twelve out of twelve or sixteen out of sixteen) replicate measurements were positive, was calculated. The standard deviation was also calculated and its value was added two times to the mean C_t value, yielding the adjusted *C_t cut-off value* (instead of the preliminarily assumed C_t cut-off value of 40). All C_t values obtained at the lowest concentration were checked to see if each of them was below the adjusted C_t cut-off value. If this was the case, the corresponding concentration was determined as the LOD. For each target (soy, celery, white mustard), two corresponding concentrations and two C_t cut-off values resulted from the three PCR runs 1A-3B (see shaded rows in Table 32 for soy (yellow), celery (orange), white mustard (blue)). As the final LOD for each target of the triplex real-time PCR assay, the higher concentration and the lower C_t cut-off value was used. The results of the determination are shown in Table 32 and the final LODs and C_t cut-off values are summarized in Table 33.

Table 32 Results obtained as part of the determination of the LODs and C_t cut-off values of the triplex real-time PCR assay. PCR run 2A and 3A shows results obtained for soy, PCR run 1A and 3B shows results obtained for celery and PCR run 1B and 2B shows results obtained for white mustard. The mean C_t value was calculated from twelve or sixteen PCR replicates.

PCR run 1A diluent soy DNA 20 µg/mL				**PCR run 1B** diluent soy DNA 20 µg/mL			
celery (FAM)				**white mustard (VIC)**			
c [mg/kg]	mean C_t value	S	C_t cut-off value	c [mg/kg]	mean C_t value	S	C_t cut-off value
0.25	37.20	0.42		0.25	37.64	0.50	
0.125	38.04	0.42	**38.87**	**0.125**	38.15	0.34	**38.84**
0.06	39.47/-[1]	0.75		0.06	39.44/-[1]	0.55	
0.03	41.15/-[1]	0.64		0.03	41.43/-[1]	0.71	
0.016	42.28/-	0.91		0.016	43.37/-[1]	0.69	
PCR run 2A diluent celery DNA 20 µg/mL				**PCR run 2B** diluent celery DNA 20 µg/mL			
soy (Cy5)				**white mustard (VIC)**			
c [mg/kg]	mean C_t value	S	C_t cut-off value	c [mg/kg]	mean C_t value	S	C_t cut-off value
0.125	37.10	0.54		0.125	37.22	0.29//0.35	
0.06	38.00	0.49	38.99	0.06	38.45	0.45	39.34
0.03	39.74/-[1]	1.13		0.03	39.55/-[1]	0.70	
0.016	40.14/-[1]/-	0.72		0.016	40.91/-[1]/-	0.78	
PCR run 3A diluent white mustard DNA 5/20 µg/mL				**PCR run 3B** diluent white mustard DNA 5/20 µg/mL			
soy (Cy5) *diluent white mustard DNA 5 µg mL*				**celery (FAM)** *diluent white mustard DNA 20 µg* mL			
c [mg/kg]	mean C_t value	S	C_t cut-off value	c [mg/kg]	mean C_t value	S	C_t cut-off value
1	n.a.	n.a.		1	34.92	0.35	
0.5	n.a.	n.a.		0.5	35.97	0.35	
0.25	36.28	0.11		0.25	37.12	0.60	
0.125	37.01	0.35		0.125	38.12	1.08	40.28
0.06	37.89	0.31	**38.51**	0.06	39.00/-[1]	0.92	
0.03	39.34/-[1]	0.73		0.03	39.11/-[1]/-	0.55	
0.016	40.21/-[1]	1.12		0.016	40.03/-[1]/-	0.64	

n.a. not analyzed
- negative PCR result: no increase of the fluorescence signal within 45 cycles
-[1] negative PCR result: mean C_t value ≥ preliminarily assumed C_t cut-off value of 40 *(used for the determination of the six individual C_t cut-off values of PCR runs 1A-3B)*

Table 33 shows the LODs and C_t cut-off values for the targets of the triplex real-time PCR assay for the simultaneous detection of soy, celery and white mustard.

Table 33 LOD and C_t cut-off value for soy, celery and white mustard of the triplex real-time PCR assay.

target	soy	celery	white mustard
LOD [mg target/kg]	0.06	0.125	0.125
C_t cut-off value	38.5	38.9	38.8

The cycle number depends on the input amount of template DNA and the desired yield of DNA product. Allergen analysis requires an increased number of cycles as minute amounts of the template need to be amplified. Therefore, 40 was selected as preliminarily assumed C_t cut-off value for the determination of LODs and C_t cut-off values of the triplex assay. At higher cycles, such as above 40, the chance of the generation of non-specific DNA products increases (for details see subchapter 4.4.1).

Since the aim of the master's thesis was to develop a qualitative assay, the amplification efficiency was not calculated. Yet, linearity between the mean C_t values of the different concentrations (1-0.016 mg target/kg) was assessed. In the soy PCR assay (see Table 32, run 2A and 3A), good linearity was achieved between the mean C_t values of the concentrations 0.25-0.016 mg target/kg with white mustard as diluent (R^2 = 0.9849). With celery as diluent, a R^2 of 0.9585 was obtained (0.125-0.016 mg/kg). In the celery PCR assay (run 1A and 3B), good linearity was achieved with both diluents, soy and white mustard. With soy as diluent, a R^2 of 0.9894 resulting from the mean C_t values of the concentrations 0.25-0.016 mg target/kg, and with white mustard as diluent, a R^2 of 0.9706 resulting from the mean C_t values of the concentrations 1-0.016 mg target/kg was obtained. In the white mustard PCR assay (run 1B and 2B), good linearity was achieved between the mean C_t values of the concentrations 0.25-0.016 mg target/kg with soy as diluent (R^2 = 0. 0.9888) and between the mean C_t values of the concentrations 0.125-0.016 mg target/kg with celery as diluent (R^2 = 0.9954). Compared to the R^2 when water was used as diluent in the soy, celery and white mustard PCR assay (see subchapter 6.3), R^2 was generally lower using soy, celery or white mustard DNA as diluent in the assays instead.

Comparing the achieved LOD for soy of the triplex assay (0.06 mg/kg) with its achieved LOD of the singleplex assay (0.016 mg/kg) shows that soy is detected slightly less sensitive in the triplex assay. Detection limits for the different targets of the triplex assay are still very low.

Since the LODs of the triplex real-time PCR assay were only investigated with DNA solutions prepared from DNA extract of SureFood® QUANTARD Allergen 40 (40 mg/kg) and serially diluted with high amounts of soy, celery or white mustard DNA, direct comparison with other real-time PCR assays investigating detection limits in model food may not be advisable. Reference materials used for validation, such as SureFood® QUANTARD Allergen 40, are not certified, which complicates comparison, too (EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA), 2014).

Therefore, investigating detection limits in different commercial food spiked with soy, celery and white mustard are highly recommended, not only to make the assay more comparable but to contribute to a better understanding of the performance of the assay.

7 Conclusion

A qualitative singleplex real-time PCR assay for the specific detection of the allergenic food soy with a very low LOD was developed and combined to a qualitative triplex real-time PCR assay for the simultaneous detection of soy, celery and white mustard in food.

The primers and the TaqMan probe, labelled with Cy5 as reporter dye and modified with a minor groove binder, of the singleplex PCR assay target a specific, short sequence within the chloroplast genome of soy. An amplicon of 133 bp is generated. Eighty-six species, both closely and not closely related to soy, were investigated for cross-reactivity. In the majority of cases, there was no increase of the fluorescence signal, meaning no cross-reactivity. In the few cases of an increase, contamination with traces of soy is strongly suspected to be the cause for this outcome. The LOD was found to be 0.016 mg soy/kg food with a calculated C_t cut-off value of 39.3. For the triplex PCR assay, the LODs and calculated C_t cut-off values for soy, celery and white mustard were found to be 0.06 mg soy/kg food and 38.5, 0.125 mg celery/kg food and 38.9 and 0.125 mg white mustard/kg food and 38.8, respectively. The requirement of adequate sensitivity for such detection methods may be becoming even more important in light of the recently updated VITAL 3.0 reference doses for allergens in foods.

The data suggests that the designed primer/probe system of the singleplex PCR assay is suitable for the specific detection of soy. However, further experiments should investigate the issue of soy contamination, e.g. by repeated DNA extraction of the species in question, if possible, in a different soy-free lab environment. The detection limits in commercial food spiked with soy (singleplex assay) and in commercial food spiked with soy, celery and white mustard (triplex assay) were not tested in the scope of the underlying master's thesis. Before the assays can be introduced to routine applications at AGES, further in-house validation regarding sensitivity is required, i.e. investigation of the LODs in real food matrices. The applicability of the assays to real food samples remains to be investigated too. This may be of special importance here as preliminary inhibition experiments with the triplex assay showed inhibitory effects on amplification caused by target background DNA, particularly in the celery and white mustard PCR assay in the case of soy being present in very high amounts. If posing a problem further on, the singleplex assay coupled with the duplex real-time PCR assay may be considered an alternative in such samples.

In future, if the two PCR assays are successfully validated, their routine application in checking compliance with food allergen labelling will save both time and resources.

8 Zusammenfassung

Die Anzahl der Fälle von Lebensmittelallergien hat in den letzten Jahrzehnten zugenommen, insbesondere in den westlichen Ländern. Darüber hinaus scheinen Kinder noch stärker betroffen zu sein als Erwachsene, was diese Art von Allergie nicht nur zu einer Last für die Wirtschaft und Sicherheit macht, sondern auch zu einer emotionalen Last. Die effizienteste Strategie zur Behandlung von Lebensmittelallergien ist der Ausschluss des betreffenden Lebensmittels aus der Ernährung, was zuallererst zuverlässige Lebensmittelkennzeichnung und Analysemethoden zur Überprüfung der Einhaltung der Kennzeichnung erfordert. Das Hauptziel dieser Masterarbeit war die Entwicklung eines qualitativen Singleplex-Real-time PCR Assays für den Nachweis von Soja (*Glycine max*) und seine Validierung als Teil eines qualitativen Triplex-Real-time PCR Assays für den gleichzeitigen Nachweis von Soja, Sellerie (*Apium graveolens*) und weißem Senf (*Sinapis alba*) in Lebensmitteln. Die Primer und die TaqMan-Sonde des Singleplex-PCR Assays, die mit Cy5 als Reporterfarbstoff markiert und mit einem Minor Groove Binder modifiziert war, zielen auf eine spezifische, kurze Sequenz innerhalb des Chloroplastengenoms von Soja ab. Es wurden sechs Primer/Sonden-Systeme getestet. Die Konzentrationen der Primer und TaqMan-Sonden wurden optimiert. 86 Arten, die mit Soja sowohl nah als auch nicht nah verwandt sind, wurden auf Kreuzreaktivität untersucht. Die Nachweisgrenze (LOD) wurde mit 0,016 mg Soja/kg Lebensmittel und einem berechneten C_t-Cut-off-Wert von 39,3 ermittelt. Die LODs und berechneten C_t-Cut-off-Werte für Soja, Sellerie und weißen Senf lagen bei 0,06 mg Soja/kg Lebensmittel und 38,5, 0,125 mg Sellerie/kg Lebensmittel und 38,9 bzw. 0,125 mg weißer Senf/kg Lebensmittel und 38,8.

9 List of abbreviations

A	adenine
AGES	Austrian Agency for Health and Food Safety
BLAST	Basic Local Alignment Search Tool
bp	base pair
C	cytosine
C_t	threshold cycle
CTAB	hexadecyltrimethylammonium bromide
DBPCFC	double-blind, placebo-controlled, food challenge
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
e.g.	exempli gratia (for example)
ED	eliciting dose
ELISA	enzyme-linked immunosorbent assay
ENGL	European Network of GMO Laboratories
et al.	et alii (and others)
EU	European Union
FAM	6-carboxyfluorescein
FPIES	food protein-induced enterocolitis syndrome
FRET	fluorescence resonance energy transfer
g	gram
G	guanine
GMO	genetically modified organisms
IC	inhibition control
IgE	immunoglobulin E
IUIS	International Union of Immunological Societies
kg	kilogram
L	liter/litre
LOAEL	lowest observed adverse effect level
LOD	limit of detection
LOQ	limit of quantification
LTPs	profilins and lipid transfer proteins
M6PR	mannose 6-phosphate receptor
mdh	mannitol dehydrogenase
mg	milligram
$MgCl_2$	magnesium chloride
min	minute

mL	milliliter
n.a.	not analyzed
ndhH	NAD(P)H-quinone oxidoreductase subunit H
ng	nanogram
NIAID	National Institute of Allergy and Infectious Diseases
nm	nanometer
NOAEL	no observed adverse effect level
OAS	oral allergy syndrome
PAL	precautionary allergen labelling
PCR	polymerase chain reaction
PDCAAS	protein digestibility corrected amino acid score
ppm	parts per million
RSD	relative standard deviation
RT	room temperature
s	second
SDS	Sequence Detection System
T	thymine
T_a	annealing temperature
TAMRA	tetramethylrhodamine
Taq	*Thermus aquaticus*
T_m	melting temperature
VITAL	Voluntary Incidental Trace Allergen Labelling
VSEP	VITAL Scientific Expert Panel
w/w	weight per weight
°C	degree centigrade
µg	microgram
µL	microliter

10 List of figures

11 List of tables

12 List of utensils

12.1 Chemicals

Agarose (BioReagent, for molecular biology, low EEO)	Sigma-Aldrich, St. Louis, USA
BSA, bovine serum albumin	Roche, Basel, Switzerland
Chloroform, contains 100-200 parts per million (ppm) amylenes as stabilizer, ≥ 99.5 %	Sigma-Aldrich, St. Louis, USA
CTAB, hexadecyl(trimethyl)azanium bromide	Sigma-Aldrich, St. Louis, USA
DNA Loading buffer, 5x, red	Biozym, Hessisch Oldendorf, Germany
EDTA, ethylenediaminetetraacetic acid, 2-[2-[bis(carboxymethyl)amino]ethyl-(carboxymethyl)amino]acetic acid	Sigma-Aldrich, St. Louis, USA
Elution buffer	Promega, Madison, USA
Ethanol, 96 %	VWR, Radnor, USA
GelRed® Nucleic Acid Stain 10000X Water	Merck, Darmstadt, Germany
GlycoBlue™ Coprecipitant, 15 mg/mL	Thermo Fisher Scientific, Waltham, USA
Herring sperm DNA, lyoph. sodium salt	Böhringer Mannheim, Mannheim, Germany
Isoamyl alcohol	Merck, Darmstadt, Germany
Lysis buffer	Promega, Madison, USA
Maxwell® RSC PureFood GMO and Authentication Kit	Promega, Madison, USA
Milli Q Water, obtained by Synergy® Water Purification System	Millipore, Burlington, USA
NaCl, sodium chloride	Merck, Darmstadt, Germany
Petroleum benzine	Merck, Darmstadt, Germany
Phenol:chloroform:isoamyl alcohol 25:24:1, saturated with 10 mM Tris, pH 8.0, 1 mM EDTA	Sigma-Aldrich, St. Louis, USA

Primers	Eurogentec, Seraing, Belgium
Propan-2-ol	Merck, Darmstadt, Germany
Proteinase K, from Tritirachium album, solution in Tris/HCl pH 7.5, 0.01 mol/L, 600 mAnson-U/mL	Sigma-Aldrich, St. Louis, USA
QuantiTect® Multiplex PCR NoROX Master Mix	Qiagen, Hilden, Germany
RNase-free water	Qiagen, Hilden, Germany
SureFood® QUANTARD Allergen 40, 40 mg/kg	R-Biopharm, Pfungstadt, Germany
TaqMan® Probes	Eurogentec, Seraing, Belgium
Tris, 2-amino-2-(hydroxymethyl)propane-1,3-diol	Sigma-Aldrich, St. Louis, USA
Tris/boric acid/EDTA (TBE) Buffer, 10x	Bio-Rad, Hercules, USA

12.1.1 Preparation of buffers and solutions

CTAB extraction buffer

20 mg/mL (2 %, w/v) CTAB (Sigma-Aldrich, USA), 0.02 M EDTA (Sigma-Aldrich, USA), 0.1 M Tris (Sigma-Aldrich, USA), 1.4 M NaCl (Merck, Germany); adjusted to pH 8.0 with 4 M HCl, autoclaved (at 110-130 °C, 1.3-1.5 bar, for at least 20 minutes)

Sodium chloride solution, 1.2 M

7 g NaCl; dissolved in 100 mL water

Precipitation solution

500 mg CTAB, 234 mg NaCl; dissolved in 100 mL water

10x RNase buffer

3 M NaCl, 100 mM Tris, 50 mM EDTA; adjusted to pH 7.4 with 4 M HCl

10x TE buffer

1.21 g Tris, 0.372 g EDTA; adjusted to 8.0 with 4 M HCl

12.2 Consumable materials

Applied Biosystems® MicroAmp® Optical 8-cap strips		Thermo Fisher Scientific, Waltham, USA
Applied Biosystems® MicroAmp® Optical 96-Well Reaction Plate	Ref: N8010560	Thermo Fisher Scientific, Waltham, USA
Applied Biosystems® MicroAmp® Optical adhesive film		Thermo Fisher Scientific, Waltham, USA
Falcon tubes	15 mL, 50 mL	VWR, Radnor, USA
Gloves		
Pipette universal fit tips, with and without filter	10 µL, 20 µL, 100 µL, 200 µL, 300 µL, 1250 µL, 5 mL, 10 mL	Biotix, Neptune Scientific, San Diego, USA
QIAxpert Slide-40		Qiagen, Hilden, Germany
Safe-lock tubes	0.5 mL, 1.5 mL, 2 mL, 5 mL	Eppendorf, Hamburg, Germany
96 PCR plate half skirt white		Sarstedt, Biedermannsdorf, Austria

12.3 Laboratory equipment

Analytical balance	research readability 0.00001 g, measurement uncertainty < 30 g 0.01 mg	Sartorius, Göttingen, Germany
Applied Biosystems® 7500 Real Time PCR System		Thermo Fisher Scientific, Waltham, USA

Centrifuge	5415 R	Eppendorf, Hamburg, Germany
Centrifuge	5810 R	Eppendorf, Hamburg, Germany
Electrophoresis equipment	Power Pac 200	Bio-Rad, Hercules, USA
Freezer		Liebherr, Ochsenhausen, Germany
Gel caster and chamber	EC330 Midicell® Primo™ Electrophoretic gel system	Thermo Fisher Scientific, Waltham, USA
Graduated Cylinders	50 mL, 100 mL, 250 mL	
Incubator Hybaid		MWG-Biotech, Ebersberg, Germany
Incubator Unihood 750 equipped with a rotator/mixer Intelli-Mixer RM-2L		Uniequip, Planegg, Germany LTF Labortechnik, Wasserburg, Germany
Laborfuge 400, Heraeus Instruments		Thermo Fisher Scientific, Waltham, USA
Maxwell® 16 Instrument (AS2000)		Promega, Madison, USA
Micro centrifuge	Fuge One	Biomedia, Pandan Loop, Singapore
PCR workstation	UV3 HEPA	LTF Labortechnik, Wasserburg, Germany
Pipettes, Research plus	10 µL, 20 µL, 100 µL, 200 µL, 300 µL, 1000 µL, 5 mL, 10 mL	Eppendorf, Hamburg, Germany
Polytron	PT 3000	Kinematica, Malters, Switzerland

QIAxpert		Qiagen, Hilden, Germany
Refrigerator		Liebherr, Ochsenhausen, Germany
Scale	readability 0.001 g, measurement uncertainty < 400 g/0.01 g	Sartorius, Göttingen, Germany
Thermomixer	comfort	Eppendorf, Hamburg, Germany
Transilluminator	Universal hood II	Bio-Rad, Hercules, USA
Volumetric Flasks	50 mL, 100 mL, 250 mL	
Vortexmixer, Vortex Genie 2		Scientific Industries, Bohemia, USA

12.4 Software programs

Applied Biosystems® 7500 System SDS Software, version 1.4.0	Thermo Fisher Scientific, Waltham, USA
CLC Genomics Workbench, version 10.1.1	Qiagen, Hilden, Germany
Maxwell RSC, version 3.0	Promega, Madison, USA
Primer Express Software for Real-Time PCR, version 3.0.1	Thermo Fisher Scientific, Waltham, USA
QIAxpert software, version 2.2.0.21	Qiagen, Hilden, Germany
Quantity One Basic Software, version 4.6.6	Bio-Rad, Hercules, USA

13 References

AAT Bioquest (2021), AAT Bioquest's interactive Spectrum Viewer: Spectrum HEX and 6-VIC. https://www.aatbio.com/fluorescence-excitation-emission-spectrum-graph-viewer, accessed 07/25/2021.

Abbas, A. K., & Lichtman, A. H. (2016), Antibodies and Antigens. In J. Merritt, & R. Gruliow (Eds.), *Cellular and molecular immunology*. 98–116. Elsevier, ninth edition.

Abbas, A. K., Lichtman, A. H., & Pillai, S. (2016), Allergy. In J. Merritt, & R. Gruliow (Eds.), *Cellular and molecular immunology*. 437–457. Elsevier, ninth edition.

Adams, P. S. (2007), Data analysis and reporting. In *Real-time PCR*. 39–62. Taylor & Francis, 1st Edition.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002), Introduction to the Cell. In *Molecular Biology of the Cell*. 3–190. Garland Science, 4th edn.

Allergen Bureau (2019). *Summary of the 2019 VITAL Scientific Expert Panel Recommendations*. https://vital.allergenbureau.net/wp-content/uploads/2021/03/VSEP-2019-Summary-Recommendations_FINAL_Sept2019.pdf.

Allergen Bureau (2021), VITAL Science. https://vital.allergenbureau.net/vital-science/, accessed 07/02/2021.

Amlot, P. L., Kemeny, D. M., Zachary, C., Parkes, P., & Lessof, M. H. (1987), Oral allergy syndrome (OAS): symptoms of IgE-mediated hypersensitivity to foods. *Clinical and Experimental Allergy*, *17*, 33–42.

André, F., André, C., Colin, L., Cacaraci, F., & Cavagna, S. (1994), Role of new allergens and of allergens consumption in the increased incidence of food sensitizations in France. *Toxicology*, *93*, 77–83.

Asero, R., Ballmer-Weber, B. K., Beyer, K., Conti, A., Dubakiene, R., Fernandez-Rivas, M., Hoffmann-Sommergruber, K., Lidholm, J., Mustakov, T., Oude Elberink, J. N. G., Pumphrey, R. S. H., Stahl Skov, P., van Ree, R., Vlieg-Boerstra, B. J., Hiller, R., Hourihane, J. O., Kowalski, M., Papadopoulos, N. G., Wal, J.-M., Mills, E. N. C., & Vieths, S. (2007), IgE-mediated food allergy diagnosis: Current status and new perspectives. *Molecular Nutrition & Food Research*, *51*, 135–147.

Ballmer-Weber, B. K., Hoffmann, A., Wüthrich, B., Lüttkopf, D., Pompei, C., Wangorsch, A., Kästner, M., & Vieths, S. (2002), Influence of food processing on the allergenicity of celery: DBPCFC with celery spice and cooked celery in patients with celery allergy. *Allergy*, *57*, 228–235.

Ballmer-Weber, B. K., & Vieths, S. (2008), Soy allergy in perspective. *Current Opinion in Allergy and Clinical Immunology*, *8*, 270–275.

Ballmer-Weber, B. K., Vieths, S., Lüttkopf, D., Heuschmann, P., & Wüthrich, B. (2000), Celery allergy confirmed by double-blind, placebo-controlled food challenge: a clinical study in 32

subjects with a history of adverse reactions to celery root. *Journal of Allergy and Clinical Immunology*, *106*, 373–378.

Bauer, T., Hammes, W. P., Haase, N. U., & Hertel, C. (2004), Effect of food components and processing parameters on DNA degradation in food. *Environmental Biosafety Research*, *3*, 215–223.

Bauer, T., Kirschbaum, K., Panter, S., Kenk, M., & Bergemann, J. (2011), Sensitive detection of soy (Glycine max) by real-time polymerase chain reaction targeting the mitochondrial atpA gene. *Journal of AOAC International*, *94*, 1863–1873.

Blom, W. M., Michelsen-Huisman, A. D., van Os-Medendorp, H., van Duijn, G., Zeeuw-Brouwer, M.-L. de, Versluis, A., Castenmiller, J. J. M., Noteborn, H. P. J. M., Kruizinga, A. G., Knulst, A. C., & Houben, G. F. (2018), Accidental food allergy reactions: Products and undeclared ingredients. *Journal of Allergy and Clinical Immunology*, *142*, 865–875.

Bock, S. A., Muñoz-Furlong, A., & Sampson, H. A. (2001), Fatalities due to anaphylactic reactions to foods. *Journal of Allergy and Clinical Immunology*, *107*, 191–193.

Bonds, R. S., Midoro-Horiuti, T., & Goldblum, R. (2008), A structural basis for food allergy: the role of cross-reactivity. *Current Opinion in Allergy and Clinical Immunology*, *8*, 82–86.

Botha, M., Basera, W., Facey-Thomas, H. E., Gaunt, B., Gray, C. L., Ramjith, J., Watkins, A., & Levin, M. E. (2019), Rural and urban food allergy prevalence from the South African Food Allergy (SAFFA) study. *Journal of Allergy and Clinical Immunology*, *143*, 662-668.e2.

Boyce, J. A., Assa'ad, A., Burks, A. W., Jones, S. M., Sampson, H. A., Wood, R. A., Plaut, M., Cooper, S. F., Fenton, M. J., Arshad, S. H., Bahna, S. L., Beck, L. A., Byrd-Bredbenner, C., Camargo, C. A., Eichenfield, L., Furuta, G. T., Hanifin, J. M., Jones, C., Kraft, M., Levy, B. D., Lieberman, P., Luccioli, S., McCall, K. M., Schneider, L. C., Simon, R. A., Simons, F. E. R., Teach, S. J., Yawn, B. P., & Schwaninger, J. M. (2010), Guidelines for the diagnosis and management of food allergy in the United States: report of the NIAID-sponsored expert panel. *Journal of Allergy and Clinical Immunology*, *126*, 1-58.

Breiteneder, H., & Radauer, C. (2004), A classification of plant food allergens. *Journal of Allergy and Clinical Immunology*, *113*, 821-30; quiz 831.

Broll, H. (2010), Polymerase Chain Reaction. In Popping, Bert; Diaz-Amigo, Carmen; Hoenicke, Katrin, *Molecular biological and immunological techniques and applications for food chemists*. 41–58. Wiley.

Bruijnzeel-Koomen, C., Ortolani, C., Aas, K., Bindslev-Jensen, C., Björkstén, B., Moneret-Vautrin, D., & Wüthrich, B. (1995), Adverse reactions to food. European Academy of Allergology and Clinical Immunology Subcommittee. *Allergy*, *50*, 623–635.

Bundesinstitut für Risikobewertung (2020). *„VITAL 3.0“: Neue und aktualisierte Vorschläge für Referenzdosen von Lebensmittelallergenen: Stellungnahme Nr. 015/2020 des BfR*.

https://www.bfr.bund.de/cm/343/vital-30-neue-und-aktualisierte-vorschlaege-fuer-referenzdosen-von-lebensmittelallergenen.pdf.

Burks, A. W., Tang, M., Sicherer, S., Muraro, A., Eigenmann, P. A., Ebisawa, M., Fiocchi, A., Chiang, W., Beyer, K., Wood, R., Hourihane, J., Jones, S. M., Lack, G., & Sampson, H. A. (2012), ICON: food allergy. *Journal of Allergy and Clinical Immunology*, *129*, 906–920.

Burks, A. W., Williams, L. W., Helm, R. M., Connaughton, C., Cockrell, G., & O'Brien, T. (1991), Identification of a major peanut allergen, I, in patients with atopic dermatitis and positive peanut challenges. *Journal of Allergy and Clinical Immunology*, *88*, 172–179.

Burks, W., & Ballmer-Weber, B. K. (2006), Food allergy. *Molecular Nutrition & Food Research*, *50*, 595–603.

Businco, L., Bruno, G., & Giampietro, P. G. (1998), Soy protein for the prevention and treatment of children with cow-milk allergy. *The American Journal of Clinical Nutrition*, *68*, 1447S-1452S.

Caldwell, J. M. (2017), Food Analysis Using Organelle DNA and the Effects of Processing on Assays. *Annual Review of Food Science and Technology*, *8*, 57–74.

Chafen, J. J. S., Newberry, S. J., Riedl, M. A., Bravata, D. M., Maglione, M., Suttorp, M. J., Sundaram, V., Paige, N. M., Towfigh, A., Hulley, B. J., & Shekelle, P. G. (2010), Diagnosing and managing common food allergies: a systematic review. *Journal of the American Medical Association*, *303*, 1848–1856.

Chapman, M. D. (1988), Purification of allergens. *Current Opinion in Immunology*, *1*, 647–653.

Chapman, M. D., Pomés, A., Breiteneder, H., & Ferreira, F. (2007), Nomenclature and structural biology of allergens. *Journal of Allergy and Clinical Immunology*, *119*, 414–420.

Christenhusz, M. J., & Byng, J. W. (2016), The number of known plants species in the world and its annual increase. *Phytotaxa*, *261*, 201.

Coen, D. M. (2012), The Polymerase Chain Reaction. In F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J.G. Seidman, J. A. Smith, & K. Struhl (Eds.), *Current protocols in molecular biology*. 15.0.1–15.0.3. John Wiley & Sons, Ltd.

Costa, J., Amaral, J. S., Grazina, L., Oliveira, M. B. P. P., & Mafra, I. (2017), Matrix-normalised real-time PCR approach to quantify soybean as a potential food allergen as affected by thermal processing. *Food Chemistry*, *221*, 1843–1850.

Costa, J., Oliveira, M. B. P., & Mafra, I. (2013), Effect of thermal processing on the performance of the novel single-tube nested real-time PCR for the detection of walnut allergens in sponge cakes. *Food Research International*, *54*, 1722–1729.

Crevel, R. W. R., Ballmer-Weber, B. K., Holzhauser, T., Hourihane, J. O., Knulst, A. C., Mackie, A. R., Timmermans, F., & Taylor, S. L. (2008), Thresholds for food allergens and their value to different stakeholders. *Allergy*, *63*, 597–609.

Daniell, H., Lin, C.-S., Yu, M., & Chang, W.-J. (2016), Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology*, *17*, 134.

Deibel, K., Trautman, T., DeBOOM, T., Sveum, W. H., Dunaif, G., Scott, V. N., & Bernard, D. T. (1997), A comprehensive approach to reducing the risk of allergens in foods. *Journal of Food Protection*, *60*, 436–441.

Demeke, T., & Jenkins, G. R. (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*, 1977–1990.

Demmel, A., Hupfer, C., Ilg Hampe, E., Busch, U., & Engel, K.-H. (2008), Development of a real-time PCR for the detection of lupine DNA (lupinus species) in foods. *Journal of Agricultural and Food Chemistry*, *56*, 4328–4332.

Druml, B., Grandits, S., Mayer, W., Hochegger, R., & Cichna-Markl, M. (2015), Authenticity control of game meat products--a single method to detect and quantify adulteration of fallow deer (Dama dama), red deer (Cervus elaphus) and sika deer (Cervus nippon) by real-time PCR. *Food Chemistry*, *170*, 508–517.

EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA) (2014), Scientific Opinion on the evaluation of allergenic foods and food ingredients for labelling purposes. *EFSA Journal*, *12*, 3894 (1-286).

Elisa Systems (2021), Elisa Kits. https://elisasystems.com/product-category/elisa_kits/, accessed 09/20/2021.

Espiñeira, M., & Santaclara, F. J. (2017), Fast Real-Time PCR Method for Detection of Soy in Foods. *Methods in Molecular Biology*, *1620*, 173–181.

Eurofins (2021), Allergen ELISA Kits. https://www.eurofins-technologies.com/products/allergens/allergen-elisa-kits.html, accessed 09/20/2021.

European Commission. Joint Research Centre. (2017), Verification of analytical methods for GMO testing when implementing interlaboratory validated methods :guidance document from the European Network of GMO Laboratories, version 2. https://op.europa.eu/en/publication-detail/-/publication/6046f778-0c84-11e8-966a-01aa75ed71a1/language-en.

European Food Safety Authority (2007), Opinion of the Scientific Panel on Dietetic Products, Nutrition and Allergies on a request from the Commission related to a notification from FEDIOL and IMACE on fully refined soybean oil and fat pursuant to Article 6, paragraph 11 of Directive 2000/13/EC - for permanent exemption from labelling. *Europäische Behörde für Lebensmittelsicherheit,* 24.10.2007. https://www.efsa.europa.eu/de/efsajournal/pub/570, accessed 06/30/2021.

European Parlament and Council (2011). *Regulation (EU) No 1169/2011*. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02011R1169-20180101&from=EN, accessed 10/31/2021.

Feldweg, A. M. (2017), Food-Dependent, Exercise-Induced Anaphylaxis: Diagnosis and Management in the Outpatient Setting. *The Journal of Allergy and Clinical Immunology: In Practice*, *5*, 283–288.

Figueroa, J., Blanco, C., Dumpiérrez, A. G., Almeida, L., Ortega, N., Castillo, R., Navarro, L., Pérez, E., Gallego, M. D., & Carrillo, T. (2005), Mustard allergy confirmed by double-blind placebo-controlled food challenges: clinical features and cross-reactivity with mugwort pollen and plant-derived foods. *Allergy*, *60*, 48–55.

Food and Agriculture Organization of the United Nations (FAO) (2021), Value of Agricultural Production. http://www.fao.org/faostat/en/#data/QV, accessed 07/01/2021.

Foucard, T., & Malmheden Yman, I. (1999), A study on severe food reactions in Sweden--is soy protein an underestimated cause of food anaphylaxis? *Allergy*, *54*, 261–265.

Fowler, H. W., Fowler, F. G., & Della Thompson (1995). *The concise Oxford dictionary of current English*. Oxford University Press; Clarendon Press, 9th ed.

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

Friedman, M., & Brandon, D. L. (2001), Nutritional and health benefits of soy proteins. *Journal of Agricultural and Food Chemistry*, *49*, 1069–1086.

Fu, L., Qian, Y., Zhou, J., Zheng, L., & Wang, Y. (2020), Fluorescence-based quantitative platform for ultrasensitive food allergen detection: From immunoassays to DNA sensors. *Comprehensive Reviews in Food Science and Food Safety*, *19*, 3343–3364.

Gallagher, S. R. (2011), Quantitation of DNA and RNA with absorption and fluorescence spectroscopy. *Current Protocols in Molecular Biology*, *93*, A.3D.1-A.3D.14.

García, M. C., Torre, M., Marina, M. L., & Laborda, F. (1997), Composition and characterization of soyabean and related products. *Critical Reviews in Food Science and Nutrition*, *37*, 361–391.

Graham, P. H., & Vance, C. P. (2003), Legumes: importance and constraints to greater use. *Plant Physiology*, *131*, 872–877.

Gryson, N. (2010), Effect of food processing on plant DNA degradation and PCR-based GMO analysis: a review. *Analytical and Bioanalytical Chemistry*, *396*, 2003–2022.

Hefle, S. L., Lambrecht, D. M., & Nordlee, J. A. (2005), Soy sauce retains allergenicity through the fermentation/production process. *Journal of Allergy and Clinical Immunology*, *115*, S32.

Hefle, S. L., Nordlee, J. A., & Taylor, S. L. (1996), Allergenic foods. *Critical Reviews in Food Science and Nutrition*, *36 Suppl*, 69-89.

Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H., & Vogt, P. H. (1997), Multiplex PCR: critical parameters and step-by-step protocol. *BioTechniques*, *23*, 504–511.

Herian, A. M., Taylor, S. L., & Bush, R. K. (1990), Identification of soybean allergens by immunoblotting with sera from soy-allergic adults. *International Archives of Allergy and Applied Immunology*, *92*, 193–198.

Herman, E. M., Helm, R. M., Jung, R., & Kinney, A. J. (2003), Genetic modification removes an immunodominant allergen from soybean. *Plant Physiology*, *132*, 36–43.

Higuchi, R., Dollinger, G., Walsh, P. S., & Griffith, R. (1992), Simultaneous amplification and detection of specific DNA sequences. *Bio/technology (Nature Publishing Company)*, *10*, 413–417.

Hirao, T., Watanabe, S., Temmei, Y., Hiramoto, M., & Kato, H. (2009), Qualitative polymerase chain reaction methods for detecting major food allergens (peanut, soybean, and wheat) by using internal transcribed spacer region. *Journal of AOAC International*, *92*, 1464–1471.

Holzhauser, T. (2018), Protein or No Protein? Opportunities for DNA-Based Detection of Allergenic Foods. *Journal of Agricultural and Food Chemistry*, *66*, 9889–9894.

Holzhauser, T., Johnson, P., Hindley, J. P., O'Connor, G., Chan, C.-H., Costa, J., Fæste, C. K., Hirst, B. J., Lambertini, F., Miani, M., Robert, M.-C., Röder, M., Ronsmans, S., Bugyi, Z., Tömösközi, S., & Flanagan, S. D. (2020), Are current analytical methods suitable to verify VITAL® 2.0/3.0 allergen reference doses for EU allergens in foods? *Food and chemical toxicology: British Industrial Biological Research Association*, *145*, 111709 (1-20).

Holzhauser, T., & Röder, M. (2015), Polymerase chain reaction (PCR) methods for detecting allergens in foods. In S. Flanagan (Ed.), *Handbook of food allergen detection and control*. 245–263. Woodhead Publishing.

Holzhauser, T., Wackermann, O., Ballmer-Weber, B. K., Bindslev-Jensen, C., Scibilia, J., Perono-Garoffo, L., Utsumi, S., Poulsen, L. K., & 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*, 452–458.

Huggett, A. C., & Hischenhuber, C. (1998), Food manufacturing initiatives to protect the allergic consumer. *Allergy*, *53*, 89–92.

Iniesto, E., Jiménez, A., Prieto, N., Cabanillas, B., Burbano, C., Pedrosa, M. M., Rodríguez, J., Muzquiz, M., Crespo, J. F., Cuadrado, C., & Linacero, R. (2013), Real Time PCR to detect hazelnut allergen coding sequences in processed foods. *Food Chemistry*, *138*, 1976–1981.

International Organization for Standardization (2005). *Foodstuffs-methods of analysis for the detection of genetically modified organisms and derived products-nucleic acid extraction.*, *67.050 General methods of tests and analysis for food products*. https://www.iso.org/standard/34616.html, accessed 10/04/2021.

Jayachandran, M., & Xu, B. (2019), An insight into the health benefits of fermented soy products. *Food Chemistry*, *271*, 362–371.

Jenkins, J. A., Griffiths-Jones, S., Shewry, P. R., Breiteneder, H., & Mills, E. N. C. (2005), Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. *Journal of Allergy and Clinical Immunology*, *115*, 163–170.

Johansson, S. G. O., Bieber, T., Dahl, R., Friedmann, P. S., Lanier, B. Q., Lockey, R. F., Motala, C., Ortega Martell, J. A., Platts-Mills, T. A. E., Ring, J., Thien, F., van Cauwenberge, P., & Williams, H. C. (2004), Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. *Journal of Allergy and Clinical Immunology*, *113*, 832–836.

Johansson, S. G. O., Hourihane, J., Bousquet, J., Bruijnzeel-Koomen, C., Dreborg, S., Haahtela, T., Kowalski, M. L., Mygind, N., Ring, J., van Cauwenberge, P., van Hage-Hamsten, M., & Wüthrich, B. (2001), A revised nomenclature for allergy: An EAACI position statement from the EAACI nomenclature task force. *Allergy*, *56*, 813–824.

Katz, Y., Gutierrez-Castrellon, P., González, M. G., Rivas, R., Lee, B. W., & Alarcon, P. (2014), A comprehensive review of sensitization and allergy to soy-based products. *Clinical Reviews in Allergy & Immunology*, *46*, 272–281.

King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G., Platts-Mills, T. A., & Thomas, W. (1994), Allergen nomenclature. WHO/IUIS Allergen Nomenclature Subcommittee. *International Archives of Allergy and Immunology*, *105*, 224–233.

Klemola, T., Vanto, T., Juntunen-Backman, K., Kalimo, K., Korpela, R., & Varjonen, E. (2002), Allergy to soy formula and to extensively hydrolyzed whey formula in infants with cow's milk allergy: a prospective, randomized study with a follow-up to the age of 2 years. *The Journal of Pediatrics*, *140*, 219–224.

Kobayashi, M. (2005), Immunological functions of soy sauce: hypoallergenicity and antiallergic activity of soy sauce. *Journal of Bioscience and Bioengineering*, *100*, 144–151.

Köppel, R., Dvorak, V., Zimmerli, F., Breitenmoser, A., Eugster, A., & Waiblinger, H.-U. (2010), Two tetraplex real-time PCR for the detection and quantification of DNA from eight allergens in food. *European Food Research and Technology*, *230*, 367–374.

Köppel, R., van Velsen-Zimmerli, F., & Bucher, T. (2012), Two quantitative hexaplex real-time PCR systems for the detection and quantification of DNA from twelve allergens in food. *European Food Research and Technology*, *235*, 843–852.

Koppelman, S. J., Lakemond, C. M. M., Vlooswijk, R., & Hefle, S. L. (2004), Detection of Soy Proteins in Processed Foods: Literature Overview and New Experimental Work. *Journal of AOAC International*, *87*, 1398–1407.

Krska, R., Welzig, E., & Baumgartner, S. (2004), Immunoanalytical detection of allergenic proteins in food. *Analytical and Bioanalytical Chemistry*, *378*, 63–65.

Kubista, M., Andrade, J. M., Bengtsson, M., Forootan, A., Jonák, J., Lind, K., Sindelka, R., Sjöback, R., Sjögreen, B., Strömbom, L., Ståhlberg, A., & Zoric, N. (2006), The real-time polymerase chain reaction. *Molecular Aspects of Medicine*, *27*, 95–125.

Kumar, P., Chatli, M. K., Mehta, N., Singh, P., Malav, O. P., & Verma, A. K. (2017), Meat analogues: Health promising sustainable meat substitutes. *Critical Reviews in Food Science and Nutrition*, *57*, 923–932.

Ladenburger, E.-M., Dehmer, M., Grünberg, R., Waiblinger, H.-U., Stoll, D., & Bergemann, J. (2018), Highly Sensitive Matrix-Independent Quantification of Major Food Allergens Peanut and Soy by Competitive Real-Time PCR Targeting Mitochondrial DNA. *Journal of AOAC International*, *101*, 170–184.

Lee, L. G., Connell, C. R., & Bloch, W. (1993), Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Research*, *21*, 3761–3766.

Lieberman, J. A., & Sicherer, S. H. (2011), Quality of life in food allergy. *Current Opinion in Allergy and Clinical Immunology*, *11*, 236–242.

M. Etesamifar, B. W. (1998), IgE-vermittelte Nahrungsmittelallergien bei 383 Patienten unter Berücksichtigung des oralen Allergie-Syndroms, 451–457.

Mayer, W., Schuller, M., Viehauser, M. C., & Hochegger, R. (2019), Quantification of the allergen soy (Glycine max) in food using digital droplet PCR (ddPCR). *European Food Research and Technology*, *245*, 499–509.

McPherson, M. J., & Møller, S. G. (2006a), An Introduction to PCR. In E. Owen (Ed.), *PCR: The basics*. 1–8. Taylor & Francis Group, 2nd ed.

McPherson, M. J., & Møller, S. G. (2006b), Reagents and instrumentation. In E. Owen (Ed.), *PCR: The basics*. 23–62. Taylor & Francis Group, 2nd ed.

McPherson, M. J., & Møller, S. G. (2006c), Real-time RT-PCR. In E. Owen (Ed.), *PCR: The basics*. 209–231. Taylor & Francis Group, 2nd ed.

McPherson, M. J., & Møller, S. G. (2006d), Understanding PCR. In E. Owen (Ed.), *PCR: The basics*. 9–22. Taylor & Francis Group, 2nd ed.

Messina, M., & Messina, V. (2010), The role of soy in vegetarian diets. *Nutrients*, *2*, 855–888.

Moore, D. D., & Dowhan, D. (2012), Preparation and Analysis of DNA. In F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J.G. Seidman, J. A. Smith, & K. Struhl (Eds.), *Current protocols in molecular biology*. 2.1.1–2.13.3. John Wiley & Sons, Ltd.

Morinaga Institute of Biological Science (2021), Food Allergen ELISA Kit Ⅱ. https://www.miobs-e.com/product/index.html#link01, accessed 09/20/2021.

Morisset, M., Moneret-Vautrin, D. A., Kanny, G., Guénard, L., Beaudouin, E., Flabbée, J., & Hatahet, R. (2003), Thresholds of clinical reactivity to milk, egg, peanut and sesame in immunoglobulin E-dependent allergies: evaluation by double-blind or single-blind placebo-controlled oral challenges. *Clinical and Experimental Allergy*, *33*, 1046–1051.

Mülhardt, C. (2013), Die Polymerase-Kettenreaktion (PCR). In *Der Experimentator Molekularbiologie/Genomics*. 81–132. Springer Berlin Heidelberg, 7., aktualisierte Auflage.

Müller, H.-J., & Prange, D. R. (2015). *PCR - Polymerase-Kettenreaktion.* 1–8, 65–76. Springer Berlin Heidelberg, 2nd ed.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., & Erlich, H. (1986), Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology*, *51 Pt 1*, 263–273.

Mullis, K. B., & Faloona, F. A. (1987), Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, *155*, 335–350.

Muraro, A., Werfel, T., Hoffmann-Sommergruber, K., Roberts, G., Beyer, K., Bindslev-Jensen, C., Cardona, V., Dubois, A., duToit, G., Eigenmann, P., Fernandez Rivas, M., Halken, S., Hickstein, L., Høst, A., Knol, E., Lack, G., Marchisotto, M. J., Niggemann, B., Nwaru, B. I., Papadopoulos, N. G., Poulsen, L. K., Santos, A. F., Skypala, I., Schoepfer, A., van Ree, R., Venter, C., Worm, M., Vlieg-Boerstra, B., Panesar, S., Silva, D. de, Soares-Weiser, K., Sheikh, A., Ballmer-Weber, B. K., Nilsson, C., Jong, N. W. de, & Akdis, C. A. (2014), EAACI food allergy and anaphylaxis guidelines: diagnosis and management of food allergy. *Allergy*, *69*, 1008–1025.

Murray, M. G., & Thompson, W. F. (1980), Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*, *8*, 4321–4325.

Niggemann, B., & Beyer, K. (2014), Factors augmenting allergic reactions. *Allergy*, *69*, 1582–1587.

Nucleotide BLAST: Search nucleotide databases using a nucleotide query (2021). https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome, accessed 08/04/2021.

Nwaru, B. I., Hickstein, L., Panesar, S. S., Roberts, G., Muraro, A., & Sheikh, A. (2014), Prevalence of common food allergies in Europe: a systematic review and meta-analysis. *Allergy*, *69*, 992–1007.

Ogawa, A., Samoto, M., & Takahashi, K. (2000), Soybean allergens and hypoallergenic soybean products. *Journal of Nutritional Science and Vitaminology*, *46*, 271–279.

Ortolani, C., Ispano, M., Pastorello, E., Bigi, A., & Ansaloni, R. (1988), The oral allergy syndrome. *Annals of Allergy*, *61*, 47–52.

Osborne, N. J., Koplin, J. J., Martin, P. E., Gurrin, L. C., Lowe, A. J., Matheson, M. C., Ponsonby, A.-L., Wake, M., Tang, M. L. K., Dharmage, S. C., & Allen, K. J. (2011), Prevalence of challenge-proven IgE-mediated food allergy using population-based sampling and predetermined challenge criteria in infants. *Journal of Allergy and Clinical Immunology*, *127*, 668-676.e1-2.

Pádua, I., Moreira, A., Moreira, P., & Barros, R. (2016), Food allergy: practical approach on education and accidental exposure prevention. *European Annals of Allergy and Clinical Immunology*, *48*, 174–181.

Paige, D. M. (2005), Lactose intolerance. In *Encyclopedia of Human Nutrition*. 113–120. Elsevier Ltd.

Parkes, H., & Saunders, V. (2007), Quantification of Total DNA by Spectroscopy. In *Analytical Molecular Biology: Quality and validation*. 47–57. Royal Society of Chemistry.

Peters, R. L., Koplin, J. J., Gurrin, L. C., Dharmage, S. C., Wake, M., Ponsonby, A.-L., Tang, M. L. K., Lowe, A. J., Matheson, M., Dwyer, T., & Allen, K. J. (2017), The prevalence of food allergy and other allergic diseases in early childhood in a population-based study: HealthNuts age 4-year follow-up. *Journal of Allergy and Clinical Immunology*, *140*, 145-153.e8.

Pfam (2021). http://pfam.xfam.org/, accessed 08/17/2021.

Platteau, C., Loose, M. de, Meulenaer, B. de, & Taverniers, I. (2011), Detection of allergenic ingredients using real-time PCR: a case study on hazelnut (Corylus avellena) and soy (Glycine max). *Journal of Agricultural and Food Chemistry*, *59*, 10803–10814.

Poms, R. E., Anklam, E., & Kuhn, M. (2004a), Polymerase chain reaction techniques for food allergen detection. *Journal of AOAC International*, *87*, 1391–1397.

Poms, R. E., Klein, C. L., & Anklam, E. (2004b), Methods for allergen analysis in food: a review. *Food Additives & Contaminants*, *21*, 1–31.

Popping, B., & Diaz-Amigo, C. (2010), Detection of Food Allergens. In Popping, Bert; Diaz-Amigo, Carmen; Hoenicke, Katrin, *Molecular biological and immunological techniques and applications for food chemists*. 175–198. Wiley.

Promega Corporation (2007–2015), Maxwell® 16 Instrument Operating Manual: Instructions for Use of Product AS2000. https://at.promega.com/resources/protocols/technical-manuals/0/maxwell-16-instrument-as2000-operating-manual/, accessed 05/19/2021.

Promega Corporation (2016–2020), Maxwell® RSC PureFood GMO and Authentication Kit Technical Manual: Instructions for Use of Product AS1600. https://at.promega.com/products/nucleic-acid-extraction/genomic-dna/maxwell-rsc-purefood-gmo-and-authentication-kit/?catNum=AS1600, accessed 05/19/2021.

Puente-Lelievre, C., & Eischeid, A. C. (2018), Development and Evaluation of a Real-Time PCR Multiplex Assay for the Detection of Allergenic Peanut Using Chloroplast DNA Markers. *Journal of Agricultural and Food Chemistry*, *66*, 8623–8629.

Qiagen (2011), QuantiTect® Multiplex PCR Handbook: QuantiTect Multiplex PCR NoROX Kit - with master mix free of ROX passive reference dye. https://www.qiagen.com/at/products/discovery-and-translational-research/pcr-qpcr-dpcr/real-time-pcr-enzymes-and-kits/probe-based-qpcr/quanti-tect-pcr-kits/?catno=204545, accessed 07/16/2021.

Qiagen (2019), QIAxpert® User Manual. https://www.qiagen.com/us/products/instruments-and-automation/quality-control-fragment-analysis/qiaxpert-system/, accessed 09/20/2021.

Rahaman, T., Vasiljevic, T., & Ramchandran, L. (2016), Effect of processing on conformational changes of food proteins related to allergenicity. *Trends in Food Science & Technology*, *49*, 24–34.

Rancé, F. (2003), Mustard allergy as a new food allergy. *Allergy*, *58*, 287–288.

Rancé, F., Dutau, G., & Abbal, M. (2000), Mustard allergy in children. *Allergy*, *55*, 496–500.

Rizzo, G., & Baroni, L. (2018), Soy, Soy Foods and Their Role in Vegetarian Diets. *Nutrients*, *10*, (43) 1–51.

Rona, R. J., Keil, T., Summers, C., Gislason, D., Zuidmeer, L., Sodergren, E., Sigurdardottir, S. T., Lindner, T., Goldhahn, K., Dahlstrom, J., McBride, D., & Madsen, C. (2007), The prevalence of food allergy: a meta-analysis. *Journal of Allergy and Clinical Immunology*, *120*, 638–646.

Sampson, H. A. (2004), Update on food allergy. *Journal of Allergy and Clinical Immunology*, *113*, 805-19; quiz 820.

Sampson, H. A. (2005), Food allergy--accurately identifying clinical reactivity. *Allergy*, *60 Suppl 79*, 19–24.

Sampson, H. A., & Burks, A. W. (1996), Mechanisms of food allergy. *Annual Review of Nutrition*, *16*, 161–177.

Sasaki, M., Koplin, J. J., Dharmage, S. C., Field, M. J., Sawyer, S. M., McWilliam, V., Peters, R. L., Gurrin, L. C., Vuillermin, P. J., Douglass, J., Pezic, A., Brewerton, M., Tang, M. L. K., Patton, G. C., & Allen, K. J. (2018), Prevalence of clinic-defined food allergy in early adolescence: The SchoolNuts study. *Journal of Allergy and Clinical Immunology*, *141*, 391-398.e4.

Saski, C., Lee, S.-B., Daniell, H., Wood, T. C., Tomkins, J., Kim, H.-G., & Jansen, R. K. (2005), Complete chloroplast genome sequence of Gycine max and comparative analyses with other legume genomes. *Plant Molecular Biology*, *59*, 309–322.

Savage, J., & Johns, C. B. (2015), Food allergy: epidemiology and natural history. *Immunology and Allergy Clinics of North America*, *35*, 45–59.

Savage, J., Sicherer, S., & Wood, R. (2016), The Natural History of Food Allergy. *The Journal of Allergy and Clinical Immunology: In Practice*, *4*, 196-203; quiz 204.

Schaafsma, G. (2000), The protein digestibility-corrected amino acid score. *The Journal of Nutrition*, *130*, 1865S–1867S.

Scharf, A., Kasel, U., Wichmann, G., & Besler, M. (2013), Performance of ELISA and PCR methods for the determination of allergens in food: an evaluation of six years of proficiency testing for soy (Glycine max L.) and wheat gluten (Triticum aestivum L.). *Journal of Agricultural and Food Chemistry*, *61*, 10261–10272.

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012), PCR inhibitors - occurrence, properties and removal. *Journal of Applied Microbiology*, *113*, 1014–1026.

Senyuva, H. Z., Jones, I. B., Sykes, M., & Baumgartner, S. (2019), A critical review of the specifications and performance of antibody and DNA-based methods for detection and quantification of allergens in foods. *Food Additives & Contaminants. Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment*, *36*, 507–547.

Shin, J., Kim, M.-J., & Kim, H.-Y. (2021), Development of triplex PCR for simultaneous detection of soybean and wheat. *Food Science and Biotechnology*, *30*, 801–806.

Sicherer, S. H. (2011), Epidemiology of food allergy. *Journal of Allergy and Clinical Immunology*, *127*, 594–602.

Sicherer, S. H., & Sampson, H. A. (2006), 9. Food allergy. *Journal of Allergy and Clinical Immunology*, *117*, 470-475.

Sicherer, S. H., & Sampson, H. A. (2014), Food allergy: Epidemiology, pathogenesis, diagnosis, and treatment. *Journal of Allergy and Clinical Immunology*, *133*, 291-307; quiz 308.

Sicherer, S. H., & Sampson, H. A. (2018), Food allergy: A review and update on epidemiology, pathogenesis, diagnosis, prevention, and management. *Journal of Allergy and Clinical Immunology*, *141*, 41–58.

Soares, S., Amaral, J. S., Oliveira, M. B. P. P., & Mafra, I. (2014), Quantitative detection of soybean in meat products by a TaqMan real-time PCR assay. *Meat Science*, *98*, 41–46.

Souci, S. W., Fachmann, W., Kraut, H., & Andersen, G. (Eds.) (2016). *Food composition and nutrition tables: Die @Zusammensetzung der Lebensmittel, Nährwert-Tabellen*. MedPharm Scientific Publishers an imprint of Wissenschaftliche Verlagsgesellschaft, 8th revised and completed edition - 8., revidierte und ergänzte Auflage - 8e édition, revue et complétée.

Sugiura, M. (2005), History of chloroplast genomics. In Govindjee, J. F. Allen, J. T. Beatty, & H. Gest (Eds.), *Discoveries in Photosynthesis*. 1057–1063. Springer.

Taylor, S. L., Crevel, R. W. R., Sheffield, D., Kabourek, J., & Baumert, J. (2009), Threshold dose for peanut: risk characterization based upon published results from challenges of peanut-allergic individuals. *Food and chemical toxicology: British Industrial Biological Research Association*, *47*, 1198–1204.

Taylor, S. L., Hefle, S. L., Bindslev-Jensen, C., Bock, S. A., Burks, A. W., Christie, L., Hill, D. J., Host, A., Hourihane, J. O., Lack, G., Metcalfe, D. D., Moneret-Vautrin, D. A., Vadas, P. A., Rance, F., Skrypec, D. J., Trautman, T. A., Yman, I. M., & Zeiger, R. S. (2002), Factors affecting the determination of threshold doses for allergenic foods: how much is too much? *Journal of Allergy and Clinical Immunology*, *109*, 24–30.

Terry, C. F., Della Shanahan, J., Ballam, L. D., Harris, N., McDowell, D. G., & Parkes, H. C. (2002a), Real-Time Detection of Genetically Modified Soya Using Lightcycler and ABI 7700 Platforms with TaqMan, Scorpion, and SYBR Green I Chemistries. *Journal of AOAC International*, *85*, 938–944.

Terry, C. F., Harris, N., & Parkes, H. C. (2002b), Detection of Genetically Modified Crops and Their Derivatives: Critical Steps in Sample Preparation and Extraction. *Journal of AOAC International*, *85*, 768–774.

Thermo Fisher Scientific (Life Technologies) (2014), Real-time PCR handbook. https://www.thermofisher.com/content/dam/LifeTech/global/Forms/PDF/real-time-pcr-handbook.pdf, accessed 08/16/2021.

Thermo Fisher Scientific (Life Technologies) (2019), PCR Primer Design Tips. https://www.thermofisher.com/blog/behindthebench/pcr-primer-design-tips/, accessed 09/21/2021.

Tortajada-Genaro, L. A., Santiago-Felipe, S., Morais, S., Gabaldón, J. A., Puchades, R., & Maquieira, Á. (2012), Multiplex DNA detection of food allergens on a digital versatile disk. *Journal of Agricultural and Food Chemistry*, *60*, 36–43.

University of Nebraska–Lincoln (2021), Food Allergy Research and Resource Program (FARRP) AllergenOnline.org database. http://www.allergenonline.org/databasebrowse.shtml, accessed 06/25/2021.

University of Portsmouth (2013), Literature searches and reviews related to the prevalence of food allergy in Europe. *EFSA Supporting Publications*, *10*, 1–343.

Untersmayr, E., & Jensen-Jarolim, E. (2006), Mechanisms of type I food allergy. *Pharmacology & Therapeutics*, *112*, 787–798.

van Loon, L. C., & van Strien, E. A. (1999), The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiological and Molecular Plant Pathology*, *55*, 85–97.

van Ree, R. (2002), Clinical importance of non-specific lipid transfer proteins as food allergens. *Biochemical Society Transactions*, *30*, 910–913.

Verhoeckx, K. C. M., Vissers, Y. M., Baumert, J. L., Faludi, R., Feys, M., Flanagan, S., Herouet-Guicheney, C., Holzhauser, T., Shimojo, R., van der Bolt, N., Wichers, H., & Kimber, I. (2015), Food processing and allergenicity. *Food and chemical toxicology: British Industrial Biological Research Association*, *80*, 223–240.

Vidal, C., Pérez-Carral, C., & Chomón, B. (1997), Unsuspected sources of soybean exposure. *Annals of Allergy, Asthma & Immunology*, *79*, 350–352.

Waiblinger, H.-U. (2010), Die Untersuchung auf gentechnische Veränderungen („GVO-Analytik"). In *Molekularbiologische Methoden in der Lebensmittelanalytik*. 143–167. Springer Berlin Heidelberg.

WHO/IUIS Allergen Nomenclature Sub-Committee (2021), Allergen Nomenclature. http://www.allergen.org/search.php?allergenname=&allergensource=&TaxSource=&TaxOrder=&foodallerg=all&bioname=&browse=Browse, accessed 07/26/2021.

Xu, J., Ye, Y., Ji, J., Sun, J., & Sun, X. (2021), Advances on the rapid and multiplex detection methods of food allergens. *Critical Reviews in Food Science and Nutrition*, 1–21.

Yeung, J. (2006), Enzyme-linked immunosorbent assays (ELISAs) for detecting allergens in foods. In S. J. Koppelman, & S. L. Hefle (Eds.), *Detecting allergens in food*. 109–124. Woodhead Publishing Limited.

Yunginger, J. W. (1988), Fatal Food-Induced Anaphylaxis. *Journal of the American Medical Association*, *260*, 1450–1452.

Zhang, W., Zhao, Y., Xu, Q., & Chen, Q. (2018), Development of a triplex real-time PCR for simultaneous detection of allergenic ingredients in processed food. *Czech Journal of Food Sciences*, *36*, 22–27.