



A proposed method for sampling, detection and quantification of maize kernels using traditional and real-time PCR

Abdulaziz M. Al-Swailem^{*}, Maher M. Shehata, Omar H. Shair, Turki A. Al-Shammari, Ibrahim O. Al-Anazi and Saeed A. Sabaan

*Biotechnology and Genetic Engineering Program, Natural Resources and Environment Research Institute (NRERI), King Abdulaziz City for Science and Technology (KACST), Riyadh 11442, P.O. 6086, Saudi Arabia. *e-mail: alswailem@kacst.edu.sa*

Received 13 September 2006, accepted 27 November 2006.

Abstract

The extent, to which sampling of raw materials and foods for detection of the presence of GMOs presents a significant problem, depends on the type of material to be sampled, the purpose of the analysis and the degree of risk that is acceptable in obtaining a wrong result. Development of internationally accepted, harmonised sampling plans based on sound scientific and statistical principles is urgently needed, and initiatives should be undertaken quickly to ensure the acceptability of GMO products, which to a great extent are traded internationally. PCR methods are available for qualitative screening and quantitation of virtually all DNA-containing food materials. Our work was planned to reflect the definition of sampling protocols for GMO detection and/or quantification, to propose a new statistical model to estimate the sampling error associated to different sampling protocols as function of both number and size of samples taken from any consignment of particulate and to demonstrate the picture of GM distribution in Saudi Arabia. Preliminary results from our simulations indicated that the pattern of convergence to the true contamination value is similar for different contamination levels and lot heterogeneity scenarios. Our results will have a remarkable impact upon the definition of sampling protocols, as it will ensure a proper sampling if non-random distribution of contaminants is observed or expected, as in the case of kernel lots. Our assay method is proved to be suitable for analytical purposes, with excellent limits of detection and quantification.

Key words: Sampling, simulation, Saudi kernels, screening, detection, quantification, GMOs, traceability, limit of detection, limit of quantification.

Introduction

Presently, genetically modified organisms (GMOs) and foods (GMFs) created using tools provided by genetic engineering are more frequently available on the market. It includes delayed ripening tomatoes, bananas or strawberries, nutritionally modified soybeans, rape, potatoes or wheat as well as herbicide resistant maize, apples or bananas. Food is also constantly modified to satisfy extravagant consumer requirements and preferences based mainly on attributes such as appearance, taste and freshness, e.g., sweeter strawberries, crisper celery or pipless grapes¹⁻³. In 1994, the first genetically modified fruit available on the market was FlavrSavr tomato purposed to soften more slowly. In 1997, eleven GMO crops, including corn, soybean, tomato, potato, chicory, squash, melons, papaya, rape seed, cotton and tobacco, were on sale mainly in the United States but also in Europe. They were modified to increase herbicide tolerance, virus and insect resistance, to obtain slower ripening and male sterility and exclusively in case of rape seed to change the oil composition and rise the level of lauric acid⁴. Presently, there are over 30 different fruits, nuts, vegetables and staples under development⁵.

The European Commission has issued directives, Council Decisions and regulations concerning the placing on the market and the mandatory labeling of food containing GMOs⁶. It is important for ethical, economic and legal reasons. Polish producers and importers usually do not inform consumers if particular food

products are prepared using genetically modified ingredients or supplements, although the article 47 introduced in the act from 22nd of June 2001 in Polish Library of Acts clearly orders indication of modified components in traded products. Consumers unaware of that fact may buy food containing transgenic components or supplements (e.g., chocolate, soybean products, cooking oils) or inadvertently feed animals with modified fodder². As products usually are processed in harsh conditions (cooking, frying, pasteurizing, autoclaving, lyophilizing, etc.), their processing may affect further detection of modified genes.

Testing of raw materials and foods containing GM components or supplements may be conducted by testing for the presence of incorporated DNA or detecting expressed novel proteins⁷. The presence of modified DNA sequences may be easily detected by application of widely known polymerase chain reaction (PCR) with primers aimed specifically at 'novel' fragments. PCR was efficiently applied for detection of genetically engineered corn in flour, meal and processed food. The reaction was e.g., designed to detect the presence of *cryIA(b)* gene which encodes *cryIA(b)* protein, a natural toxin of potential insecticidal activity produced by *Bacillus thuringiensis* or 35S promoter and NOS terminator genes in samples from soybean and maize flour collected in different countries⁶.

In genetically modified foods, an uniformly distributed intrinsic component of the food has been changed and the distribution in most cases will be no more heterogeneous compared to the same component in the non-GM food. Thus, in the simplistic situation, where either all the food is GM or all the food is not GM, then a qualitative method could be employed and the manner of sampling would be essentially immaterial. Indeed, it could be argued that unequivocal proof of GM material could be obtained by multiple replicate analysis of single maize kernels or single soya beans. The problem in interpretation of the results would occur if only a small proportion of the kernels or beans were found to contain GM material. The question would then arise, how representative were those kernels/beans that were selected for analysis of the whole batch of the material. This then becomes a classical sampling problem with the same considerations as apply in other areas where there is heterogeneous analyte distribution. Kernel lot sampling is a complex multi-stage procedure that should reduce a lot to an analytical sample of suitable working size, representing the lot. Most kernel sampling plans are based upon the assumption of random distribution of GMOs so that the mean, the standard deviation and both the producer and consumer risks can be estimated according to the binomial or the Poisson distributions. Given the high likelihood of non-detectable strata of GM material in kernel lots⁸, assuming randomness is very risky because even modest deviations from randomness have a strong effect on the accuracy (GMO %) and precision (variance of GMO %) of GMO estimates⁹.

Development of sampling plans should be regarded as an integral part of method development, which follows logically after the stage of establishment of adequately sensitive and precise methodology. Adequate methods for analysis and sampling are essential for the appropriate labeling of foods to increase transparency on production processes and to facilitate traceability thus contributing to strengthen food safety systems. The conventional wisdom in trace analysis is that sampling and sub-sampling errors are significantly greater than analytical errors and that the error increases dramatically with decrease in analyte concentration. The reasoning is based on the assumption which holds true in most cases that analytes (frequently contaminants) are heterogeneously distributed in foods making the taking of a representative sample difficult. Sampling frequently therefore requires the taking of large samples comprising multiple sub-samples, which are subsequently homogenised and mixed, prior to taking of the analytical sample. The extent, to which sampling of raw materials and foods for detection of the presence of GMOs presents a significant problem, depends on the type of material to be sampled, the purpose of the analysis and the degree of risk that is acceptable in obtaining a wrong result¹⁰.

In order to develop a sampling plan based on statistical considerations it is necessary to understand something of the nature of the homogeneity of the batch of material being sampled. The distribution of the analyte to be measured can be described in mathematical terms e.g., a negative binomial distribution is frequently used to describe mycotoxin contamination of groundnuts¹¹. From this distribution curve, a sampling plan can be proposed the characteristics of which are typically described in the form of an operating curve¹². An operating curve indicates the 'Producer Risk' which is the risk (expressed as a probability) of the producer rejecting a batch of material which is within a

desired specification, and the 'Consumer Risk' which is the risk of accepting a batch of material which does not conform with the desired specification.

Although several guidelines defining sampling strategies for quality and purity analyses are available and currently adopted for GMO surveys, their suitability for GMO detection is questionable because of their stringent statistical assumptions with respect to the possible distribution of the contaminants. Indeed, most of these guidelines recognize that the procedures are not effective in the sampling of non-random distributions, a likely situation with respect to the adventitious presence of GM material in bulk commodities, due to segregation during transportation and handling. Yet, given the lack of suitable statistical approaches with respect to situation where sampling procedures based on standard statistical models cannot be applied, currently available sampling protocols continue to be used without verification of their assumptions. Understanding distribution patterns in kernel lots is a pre-requisite to develop and recommend suitable sampling plans with respect to GMO legislative requirements.

From Dec 2001, the Saudi Ministry of Commerce has issued a directive requiring all foodstuffs and pet foods imported into Saudi Arabia to be labelled whether or not the product contains GMOs. They must be marked with a triangle and a warning in both Arabic and English¹³. The Ministerial Directive No 166 placed a total ban on the import of foodstuffs containing GE animal products to the country. GMOs or GE products which are exported to Saudi Arabia must be accompanied by a health certificate by the GMO licensing government agency in the country of origin stating that the GMO ingredients are approved for human consumption¹⁴. In March 2003, a decree was issued that also requires the labelling of all imported and locally produced GE animal feed, planting seed, fruits and vegetables effective from the end of January 2004¹⁵.

The definition of internationally harmonized strategies for the evaluation of GMO safety is a priority and there is a strong interest in sampling schemes to ensure accuracy and precision of GM surveys. Our work is planned to reflect the definition of sampling protocols for GMO detection and/or quantification, to propose a new statistical model to estimate the sampling error associated to different sampling protocols as function of both number and size of samples taken from any consignment of particulate, to improve and develop the qualitative and quantitative assays to detect genetically modified samples and to demonstrate the picture of GM distribution in Saudi Arabia.

Materials and Methods

Plant materials and sampling: Maize kernels were obtained from Ministry of Agriculture. Standardized initial sampling protocol for kernels was used to get the initial sample according to GIPSA. Then, the acceptance of sampling and testing plans were used to control and detect this initial sample, respectively. Six kg maize kernels were collected, grinded, homogenized and distributed to form 2 groups (3 kg each one). Then, each group was subgrouped into 6 subgroups (500 g each one). The first group was used as a control. The other groups were mixed with GM to form different concentrations. After homogenization, each subgroup was re-distributed several times in successive steps to reach to 10 g. This weight was used for the preparation of the test samples (Table 1).

Table 1. Quantification of *CryIA(b)* gene in simulated samples of Saudi maize kernels.

| Sample | Test sample | | | Relative conc. of <i>CryIA(b)</i> |
|---------|-------------|-----------|-------|-----------------------------------|
| | Size (g) | Ratio | % | |
| Control | 0.01 | 0.01/1000 | 0.001 | - |
| | 0.05 | 0.05/1000 | 0.005 | - |
| | 0.1 | 0.1/1000 | 0.01 | - |
| | 0.2 | 0.2/1000 | 0.02 | - |
| | 0.3 | 0.3/1000 | 0.03 | - |
| | 0.4 | 0.4/1000 | 0.04 | - |
| | 0.5 | 0.5/1000 | 0.05 | - |
| | 1 | 1/1000 | 0.1 | - |
| 0.1% GM | 0.01 | 0.01/1000 | 0.001 | 0.002 |
| | 0.05 | 0.05/1000 | 0.005 | 0.004 |
| | 0.1 | 0.1/1000 | 0.01 | 0.006 |
| | 0.2 | 0.2/1000 | 0.02 | 0.010 |
| | 0.3 | 0.3/1000 | 0.03 | 0.008 |
| | 0.4 | 0.4/1000 | 0.04 | 0.007 |
| | 0.5 | 0.5/1000 | 0.05 | 0.005 |
| | 1 | 1/1000 | 0.1 | 0.004 |
| 0.5% GM | 0.01 | 0.01/1000 | 0.001 | 0.026 |
| | 0.05 | 0.05/1000 | 0.005 | 0.044 |
| | 0.1 | 0.1/1000 | 0.01 | 0.096 |
| | 0.2 | 0.2/1000 | 0.02 | 0.143 |
| | 0.3 | 0.3/1000 | 0.03 | 0.143 |
| | 0.4 | 0.4/1000 | 0.04 | 0.094 |
| | 0.5 | 0.5/1000 | 0.05 | 0.088 |
| | 1 | 1/1000 | 0.1 | 0.035 |
| 1% GM | 0.01 | 0.01/1000 | 0.001 | 0.212 |
| | 0.05 | 0.05/1000 | 0.005 | 0.326 |
| | 0.1 | 0.1/1000 | 0.01 | 0.418 |
| | 0.2 | 0.2/1000 | 0.02 | 0.452 |
| | 0.3 | 0.3/1000 | 0.03 | 0.488 |
| | 0.4 | 0.4/1000 | 0.04 | 0.310 |
| | 0.5 | 0.5/1000 | 0.05 | 0.201 |
| | 1 | 1/1000 | 0.1 | 0.199 |
| 2% GM | 0.01 | 0.01/1000 | 0.001 | 0.474 |
| | 0.05 | 0.05/1000 | 0.005 | 0.496 |
| | 0.1 | 0.1/1000 | 0.01 | 0.664 |
| | 0.2 | 0.2/1000 | 0.02 | 0.735 |
| | 0.3 | 0.3/1000 | 0.03 | 0.764 |
| | 0.4 | 0.4/1000 | 0.04 | 0.610 |
| | 0.5 | 0.5/1000 | 0.05 | 0.581 |
| | 1 | 1/1000 | 0.1 | 0.423 |
| 5% GM | 0.01 | 0.01/1000 | 0.001 | 0.712 |
| | 0.05 | 0.05/1000 | 0.005 | 0.880 |
| | 0.1 | 0.1/1000 | 0.01 | 0.986 |
| | 0.2 | 0.2/1000 | 0.02 | 1.580 |
| | 0.3 | 0.3/1000 | 0.03 | 0.901 |
| | 0.4 | 0.4/1000 | 0.04 | 0.846 |
| | 0.5 | 0.5/1000 | 0.05 | 0.776 |
| | 1 | 1/1000 | 0.1 | 0.692 |

Table 2. Primers used in this study.

| Primer | Sequence | Gene | Amplicon (bp) | Reference |
|--------------------|---|-----------------|---------------|------------------------------------|
| 35S-1 | 5'-GCT CCT ACA AAT GCC ATC A-3' | 35S promoter | 195 | Lipp <i>et al.</i> ²⁰ |
| 35S-2 | 5'-GAT AGT GGG ATT GTG CGT CA-3' | | | |
| NOS-1 | 5'-GAA TCC TGT TGC CGG TCT TG-3' | NOS terminator | 180 | Lipp <i>et al.</i> ²⁰ |
| NOS-2 | 5'-TTA TCC TAG TTT GCG CGC TA-3' | | | |
| <i>cryIA(b)</i> -1 | 5'-ACC ATC AAC AGC CGC TAC AAC GAC C-3' | Delta-Endotoxin | 184 | Ehlers <i>et al.</i> ³⁰ |
| <i>cryIA(b)</i> -2 | 5'-TGG GGA ACA GGC TCA CGA TGT CCA G-3' | | | |
| <i>ivr1</i> -1 | 5'-CCG CTG TAT CAC AAG GGC TGG TAC C-3' | Invertase | 226 | Ehlers <i>et al.</i> ³⁰ |
| <i>ivr1</i> -2 | 5'-GGA GCC CGT GTA GAG CAT GAC GAT C-3' | | | |

Powdered certified reference materials (CRM) of Event 176 GM-maize (Novartis/Ciba-Geigy, USA) with 0, 0.1, 0.5, 1, 2 and 5% (w/w) GMO contents were obtained from the Institute for Reference Materials and Measurements (IRMM) and commercialized by Fluka (Buchs, Switzerland).

Kits and PCR primers: High Pure GMO Sample Preparation, LightCycler GMO Screening and LightCycler GMO Maize Quantification Kits were purchased from Roche Applied Science (Germany). PCR primers listed in Table 2 were used in this study. The primers of 35S and NOS were used to detect 35S-promoter and NOS-terminator. The primers of *cryIA(b)* and *ivr1* are specific to structure gene region – in Delta-Endotoxin gene – existed in Event 176 GM-maize, and Invertase gene, respectively.

Equipments: Spectrophotometer Gene Quant (Pharmacia Biotech), two thermal cyclers (Techne and Clemens), LightCycler system or Real-time quantitative-PCR (RTQ-PCR) (Roche Applied Science), submarine gel electrophoresis (Pharmacia Biotech) and Gel Doc 2000 (Bio Rad) were used.

DNA extraction and quantitation: DNA was extracted from the raw materials of maize and certified standards according to the instruction's manual of High Pure GMO Sample Preparation Kit with slight modifications. DNA concentration was quantified using the spectrophotometer. DNA purity was assessed by OD_{260/280} measurements. The integrity of the isolated DNA was further checked by agarose (0.5%) (Solon, Ohio, USA) gel electrophoresis and ethidium bromide staining¹⁶.

Conventional PCR reactions and products analysis: Polymerase chain reactions were carried out on two thermal cyclers (Techne and Clemens). All PCR reactions were performed in final volumes of 25 µl in 0.5 ml tubes containing 50 ng of DNA, 1x reaction buffer (50 mM KCl; 10 mM Tris-HCl, pH 8.2; 0.2 mg BSA), 1.5 mM MgCl₂, 0.2 mM each of dATP, dGTP, dTTP and dCTP; 0.5 mM of each primer and 1 U of *Taq* polymerase (Pharmacia Biotech, Germany). The DNA was incubated in the thermocyclers under the following program: 95°C for 5 min followed by 95°C for another 20 sec, 57 and 63°C for 40 sec (57°C for 35S and NOS primers; while 63°C for *cryIA(b)* and *ivr1* primers), at 72°C for 1 min (in total, 40 cycles of above program was performed) and finally at 72°C for 3 min. The PCR products were separated using a 2% agarose gel electrophoresis in 1x TBE buffer (0.045 mM Tris-borate; 1 mM

EDTA, pH 8.2), and visualised by UV after staining for 15 min in a 1 mg ethidium bromide solution. The PCR products of 35S, NOS, *cryIA(b)* and *ivr1* (reference gene) has a size of 195, 180, 184 and 226 bp, respectively.

RTQ-PCR (Real-time quantitative PCR): Real-time screening, qualitative and quantitative detection was performed on the LightCycler instrument according to the instruction's manual with slight modifications in a total volume of 20 μ l in the presence of 5 μ l of DNA (or water). Screening for GM was performed with the LightCycler GMO Screening Kit. The amplicons of the 35S promoter and the NOS terminator were simultaneously detected by fluorescence measurements, using specific pairs of hybridization probes. For the detection of the reference plant gene, a specific pair of probes, one labeled with LightCycler-Red 640 and the other with fluorescein, were utilized. The reaction product serves as both a control for DNA integrity and PCR inhibitors. The threshold cycle (C_T) values were determined. The determination of the copy number of *ivr1* and *cryIA(b)* genes was based on the molecular weight of corn genome and assuming that both *ivr1* and *cryIA(b)* genes are single-copy genes. To quantify DNA by RTQ-PCR, an increase in fluorescence emission during PCR proportional to the initial copy number of the target gene is generated and detected. Fluorescent dsDNA-binding dyes and fluorogenic probes were applied. Samples were quantified by interpolation in a standard regression curve of C_T values generated from DNA samples of known concentrations (positive control and CRM) using ten-fold dilution series from 10,000 to 1 for reference gene and from 3,500 to 0.35 for target gene.

Confirmation of results: Repeatability was tested using the same machine, condition, technician and lab. To confirm that the observed bands or peaks were amplified genomic DNA and not primer artifacts, the exact PCR procedures were carried out on a replicate for each sample used for all primers. Reproducibility was tested using two PCR machines in two different labs (Biotechnology labs at KACST) by three different researchers. To confirm the detection and quantification of GM results, certified reference materials were used.

Results and Discussion

The differentiation of GMO and non-GMO products is a prerequisite to keep the content of undesirable GM material below the allowed threshold in the non-GM food production chain, or to keep and commercialize GMOs separately due to their high added value such as in consumer-directed second generation biotech products. The demand for traceability of GMOs has its particular implications for each step of the food chain. Moreover, these implications should be considered in a global perspective for the GMO sector, mostly since GMO maize and soybean varieties have been largely adopted in the US, and a large part of these US products is exported worldwide.

In this paper we present the heterogeneity model and the preliminary results of our investigation of the sampling simulation and the stability of heterogeneity parameters. In the process of defining a sampling plan for bulk products, the main parameters that should be statistically taken into account include lot size and uniformity, accepted risks (tolerances) and adopted testing methods, while parameters to be settled include increment size,

rate of increment sampling and preparation of the sample prior to the analysis¹⁷. Preliminary results from our simulations done with KeSTE (Kernel Sampling Technique Evaluation)⁹ indicated that the pattern of convergence to the "true" contamination value is similar for different contamination levels and lot heterogeneity scenarios. Specifically, the spread of the simulation results (SD = standard deviation of the estimates, x = number of primary samples) for each sampling number of primary samples taken from a lot provides an indication of the possible sampling error (SE). When a given population is sampled multiple times for an increasing number of primary samples, the error associated to the contamination estimates in the bulk sample decreases. The presence of a common convergence pattern allows the definition of a model to estimate the maximum possible SE associated to any sampling number. We found that a negative exponential model best describes the decreasing trend of SD as a function of the increasing number of primary samples. The parameter h is indicative of lot heterogeneity and s is associated to primary samples characteristics. The level of lot heterogeneity will determine how large is the error and how rapidly SD converges to 0. Similarly to SD, also the largest values of contamination estimates (dominant values) obtained performing repeated sampling at each sampling number, show a constant pattern for different contamination and heterogeneity scenarios, converging asymptotically to the true lot contamination value. Such pattern is described by the exponential model $y = y_0 + he^{(-svx)}$, where y_0 = true lot contamination. The exponential SD curve is more robust compared to the exponential contamination curve given that it converges to 0, implying independency from the y_0 parameter, and it does not require selection of dominant values at each sampling number. These results are in harmony with those obtained by other researchers^{8,17}. They concluded that lot characteristics were a priority defined in simulations, and anyone could impose the known value of y_0 and estimate parameters h and s for both the SD and contamination curve, for a series of different contamination and heterogeneity scenarios.

According to Codex Alimentarius, the total number of primary samples will vary according to lot size. For statistical reasons, the total amount of grain material sampled should be collected as primary samples of 0.5 kg and should not be less than 0.01% of the total lot size. In case of large lots, the collection of 100 primary samples should be sufficient for the purpose of the present sampling protocol (even if it constitutes less than 0.01% of the total lot size). Primary samples should be analyzed individually, according to the sequential analysis protocol. In our simulation, primary samples characteristics were maintained constant. As a result, we could use the average value of s (calculated over a broad range of heterogeneity and contamination conditions) to re-estimate h . This improves the precision of h estimates because h and s show correlation in both exponential models. Once the two series of h parameters are estimated, h_y can be expressed as a function of h_{SD} and, using the average value of s_y , y_0 can be estimated with the greatest precision. Our results indicate that we can analyse samples less than 0.01% (0.005) (Table 1 and Fig. 1).

In a single-sampling plan, N groups of n grains are separately ground and analysed to determine if they are GMO-positive or not. If there are X positive out of the N groups the GMO proportion may be estimated by X/N and the lot is accepted if $X \leq A$, rejected if $X > A$, where A is a predetermined acceptance threshold. Thus, the

set of parameters defining the plan is (N, n, A) . In the case of grains, the double sampling is defined as follows. First, N_1 groups of n_1 grains are assayed. The lot is accepted if $X_1 \leq A_1$, where X_1 is the number of positive groups and A_1 a predetermined threshold. If $X_1 \geq R_1$, where R_1 is a predetermined rejection threshold, the lot is rejected. In between, that is if $A_1 < X_1 < R_1$, N_2 new groups are assayed. The lot is then accepted if $X_2 \leq A_2(X_1)$, where X_2 is the number of positive among the N_2 new groups and $A_2(X_1)$ a predetermined threshold. The function $A_2: X_1 \rightarrow A_2(X_1)$ giving for each value of X_1 the acceptance threshold at the second step must be a decreasing function of X_1 because the bigger is X_1 , the smaller must be X_2 to compensate. In classical quality control, this function has the form $A_2 - X_1$ where A_2 is a fixed number. The lot is therefore accepted at the second step if $X_2 + X_1 \leq A_2$ and rejected if $X_2 + X_1 \geq R_2 = A_2 + 1$. Since we allow here the subsamples examined at steps 1 and 2 to have possibly different numbers n_1 and n_2 of grains, it seems natural as a consequence to adopt a larger frame making possible the use of something less symmetric than the sum $X_1 + X_2$ to base the acceptance at the second step. So the set of parameters to determine in that case is $T = \{N_1, n_1, A_1, R_1, N_2, n_2, A_2(\cdot)\}$. The brackets following A_2 are put to remind that A_2 is a function.

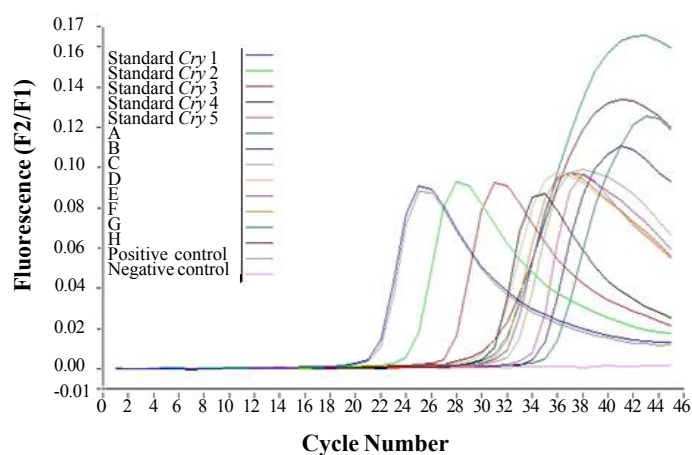


Figure 1. Quantification of *cryIA(b)* modified gene. Samples 1-5: Standard serial dilution series of *cryIA(b)* gene. Samples 6-13 (A-H): different percentages of test samples of 0.5% GM maize kernels (A: 0.001, B: 0.005, C: 0.01, D: 0.02, E: 0.03, F: 0.04, G: 0.05 and H: 0.1%, respectively). Samples 14 and 15: positive and negative control of *cryIA(b)*, respectively.

The evaluation of the GM materials present in a lot entails relevant implications for the trade and food production in view of the differences in legislation and in GMO acceptance worldwide. Determination of the content of GMOs in raw materials is subjected to errors during the various stages of the “diagnostic chain” (sampling, sub-sampling and analysis). Since, in most cases, GMOs are non-homogeneously distributed in the bulk, the variance associated with the sampling step is likely to represent the major contribution to the overall variance. The lower the GMO concentration is, the more relevant the effect of different sampling strategies will be. The experience with sampling methodologies for the analysis of mycotoxins provides the basis for provisional sampling schemes that can be applied to GMOs and in many European countries. Directive 98/53 on sampling and analysis of

certain contaminants in foodstuffs was the sampling plan first suggested and still used as applicable to GMOs¹⁸. Certified reference materials (CRMs) are an essential tool in the quality assurance of analytical measurements. They are produced, certified and used in accordance with relevant ISO (International Organization for Standardization) and BCR (Community Bureau of Reference) guidelines. The Institute for Reference Materials and Measurements (IRMM; Geel, Belgium) has produced the first powdery genetically modified organism (GMO) CRMs in cooperation with the Institute for Health and Consumer Protection (Ispra, Italy). Until now, different weight percentages in the range of 0-5% for 4 GMOs in Europe were produced and certified: Bt (*Bacillus thuringiensis*)-11 and Bt-176 maize, Roundup Ready soybean and MON810 maize. Bt-11 and Bt-176 maize and Roundup Ready soybean were produced by IRMM on behalf of Fluka Chemie AG (Buchs, Switzerland).

GMO PCR quantification is based on the parallel amplification of a region of the transgene and a species-specific endogenous gene, which gives an estimation of the total amount of amplifiable DNA from the species of interest present in the sample. CRM validated by the IRMM are usually used as standards for amplification, with genomic DNA obtained from these certified materials used to construct the quantification calibration curves. One major advantage of using CRM DNA as calibration standards is that the transgene and the species-specific target gene are in a constant relative amount, which minimizes errors in relative concentrations. However, calibration curves obtained with certified materials have the disadvantage that at high amounts of genomic DNA as required in the upper limits of the standard curve, signal inhibition occurs which narrows down the linear range of the obtained calibration curves¹⁹.

It has been shown empirically that the concentration of DNA in the real-time PCR reaction is proportional to PCR cycle number during the exponential phase of the PCR reaction. Therefore, if the number of cycles can be determined that it takes for a sample to reach the same point in its exponential growth curve, it is possible to calculate the precise content of genetically modified DNA. The real-time PCR method makes use of these principles to provide precise quantification of the GMO content of agricultural products. Each series of analyses includes the analysis of a full set of standards, giving rise to a standard curve. The results obtained for individual unknown samples are compared to the standard curve to determine the GMO content of those unknowns. Most real-time systems of instrumentation automate this analytical procedure. In real-time analyses the amount of product synthesised during PCR is estimated directly by measurement of fluorescence in the PCR reaction. Several types of hybridization probes are available that will emit fluorescent light corresponding to the amount of synthesized DNA. However, the amount of synthesised product can also be estimated with fluorescent dyes, e.g. SYBR Green I that intercalates double-stranded DNA. With the latter, it is not possible to distinguish between the specific product and non-specific products, and consequently the use of specific hybridization probes is normally preferred. As with double competitive PCR, the quantitative estimate is based on extrapolation by comparison of the GMO sequence relative to the reference of interest. The idea is that with the use of fluorescence it becomes possible to measure exactly the number of cycles that are needed to produce a certain amount of PCR product. This

amount corresponds to the amount producing a fluorescence signal clearly distinguishable from the background signal and measured well before the plateau effect becomes a problem. The number is called the C_T value. Then by comparison of C_T -values for the GMO target sequence, e.g. *cryIA(b)*, and the reference gene, e.g. *ivr1*, it becomes possible to estimate the ratio of the GMO target sequence to the reference sequence in terms of difference in number of cycles needed to produce the same quantity of product (Fig. 1 and Table 1).

The promoter and terminator elements used to transform most of the currently approved genetically modified plants are the Cauliflower Mosaic Virus promoter (P-35S) and the *Agrobacterium tumefaciens* nopaline synthase terminator (T-NOS). Although, other promoters and terminators have also been used, almost all GM plants contain at least one copy of the P-35S, T-35S and/or the T-NOS as a part of the gene construct integrated in its genome. Consequently, methods detecting one of these elements are popular for screening purposes. One problem with these methods is that the elements they detect are from naturally occurring virus and bacteria which are often present in fresh vegetables or the environment in which they are grown. Such elements therefore pose a significant risk of yielding false positive results²⁰. However, by performing a CaMV-specific PCR based on genes normally not present in GMOs, false positives, as a result of virus infected plants, can be eliminated. Our results exclude the false positive results (Table 2) using the previous primers.

Although amplification of DNA from highly processed food products may be related to the absolute amount of DNA present in the sample, profound degradation of DNA will make the interpretation of analytical results difficult if the housekeeping gene and the transgene are not degraded equally or if degradation resulted in DNA fragments that are below the amplifiable size²¹. So, failure of the PCR reaction to amplify the target DNA sequence could be the result of degraded DNA, insufficient target DNA or DNA contaminated by inhibitors^{22,23}. DNA could not be detectable in highly heat-treated food products, hydrolysed plant proteins, purified starch derivatives and refined oils derived from GMO²⁴. However, failure of the PCR reaction to amplify the target DNA sequence in simulated sampling plans could be the result of sampling risk²⁵. In our study, all tested samples give good results – using both conventional and real-time PCR (Fig. 1) – whatever the size of samples. Conventional PCR yielded positive results of 35S, NOS, *cryIA(b)* and *ivr1* (reference gene) has a size of 195, 180, 184 and 226 bp, respectively. These results indicate that both conventional PCR and real-time PCR are sensitive and suitable for routine GMO analysis and there is no sampling risk of our simulated sampling plan. These results are in accordance with the results obtained by Al-Swailem *et al.*²³, who analysed genetic modification of maize and some of its product and indicated that real-time was more effective than conventional PCR.

The reliable detection, characterisation and quantification of GMO requires extensive validation of test methods involved. Estimates for the Limit of Detection and Quantitation for all the tests must be determined. The robustness and reproducibility of methods, areas of measurement variation and the effects of diverse sample matrices on the test methods must be examined to achieve the most accurate, precise results. In general, PCR-based methods have a threshold detection of 0.01%. This limit is caused by the amount of DNA that is introduced into the reaction. If a sample of

maize genomic DNA contains 0.01% GMO material, the number of GMO targets in a PCR reaction will on average be very small (one to four target molecules). The total number of maize genomes in each PCR reaction causes the reason for this. The size of the maize genome is about 4.5×10^9 base pairs and the amount of sample DNA introduced into PCR reactions is normally around 25 to 75 μg (sometimes up to 200 μg). Using Avogadro's number (6.023×10^{23}), one can calculate from these values that the actual number of maize genomes present in a PCR reaction will range from around 10,000 (50 μg sample DNA) to around 40,000 (200 μg sample DNA). If the concentration of GMO genomes in the sample is 0.01% (=1 in 10,000) the number of GMO genomes (the number of GMO target molecules) would be 1 for a 50 μg DNA sample to 4 for a 200 μg DNA sample. Because the soybean genome is smaller than that of maize (2.5×10^9 base pairs), sampling statistics are more favorable little better for soybeans than for maize. Even though 0.01% is the limit of detection using PCR, quantitative analysis is not possible in this concentration range. In samples from a DNA preparation whose actual composition is 0.01% GMO material, the number of GMO targets in any given sample could be zero, one, two, three, four or more. Amplifying these samples will lead to results having substantial differences in signal intensity. These differences will not be related to the actual GMO content of the original sample but will be due to the statistical variations related to sampling of that DNA preparation. Thus, differences in signal intensity cannot be correlated with quantitative differences in GMO content in samples that contain low GMO levels such as 0.01%. Therefore, most laboratories set the limit of quantification ten-fold higher at 0.1% to avoid the problems with precision that occur near the limit of detection²⁶. Furthermore, to reliably quantify at the 0.1% level (1 in 1,000), thorough statistics require that at least 10,000 seeds should be homogenised and thoroughly mixed, and duplicate samples of this homogeneous powder subjected to DNA analysis. Several companies have shown that it is a practical reality to operate at this threshold.

We have determined the limit of detection (LOD) and the limit of quantification (LOQ) (Table 1). The estimated number of target molecules ranging from 4×10^6 to 1, the slopes of the linear regression curve analyses indicated very efficient amplification rates. Correlation coefficient was about 0.99. Theoretical LOQ was calculated taking into consideration the error associated with the serial dilutions of the DNA. The range of target copies with a 95% probability was present in each reaction. In the higher dilution ranges (below 40 copies of target) confidence intervals overlap and reliable quantification is not possible. Therefore, based on these statistical and experimental considerations, the LOQ was determined to be 40 copies. Our results are in harmony with those obtained by Hubner *et al.*²⁷, who indicated that the sensitivity of both LOD and LOQ are dependent on the plant genome size and the number of transgenes. Since the detection limits of our testing methodologies are very low, the threshold limit for unavoidable contamination represents one of the most critical points for traceability. In addition, similar results were observed by Miraglia *et al.*²⁸, who reported that, the LOD and LOQ are method specific but also depend on the sample that is being analysed. According to Berdal and Holst-Jensen¹⁹, there are three types of detection and quantification limits: 1) the absolute limits, i.e. the lowest number of copies that must be present at the beginning of the first cycle to obtain a probability of at least 95% of detecting/

quantifying correctly, 2) the relative limits, i.e., the lowest relative percentage of GM materials that can be detected/quantified under optimal conditions and 3) the practical limits, i.e., the limits applicable to the sample that is being analysed (taking into consideration the actual contents of the DNA sample and the absolute limits of the method).

It is important to realize that, if samples are to be analysed by a commercial company, it is very important to ask very specifically how the laboratory handles samples. Many labs will analyse only a small portion (not statistically representative) of the sample submitted for analysis, regardless of the size of the original sample that they received. Small sample sizes increase the possibility for false negatives (i.e., finding a sample as GMO-free even it actually contained such material). Thus, a sample obtained via a valid sampling plan may be improperly handled by the laboratory, resulting in an analytical result that does not accurately reflect the GMO content of the sample submitted for analysis or the GMO content of the larger lot of material from which the sample was taken. In addition, repeated analytical samples drawn from a homogenised field sample may not have identical proportions of GMO versus non-GMO copies, especially when tiny GMO level is present. It is also important to consider aspects of cross-contamination when preparing samples. For example, when grinding large samples, the sample preparation area requires environmental control. Several of the analytical methods are very sensitive and dust drift can cause cross-contamination (false positive). Also, the sampling device (grinding and mixing equipment) must be analytically clean and free of material before collecting the next sample as just a few GMO left in the device will cause a positive result in the next sample.

The quantity and quality of template DNA is critical in PCR. Our data demonstrate that the increased volumes of samples and concentrations of DNA had substantial effect on the conventional PCR banding pattern and real-time peaks and GM relative concentration (Table 1 and Fig. 1). According to Rasmussen²⁹, 1 ~ 100 ng is sufficient for plant and human genomic DNA, while 1 ~ 10 ng is sufficient for bacterial genomic DNA, and 0.1 ~ 1 ng is enough for plasmid DNA. All of the other factors tested had minimal effects on pattern reproducibility. Thus, our results suggest that variations in DNA concentrations in PCR reactions are the most likely cause of nonreproducible patterns, and that small volumes of samples and low concentrations of DNA are recommended for GM detection and quantification.

In conclusion, it is possible at the present time to provide definitive recommendations on sampling and to give some general indications of the issues related to sampling and a few pointers as to what will need to be done in the future. Our results will have a remarkable impact upon the definition of sampling protocols, as it will ensure a proper sampling if non-random distribution of contaminants is observed or expected, as in the case of kernel lots. Our assay method is proved to be effective and suitable for analytical purposes, with excellent limits of detection and quantification.

Several challenges must be met by laboratories testing for biotech events in food crops or processed products. Although these challenges generally fall into technical and non-technical categories, they cannot be completely separated since they closely lack of standard and approved tests for various biotech events, need for consistent results, need for improved test methods, need

for communicating nature of testing process and its relationship to issues of accuracy and labeling sequence information such that it is used responsibly are crucial.

References

- ¹MacCormick, C. A., Griffin, H. G., Underwood, H. M. and Gasson, M. J. 1998. Common DNA sequences with potential for detection of genetically manipulated organisms in food. *J. Appl. Microbiol.* **84**:969-980.
- ²Pozorski, P., Pozorska, A. and Twardowski, T. 1999. Nowa żywność - perspektywy i aspekty prawne. [Novel food- Perspectives and regulatory aspects]. *Przem. Spoż.* **1**:7-9.
- ³Hollricher, K., Morris, P. and Welters, P. 2000. Novel foods conquer the market. *Bioforum Int.* **4**:39-41.
- ⁴Gachet, E., Martin, G. G., Vigneau, F. and Meyer, G. 1999. Detection of genetically modified organisms (GMOs) by PCR: A brief review of methodologies available. *Trends Food Sci. Technol.* **9**:380-388.
- ⁵Roller, S. 2001. Genetically modified foods: Threat or opportunity. *Food Technol. Biotechnol.* **39**:259-263.
- ⁶Anklam, E. 1999. The validation of methods based on polymerase chain reaction for the detection of genetically modified organisms in food. *Anal. Chim. Acta* **393**:177-179.
- ⁷Agersborg, A., Dahl, R. and Martinez, I. 1997. Sample preparation and DNA extraction procedures for polymerase chain reaction identification of *Listeria monocytogenes* in seafoods. *Int. J. Food Microbiol.* **35**:275-280.
- ⁸Lischer, P. 2001. Sampling procedures to determine the proportion of genetically modified organisms in raw materials. Part I: Correct sampling, good sampling practice. *Mitt. Lebensm. Hyg.* **92**:290-304.
- ⁹Paoletti, C., Donatelli, M., Kay, S. and Van den Eede, G. 2003. Simulating kernel lot sampling: The effect of heterogeneity on the detection of GMO contamination. *Seed Sci. Technol.* **31**(3):629-638.
- ¹⁰Gilbert, J. 1999. Sampling of raw materials and processed foods for the presence of GIPSA. *Grain Inspection Handbook, Book 1, Grain Sampling.* <http://www.usda.gov/gipsa/pubs/pubs.htm>.
- ¹¹Knutti, R. and Schlatter, C. 1982. Distribution of aflatoxin in whole peanut kernels-Sampling plans for small samples. *Z. Lebensmittel-Untersuchung und -Forschung* **174**:122-128.
- ¹²Whitaker, T. B. 1977. Sampling granular foodstuffs for aflatoxin. *Pure and Applied Chemistry* **49**:1709-1717.
- ¹³USDA GAIN report #SA0021 (18 December 2000) Saudi Arabia Biotechnology – Saudi Arabia Bans Imports of GMO Animal Products, revises GMO Labelling and Extends Grace Period.
- ¹⁴USDA GAIN report (18 October 2002) Exporter Guide Annual Saudi Arabia.
- ¹⁵USDA Gain report #SA3005 (24 March 2003) Saudi Arabia Biotechnology - Saudi Ministry of Agriculture Issues a GMO Labeling Requirements.
- ¹⁶Sambrook, J. and Russel, D. W. 2001. Detection of DNA in agarose gels. In *Molecular Cloning. A Laboratory Manual.* 3rd ed., Cold Spring Harbor Laboratory Press, New York, pp. 5.14–5.17.
- ¹⁷Kay, S. and Paoletti, C. 2002. Sampling Strategies for GMO Detection and/or Quantification. EUR 20239 EN. Joint Research Centre, European Commission, Ispra. http://biotech.jrc.it/doc/EuroReport_sampling_strategies.
- ¹⁸European Commission 1998b. Commission Directive 98/53/EC of 16 July 1998 laying down the sampling methods and the methods of analysis for the official control of the levels for certain contaminants in foodstuffs. *Official Journal of the European Communities* L201:93-101.
- ¹⁹Berdal, K. G. and Holst-Jensen, A. 2001. RoundupReady® soybean event specific real-time quantitative PCR assay and estimation of the practical detection and quantification limits in GMO analyses. *Eur. Food Res. Technol.* **213**:432–438.

- ²⁰Lipp, M., Brodmann, P., Pietsch, K., Pauwels, J. and Anklam, E. 1999. IUPAC collaborative trial study of a method to detect genetically modified soybeans and maize in dried powder. *J. AOAC Int.* **82**:923-928.
- ²¹Van Duijn, G., Van Biert, R., Bleeker-Marcelis, H., Van Boeijen, I., Adan, A.J. and Jhakrie, S. 2002. Detection of genetically modified organisms in foods by protein-and DNA-based techniques: Bridging the methods. *J.AOAC Int.* **85**(3):787–791.
- ²²Holden, M. J., Blasic, J. R., Bussjaeger, L., Kao, C., Shokere, L. A. and Kendall, D. C. 2003. Evaluation of extraction methodologies for corn kernel (*Zea mays*) DNA for detection of trace amounts of biotechnology-derived DNA. *J. Agric. Food Chem.* **51**:2468–2474.
- ²³Al-Swailem, A.M., Shehata, M. M., Shair, O.H., Sabaan, S. A., Al-Anazi, I.O. and Al-Shammari, T. A. 2005. An efficient method for identification and quantification of genetic modification of local, imported and food products of maize in Saudi Arabia using PCR-based markers and real-time PCR. *J. Food, Agric. Environ.* **3**(2):14-19.
- ²⁴Kuiper, H.A. 1999. Summary report of the ILSI Europe workshop on detection methods for novel foods derived from genetically modified organisms. *Food Cont.* **10**:339–349.
- ²⁵Kobilinskya, A. and Bertheaub, Y. 2005. Minimum cost acceptance sampling plans for grain control, with application to GMO detection. *Chemomet. and Intell. Lab. Syst.* **75**:189-200.
- ²⁶Kay, S. and Van den Eede, G. 2001. The limits of GMO detection. *Nature Biotech.* **19**:405-415.
- ²⁷Hubner, P., Waiblinger, H. U., Pietsch, K. and Brodmann, P. 2001. Validation of PCR methods for quantitation of genetically modified plants in food. *J. Assoc. Offic. Analyt.Chem. Inter.* **84**:1855–1864.
- ²⁸Miraglia, M., Berdal, K. G., Brera, C., Corbisier, P., Holst-Jensen, A., Kok, E. J., Marvin, H. J. P., Schimmel, H., Rentsch, J., Van Rie, J. P. P. F. and Zagon, J. 2004. Detection and traceability of genetically modified organisms in the food production chain. *Food and Chemic. Tox.* **42**:1157-1180.
- ²⁹Rasmussen, R. 2001. Quantification on the LightCycler instrument. In Meuer, S., Wittwer, C. and Nakagawara, K. (eds). *Rapid Cycle Real-time PCR: Methods and Applications*. Springer Press, Heidelberg, pp. 21–34.
- ³⁰Ehlers, V. B., Strauch, E., Goltz, M., Kubsch, D., Wagner, H., Maidhof, H. *et al.* 1997. Nachweis genetischer Veränderungen in Mais mittels PCR. *Bundesgesundheitsblatt* **4**:118–121.