

Increased efficacy for in-house validation of real-time PCR GMO detection methods

I. M. J. Scholtens · E. J. Kok · L. Hougs · B. Molenaar ·
J. T. N. M. Thissen · H. van der Voet

Received: 2 July 2009 / Revised: 4 November 2009 / Accepted: 15 November 2009 / Published online: 9 December 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract To improve the efficacy of the in-house validation of GMO detection methods (DNA isolation and real-time PCR, polymerase chain reaction), a study was performed to gain insight in the contribution of the different steps of the GMO detection method to the repeatability and in-house reproducibility. In the present study, 19 methods for (GM) soy, maize canola and potato were validated in-house of which 14 on the basis of an 8-day validation scheme using eight different samples and five on the basis of a more concise validation protocol. In this way, data was obtained with respect to the detection limit, accuracy and precision. Also, decision limits were calculated for declaring non-conformance ($>0.9\%$) with 95% reliability. In order to estimate the contribution of the different steps in the GMO analysis to the total variation variance components were estimated using REML (residual maximum likelihood method). From these components, relative standard deviations for repeatability and reproducibility (RSD_r and RSD_R) were calculated. The results showed that

not only the PCR reaction but also the factors ‘DNA isolation’ and ‘PCR day’ are important factors for the total variance and should therefore be included in the in-house validation. It is proposed to use a statistical model to estimate these factors from a large dataset of initial validations so that for similar GMO methods in the future, only the PCR step needs to be validated. The resulting data are discussed in the light of agreed European criteria for qualified GMO detection methods.

Keywords In-house validation · Real-time · PCR · GMO · REML

Introduction

Within the European Union, products that are genetically modified organisms or derived thereof, in whole or in part, should be labelled as such (European regulation 1829/2003). The exception to this rule is when the presence of the GMO component is adventitious and unavoidable. In the latter case, the GMO component is allowed up to the level of 0.9% per ingredient. If the GMO component exceeds this threshold value, then the product needs to be labelled as containing GMO materials. It is also laid down in Regulation 1829/2003 that producers have to supply an event-specific detection method for each new GMO variety. This detection method, as well as the associated reference materials is an integral part of the approval dossier. The Joint Research Centre (JRC) of the European Commission in Ispra, Italy, being the Community Reference Laboratory for GMO detection methods, will subsequently test the method and, in a first phase, compare the results with the agreed Method Acceptance Criteria for Analytical Methods of GMO Testing, as formulated by the European Network

Electronic supplementary material The online version of this article (doi:10.1007/s00216-009-3315-6) contains supplementary material, which is available to authorized users.

I. M. J. Scholtens (✉) · E. J. Kok · B. Molenaar
RIKILT—Institute of Food Safety,
Wageningen University and Research Centre,
P.O. Box 230, 6700 AE Wageningen, The Netherlands
e-mail: ingrid.scholtens@wur.nl

J. T. N. M. Thissen · H. van der Voet
BIOMETRIS—Wageningen University and Research Centre,
P.O. Box 9101, 6700 HB Wageningen, The Netherlands

L. Hougs
The Danish Plant Directorate,
Skovbrynet 20, 2800 Kgs.,
Lyngby, Denmark

of GMO Laboratories [4]. If the method meets these requirements, it will, in a second phase, be validated in a European ring trial, and the results are evaluated for reproducibility and trueness (Method Performance Requirements). To this end, the JRC will be assisted by the ENGL, as is stipulated in EU Regulation 1981/2006 [6]. To date, over 45 methods for GMO analysis have been tested in a EU ring trial (<http://gmo-crl.jrc.ec.europa.eu/>). These ring trials were, in some cases, carried out on the basis of plant materials (then the DNA isolation was included in the ring trial), but in most cases with purified DNA.

In the in-house validation studies presented in this paper, RIKILT standard operating procedures for validation were followed. The PCR efficiency and slopes of the calibration curves, the repeatability, in-house reproducibility, bias and detection limit of the methods were determined. The results were compared with the following ENGL criteria [4]:

- Amplification efficiency: the average value of the slope of the standard curve should be in the range of $-3.1 \geq \text{slope} \geq -3.6$, corresponding to an apparent efficiency between 90% and 110%.
- R^2 coefficient: the average value of the squared coefficient of correlation (R^2) should be ≥ 0.98 .
- Trueness: the trueness (expressed as lack of agreement or bias) should be within $\pm 25\%$ of the accepted reference value over the whole dynamic range.
- Reproducibility standard deviation (RSD_R): the relative reproducibility standard deviation should be below 35% at the target concentration and over the entire dynamic range. An $\text{RSD}_R < 50\%$ is acceptable for concentrations below 0.2%.

The EU validation studies include methods for maize, soy, cotton, potato and rapeseed. The results of the EU ring trials of the event-specific methods have been published by the JRC in individual reports. In these reports, the bias, repeatability and reproducibility of the methods are presented for DNA from reference materials of different percentages. These reports of EU validation studies provide important data on the performance of the respective methods in different laboratories within the European Union.

In addition a number of element-specific and construct-specific methods are available that use plasmid calibration curves instead of genomic DNA isolated from reference materials with certified GMO percentages on a weight/weight basis [10, 15]. These methods have been validated in an international ring trial in Asia.

For application of the method for maintenance of EU GMO regulations, it is necessary to perform an additional in-house validation study to establish the performance characteristics of the methods under in-house laboratory conditions.

For this in-house validation study, European GMO labs have two options: (1) use the same or similar conditions and samples as have been used in the EU method validation study in order to be able to compare the results of the in-house validation series with the results of the ring trial, or (2) validate the performance of the whole-routine GMO analysis procedure, including the DNA analysis, and with different types of real-life, practical samples and generate in this way additional data with increased practical relevance compared to the international validation study.

The GMO analyses all consist of the same steps: sample preparation, DNA isolation and real-time PCR. The most important differences between the different methods are the sequences of primers and probes. The different steps and their connections are illustrated in Fig. 1. At the RIKILT—Institute of Food Safety, depending on the type of sample, one of the three DNA isolation methods is chosen. This choice is independent of the subsequent PCR tests. The choice of PCR tests depends on the matrix, e.g., soy, maize or a mixed material. Usually, a sample is first screened for the presence of the 35S promoter and/or the nos terminator. If one of these tests is positive or if both of these tests are positive then event-specific tests are carried out to identify the specific GMO. In the example in Fig. 1, three samples are analysed: two DNA isolations are done with the Qiagen Plant Mini Kit and one DNA isolation is done with the Promega Kit. For each sample, a different set of PCR tests is carried out. This is using the idea of modularity which has been described by Holst-Jensen and Berdal [7].

Both the DNA isolation step and the PCR step contribute to the total variability of the method results. Moreover, variation is typically larger when PCR analyses are spread over a period of days than under strict repeatability conditions. The goal of our studies was to determine the relative contributions of the factors ‘PCR day’ (variation between runs on different days), ‘DNA isolation’ (variation between isolations) and ‘PCR’ (PCR to PCR variation in the same run) to the total variation of the GMO analysis.

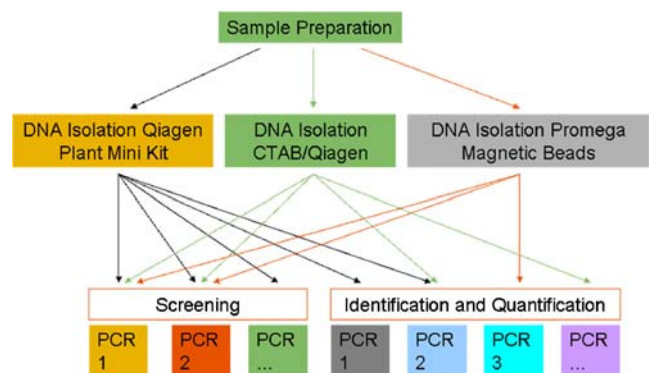


Fig. 1 Modular presentation of routine GMO testing on three samples with different matrices

The aim of the study was to test whether it is sufficient to only validate in-house the ‘PCR’ step of the method and calculate the repeatability and reproducibility of the GMO method using the whole dataset of all previously validated methods to estimate the contribution of the factors ‘DNA isolation’ and ‘PCR day’. This approach would make the whole procedure of validation of new GMO detection methods not only more efficient, but indeed more accurate as the results will be based on a larger dataset. This will especially be helpful for new GMO methods that maintenance labs have to implement as soon as possible.

In-house validation studies were performed for event-specific as well as construct-specific and element-specific methods in GM soy, maize, canola and potato. Event-specific methods positively identify a specific event, while the construct-specific methods identify a specific part of the construct, comprising more than one genetic element but not the event as such. Thirdly, the element-specific method identifies only a specific element of the GMO, but not the construct and/or the event. The same construct and/or element may be used to generate more than one event. The data of this series of validation studies are compared to the ENGL method performance requirements and the results are analysed to determine the relative contribution of the different aspects of the test procedure to the variation. Also, now that an increasing number of detection methods are available, it proves impractical to use many different reference genes per crop. It is therefore necessary to reduce the number of reference genes per crop and indeed select the best performing reference gene. To this end, some data are presented on the comparison of the maize *adh1* and

hmg reference genes. Moreover, the results are discussed in the light of day-to-day practice and in improving the efficacy of in-house validation studies and thereby the maintenance of EU GMO regulations.

Materials and methods

Our routine GMO analysis results are based on the mean value of 2 DNA isolations and 2 PCRs on the same day with each DNA (4 PCRs). Therefore this mean value was considered as the measurement to be validated.

Nineteen soy, maize, canola and potato GMO detection methods were validated in-house. These methods were in general terms described on <http://gmo-crl.jrc.ec.europa.eu/statusofdoss.htm> and in Kuribara et al. [10] and their detailed implementation is described in RIKILT standard operating procedures. The first 15 methods were validated in-house using at least eight different samples. These samples consisted of real feed samples, if available or mixed samples of GeMMA proficiency tests and certified reference materials. If no other samples were available only (mixtures of) reference materials were used. For calculation of the limit of detection, 0.1% samples were prepared. Detailed information on the samples used is available as [electronic supplementary material](#).

For all of the construct-specific methods and five event-specific maize methods, the validations were carried out on eight different days. Each sample was tested on at least two different days, with one or two DNA isolations of this sample per day and with two PCR analyses per DNA isolation (Table 1).

Table 1 Standard validation scheme

Validation parameter	Sample	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
Repeatability and reproducibility of the method	Sample 1	xx xx				xx			
	Sample 2		xx xx				xx		
	Sample 3			xx xx				xx	
	Sample 4				xx xx				xx
	Sample 5	xx				xx xx			
	Sample 6		xx				xx xx		
	Sample 7			xx				xx xx	
	Sample 8				xx				xx xx
Detection limit	0.1% reference material	xx	xx	xx	xx	xx	xx	xx	xx
Accuracy	1% reference material	xx	xx	xx	xx	xx	xx	xx	xx
Specificity control reference material	0% reference material	xx	xx	xx	xx	xx	xx	xx	xx
Contamination control DNA extraction	extraction control	xx	xx	xx	xx	xx	xx	xx	xx
Contamination control PCR	water	xx	xx	xx	xx	xx	xx	xx	xx

xx one DNA isolation and two PCRs, xx xx two DNA isolations and on each DNA two PCRs

The last five validation studies, for DAS 59122 maize event, RT73 canola event, EH92-527-1 potato event, H7-1 sugarbeet event and MIR604 maize event, respectively, were carried out with PCRs on, in principle, 4 days and with only one or two DNA isolations. For these validations, the ‘DNA isolation’ and ‘PCR day’ variance components were calculated on the basis of an overall model.

The Roundup Ready construct method has been used for 3 years and also 2008 data from nine routine samples and of 1% control charts were included in the validation.

The relative standard deviations for repeatability and reproducibility (RSD_r and RSD_R) were calculated from variance components estimated using the residual maximum likelihood (REML) method using all samples except those at the 0.1% level.

DNA isolation

IRMM and GeMMA Proficiency test materials were used for DNA isolation without further preparation. Routine samples of animal feed were milled through a 1-mm sieve on a Retsch ZM200 mill.

DNA was isolated from 100-mg dry material of each sample using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturers’ protocol. For maize samples, the lysis step was carried out with CTAB buffer (20 g/L CTAB, 1.4 M NaCl, 0.1% Tris, 20 mM Na₂EDTA, pH 8.0), instead of the manufacturers’ API buffer. During incubation, 20 mg/ml Proteinase K is added to the isolation. DNA isolations from maize gluten samples were done with the Promega Wizard Magnetic DNA Purification System for Food according to the manufacturer protocol. The DNA concentrations were measured on a Thermo Scientific NanoDrop™ Spectrophotometer.

In case of the comparisons of endogenous maize reference genes, the DNA was purified by NucleoSpin Food kit (Macherey Nagel, Düren, Germany) according to the manufacturers’ protocol.

Preparation of 1% reference material

If 1% reference material was not commercially available, it was made by mixing a higher percentage reference material with the corresponding 0% reference material on weight basis. After adding the two reference materials together, it was vortexed and mixed overnight before DNA isolation. For example, 10 mg of 10% GMO mixed with 90 mg of 0% GMO resulted in 100 mg of 1% GMO.

Real-time PCR

Real-time PCRs were performed on BioRad i-Cyclers iQ and MyiQ with Optical System software version 3.1 or iQ5 Optical systems software version 2.

The methods were carried out according to the JRC protocols with some standardised conditions: 50 ng of sample DNA per well were used, all reaction volumes were 25 µl and for all GMO detection methods, the same master mix was used (Diagenode, Belgium). In case of the comparison of endogenous maize reference genes the real-time PCR methods were performed on a Mx3000 (Stratagene) with MxPro software version 4. The mastermix used was JumpStart (Sigma) and the concentration of primers was 800 nM and the concentration of probes was 200 nM. All validations of event-specific real-time GMO detection methods were carried out with two calibration curves, one for the endogenous gene (the gene specified in the ENGL validated method, except for the MIR604 maize method where *hmg* was used instead of *adh1*) and one for the GMO event. The calibration curves were prepared as dilutions from one standard.

Statistical analysis

All data were ln-transformed and analysed as described below.

For each GMO test method, a mixed model was fitted to the data of the samples with a level higher than 0.1%, using the REML (residual or restricted maximum likelihood) method, first introduced by Patterson and Thompson [13]. The model had one fixed factor (sample) and three random factors (day, isolation and PCR). In this analysis three, variance components are estimated corresponding to the nested random structure specification day/isolation/PCR (read this as: PCR nested within isolation nested within day). The respective variance components are:

1. v_d : quantifies variation due to differences between days, averaged over isolations and PCRs (PCR day)
2. v_i : quantifies variation due to differences between isolations, averaged over PCRs (DNA isolation)
3. v_p : quantifies variation due to differences between PCRs of the same isolation on the same day (PCR)

These analyses provide estimates for each DNA test separately. All statistical analyses were performed with the statistical package GenStat Release 11 [14].

From these variance components, the next statistics are calculated:

Relative standard deviation for repeatability and reproducibility (RSD_r and RSD_R)

Repeatability is the closeness of agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement [17]. This means results obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within a short interval

of time [8]. Hence, the repeatability variance for single measurements (one isolation, one PCR) is the sum $v_i + v_p$. In our laboratory, a measurement result is defined as the average of four determinations from two PCRs on each of two isolations. Consequently, the repeatability variance for such measurement results is lowered to $v_i/2 + v_p/4$.

Reproducibility is the closeness of agreement between the results of measurements carried out under changed conditions of measurement. A valid statement of reproducibility requires specification of the conditions that are changed [18]. In this validation study, the work was done on different days in the same laboratory. The reproducibility variance for measurement results of the 2×2 type then is the sum $v_d + v_i/2 + v_p/4$.

The repeatability and reproducibility variances at the ln scale can then be translated to relative standard deviations by $RSD = \sqrt{e^{\text{var}} - 1}$ [12].

Decision limit

Standard concepts of decision and detection limits [3] can be adapted to inspection procedures with non-zero legally permitted maximum values M by considering M rather than 0 as the central value of the error distribution [16]. When applied to data with an assumed lognormal distribution this leads to the equation

$$\ln(\text{limit}) = \ln(M) + 1.645 \times S$$

where S is the relevant standard deviation on the ln scale. A decision limit for declaring non-conformance ($>0.9\%$) with 95% reliability is therefore calculated as the upper 95% confidence limit for a 2×2 -type measurement of a 0.9% sample, based on the assumption of log-normality:

$$\text{limit} = e^{\ln(0.9) + 1.645 \cdot \sqrt{\ln(RSD_R^2 + 1)}}$$

Trueness is quantified as the relative systematic deviation of reference samples. To estimate the trueness of each test method, extra REML analyses are performed for the measurements on the 1% reference samples only. These measurements were made on eight separate days, on each day with new DNA isolations, with two PCRs per isolation (for the first 14 GMO detection methods, Table 1).

In the absence of any bias, the ratio expected/ref should be 1 and the natural logarithm of this ratio is therefore 0. The data were analysed by a REML analysis on ln (expected/ref) with random factors day, isolation and PCR. Significance of the observed mean log-ratio m was determined using a Wald test (observed mean divided by estimated standard error of mean referred to a normal distribution). The relative systematic deviation was quantified as $rdev = (gm - ref)/ref$, where $gm = \exp(m) \times ref$ is the geometric mean of the measured concentrations, and ref is the reference value (1%).

For the Roundup Ready soy/lectin method, the bias was also calculated for the 1% data from the routine analyses (2005–2008, ERM-BF410d, data not shown).

The Japanese 35S, nos, Roundup Ready soy construct and four different maize constructs are element- or construct-specific, with plasmid calibration curves. Therefore, the results are expressed as percentage GMO element or GMO construct relative to the endogenous gene. The results of the maize construct methods cannot be used to verify the accuracy, when compared to 1% IRMM mass-certified reference materials, because maize is a heterozygote and because of the complex structure of maize seeds.

The detection limit or limit of detection (LOD) is formally defined as the lowest level that can be reliably detected [4, 9]. To estimate the detection limit of each method, extra REML analyses are performed for the measurements on the 0.1% reference samples only. The empirical validation requires the analysis of large numbers of samples at a specified low level and counting the number of positives. Because this is impractical, an alternative method has been used based on the definitions developed by international organisations like IUPAC and ISO [3]. The LOD is operationally defined as three times the reproducibility standard deviation at a low level (NEN 7777, 2003). This was calculated from the measurements on the 0.1% reference samples. These measurements were made on eight separate days, on six to eight fresh isolations, with two to four PCRs per isolation (Table 1).

The fraction of positive measurement results was determined to check the assumption that it was in most cases possible to obtain a numerical result for samples at this low level.

From the ln-transformed positive measurements, the variance components between-days, between-isolations and between-PCRs were estimated by REML. The reproducibility variance for measurement results of the 2×2 type then is the sum $v_d + v_i/2 + v_p/4$. This was translated to a relative standard deviation by $RSD = \sqrt{e^{\text{var}} - 1}$, and then to a standard deviation $s = RSD \times \text{mean}(\text{conc})$, and a detection limit $lod = 3 \times s$. Note that this is the detection limit calculated for the routine 2×2 determination. The traditional theory of detection limits [3] assumes a normal error distribution around 0, which precludes the use of log transformation. Therefore, we calculated a detection limit by the traditional 3-s approach on the original percentage scale, but using s derived from a model fitted to the ln-transformed data.

Detection and decision limits are only valid when the trueness is sufficiently good. All calculations were done in GenStat Release 11.1.

Endogenous maize reference gene

To test whether the same endogenous reference gene could be used for all GMOs of the same species a small comparison experiment was set up. MON810 and MON863 maize were analysed with two different reference genes, one method detecting ADH (from the MON863 event-specific method) and the other detecting hmg (from the MON810 event-specific method). From standard curves of 5% or 10% reference material, 1% reference materials were analysed.

Results

In 19 validation studies, about 170 samples of several different crop species and/or GM crop varieties were analysed. The first 14 methods were validated on the basis of the scheme in Table 1, the last five were validated with one or two DNA isolations and four PCR runs on four different days. The results were subsequently evaluated for the different aspects described here.

PCR calibration curves

To check whether the different GMO detection methods comply with the ENGL criteria [4] mean PCR efficiencies, squared coefficient of correlations R^2 and slopes of the

calibration curves were calculated for all GMO detection methods. The results are shown in Table 2. All values are the mean of a minimum of eight PCR runs, except for the last four GMO-detection methods where the values listed are the mean values of four PCR runs. The average PCR efficiencies were mostly over 90% (86.6–105.0) for all endogenous and event-specific GMO detection methods (Table 2). The efficiencies of the H7-1 sugar beet event (88.6) and the *SSIIB* method (averaged over six validation studies 88.6% (86.6–90.8)) were slightly lower than the ENGL criterion of 90%. The average squared coefficient of correlation R^2 of the calibration curves was over 0.99 for all methods tested.

Basically, the real-time PCR conditions as stated in the validated methods were used, but for convenience in routine laboratory only one brand of master mix was used. Theoretically, it is possible that this master mix did not work equally well for all methods, because the methods have been designed and ENGL validated with a different master mix. From the results in Table 2 it can, however, be concluded that our master mix worked well for all methods. This can be concluded because the efficiencies and average slopes of the calibration curves agreed with the ENGL Method Performance criteria [4]. Only the average slope for the *SSIIB* gene (averaged over the six validated methods) was -3.64 , which is slightly lower than the ENGL criteria (-3.1 to -3.6). For all other slopes the requirements were met.

Table 2 Efficiency (E), squared coefficient of correlation, slope and intercept for reference gene PCRs and GMO PCRs

GMO detection method	E (%) endo gene	Squared coefficient of correlation	Slope	E (%) GMO	Squared coefficient of correlation	Slope
Roundup ready soy construct/lectin	97.1	0.996	-3.400	92.9	0.998	-3.508
35S promoter/lectin	95.8	0.998	-3.430	95.4	0.997	-3.438
Nos terminator/lectin	95.8	0.998	-3.430	91.8	0.997	-3.540
35S promoter/ <i>SSIIB</i> gene	86.6	0.994	-3.709	93.6	0.995	-3.491
Nos terminator/ <i>SSIIB</i> gene	86.6	0.994	-3.709	93.0	0.998	-3.506
Bt11 maize construct/ <i>SSIIB</i> gene	90.8	0.996	-3.569	95.5	0.999	-3.435
Bt176 maize construct/ <i>SSIIB</i> gene	88.2	0.999	-3.643	95.5	0.999	-3.437
Mon810 maize construct/ <i>SSIIB</i> gene	88.5	0.998	-3.634	95.5	0.999	-3.435
GA21 maize construct/ <i>SSIIB</i> gene	90.8	0.997	-3.570	94.3	0.999	-3.468
TC1507 maize event/ <i>Hmg</i> gene	89.5	0.999	-3.605	95.3	0.994	-3.447
Mon810 maize event/ <i>Hmg</i> gene	95.6	0.998	-3.440	94.2	0.990	-3.480
GA21 maize event/ <i>Adh1</i> gene	100.5	0.995	-3.315	92.5	0.992	-3.525
NK603 maize event/ <i>Adh1</i> gene	105.0	0.995	-3.220	89.8	0.990	-3.600
Mon863 maize event/ <i>Adh1</i> gene	98.8	0.997	-3.350	97.2	0.996	-3.400
RT73 canola event/ <i>FatA</i> gene	97.6	0.998	-3.385	96.3	0.999	-3.405
DAS 59122-7 maize event/ <i>Hmg</i> gene	91.1	0.998	-3.559	98.1	0.992	-3.399
EH92-527-1 potato event/ <i>UGPase</i> gene	98.7	0.999	-3.362	98.5	0.996	-3.377
H7-1 sugarbeet event/ <i>GS</i> gene	93.9	0.999	-3.480	88.1	0.999	-3.645
MIR604 maize event/ <i>Hmg</i> gene	98.6	0.999	-3.355	99.5	0.999	-3.333

Repeatability, reproducibility and decision limit

Repeatability and in-house reproducibility were calculated on the basis of the mean value of 2×2 PCR results (two DNA isolations and two PCR reactions with each DNA). The difference between the repeatability and in-house reproducibility was in these series the factor time: different analyses (DNA isolation and PCR) on different days. The repeatability and reproducibility were calculated for each individual GMO detection method. These methods have been validated according to the scheme in Table 1, in a time period of at least 2 months, with new DNA isolations on each day. The results can be seen in Tables 3, 4 and 5.

Table 3 shows the number of data that have been used for the calculations. The variance components at a ln scale, the repeatability and reproducibility relative standard deviations as well as the decision limit are shown in Table 4.

The repeatability, reproducibility and decision limit for declaring non-conformance for all samples could be calculated with overall values for v_d and v_i . These calculations are more precise than the calculations in Table 4, because they are based on a larger number of data.

Repeatability, reproducibility and limit of detection for 0.1% samples

The repeatability and reproducibility relative standard deviations (for a result from 2 DNA isolations×2 PCRs at

the same day) as well as the limit of detection based on 0.1% reference materials are shown in Table 5. In these validation series, 0.1% GMO (50 ng DNA) could be readily detected for all GMO detection methods (Table 5 columns 5 and 6), but occasionally no numerical result was obtained due to aberrant amplification curves.

Bias

The trueness was validated in-house with 1% certified reference material, if available. In other cases, the 1% sample was mixed in-house. The results are shown in Table 6. For the construct-specific methods with plasmid calibration curves the ‘bias’ is also shown, but this is difficult to interpret, because the percentage of the reference material on a haploid genome basis is not known (except for the MON810 maize) and because the exact number of construct copies in the genome is not known.

In order to estimate the relevance of these figures, the results were compared with the results as obtained in GeMMA proficiency tests.

Endogenous maize reference genes

The results of the MON810 and MON863 maize quantifications with the use of both *adh1* and *hmg* as a reference gene are shown in Table 7.

MON810 and MON863 maize were analysed with two different reference genes; one method detecting

Table 3 Numbers of data for estimating precision

GMO method	<i>n</i> samples	<i>n</i> days	<i>n</i> isolations total (per sample)	<i>n</i> PCRs total (per isolation)
Roundup ready soy construct/lectin	14	17	72 (3–7)	148 (1–3)
35S promoter/lectin	8	13	54 (5–7)	113 (1–3)
Nos terminator/lectin	8	13	54 (5–7)	114 (1–3)
35S promoter/ <i>SSIb</i> gene	8	17	23 (2–3)	50 (1–3)
Nos terminator/ <i>SSIb</i> gene	8	11	24 (3)	52 (1–3)
Bt11 maize construct <i>SSIb</i> gene	8	8	32 (4)	132 (4–12)
Bt176 maize construct/ <i>SSIb</i> gene	8	8	32 (4)	134 (4–14)
Mon810 maize construct/ <i>SSIb</i> gene	8	8	32 (4)	134 (3–12)
GA21 maize construct/ <i>SSIb</i> gene	8	8	42 (7×4+2×6+1×2)	134 (1–4)
TC1507 maize event/ <i>Hmg</i> gene	7	8	40 (6×3+2×11)	55 (1–2)
Mon810 maize event/ <i>Hmg</i> gene	8	8	47 (7×3+11+7+8)	64 (1–2)
GA21 maize event/ <i>Adh1</i> gene	7	8	40 (6×3+2×11)	54 (1–2)
NK603 maize event/ <i>Adh1</i> gene	7	8	40 (6×3+2×11)	58 (2)
Mon863 maize event/ <i>Adh1</i> gene	7	8	40 (6×3+2×11)	58 (2)
RT73 canola event/ <i>FatA</i> gene	7	5	17 (7×2+1×3)	80 (2–12)
DAS 59122 maize event/ <i>Hmg</i> gene	9	4	9 (1)	86 (2–8)
EH92-527-1 potato event/ <i>UGPase</i> gene	8	8	6 (1)	138 (2–6)
H7-1 sugarbeet event/ <i>GS</i> gene	7	4	3 (1)	74 (2–8)
MIR604 maize event/ <i>Hmg</i> gene	8	4	2 (1)	84 (2–8)

Table 4 Precision: variance components at ln scale, relative standard deviations of repeatability and within-laboratory reproducibility, and decision limit for declaring non-conformance (>0.9%) with 95% reliability

	v_d (days)	v_i (isolations)	v_p (PCRs)	RSD_r^* (%)	RSD_R^* (%)	Decision limit ^a (% cp/cp or % w/w)
Roundup ready soya construct/lectin	0.036	0.091	0.148	29	36	1.59 (% cp/cp)
Roundup ready soya construct/lectin(2)	0.038	0.000	0.097	16	25	1.36 (%cp/cp)
35S promoter/lectin	0.008	0.134	0.059	29	31	1.47 (% cp/cp)
Nos terminator/lectin	0.265	0.116	0.141	31	66	2.41 (% cp/cp)
35S promoter/SSIIB gene	0.147	0.000	0.146	19	45	1.82 (% cp/cp)
Nos terminator/SSIIB gene	0.035	0.000	0.093	15	24	1.34 (% cp/cp)
Bt11 maize construct SSIIB gene	0.153	0.015	0.032	12	43	1.77 (% cp/cp)
Bt176 maize construct/SSIIB gene	0.089	0.005	0.031	10	32	1.51 (% cp/cp)
Mon810 maize construct/SSIIB gene	0.149	0.002	0.035	10	41	1.73 (% cp/cp)
GA21 maize construct/SSIIB gene	0.012	0.000	0.040	10	15	1.15 (% cp/cp)
TC1507 maize event/Hmg gene	0.000	0.076	0.017	21	21	1.26 (% w/w)
Mon810 maize event/Hmg gene	0.207	0.000	0.076	14	50	1.97 (% w/w)
GA21 maize event/Adh1 gene	0.055	0.024	0.171	24	34	1.55 (% w/w)
NK603 maize event/Adh1 gene	0.070	0.000	0.094	15	31	1.49 (% w/w)
Mon863 maize event/Adh1 gene	0.048	0.000	0.029	9	24	1.32 (% w/w)
RT73 canola event/ <i>Fata</i> gene	0.000	0.008	0.054	13	13	1.12 (% w/w)
DAS 59122 maize event/Hmg gene	0.016	0.095 ^b		15–22 ^c	20–26 ^c	1.25–1.36b (% w/w)
EH92-527-1 potato event/UGPase	0.099	0.081 ^b		14–20 ^c	36–39 ^c	1.59–1.66b (% w/w)
H7-1 sugarbeet event/GS gene	0.010	0.019 ^b		7–10 ^c	12–14 ^c	1.10–1.13b (% w/w)
MIR604 maize event/Hmg gene	0.000	0.064 ^b		13–18 ^c	13–18 ^c	1.11–1.21b (% w/w)

^a Valid for routine method with a result from 2 isolations × 2 PCRs at the same day

^b v_i and v_p are aliased: only their sum can be estimated

^c Range of values: low (high) values are estimates assuming variation between PCRs (isolations) are dominating the aliased sum of variance components.

ADH (from the MON863 method) and the other detecting *hmg* (from the MON810 method). From standard curves of 5% MON810 (BF413f) or 10% MON863 (BF416d) reference material 1% reference materials were analysed and the results compared in Table 7. The same dilution of reference material was added in duplicate to detect *adh1*, *hmg* and the GMO gene. The GMO contents were calculated from the GMO gene using either *adh1* or *hmg* as the reference gene. Even though there was 1 to 1.5 Ct values between the signals from the reference genes it resulted in the same GMO percentage after the calculation. This was true for both MON810 and MON863 (Table 7).

The trueness of the *hmg* method was tested with 1% and 10% certified reference material from a number of different GMO events. The mean GMO percentage was calculated for 20 controls of 1% IRMM reference material from the events MON810 (BF413d), MON863 (BF416c), NK603 (BF415d), MIR604 (BF423c), DAS59122 (BF424c), Bt11 (BF412d), and TC1507 (BF418c). For the 10% level reference material from MON863 (BF416d), MIR604

(BF423d), DAS59122 (BF424d), and TC1507 (BF418d) was used. The relative standard deviation was calculated for both GMO levels. All PCRs were done in duplicates (Table 8).

Discussion and conclusions

All internationally validated GMO detection methods need to be validated in-house before they can be used by maintenance labs in the EU member states.

According to ISO 17025, Directive 96/23/EC [5] and the ENGL Method Performance Criteria [4], new methods that have already been internationally ring-tried need to be validated in-house with respect to the in-house repeatability, in-house reproducibility, accuracy and detection limit of the method. Also, after 3 years of routine use, these data should be updated and extended with extra data obtained during routine analysis.

Requirements for in-house validations have also been discussed by Žel et al. [19].

Table 5 Precision of routine methods (result from 2 isolations×2 PCRs at the same day and limit of detection (based on reference samples 0.1%)

GMO method	RSD _r (repeatability relative standard deviation) (%)	RSD _R (reproducibility relative standard deviation) (%)	Limit of detection (3s _R) (%)	Number of PCR reactions	Number of positive PCR reactions
Roundup ready soy construct/lectin	23	98	0.35	50	46
35S promoter/lectin	28	67	0.16	12	11
Nos terminator/lectin	30	74	0.24	12	11
35S promoter/ <i>SSIIB</i> gene	33	73	0.60	19	17
Nos terminator/ <i>SSIIB</i> gene	11	86	0.30	20	17
Bt11 maize construct <i>SSIIB</i> gene	50	50	0.21	28	28
Bt176 maize construct/ <i>SSIIB</i> gene	34	34	0.19	26	26
Mon810 maize construct/ <i>SSIIB</i> gene	32	32	0.19	24	24
GA21 maize construct/ <i>SSIIB</i> gene	15	25	0.30	24	24
TC1507 maize event/ <i>Hmg</i> gene	24	48	0.10	22	22
Mon810 maize event/ <i>Hmg</i> gene	32	32	0.19	128	126
GA21 maize event/ <i>Adh1</i> gene	15	25	0.30	22	21
NK603 maize event/ <i>Adh1</i> gene	24	48	0.10	22	22
Mon863 maize event/ <i>Adh1</i> gene	18	35	0.12	22	22
RT73 canola event/ <i>FatA</i> gene	— ^a	— ^a	— ^a	24	24
DAS 59122 maize event/ <i>Hmg</i> gene	— ^a	— ^a	— ^a	24	23
EH92-527-1 potato event/ <i>UGPase</i> gene	— ^a	— ^a	— ^a	48	43
H7-1 sugarbeet event/ <i>GS</i> gene	— ^a	— ^a	— ^a	24	24
MIR604 maize event/ <i>Hmg</i> gene	— ^a	— ^a	— ^a	24	21

^a Not calculated for these methods from individual data, because for these methods only the individual PCR steps were validated.

Repeatability and reproducibility

To estimate the repeatability and reproducibility of a routine GMO analysis method, it is advised to isolate DNA from at least eight different samples (with different matrix and different GMO percentage), on two different days, over a period not shorter than 2 months (according to the scheme in Table 1 [11]). It is important to first define what is considered to be the routine GMO analysis. Here, the routine analysis is considered to be two DNA isolations from a 100-mg sample each and two PCR reactions with each isolated DNA (50 ng of DNA in a 25 µl reaction). If more DNA isolations per sample are done or more PCR reactions per DNA, this would influence the repeatability and reproducibility for the resulting average. Also, when DNA is isolated from a different starting amount or with a different method, the repeatability and reproducibility can be influenced.

To validate all real-time GMO detection methods in this way is very laborious and time-consuming, as methods for new GMO events currently become available at a regular pace. Moreover, it should not be necessary because the most important differences between the methods are the sequences of the primers and probes. The choice for a

certain DNA isolation method is based on the matrix, not on the subsequent GMO detection methods (Fig. 1). For the validation series described in the present study, two different DNA isolation methods have been used, with slight modifications. It is not possible in practice to do all DNA isolations in different matrices with exactly the same method so if the validation study of a method is performed with just a single DNA isolation method, this will not reflect the real-life situation. Therefore, a model was developed to estimate the contribution of the different steps in the GMO analysis to the total repeatability and reproducibility (Table 4). In this way, it is possible to estimate how much the ‘DNA isolation’ and ‘PCR day’ variance components contribute to the total variance of the individual GMO analysis/real-time GMO detection methods. The ‘PCR’ variance component was expected to be much larger than the ‘DNA isolation’ and ‘PCR day’ variance, but from Table 4 it can be seen that the factors ‘DNA isolation’ and ‘PCR day’ contribute substantially to the reproducibility. Although the actual values for the repeatability relative standard deviation RSD_r and the in-house reproducibility relative standard deviation RSD_R vary, the reliability of the individual values is not so high since they are based on a limited number of data. For some

Table 6 Trueness of methods for reference samples 1% (w/w)

GMO detection method	1% (w/w) Reference material			
	1% Reference material	n_{tot}	Relative deviation (%)	Significance (p value)
Roundup ready soy construct/lectin	1% Roundup ready soy IRMM-410S-4, ERM BF410d ^a	21/169 ^a	-36, -21 ^a	0.01
35S promoter/lectin	1% Roundup ready soy IRMM-410S-4	11	-31	0.01
Nos terminator/lectin	1% Roundup ready soy IRMM-410S-4	11	-47	0.01
35S promoter/ <i>SSIb</i> gene	1% Bt11 maize IRMM-BF-412-4	14	-13	0.39
Nos terminator/ <i>SSIb</i> gene	1% Bt11 maize IRMM-BF-412-4	13	-15	0.09
Bt11 maize construct <i>SSIb</i> gene	1% Bt11 maize IRMM-BF-412-4	20	+6	0.23
Bt176 maize construct/ <i>SSIb</i> gene	1% Bt176 maize ERM-BF411d	22	+116	<0.001
Mon810 maize construct/ <i>SSIb</i> gene	1% MON810 maize ERM-BF412d	24	+79	<0.001
GA21 maize construct/ <i>SSIb</i> gene	0.99% GA21 maize IRMM 414-3	22	+394	<0.001
TC1507 maize event/ <i>Hmg</i> gene	0.99% TC1507 maize ERM-BF418c	22	-23	<0.001
Mon810 maize event/ <i>Hmg</i> gene	1% MON810 maize ERM-BF413d	22	-20	<0.001
GA21 maize event/ <i>Adh1</i> gene	0.99% GA21 maize ERM-BF414d	21	-43	<0.001
NK603 maize event/ <i>Adh1</i> gene	0.98% NK603 maize ERM-BF415d	22	-43	<0.001
Mon863 maize event/ <i>Adh1</i> gene	0.98% MON863 maize ERM-BF416c	22	-31	<0.001
RT73 canola event/ <i>FatA</i> gene	1% Homemade from 100% RT73 canola AOCS 0304B and 0% RT73 canola AOCS 0304A	23	-20	<0.001
DAS 59122 maize event/ <i>Hmg</i> gene	0.99% DAS 59122-7 maize ERM-BF424c	31	-17	<0.001
EH92-527-1 potato event/ <i>UGPase</i> gene	1% Homemade from 100% EH92-527-1 potato ERM-BF421b and 0% EH92-527-1 potato ERM-BF421a	45	-25	0.22
H7-1 sugarbeet event/ <i>GS</i> gene	1% homemade from 100% ERM-BF419b and 0% ERM-BF419a	29	+39	<0.001
MIR604 maize event/ <i>Hmg</i> gene	1% MIR604 maize ERM-BF423c	31	-23	<0.001

^a Bias calculated from 1% reference material control charts routine testing ERM BF410d.

PCR tests, the v_i values are low, but for some PCR tests higher values were found. We tried to find an explanation for these differences in v_i and presumed that perhaps higher v_i values were caused by the use of difficult samples for DNA isolation. Indeed, if the maize gluten samples were removed from the dataset then the variance caused by the DNA isolation became lower (data not shown). Also, for the soy methods RRS/lectin, 35S/lectin, nos/lectin, relatively high values for v_i were found (Table 4). After repeating the validation of RRS/lectin(2) with a set of routine soy feed samples (Electronic supplementary material, Table S1) the v_i was 0 (Table 4). We therefore conclude that a large v_i can be explained by the use of samples with a difficult matrix for DNA isolation or a low GM copy number in the sample.

Differences in v_d for the various validations could be expected because the validations were carried out over longer or shorter periods of time and during a time period of 3 years. The estimations of v_d and v_i are probably not very accurate for the individual tests. It could therefore be more accurate to use overall estimates, using all obtained data, for v_d and v_i , and add the v_p for the individual GMO tests.

It is difficult to directly compare the repeatability and reproducibility data of our validation studies to the data obtained in the ring trials. In the ring trials, the repeatability and reproducibility were generated by many different labs and calculated per individual GMO concentration and for a single matrix for the different crops, namely meal. Also, in

Table 7 Comparison of two endogenous maize reference genes

Event	<i>adh1</i> GMO % (w/w)	<i>adh1</i> Ct highest standard	<i>hmg</i> GMO % (w/w)	<i>hmg</i> Ct highest standard
MON810 (BF413d)	1.01	24.23	1.04	26.61
MON863 (BF416c)	1.34	23.76	1.21	24.64

Table 8 Trueness of one endogenous reference gene (*hmg*) tested for several events

	Number of samples	RSD (%)	Mean GMO % (w/w)
1% reference material IRMM	20	20.72	1.019
10% reference material IRMM	9	16.60	10.08

most ring trials, the labs received DNA rather than meal material, so in these cases the DNA isolation step was not included in the ring trial. We have compared our reproducibility results to the data from four ring trials where the DNA isolation was also included (Table 9). In the ring trials, two DNA isolations were carried out per sample and three to four PCRs with each DNA. For comparison, our results were also recalculated to 2×3 and 2×4 set-up and RSD_R for the methods were calculated for the CRL ring trial samples with GMO percentages higher than 0.1%. It can be seen that both RIKILT and ring trial RSD_R are in the same range. For MON810 maize, our reproducibility for the 2×3 GMO analysis is 50% and the ring trial value is 38%.

In the ENGL validation studies, the repeatability and reproducibility were calculated for distinct levels of GMO meal between 0.1% and 5%. In this study, the repeatability and in-house reproducibility were calculated using samples of different feed matrices and including GMO percentages between 0.1% and 100% (Electronic supplementary material, Table S1). The assumption is that variance components on the ln scale are constant over this range. Further work is necessary to investigate whether this is a reasonable assumption. It might be found useful to use a more restricted range of levels in future validation studies.

From Table 4, it can be seen that the repeatability standard deviation of our routine method is usually below the ENGL criterion of 25%. The reproducibility standard deviation is in most cases below the criterion of 35%, only for the Roundup Ready soy construct method is it slightly higher, i.e. 37%. Considering the fact that they are based on a wide variety of sample matrices, including real-life samples, these results do not deviate from what can be expected on the basis of the EU validation data and can be considered satisfactory.

Detection limits

In this study, the limit of detection for quantitative results of the 2×2 standard analysis, based on 0.1% reference material, varied from 0.1% to 0.6% depending on the GMO detection method (Table 5, column 4). In this approach, the lowest achievable value for the LOD is in fact 0.1%. It can therefore be seen as a very conservative estimation. This could be caused by differences in efficiencies of the GMO detection methods (Table 2). There are also other reasons why the detection limits may vary. For maize 50 ng DNA of 0.1% GMO maize reference

material, contains on a weight/weight basis 18,349 copies of endogene and 18 copies of GMO event ($1C=2.725$ pg [1]) using calibration curves from genomic DNA. On a haploid genome basis, however, there are only around nine GMO copies in 50 ng DNA of 0.1% GMO maize reference material, but this also varies with the maize variety and the part of the maize plant from which the DNA is extracted [17]. For soy ($1C=1.13$ pg [1]) 50 ng DNA 0.1% GMO soy contains on a weight/weight basis 44,248 copies of endogene and 44 copies of GMO. Since soy is homozygous for the GMO event, this is also the case on a haploid genome basis. For these reasons, it is easier to quantify 0.1% GMO soy than to quantify 0.1% GMO maize in the same amount of DNA. For the construct-specific GMO detection methods, another complication is that the constructs may be present in multiple copies. Although LODs for the standard 2×2 analyses were generally found above 0.1%, empirical results showed in most cases a probability of at least 95% for a positive PCR (Table 5 lanes 5 and 6).

According to the definition of the ENGL, the LOD is the lowest amount or concentration of an analyte in a sample which can be reliably detected but not necessarily quantified, as demonstrated by single laboratory validation. Qualitative methods should detect the presence of the analyte at least 95% of the time at the LOD. In practice, the definition of a positive sample is ambiguous: it may be a positive single PCR test or a positive result in either, e.g. two or four parallel PCR tests. In these studies, all four PCR results (from two DNA isolations and two PCRs with each DNA) should be positive in order to score the sample as positive.

Detection limits for the quantitative GMO detection methods were calculated for the 2×2 standard analyses using 0.1% reference materials (50 ng DNA per reaction, Table 5, lane 4). Routinely, 50 ng of DNA were used per reaction to exclude the possibility of inhibition. At a DNA concentration of 100 ng per reaction, inhibition was observed for some matrices (data not shown).

Bias

To estimate the bias of the methods 1% certified reference materials were used if available. For the construct-specific maize methods, it was not possible to compare the bias (Table 6) to the data of the CRL ring trials (Table 10) because in these GMO detection methods, plasmid standards were used which generate a GMO construct percentage

Table 9 Comparison of RIKILT and ring trial reproducibility data

GMO test	CRL ring trial							RIKILT		
MON863 maize ^b	% (w/w)		0.1	1	5	10				
	RSD _R		34.51	17.81	17.73	20.64		23.71 ^{a,b}	23 ^{a,b}	
MON810 maize ^c	% (w/w)	0.02	0.1	0.5	1	2	5			
	RSD _R	83.03	45.43	43.50	31.86	31.78	36.64		38.27 ^{a,c}	50 ^{a,c}
GA21 maize ^b	% (w/w)		0.1	0.49	0.98	1.3	1.71	4.26		
	RSD _R		43.94	34.72	29.19	30.79	27.13	30.24		30.51 ^{a,b}
NK603 maize ^b	% (w/w)		0.1	0.49	0.98	1.96	4.91			
	RSD _R		37.08	34.45	25.43	26.06	31.04			29.48 ^{a,b}

^a Calculated for all samples with GMO > 0.1% w/w

^b Two DNA isolations and four PCRs with each DNA

^c Two DNA isolations and three PCRs with each DNA

on a haploid genome basis. Also, the copy number of the construct was not exactly known. For quantifications of maize GMOs, the event-specific tests are used. For soy, the percentage on a haploid genome basis is the same as the percentage on weight/weight basis.

A bias between -43% and +37% was found for the event-specific methods in these in-house validations (Table 6). The results were compared to the results of the EU ring trials (Table 10). In the EU ring trials, either extracted DNA or flour samples of reference material were distributed and therefore only four of the in-house validations could be compared to ring trial results (marked in yellow in Table 10). The bias on the 1% level varied from about -16.7% to +46.5% for the four maize-ring trials in which flour samples were distributed and DNA was extracted by the laboratories (Table 10). When DNA was used in the EU ring trials the bias, at 0.9% GMO, was lower, between -1 and +9%. (Table 10).

This means that, as can be expected, the bias is higher when DNA isolation plus real-time PCR are validated compared to the validation of just the PCR step.

The bias in our studies is not in all cases between the set criteria of +25% and -25%. For the NK603, GA21 and MON863 maize the *Adh1* gene is used as a reference gene. It was shown by Broothaerts et al. [2] that in the test sequence of the *Adh1* reference gene, a single nucleotide polymorphism hampers the binding of the reverse primer. Because of this, the copy number of the reference gene will be positively or negatively biased, depending on the *Adh1* genotype of sample and calibrant. Indeed, a relatively high bias was found for these methods: NK603 -43%, GA21 -43% and MON863 -31%. For homemade 1% reference material (RT73 canola event, EH92-527-1 potato event, H7-1 sugar beet), part of the bias may also be caused by the in-house preparation of the material.

The biases were also compared with the results that were obtained in proficiency tests (data not shown). An impor-

tant aspect of proficiency tests is that it enables a comparison of the data from one lab with the results in other labs. In case there is no 1% certified reference material, this is the only way to monitor the accuracy of the method. A limitation is that the participating labs may use different methods for DNA isolation and real-time PCR. If several labs use a method for DNA isolation and/or PCR that causes bias this will influence the assigned value of the proficiency test samples. From this comparison, it can be seen that the results of our lab in the proficiency tests using the methods described in the present study are in line with the outcome of other labs analysing the same samples.

The possibility to use the same endogenous reference gene for all GMOs of the same species was investigated by using different reference genes for individual GMOs and comparing the results. No detectable difference was observed between using *adh1* or *hmg* as a reference gene for MON810 and MON863 when the 1% reference material was measured (Table 7). The Ct value for the two systems was different, but this did not influence the GMO percentage. The *adh1* system from the MON863 validation is known to fail to detect one allele of the gene. As stated above, this can lead to an overestimation of the GMO content in samples where this allele is present. This makes it very relevant to be able to use alternative reference genes. Table 8 shows a trueness of 2% for 1% reference material and of 0.8% for 10% reference material. This is below the 25% which is in the ENGL criteria [4].

As stated earlier, for routine use, it is not convenient to use all the different endogenous genes prescribed in the different methods. In practice, using the same reference gene is also more cost-effective in routine laboratories. We therefore advocate to select one endogenous gene for soy, one for maize, etc., or if this is not possible as few as possible endogenous genes for all varieties of a crop.

Table 10 Comparison of the bias of RIKILT in-house validations and CRL ring trials

Matrix RIKILT	Matrix CRL	GMO % (w/w)	1	0.1	0.4	0.5	0.5	0.9	1.0	1.0	1.3	1.7	2.0	2.0	2.2	2.5	4.3	4.5	4.9	5.0	5.5	6.0	
		GMO detection method	Bias % RIKILT																				
Flour	Flour	MON863 maize event	-31	28.0					20.0														0.0
Flour	Flour	MON810 maize event	-29	2.3		-7.7			-16.7				-10.9										-9.7
Flour	Flour	GA21 maize event	-43	55.0	35.9		17.7	20.9		26.2						8.9							
Flour	Flour	NK603 maize event	-43	83.0	72.9		46.5					14.0											22.1
Flour	DNA	DAS59122 maize event	-17	29.0	15.0			9.0					7.0				-1.0						
Flour	DNA	EH92-527 -1 potato event	-6	4.9	5.1			8.2						4.2									4.3
Flour	DNA	TC1507 maize event	-23	6.0		-4.0		3.7					-1.7										8.4
Flour	DNA	H7-1 sugar beet event	39	-7.4		10.2		6.3					9.1										10.0
Flour	DNA	MIR604 maize	-23	3.6	3.1			-1.0							0.7								-3.6

Only the values in italics can be compared because in these cases the DNA isolation was included in both validation studies

Maintenance of EU GMO regulations

Finally, a decision limit was calculated for each method. The decision limit is the threshold value for the decision whether a sample contains the detected event in a percentage higher than 0.9% with a statistical reliability of 95%. The decision limit varied between 1.10% and 2.41% for the individual methods (Table 4). In this calculation, the bias was not included because there were no sufficient data on the reference materials and/or proficiency samples. In the future, when other well-characterised reference materials, e.g. plasmids are available, it may be possible to include the bias as well. It furthermore needs to be stressed that only the intralab variation is included in these decision limits and not the interlab variation. This latter fact may be of relevance in the case of the analysis of contra-expertise samples to confirm quantified GMO events in a specific sample in a second laboratory.

In conclusion, it can be stated that the present study shows that variations between PCR days (runs), DNA isolations and the PCRs all contribute to the total variance. In most cases, the DNA isolation contributes very little to the total variance, but with difficult samples like maize gluten there can be a large v_i (variation due to differences between isolations, averaged over PCRs (DNA isolation)). Although the data showed a lot of variation between the validations with respect to the values of the v_d (variation due to differences between days, averaged over isolations and PCRs (PCR day) and v_i (variation due to differences between isolations, averaged over PCRs (DNA isolation)), in our opinion, this is because only few samples were used per validation and because in different validations different sets of sample matrices were used. The main technical difference between the GMO methods is the sequence of the primers and probes. For in-house validation of new CRL-validated PCR tests, we therefore propose to use an initial set of in-house validation data using levels in a restricted dynamic range, matrices of comparable complexity and experiments of comparable duration to determine overall values for the v_d and v_i that can then be used for additional validation studies. This approach will shorten the timeframe of the validation study and the number of analyses that will need to be performed and thereby improve the efficacy of the validation procedures. At the same time, it will increase the accuracy of the approach as the overall v_d and v_i values will be based on a larger number of analyses as will be the case in most single validation studies. This improved approach for in-house validation studies will aid the European maintenance labs to keep pace with the entering of new GMO varieties onto the European market.

Acknowledgements We like to thank P. de Heer, J. Busscher-Lange, M.M. Voorhuijzen-Harink, P. Boleij and L. Rutten for carrying out some of the validation experiments.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Arumaganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 9:208–218
2. Broothaerts W, Corbisier P, Schimmel H, Trapmann S, Vincent S, Emons H (2008) A single nucleotide polymorphism (SNP839) in the *Adh1* reference gene affects the quantification of genetically modified maize (*Zea mays* L.). *J Agric Food Chem* 56:8825–8831
3. Currie LA (1999) Detection and quantification limits: origins and historical overview. *Anal Chim Acta* 391:127–134
4. Definition of minimum performance requirements for analytical methods of GMO testing, European network of GMO laboratories (ENGL), Version 25-01-2005, replaced by a new version 18 October 2008 and date of application 13 April 2009 http://gmo-crl.jrc.ec.europa.eu/doc/Min_Perf_Requr_Analyt_methods_131008.pdf
5. Directive 96/23/EC, Official Journal of the European Communities, 17.8.2002, Commission Decision of 12 august 2002 implementing Council directive 96/23/EC concerning the performance of analytical methods and the interpretation of results
6. Commission regulation (EC) No 1981/2006 of 22 December 2006 on detailed rules for the implementation of Article 32 of Regulation (EC) No 1829/2003 of the European Parliament and of the Council as regards the Community reference laboratory for genetically modified organisms
7. Holst-Jensen A, Berdal KG (2004) The modular analytical procedure and validation approach and the units of measurement for genetically modified materials in foods and feeds. *J AOAC Int* 87:927–936
8. ISO 3534-1 (1993) Statistics—vocabulary and symbols. Part 1: probability and general statistical terms. ISO, Genève
9. IUPAC. Compendium of chemical terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell, Oxford (1997) XML on-line corrected version: <http://goldbook.iupac.org> (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins. ISBN 0-9678550-9-8. doi:10.1351/goldbook
10. Kuribara H, Shindo Y, Matsuoka T, Takubo K, Tuto S, Aoki N, Hirao T, Akiyam H, Goda Y, Toyoda M, Hino A (2002) Novel reference molecules for quantitation of genetically modified maize and soybean. *J AOAC Int* 85:1077–1089
11. NEN 7777 (2003) Environment—performance characteristics of measurement methods. Dutch Standard
12. NEN 7779 (2008) Environment—measurement uncertainty. Dutch Standard
13. Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* 58:545–554
14. Payne RW, Harding SA, Murray DA, Soutar DM, Baird DB, Glaser AI, Channing IC, Welham SJ, Gilmour AR, Thompson R, Webster R (2008) The guide to GenStat release 11, part 2: statistics. VSN, Hemel Hempstead
15. de Ruig WG, van der Voet H (1995) Quality in the analytical laboratory: From 'I' to 'we'. In: Parkany M (ed) Quality assurance and TQM for analytical laboratories. The Royal Society of Chemistry, Cambridge, pp 206–214

16. Shindo Y, Kuribara H, Matsuoka T, Futo S, Sawada C, Shono J, Akiyama H, Goda Y, Toyoda M, Hino A (2002) Validation of real-time PCR analyses for line-specific quantitation of genetically modified maize and soybean using new reference molecules. *J AOAC Int* 85:1119–1126
17. Trifa Y, Zhang D (2004) DNA content in embryo and endosperm of maize kernel (*Zea mays* L.): impact on GMO quantification. *J Agric Food Chem* 52:1044–1048
18. VIM (1993) International vocabulary of basic and general terms in metrology, 2nd edn. ISO, Genève
19. Žel J, Mazzara M, Savini C, Cordeil S, Camloh M, Štebih D, Cankar K, Gruden K, Morisset D, Van den Eede G (2008) Method validation and quality management in the flexible scope of accreditation: an example of laboratories testing for genetically modified organisms. *Food Anal Methods* 1:61–72