Collaboration

Validation of the real-time PCR method for detecting the allergen soya in foods

Mª del Camino Martín Forero, Silvia Gil, Mercedes Fernández de la Puebla*, Marisa Salvador, Mª Isabel Rodriguez, Mirian Gómez* and Mª Isabel Prieto
Biotechnology Service, National Food Centre, Spanish Agency for Consumer Affairs, Food Safety and Nutrition
*Tragsatec

Abstract

Soya *(Glycine max)* is considered one of the "classic food allergens" responsible for causing reactions in children and adults. Soya is one of the substances causing allergies or intolerances listed in annex II of Regulation (EU) No 1169/2011.

The use of molecular biology methods to detect allergens has developed and spread in recent years to the extent that they compete with immunological methods, until now considered the methods of choice. In this respect, regulations UNE-EN 15842:2010 and UNE-CEN/TR 16338 IN.:2012 provide information on the criteria which must be met by molecular biology methods in the detection of food allergens.

The method chosen for detecting soya residue in foods is based on the detection of an 81 bp (base pair) fragment of the lectin (*le1*) gene using the real-time polymerase chain reaction (real-time PCR) technique.

The aim of the study was to fine-tune and subsequently validate the selected method. The validation studies carried out demonstrated that the method is sufficiently sensitive and specific for use in the detection of traces of soya in raw and processed foods. It was possible to detect 10 copies of the soya haploid genome and allowed the detection at greater than 95 % sensitivity of 5 mg/kg soya in food samples.

Key words

Allergens, soya, processed foods, real-time PCR.

Translated from the original published in the journal: Revista del Comité Científico de la AECOSAN, 23, pp: 45-57

1. Introduction

Soya is considered a "classic food allergen" and is one of the foods responsible for causing allergic reactions in children and adults. Allergic reactions to soya mainly take the form of skin and stomach problems, but they also include respiratory symptoms that can at times lead to severe allergic reactions.

Therefore, soya is one of the substances listed under annex II of Regulation (EU) No 1169/2011, requiring all substances that may cause allergies or intolerances to be labelled (EU, 2011).

Soya (*Glycine max*) is an edible legume that belongs to the *Fabaceae* family; the seed contains around 20 % oil and between 38-40 % protein. Consumption of soya is widespread in Asia and the United States, where the incidence of allergies to soya is 2.7 % in children under 3 years of age and between 0.1 and 1.8 % in adults (Greenhawt et al., 2009). The incidence is presently much lower in Europe, but it must be taken into consideration that soya consumption has increased considerably in recent years; this is due to its technological and nutritional power, making it possible for us to find soya as part of natural seasonings, stock, starch and chewing gum of plant origin, sauces, etc., in addition to the fact that it is considered a food that provides several health benefits, and without forgetting about the increasingly frequent use of vegetarian cooking.

Soya consumption is also frequent in children when soya-based formulas are introduced into childhood nutrition and utilised as an alternative in the treatment of allergies to milk proteins or lactose intolerance. This rise in soya consumption has led to an increase in the appearance of individuals who are allergic to this legume, thus making it increasingly necessary to have safe and reliable techniques for its detection in foods.

The use of molecular biology methods for detecting allergens has developed and spread in recent years, to the extent that they compete with immunological methods which were considered the methods of choice until now. In the specific case of soya, many studies have been carried out to demonstrate the use of polymerase chain reaction (PCR) methods in detecting soya DNA and which show their effectiveness in detecting soya in processed and unprocessed foods (Soares et al., 2010).

In this regard, Standard UNE-EN 15634-1:2009 provides the broad outline on the characteristics that must be met by molecular biology methods used to detect sequences corresponding to species containing allergens, by using PCR in food matrices.

The aim of this study has therefore been to fine-tune and subsequently validate a method for detecting soya in food using real-time PCR.

2. Materials and methods

2.1 Materials

With the aim of assessing the possible use of the selected method for the detection of soya in complex foods susceptible to containing residual soya, the behaviour of the method was tested on different matrices.

The matrices were chosen bearing in mind the following characteristics:

- Foods which, by their nature, may be involved in allergic reactions to soya owing to contamination or adulteration.
- Foods that are processed and formed by complex mixtures, in such a way that it would be of use to us in testing the effectiveness of the method and in guaranteeing that no disturbances occurred during the trial.
- Foods selected in accordance with their labelling.

Bearing in mind the characteristics described, the samples chosen were divided into four types of products.

- Meat products (cooked ham, sausages, hamburger and chorizo).
- Bread products and baked goods.
- Liquid and semi-liquid products (milkshakes, yogurt, juices…).
- Products such as sauces, prepared soups, balsamic vinegar, margarine.

A total of 26 samples were analysed in order to carry out the validation trials. From the types that were described, 10 labelled as containing soya or traces of soya among their ingredients were chosen, and 16 were chosen which did not list soya among their ingredients.

As no certified reference material is currently available, the fortified samples for determining the limit of detection were prepared using a sample of 100 % soya flour as reference material for the fortification.

The loading buffer was prepared in a concentration of 1 mg/ml in a 6M urea buffer (6M Ultrapure urea, 20 mM Phosphate-buffered saline (PBS), 0.1 % Tween 20). The fortification in the desired concentration was performed on 5 grams of blank matrix.

2.2 Methods

The analytical process consisted in extracting DNA from the samples, then subsequently amplifying a specific target sequence of soya using Real-time PCR.

2.2.1 DNA Extraction

To perform DNA extraction, a representative amount from each of the samples was crushed and homogenised in order to achieve the correct distribution of the components in the amount in question.

DNA extraction and purification was carried out on 1 gram of the homogenised sample, following the extraction procedure based on the CTAB method (Murray and Thompson, 1980).

Each of the samples was extracted twice and the purified DNA was diluted with 100 µl of TE buffer solution (10 mM Tris-ClH, 1 mM EDTA, pH 8.0).

The quantification and purity of the DNA obtained was analysed by means of the absorbance measurement at 280 nm and the A260/A280 ratio calculation using a NanoDrop 2000 spectrophotometer.

In the majority of the samples, the results obtained for concentration and purity were considered

satisfactory to continue with the analytical process and to proceed to the amplification of the target sequence by means of PCR.

In the cases for which the results obtained for concentration and purity in the DNA extraction by means of the CTAB method were not suitable enough to continue the trials, we used, as an alternative extraction method, the commercial Kit (SureFood® PREP Avanced Art. No S1053, R-Biopharm), especially recommended for DNA extraction in highly processed food.

The soya flour used as reference material was extracted at the same time as the samples using the CTAB method.

The DNA solution obtained was diluted when necessary with TE buffer solution in order to achieve a working solution at a concentration of 50 ng/µl.

2.2.2 Detection of soya using real-time PCR

The method chosen for detecting soya residue in foods is based on the detection of an 81 bp (base pair) fragment of the lectin (*le1*) gene using the real-time polymerase chain reaction (real-time PCR) technique.

The lectin (*le1*) gene was chosen due to the fact that it is an endogenous gene present in all varieties of soya (*Glycine max*).

The PCR reaction was prepared in a final volume of 25 µl, for which 12.5 µl of 2x TaqMan® Universal Master Mix (Applied Biosystems), 5 µl of the DNA extracted in a concentration of 50 ng/µl, 500 nM of each of the primers and 200 nM of probe were mixed. The PCR reactions were conducted in the Applied Biosystems 7500 system, using the following programme: 2 minutes at 50 ºC, 10 minutes at 95 ºC and 45 cycles of 15 seconds at 95 ºC and 1 minute at 60 ºC.

The oligonucleotide sequences used as primers and the probe marked with TAMRA and FAM fluorophores were synthesised by the Metabion GMBH trade house (Table 1).

Table 1. Sequences used in the amplification of the lectin (*le1*) gene

Gene (lectin *Le1*)	**Sequence**
TM-Lectin-F	5´-TCC ACC CCC ATC CAC ATT T-3´
TM-Lectin-R	5´-GGC ATA GAA GGT GAA GTT GAA GGA-3´
Lectin-p	FAM 5´-AAC CGG TAG CGT TGC CAG CTT CG-3´TAMRA

3. Validation requirements

The criteria and pre-requisites necessary for the soya detection method in food samples using PCR were established bearing in mind Standards UNE-CEN/TR 16338 IN.:2012 and UNE-EN 15842:2010 (standards referring to the criteria that must be met by molecular biology methods for the detection of allergens):

- Applicability. Demonstrate that the method is effective for the intended purpose. In this case, specifically, it should be demonstrated that the method is useful for the detection of traces of soya in food samples.

- Specificity. An attribute of the method to respond exclusively to the analyte of interest. The method must comply with specificity 100 %. The specificity of the target sequence should be verified both theoretically and experimentally.
- Sensitivity. The sensitivity calculated as the coefficient of the number of true positive samples divided by the number of known positive samples. Sensitivity must be ≥95 % (UNE-CEN/TR 16338 IN.:2012).
- Limit of detection (LD). The minimum amount of analyte that can be detected with a confidence level greater than or equal to 95 %. The limit of detection must be appropriate for the intended purpose.

3.1 Experimental design for validation of the method

The validation studies were designed in order to test the effectiveness of the method for the intended purpose.

3.1.1 Method applicability

The method applicability was demonstrated through the analysis of the 10 selected samples, based on their content of soya or traces of soya, in order to demonstrate that the extraction process and subsequent amplification of the selected target could be considered adequate for the detection of soya in most types of foods that are possibly involved in the soya allergic reaction (Table 2).

In order to do so, the concentration and purity of the DNA obtained were tested as indicated in section 2.2.1. The amplification reaction was subsequently carried out using real-time PCR as described in section 2.2.2.

The control used for the reaction was DNA extracted from soya flour with a concentration of 0.001 % soya DNA in the PCR reaction.

3.1.2 Specificity

The specificity of the chosen target sequence is evidenced and published in the methods validated by the European Reference Laboratory for Genetically Modified Organisms (EURL-GMFF, 2016). According to these methods, the selected sequence has been designed to recognise a sequence in the specific soya gene (*le1*) as the target, described in the GenBank database. No similarities have been found between the target sequence and DNA sequences of other cultivated plants (legumes, vegetables and cereals) during the searches performed in databases such as NCBI, BlastN and EMBL, thus the reason why it was not considered necessary to conduct theoretical studies to verify the specificity.

However, to ensure that no interferences or non-specific reactions occurred with the rest of the ingredients or components that form part of the complex samples, the specificity was experimentally confirmed through the analysis of the 16 food samples chosen to carry out the validation process based on the absence of soya on food labels (Table 3).

Likewise, it was confirmed that no amplification occurred with other plant varieties that are closely-related or susceptible to cause cross-reactivity with soya (Table 3).

3.1.3 Limit of detection

The limit of detection was determined under two aspects; as the absolute limit of detection, understood as the number of copies of the soya haploid genome that the trial is able to detect, and as the real or experimental limit of detection, considered to be the minimum amount of soya that the trial is capable of detecting in a certain sample, expressed as mg of soya per kg of food.

3.1.3.1 Absolute limit of detection

The absolute limit of detection was determined by verifying the minimum amount of soya DNA which is capable of producing an amplification signal in real-time PCR below cycle 40.

In order to do so, five independent amplification trials were conducted, under conditions of reproducibility, making two PCR replicas for each point, on a series of dilutions obtained from the DNA extracted from soya flour, so that the DNA concentration in the PCR would be between 1 000 pg and 10 pg.

The transformation of pg of DNA into a number of copies was carried out bearing in mind that a copy of the soya genome corresponds to 1.13 pg of DNA (Arumuganathan and Earle, 1991), in such a way that we would have values between 885 and 8.8 copies of the soya haploid genome in the amplification reaction.

At the same time, a graphic representation was made for the values of the corresponding cycle of amplification (Ct) compared to the logarithm of the concentration in the number of copies in order to verify that the reaction took place with sufficient linearity and efficiency.

The amplification reaction was considered to be acceptable when:

- $R^2 \geq 0.98$
- The efficiency calculated from the slope of the line was between 80 and 110 %

$$\text{Efficiency} = [10^{(-1/\text{slope})}] - 1 \times 100$$

3.1.3.2 Experimental limit of detection

The experimental limit of detection, expressed as mg of allergenic constituent per kg of food, was verified by analysing each of the 16 blank matrices selected, fortified at low concentration levels as indicated in section 2.1.

Duplicate analyses were performed in two independent analytical processes, under conditions of internal reproducibility, and making 10 PCR replicas in each one.

The real limit of the method was considered to be the lowest concentration of the analyte that the method is capable of detecting in the different sample types with a 95 % confidence interval.

3.1.4 Sensitivity

Sensitivity was determined from the total results obtained for each of the four types of matrices fortified at the limit of detection of the trial, i.e. from 80 results for each type of matrix, from 10 replicas, 2 independent trials, 4 different samples per matrix group.

Sensitivity was calculated according to Standard UNE-CEN/TR 16338 IN.:2012 as the coefficient of the number of confirmed positive samples divided by the number of known positive samples. Sensitivity should be ≥95 %.

4. Results

4.1 Method applicability

The results obtained in the 10 samples labelled as containing soya or traces of soya that are reflected in Table 2 demonstrate that the extraction process is appropriate for its application to those types of matrices, as the concentration was superior to the concentration established for use as the working concentration in the majority of the cases, 50 ng/µl, and the purity was within the accepted limits.

It was also confirmed that amplification of soya DNA was detected in almost all cases at amplification cycle values below the accepted maximum, cycle 40; we could therefore consider the amplification reaction to be appropriate, and no interferences or inhibitions owing to complex matrices were observed.

Table 2. Results obtained in the extraction of soya samples

Sample	Labelling	DNA Concentration (ng/µl)	DNA Purity A260/A280	Amplification Cycle (*Le1* gene)
Sausages	Contains soya	398.1	1.9	32.22
Chicken soup	May contain traces of soya	121.1	1.85	36.51
"Special Line" cookies	Soya protein	120.8	1.9	31.36
Beef hamburger	Soya protein	223.3	1.8	29.15
Yoghurt	Soya	27.3	1.5	32.25
Wheat cereal with honey	May contain soya	156.1	1.6	35.27
Rice snack	Contains soya sauce	185.7	1.7	37.58
Chocolate drink	Contains soya	85.1	1.6	35.42
Muffins	May contain traces of soya	137.4	1.8	36.12
Bread with seeds	May contain traces of soya	269.9	1.9	35.27

4.2 Specificity

Theoretical specificity was confirmed by analysing 16 samples for which soya was not listed on the food label, as well as in plant varieties that are closely-related or others that may cause cross-reactivity.

The results showed that no amplification was detected in the absence of soya, and that no non-specific reactions occurred as a result of the ingredients present in the samples that may have led to the detection of false positives (Table 3).

As we can see in Table 3, the amount of DNA extracted from the samples was considered to be sufficient to be able to continue with the analytical process. In most of the cases the concentration was superior to the working concentration, 50 ng/µl. In highly-processed samples in which a concentration of 50 ng/µl was not reached, the concentrations obtained were close to 20 ng/µl and were thus considered acceptable for continuing on with the trials.

Table 3. Results obtained in samples for which soya was not listed as an ingredient

	Sample	**Method**	**DNA Concentration (ng/µl)**	**DNA Purity A280/A260**	**Amplification Cycle (*Le1* gene)**
Meat products	Chorizo	CTAB	908.4	1.89	n.d.[1]
	Hamburger	CTAB	116.7	1.85	n.d.
	Cooked ham	CTAB	105.5	1.87	n.d.
	Sausages	CTAB	453.4	1.91	n.d.
Baked goods	Butter sponge cakes	SureFood	66.8	1.64	n.d.
	Muffins	CTAB	137.4	1.79	n.d.
	Sponge-finger cakes	CTAB	228.5	1.87	n.d.
	Sliced bread	CTAB	427.9	1.89	n.d.
Various	Cream of asparagus	CTAB	133.7	1.84	n.d.
	Chocolate ice cream	SureFood	38.9	1.8	n.d.
	Margarine	SureFood	24.7	1.5	n.d.
	Pink sauce	CTAB	19.6	1.36	n.d.
Liquids	Chocolate milkshake	CTAB	18.9	1.39	n.d.
	Yoghurt	SureFood	18.4	1.17	n.d.
	Pudding	CTAB	34.5	1.73	n.d.
	Fruit juice	SureFood	72.2	1.47	n.d.

[1]n.d.: not detected.

The absence of amplification was similarly confirmed in those plant varieties that are closely-related or a possible cause of cross-reactivity by immunological methods (Table 4).

Table 4. Results of closely-related plant varieties	
Plant variety	**Amplification (*Le1* gene)**
Peanut	n.d.[1]
Peas	n.d.
Wheat	n.d.
Barley	n.d.

[1]n.d.: not detected.

To ensure the results, in all samples it was confirmed that the negative amplification result was due to the lack of said component in the sample and was not a result of inhibition processes by some matrix component; an inhibition trial was thus conducted on each of the matrices by adding 1 µl of DNA extracted from soya flour onto the DNA extracted from the sample, and then by conducting a new amplification trial.

Amplification was detected in all cases, and the possibility that the negative result was due to the presence of inhibiting substances in the sample was thus ruled out (data not shown).

4.3 Limit of detection of the trial

4.3.1 Absolute limit of detection (number of copies)

The results obtained in the five independent trials conducted over the series of dilutions between 1 000 pg and 10 pg, prepared from DNA extracted from soya flour, demonstrated that the trial was capable of amplifying -at amplification cycle values lower than cycle 40- a concentration of 10 pg of DNA in the real-time PCR reaction; this would correspond to 8.8 copies bearing in mind that a copy of the soya genome is equivalent to 1.13 pg (Table 5).

Table 5. Results for the absolute limit of detection

Concentration		**Curve 1**	**Curve 2**	**Curve 3**	**Curve 4**	**Curve 5**
pg (PCR)	**Copies of GHP**					
1 000	885	29.12	28.8	28.87	29.34	29.3
200	177	31.31	31.27	30.87	31.87	31.47
40	35.4	34.1	33.58	33.24	33.94	34.76
20	17.7	35.59	34.353	35.3	35.94	34.78
10	8.8	36.97	35.45	36.33	36.22	36.16
Slope		-3.93	-3.28	-3.77	-3.55	-3.45
R^2		0.99	0.99	0.98	0.984	0.98
Efficiency		80.5	101.8	84.2	91.3	94.9

This result allows us to determine that the method is capable of reliably detecting concentrations of 10 copies of the soya genome.

From the graph of amplification cycle values in relation to the logarithm of the number of copies, we can verify that the amplification reaction takes place with sufficient linearity and efficiency, as R^2 in all cases was ≥0.98 and efficiency fell between the optimum values of 80 and 110 % (Figure 1).

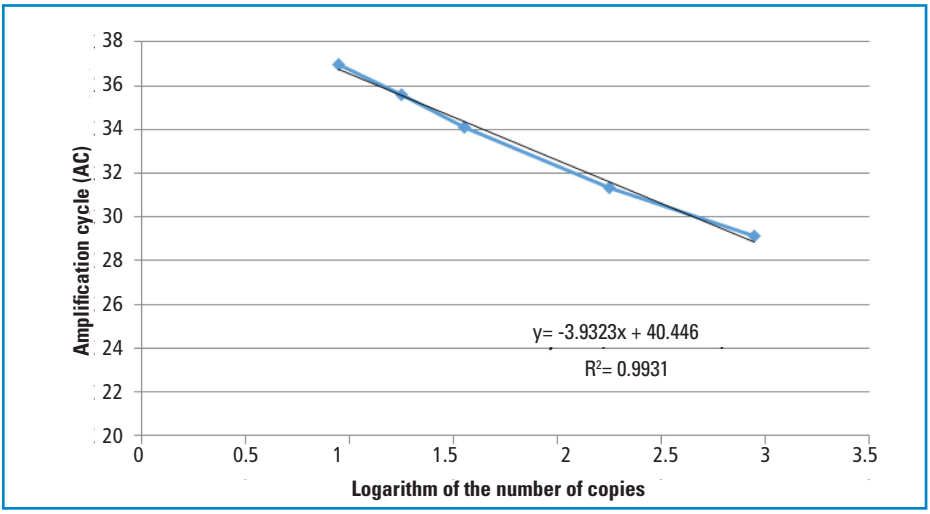

Figure 1. Example of the graph of curve 1.

4.3.2 Experimental limit of detection

An initial trial conducted on the 16 blank samples selected, fortified at two concentration levels (2.5 mg/kg and 5 mg/kg), demonstrated that amplification in the samples fortified with 2.5 mg/kg of soya was not detected on all occasions.

Nonetheless, it was possible to guarantee sufficient reliability in the detection when the samples were fortified with 5 mg/kg of soya (data not shown).

Confirmation of the experimental limit of detection in the 5 mg/kg concentration was carried out with the analysis of the 16 samples chosen from among the four types of matrices that were fortified for that concentration level; two independent trials were conducted under conditions of internal reproducibility using 10 PCR replicas in each trial.

The review of the results obtained for each type of matrix guaranteed that the method is capable of detecting 5 mg/kg of soya in food samples (Table 6).

Table 6. Results obtained for identification of the 5 mg/kg experimental LD

	Sample fortified with 5 mg/kg of soya	**Extraction 1 Ct ± Sd (10 replicas)**	**Extraction 2 Ct ± Sd (10 replicas)**
Meat products	Chorizo	37.50±0.93	36.40±0.65
	Hamburger	37.54±1.30	37.43±0.64
	Cooked ham	36.91±0.85	36.58±1.40
	Sausages	37.42±0.90	35.32±0.65
Baked goods	Butter sponge cakes	37.51±0.66	36.15±0.54
	Muffins	32.50±0.28	35.31±0.60
	Sponge-finger cakes	34.52±0.33	35.63±0.68
	Sliced bread	34.30±0.36	36.25±0.48
Various	Cream of asparagus	36.28±0.76	36.13±0.85
	Chocolate ice cream	36.42±0.56	36.39±1.21
	Margarine	36.13±0.34	38.29±0.72
	Pink sauce	32.36±0.22	35.19±0.74
Liquids	Chocolate milkshake	34.49±0.32	37.49±0.89
	Yoghurt	36.04±1.40	37.54±1.05
	Pudding	32.89±0.23	37.58±1.50
	Fruit juice	34.17±0.51	35.33±0.24

As we can see in the results shown in Table 6, average amplifications obtained in the 10 replicas conducted on the DNA extracted from the samples fortified with 5 mg/kg confirm that amplification was detected at values lower than cycle 40 in all cases, thus making it possible to guarantee that the method is capable of detecting 5 mg/kg soya concentrations with sufficient reliability. The results were confirmed in the two independent trials conducted under conditions of internal reproducibility, thus allowing us to evaluate the method as being compliant with the requirement of internal reproducibility.

4.4 Sensitivity

The results show that the trial sensitivity calculated according to Standard UNE-CEN/TR 16338 IN.:2012 as the coefficient of the number of true positive samples divided by the number of known positive samples (true positives/true positives+false negatives) was within the accepted level of ≥95 % (Table 7) in all cases.

Table 7. Sensitivity of the 5 mg/kg limit of detection according to matrix type

Matrix type fortified with 5 mg/kg of soya	Number of known positives	True positives	False negatives	Sensitivity
Meat products	80	78	2	97.5 %
Baked goods	80	79	1	99.0 %
Various	80	79	1	99.0 %
Liquids	80	76	4	95.0 %
Total	320	312	8	97.5 %

5. Conclusions

1. The analytical method can be utilised for the detection of soya or traces of soya in most of the matrices which may be involved in food allergies to soya.
2. The absolute limit of detection of the trial has been determined at 10 picograms of soya DNA which would theoretically correspond to 10 copies of the soya haploid genome.
3. The real limit of detection of the trial has been established at 5 mg/kg of soya in food.
4. Trial sensitivity is greater than 95 %.
5. The trial method and validation process were presented to Spain's National Accreditation Body (ENAC), for accreditation in May 2015, achieving accreditation -without comments- in accordance with Standard UNE 17025. The method has been included in the Scope of Accreditation of the National Centre for Food (CNA) (Technical Annex 178_LE397 Rev_30).

References

Arumuganathan, K. and Earle, E.D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*, 9, pp: 210-220.

EURL-GMFF (2016). European Union Reference Laboratory for GMO Food & Feed. Compendium of reference methods for GMO analysis. Available at: http://gmo-crl.jrc.ec.europa.eu/ [accesed: 29-06-16].

Greenhawt, M.J., Singer, A.M. and Baptist, A.P. (2009). Food allergy and food allergy attitudes amongh collegue students. *Journal of Allergy and Clinical Immunology,* 124, pp: 323-327.

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

Soares, S., Mafra, I., Amaral, J.S. and Oliveira, M.B.P.P. (2010). A PCR assay to detect trace amounts of soybean in meat sausages. *International Journal of Food Science & Technology*, 45, pp: 2581-2588.

UE (2011). Regulation (EU) No 1169/2011 of the European Parliament and of the Council of 25 October 2011 on the provision of food information to consumers. OJ L 304 of 22 november 22nd, 2011, pp: 18-63.

UNE-EN 15634-1:2009. Productos alimenticios. Detección de alérgenos mediante métodos biológicos moleculares. Parte1: Consideraciones generales.

UNE-EN 15842:2010. Productos alimenticios. Detección de alérgenos alimenticios. Consideraciones generales y validación de los métodos.

UNE-CEN/TR 16338 IN.:2012 Productos alimenticios. Detección de alérgenos alimenticios. Formato para proporcionar información sobre métodos inmunológicos y métodos de biología molecular.