

# The impact of analytical quality criteria and data evaluation on the quantification of genetically modified organisms

W. Meyer · M. Caprioara-Buda · B. Jeynov ·  
P. Corbisier · S. Trapmann · H. Emons

Received: 10 May 2012/Revised: 4 July 2012/Accepted: 9 July 2012/Published online: 27 July 2012  
© Springer-Verlag 2012

**Abstract** The quantification of genetically modified organisms (GMOs) in food and feed is commonly carried out by event-specific quantitative real-time PCR (qPCR) requiring the use of certified reference materials (CRMs) for the calibration of the qPCR step. Those same CRMs can also be included in the measurement process as quality control samples. The CRMs ERM-BF415e, ERM-BF427c and ERM-BF425c are mixtures of dried conventional and genetically modified maize NK603, maize 98140 and soya 356043 seed powders, respectively. Based on a gravimetric approach, they were prepared and certified for their GMO content expressed as a mass fraction. In a later certification campaign, which required the establishment of analytical quality criteria for the quantification of GMOs described here, they were certified additionally for their genetically modified (GM)-DNA copy numbers in relation to target taxon-specific DNA copy numbers calculated in terms of haploid genomes (DNA copy number ratio). For the three certification campaigns of ERM-BF415e, ERM-BF427c and ERM-BF425c in which the GMO content was measured as DNA copy number ratio, interlaboratory comparisons were run with a large number of expert laboratories. Quality criteria for the evaluation and scrutiny of qPCR raw data were established, such as cut-off limits for the coefficient of determination ( $R^2$ ), PCR efficiency and reproducibility. In this manuscript, different data evaluation approaches for qPCR raw data from the three interlaboratory comparisons

were compared and the impact on the mean DNA copy number ratio and its standard deviation has been assessed. The importance of raw data evaluation and data filtering is discussed and its relevance not only for GMO reference material certification campaigns but also for routine GMO measurements is shown.

**Keywords** Genetically modified organism · Certified Reference Material · DNA extraction · Quantitative real-time PCR · qPCR · Raw data evaluation · Cycle threshold · PCR efficiency · Coefficient of determination · Relative standard deviation · Reproducibility

## Introduction

Legislation in the European Union demands the labelling of food products consisting of or containing “more than 0.9 % genetically modified organisms” (GMOs), provided that the GMO has been placed on the market in accordance with Community legislation [1]. This enforces the necessity to develop and validate reliable GMO-specific quantification methods. Most methods for the detection and quantification of GMOs are based on event-specific quantitative real-time PCR (qPCR) targeting one of the junction regions of the inserted transgenic DNA. These event-specific qPCR methods are developed and proposed by the biotech companies that request authorisation of the GMO event in Europe. As a ratio between genetically modified (GM)-specific and species-specific entities has to be determined, a DNA region around the integration site (the so-called 5' or 3' insert-to-plant junction) and a taxon-specific reference sequence are targeted. The performance of the event-specific qPCR method is further evaluated and approved by the European Union Reference Laboratory for GM Food and Feed (EURL-GMFF) for being suitable for

W. Meyer · M. Caprioara-Buda · B. Jeynov ·  
P. Corbisier (✉) · S. Trapmann · H. Emons  
European Commission, Joint Research Centre,  
Institute for Reference Materials and Measurements (IRMM),  
Reference Materials Unit (RM), Retieseweg 111,  
2440 Geel, Belgium  
e-mail: Philippe.CORBISIER@ec.europa.eu

the quantification of GMOs in view of (EC) No 1830/2003 [1]. Those qPCR methods are preferably applied by GMO testing laboratories including European National Reference Laboratories (NRLs). In parallel, reference materials (RMs) for the calibration and quality control of the qPCR methods are requested by the EU legislation [2]. The Institute for Reference Materials and Measurements (IRMM) provides a number of certified reference materials (CRMs) suitable for the quantification of GMOs such as ERM-BF415, ERM-BF427 and ERM-BF425 for the GM events maize NK603, maize 98140 and soya 356043, respectively. The CRMs ERM-BF415 [3] and ERM-BF427 [4] are a series of materials consisting of mixtures of GM NK603 and 98140 maize powders with conventional maize powders, respectively. The CRM ERM-BF425 [5] is a series of materials composed of GM modified soya 356043 seed powder and conventional soya seed powder. The mass fraction CRMs are intended to be used for quality control, for example, to check for any bias in the measurement or for calibration of the qPCR measurement. Beside the GM content expressed as mass fraction, one of the CRM of each series (namely ERM-BF415e [6], ERM-BF427c [7] and ERM-BF425c [8]) is additionally certified for their GM content expressed as DNA copy number ratio.

The IRMM is committed to produce CRMs based on most advanced metrological concepts and a certified value with an uncertainty fit for the intended purpose. Therefore, three independent interlaboratory comparisons were launched for the quantification and certification of the GM content of ERM-BF415e, ERM-BF427c and ERM-BF425c expressed in DNA copy number ratios. Laboratories were asked to extract DNA from those CRMs, to determine the total DNA concentration and to perform the qPCR measurements. At the time when certification studies are carried out, the GM events can still be in the authorisation process in Europe. This means that the sequences for the primers and probes for the detection of the GM event are still confidential. As a consequence, the laboratories have been asked to analyse the samples without knowing which event they analysed. They were requested to submit raw data in the form of cycle threshold (Ct) values applying automatic threshold settings. The subsequent scrutiny of the data and its evaluation was done by IRMM.

There is typically one validated qPCR method for the quantification of the GM content available per event, and this method is used as reference method as defined in Regulation (EC) No 882/2004 [9]. For qPCR measurements, usually repeatability (relative repeatability standard deviation ( $RSD_r$ )) of up to 25 % and reproducibility (relative reproducibility standard deviation ( $RSD_R$ )) of below 35 % are accepted [10]. In order to minimise the influence of repeatability and reproducibility effects in the characterisation study, a large number of up to 42 data sets were created per interlaboratory comparison. For each data set, the mean

measured GM content expressed as DNA copy number ratio was calculated. Finally, the GM content of ERM-BF415e, ERM-BF427c and ERM-BF425c was established as the mean of means.

In order to find out the best approach for the establishment of the certified values and their uncertainties, the mean of means and its standard deviation were calculated following different approaches: by arithmetic mean from all data sets, by robust statistical procedures and by arithmetic mean from data sets after scrutiny of raw data.

The analysis of raw data from the three large interlaboratory comparisons allowed to investigate the best approach for the certification of matrix CRMs. This study defines a number of measurement parameters that guarantee the quality of the qPCR raw data that are selected to derive a certified value and its related uncertainty.

For the first time, the impact of such quality criteria and data assessment on reported results has been evaluated on a large number of data, and their applicability and relevance not only for RM certification campaigns but also for routine GMO measurements are discussed in this manuscript.

## Materials and methods

### Experimental set-up of interlaboratory comparisons

Three independent interlaboratory comparisons were carried out on two maize CRMs ERM-BF415e and ERM-BF427c and one soya CRM ERM-BF425c, respectively. The CRMs are each composed of a mixture of genetically modified and conventional seed powders. The CRMs were initially certified for their GM content expressed as a mass fraction with certified values and expanded uncertainty of  $19.6 \pm 0.9 \text{ g kg}^{-1}$  for ERM-BF415e,  $10.0 \pm 1.1 \text{ g kg}^{-1}$  for ERM-BF425c and  $20.0 \pm 0.8 \text{ g kg}^{-1}$  for ERM-BF427c. Afterwards, they were additionally certified for their GM content expressed as a DNA copy number ratio based on the interlaboratory comparison data and their evaluation described here. Expert laboratories with proven experience and quality management systems in place provided in total 42 data sets for ERM-BF415e and ERM-BF425c, respectively, and 35 data sets for ERM-BF427c. Three different DNA extraction methods were applied per CRM. For ERM-BF415e, DNA was extracted either by applying a modified cetyltrimethylammonium bromide (CTAB) extraction method (adopted from ISO21571:2005 [11]), the modified CTAB method combined with a Genomic-tip 20/G purification step (Qiagen, Benelux B.V., Venlo, NL) or the GENESpin extraction method (GeneScan Analytics GmbH, Freiburg, DE) [6]. For ERM-BF427c, DNA was extracted by using a modified CTAB method [12], the DNeasy plant mini kit (Qiagen, Benelux B.V., Venlo, NL) or the GENESpin DNA kit (GeneScan Analytics GmbH,

Freiburg, DE) [7]. The DNA from ERM-BF425c was extracted by using a modified protocol based on the DNAExtractor kit (Eurofins GeneScan GmbH, Freiburg, DE) with a Genomic-tip 20/G purification step, the CTAB method combined with a Genomic-tip 20/G purification step, originally validated for cotton 3006-210-23/281-24-236 [13] or a slightly modified protocol of the GENESpin extraction kit (Eurofins GeneScan GmbH, Freiburg, DE) [8].

For each independent analysis, two bottles of the matrix CRMs ERM-BF415e, ERM-BF427c or ERM-BF425c were provided. The laboratories did not know which specific GM event was analysed. From each bottle, three analytical samples were extracted and analysed ( $N = 2, n = 3$ ), further referred to as “unknowns” (U1–U3 for bottle 1, U4–U6 for bottle 2). U1–U3 were analysed on day 1, and U4–U6 were analysed on day 2, that is, the complete analysis of one data set had to be spread over at least 2 days. The total DNA concentration as well as the quality of the DNA extract was estimated by UV spectrometry at 280, 260 and 230 nm or by fluorometry. The DNA concentration of each extract was adjusted to 20 ng/μL with nuclease-free water (referred to as “undiluted extracts”). Afterwards, all extracts were diluted: 1.5-, 3- and 4-fold for ERM-BF427c samples (Fig. 1), 1.5-, 3-, 4- and 6-fold for ERM-BF415e samples and 3-, 6-, 9 and 12-fold for ERM-BF425c samples. The dilution series were analysed by qPCR in triplicate for ERM-BF415e and ERM-BF427c and in quadruplicate for ERM-BF425c according to the EURL-GMFF published event-specific qPCR methods [11, 14, 15]. For the quantification of the maize NK603 event in ERM-BF415e, *hmg* served as reference gene instead of targeting *adh1*. This change of the reference gene avoids an

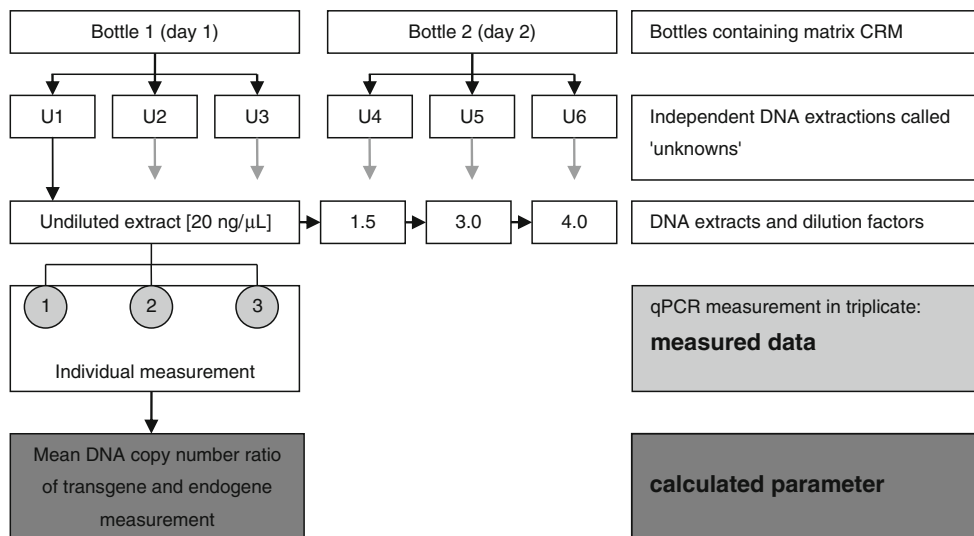
erroneous quantification of the copy number as reported previously by Broothaerts et al. [16]. Event-specific primer pairs and labelled TaqMan® probes for the three studies were provided by IRMM as given in Table 1.

Plasmid DNA calibrant CRMs ERM-AD415, ERM-AD427 and ERM-AD425 with known DNA copy number ratios were used for the calibration of qPCR experiments [17–19]. Dilution series were prepared by the participating laboratories for the calibrants ERM-AD415 and ERM-AD427 with concentrations ranging from 10 to 10<sup>5</sup> copies μL<sup>-1</sup> and for ERM-AD425 with concentrations ranging from 5 to 2 × 10<sup>5</sup> copies μL<sup>-1</sup>. The qPCR experiments using TaqMan Universal mastermix were carried out according to the manufacturer’s instructions (Applied Biosystems, Foster City, CA, USA) on 96-well microtiter plates with all runs performed for 45 cycles. One laboratory adapted the set-up for the performance in 384-well plates for ERM-BF415e and ERM-BF425c. The laboratories were requested to submit raw data in the form of Ct values. These data were scrutinised and evaluated by IRMM. The mean Ct values for the undiluted DNA extracts and their dilutions for each unknown were converted into transgene and endogene copy numbers via the calibration curves by the following equation:

$$\text{DNA copy number} = V \cdot 10^{\frac{Ct_{\text{mean}} - y}{m}}$$

where  $V$  = sample volume of PCR,  $Ct_{\text{mean}}$  = mean cycle threshold from triplicate or quadruplicate measurement,  $y$  = ordinate intercept,  $m$  = slope.

Copy number ratios were calculated from transgene and endogene copy numbers for each unknown for each day.



**Fig. 1** Example for the analysis scheme followed by the participating laboratories during the copy number study. Undiluted extract U1 is given as an example. The same steps have been followed for all other samples and their dilutions. One data set comprises of six DNA

extractions and their dilutions analysed in triplicates (or quadruplicates (not shown)) by qPCR, using the plasmid DNA calibrant CRMs ERM-AD415, ERM-AD427 or ERM-AD425

**Table 1** Primer and probe sequences for event-specific qPCR methods

Target	Code	Sequence 5'–3'	Ref
NK603	NK603 primer F	ATGAATGACCTCGAGTAAGCTTGTTAA	[11]
	NK603 primer R	AAGAGATAACAGGATCCACTCAAACACT	[11]
	NK603 probe	(6-FAM)-TGGTACCACGCGACACACTTCCACTC-TAMRA	[11]
98140	DP098-f6	GTGTGTATGTCTCTTTGCTTGGTCTT	[14]
	DP098-r2	GATTGTTCGTTTCCCGCCTTC	[14]
	DP098-p5	(6-FAM)-CTCTATCGATCCCCCTCTTTGATAGTTTAAACT-TAMRA	[14]
356043	DP356-f1	GTCTGATAGGCTAGGTTTACGAAAAA	[15]
	DP356-r1	TTTGATATTCTTGGAGTAGACGAGAGTGT	[15]
	Dp356-p	(6-VIC)-CTCTAGAGATCCGTCAACATGGTGGAGCAC-TAMRA <sup>a</sup>	[15]
<i>hmg</i>	MaiJ-F2	TTGGACTAGAAATCTCGTGCTGA	[14, 16]
	mHMG-rev	GCTACATAGGGAGCCTTGTCCT	[14, 16]
	mHMG-probe	(6-FAM)-CAATCCACACAAACGCACGCGTA-TAMRA	[14, 16]
<i>le1</i>	Lec for2	CCAGCTTCGCCGCTTCCTTC	[15]
	GMO3-126 Rev	GAAGGCAAGCCCATCTCGAAGCC	[15]
	Lec probe	(6-FAM)-CTTCACCTTCTATGCCCTGACAC-TAMRA	[15]

<sup>a</sup> Method validation was done with a FAM labelled probe

Only the copy number ratios of the undiluted extracts were used for each accepted unknown to calculate the mean GM content expressed as relative DNA copy number ratio.

#### Data acceptance criteria

All data sets needed to fulfil the minimum performance requirements in order to be accepted (Fig. 2): (1) the calibration curves had to consist of 3 or more accepted points from each dilution; (2) Ct values for DNA extracts and their dilutions needed to fall inside the working interval of the calibration curve; (3) the copy number ratio measured on an undiluted extract was used for the calculation of the mean GM content if at least one more copy number ratio measured in its dilution series could be calculated; and (4) a minimum of two DNA copy number ratios per day was needed to fulfil this criterion, that is, a minimum of four copy number ratios per data set were accepted.

For the calculation of the certified values, additional data acceptance criteria were established. The data had also to pass the following internal controls: (1) only data sets with negative results on non-template controls (NTC) were accepted; (2) data points were rejected if any inconsistency in the results within a dilution series was noticeable in the trend of reported Ct values, that is, the more diluted a sample is, the higher the Ct values should be; and (3) means of the triplicate or quadruplicate measurement results were not accepted if the Ct variation between one triplicate or quadruplicate result was higher than 1.5 of the mean Ct of the remaining data for replicates at the particular dilution point ( $\Delta$ Ct). Dilution points were also

excluded if one single amplification signal was absent among the replicates.

The calibration curves were needed to fulfil the minimal performance requirements stipulated by the EURL-GMFF in the method validation guidelines [10]. In particular, these selection criteria apply to the linearity and slope, that is, PCR efficiencies in the calibration curves. Calibration curves with a  $R^2$  above 0.98 were accepted. For the certification campaigns, control limits for PCR amplification efficiencies ( $\epsilon$ ) were established in a method-dependent approach. For each event, transgene and endogene, mean PCR efficiencies and their standard deviations were calculated. Genomic DNA extracted from young plant leaves was also used as calibrant to allow a comparison between plasmid and genomic DNA calibrant, but those data have been submitted in a separate paper [20]. Control limits have been calculated for both calibrants per transgene and endogene as the mean plus or minus one standard deviation, that is, four intervals were generated. The minimum and maximum values of all intervals were then applied as the lower and higher limits of the PCR efficiency per interlaboratory study. The resulting cut-off limits for PCR efficiency applied in the interlaboratory comparisons are given in Table 2.

Data sets considered for the certified value also needed to fulfil the minimum requirements regarding precision. Two types of relative standard deviations (RSD) were considered: (1) the relative repeatability standard deviation ( $RSD_r$ ) between the results on three analytical samples within one day and (2) the RSD for intermediate precision ( $RSD_{ip}$ ) obtained from one laboratory on two bottles over two days.  $RSD_r$  and  $RSD_{ip}$  had both to be below 25 %.

**Fig. 2** Acceptance criteria applied for data evaluation.  $C_t$  cycle threshold,  $R^2$  coefficient of determination,  $\varepsilon$  PCR efficiency,  $RSD_r$  relative repeatability standard deviation and  $RSD_{ip}$  relative standard deviation for intermediate precision

	Acceptance criterion	Applicable to
<b>1. Minimum requirements</b>	calibration curves consist of minimum 3 accepted dilution points	the measured data of calibration curves
	unknowns fitting in the working range of calibration curve	the measured data of unknowns
	DNA copy number ratios of the highest and one other dilution point accepted	the calculated parameter of unknowns
	minimum 2 (from 3) DNA copy number ratio replicates per day accepted	
<b>2. Internal controls</b>	negative non-template controls	all data sets
	higher dilution is reflected in higher $C_t$ values	the measured data of calibration curves and unknowns
	$\Delta C_t$ between replicate measurements $< 1.5$	
<b>3. Calibration Curve</b>	$R^2 > 0.98$	the measured data of calibration curves
	$\varepsilon_{\min} < \varepsilon < \varepsilon_{\max}$	
<b>4. RSD</b>	$RSD_r < 25\%$ and $RSD_{ip} < 25\%$	the calculated parameter of unknowns

**Table 2** Lower and higher limits of PCR efficiency ( $\varepsilon$ ) applied as acceptance criteria for transgene and endogene calibration curves

	ERM-BF415e	ERM-BF427c	ERM-BF425c	Set by ENGL <sup>b</sup>
$\varepsilon$ Interval	77–99 %	87–106 %	76–110 %	90–110 %

<sup>b</sup> Definition of minimum performance requirements for analytical methods of GMO testing (European network of GMO testing laboratories (ENGL), 13 October 2008 [10])

### Calculation of means

Means were calculated following three different approaches. Firstly, the arithmetic means from all available data sets for a given measurement target were calculated, in this manuscript referred to as “mean<sub>all</sub>”. Secondly, as large numbers of data sets were available, the GM content was estimated by determining the median and calculating a robust mean. The Huber’s H15 estimator for robust mean (referred to as “robust mean” hereafter) was chosen as the data sets were of medium size, roughly symmetric and unimodal but contaminated with outliers [21]. As a third approach, the raw data, that is, Ct values provided by the laboratories, were subjected to a data filtering using the quality criteria described here. That means only data based on pre-set values for the coefficient of determination ( $R^2$ ), method-dependent control limits for PCR efficiencies for the calibration curves and a maximum threshold for repeatability in the measurement of the GM content were included in the calculation. As a consequence, extreme values were not taken on board. The arithmetic mean was only calculated from those data sets that fulfilled all quality criteria. In this manuscript, we refer to this value as “mean<sub>scrutinised</sub>”.

Arithmetic means, medians, robust means and standard deviation ( $s$ ) were calculated using Microsoft Excel and the Microsoft Excel Add-in for robust statistics [22]. For the calculation of mean<sub>all</sub>, mean<sub>log</sub>, robust mean, robust mean<sub>log</sub> and median, a total of 37, 35 and 39 qPCR data sets, respectively, were considered for ERM-BF415e, ERM-BF427c and ERM-BF425c. From them, 21, 20 and 15 data sets for ERM-BF415e, ERM-BF427c and ERM-

BF425c, respectively, were accepted for the calculation of the mean<sub>scrutinised</sub>.

### Results

The comparison presented in this manuscript is based on different approaches of calculating the GM content expressed as DNA copy number ratios of the matrix CRMs ERM-BF415e, ERM-BF427c and ERM-BF425c.

#### Impact on calibration

The impact of applying data acceptance criteria on the calibration curves was investigated. Mean  $R^2$  of the calibration curves was always above 0.98 regardless of applying data quality criteria or not (data not shown). Few calibration curves were rejected for technical reasons, in the cases that a calibration curve consisted of less than three dilution points or that a higher dilution was not reflected in higher Ct values (data not shown).

Looking at the mean  $\varepsilon$ , a total of 66 out of 476 calibration curves were rejected when the data quality criteria were applied, which are described in Table 2 and Fig. 2. The mean  $\varepsilon$  and  $s$  increased slightly for the PCR targets *hmg* and 356043, while for all other investigated targets, mean  $\varepsilon$  decreased when the quality criteria were applied (Table 3). A mean  $\varepsilon$  above 100 % was observed for the *le1* PCR target when no limits for  $\Delta Ct$ ,  $R^2$  and  $\varepsilon$  were applied. PCR efficiencies above 100 % are an indication for potential inhibition of the polymerase during the amplification process [23]. However, when the quality criteria are applied, the mean  $\varepsilon$  moved below 100 % for the *le1* target. In all cases,  $s$  has been much lower when  $\Delta Ct$ ,  $R^2$  and  $\varepsilon$  were within the established limits (Table 3).

#### Impact on mean measured copy number ratios and $s$

The above comparison demonstrated that mean  $R^2$  and  $\varepsilon$  were all acceptable even if no acceptance criterion regarding

**Table 3** Mean PCR efficiencies ( $\varepsilon$ )  $\pm s$ , Ct is the cycle threshold,  $R^2$  is the coefficient of determination,  $\varepsilon_{\min}$  and  $\varepsilon_{\max}$  are given in Table 1,  $n$  = number of accepted calibration curves

	ERM-BF415e		ERM-BF427c		ERM-BF425c	
	NK603 $n = 84$	<i>hmg</i> $n = 84$	98140 $n = 70$	<i>hmg</i> $n = 70$	356043 $n = 84$	<i>le1</i> $n = 84$
No quality criteria applied	89.1 $\pm$ 8.6 ( $n = 82$ )	96.7 $\pm$ 9.1 ( $n = 84$ )	96.2 $\pm$ 9.6 ( $n = 70$ )	95.4 $\pm$ 9.1 ( $n = 70$ )	95.1 $\pm$ 9.4 ( $n = 84$ )	102.1 $\pm$ 8.2 ( $n = 84$ )
Selection criteria for $\Delta Ct < 1.5$ , $R^2 > 0.98$ and $\varepsilon_{\min} < \varepsilon < \varepsilon_{\max}$ applied	87.6 $\pm$ 3.5 ( $n = 74$ )	94.9 $\pm$ 2.4 ( $n = 69$ )	95.5 $\pm$ 4.3 ( $n = 62$ )	96.9 $\pm$ 2.9 ( $n = 61$ )	95.5 $\pm$ 6.0 ( $n = 78$ )	99.5 $\pm$ 5.3 ( $n = 66$ )

All calibration curves had to consist of 3 or more accepted dilution points

**Table 4** DNA copy number ratios  $\pm s$

	ERM-BF415e ( $N_d = 42$ )	ERM-BF427c ( $N_d = 35$ )	ERM-BF425c ( $N_d = 42$ )
Minimum requirements fulfilled only <sup>a</sup> (mean <sub>all</sub> )	1.15 $\pm$ 1.01 ( $N_d = 37$ )	1.92 $\pm$ 0.79 ( $N_d = 35$ )	0.87 $\pm$ 0.27 ( $N_d = 39$ )
RSD <sub>r</sub> and RSD <sub>ip</sub> <25 % <sup>a</sup>	0.98 $\pm$ 0.18 ( $N_d = 33$ )	1.92 $\pm$ 0.74 ( $N_d = 29$ )	0.83 $\pm$ 0.17 ( $N_d = 26$ )
$R^2 \geq 0.98$ , 90 % $\leq \varepsilon \leq 110$ %, RSD <sub>r</sub> $\leq 25$ % <sup>a, b</sup>	1.07 $\pm$ 0.31 ( $N_d = 4$ )	1.81 $\pm$ 0.48 ( $N_d = 23$ )	0.80 $\pm$ 0.20 ( $N_d = 12$ )
Sequential application of quality criteria to data set fulfilling the minimum requirements			
Quality criteria regarding internal controls <sup>a, c</sup>			
$R^2 > 0.98$ <sup>a, c</sup>	1.12 $\pm$ 0.97 ( $N_d = 33$ )	1.90 $\pm$ 0.77 ( $N_d = 29$ )	0.87 $\pm$ 0.23 ( $N_d = 29$ )
$R^2 > 0.98$ , $\varepsilon_{\min} < \varepsilon^a$ , $\varepsilon^c$	0.94 $\pm$ 0.13 ( $N_d = 29$ )	1.93 $\pm$ 0.77 ( $N_d = 28$ )	0.86 $\pm$ 0.22 ( $N_d = 26$ )
$R^2 > 0.98$ , $\varepsilon_{\min} < \varepsilon < \varepsilon_{\max}^a$ , $\varepsilon^c$	0.94 $\pm$ 0.13 ( $N_d = 29$ )	1.72 $\pm$ 0.26 ( $N_d = 22$ )	0.84 $\pm$ 0.21 ( $N_d = 25$ )
$R^2 > 0.98$ , $\varepsilon_{\min} < \varepsilon < \varepsilon_{\max}^a$ , $\varepsilon^c$	0.96 $\pm$ 0.11 ( $N_d = 23$ )	1.75 $\pm$ 0.25 ( $N_d = 20$ )	0.87 $\pm$ 0.19 ( $N_d = 20$ )
All quality criteria applied (mean <sub>scrutinised</sub> ) <sup>a, c, d</sup>	0.95 $\pm$ 0.10 ( $N_d = 21$ )	1.75 $\pm$ 0.25 ( $N_d = 20$ )	0.85 $\pm$ 0.13 ( $N_d = 15$ )

$N_d$  = Number of accepted data sets, RSD is the relative standard deviation with RSD<sub>r</sub> for repeatability and RSD<sub>ip</sub> for intermediate precision,  $\varepsilon$  is the cycle threshold,  $R^2$  is the coefficient of determination,  $\varepsilon$  is PCR efficiency, and  $\varepsilon_{\min}$  and  $\varepsilon_{\max}$  are given in Table 1

<sup>a</sup> Data sets were only accepted if they fulfil the minimum requirements, that is, (1) the calibration curve consisted of 3 or more dilution points, (2) unknowns were within the working range of the calibration curve, (3) DNA copy number ratios of the highest and one other dilution point were accepted, and (4) minimum 2 copy number ratios per day could be calculated

<sup>b</sup>  $R^2 \geq 0.98$ , 90 %  $\leq \varepsilon \leq 110$  %, RSD<sub>r</sub>  $\leq 25$  % are set by the ENGL [10]

<sup>c</sup> Data sets were accepted if the NTCs were negative,  $\Delta Ct$  between triplicate or quadruplicate measurements was below 1.5, and the increasing dilution was reflected in higher Ct values

<sup>d</sup>  $R^2 > 0.98$ ,  $\varepsilon_{\min} < \varepsilon < \varepsilon_{\max}$ , RSD<sub>r</sub> and RSD<sub>ip</sub>  $\leq 25$  %. This approach was followed for obtaining the certified values

linearity was applied to improve the data sets. However, the data evaluation of the three interlaboratory comparisons implied that the failure of a single calibration curve for either the transgene or the endogene on one day triggered the omission of the whole data set. Therefore, filtering the data by using the impact of quality criteria described in the “Materials and methods” section on the mean DNA copy number ratios and  $s$  was assessed as follows.

Firstly, data sets needed to fulfil the minimum requirements only (Fig. 2). Following this approach, all 35 data

sets were retained for ERM-BF427c, and 37 and 39 data sets (from 42 each) were accepted for ERM-BF415e and ERM-BF425c, respectively (Table 4). Calculated mean DNA copy number ratios (mean<sub>all</sub>) are given in Table 4. Large standard deviations were obtained (Table 4), corresponding to values for relative reproducibility standard deviations (RSD<sub>R</sub>) of 88, 41 and 31 % for ERM-BF415e, ERM-BF427c and ERM-BF425c, respectively (Table 5). Hence, under these evaluation criteria, the quantification methods for the maize events NK603 and 98140 did not

**Table 5** Comparison of the relative reproducibility standard deviation (RSD<sub>R</sub>) calculated for mean<sub>all</sub>, for data sets filtered by ENGL requirements, for mean<sub>scrutinised</sub> (the certified values), robust means, medians and RSD<sub>R</sub> as reported in the EURL-GMFF validation reports [32–34]

	RSD <sub>R</sub>					
	Mean <sub>all</sub>	For data fulfilling ENGL requirements	Certified value	Robust mean	Median	From EURL-GMFF validation report
NK603	88 %	29 %	12 %	17 %	13 %	26 % <sup>a</sup>
[ERM-BF415]	$N_d = 37$	$N_d = 4$	$N_d = 21$	$N_d = 37$	$N_d = 37$	$N_d = 12$
98140	41 %	27 %	14 %	26 %	15 %	13 % <sup>b</sup>
[ERM-BF427]	$N_d = 35$	$N_d = 23$	$N_d = 20$	$N_d = 35$	$N_d = 35$	$N_d = 12$
356043	31 %	25 %	15 %	28 %	19 %	12 % <sup>c</sup>
[ERM-BF425]	$N_d = 39$	$N_d = 12$	$N_d = 15$	$N_d = 39$	$N_d = 39$	$N_d = 12$

$N_d$  is the number of accepted data sets

<sup>a</sup> RSD<sub>R</sub> for matrix material, GM target value 1.95 % mass fraction. Method validation comprised of DNA extraction and qPCR measurements. *adh1* was used as reference gene

<sup>b</sup> RSD<sub>R</sub> for matrix material, GM target value 2.5 % mass fraction. Method validation comprised of qPCR measurement using DNA extracts provided by Pioneer Overseas

<sup>c</sup> RSD<sub>R</sub> for matrix material, GM target value 0.9 % mass fraction. Method validation comprised of qPCR measurement using DNA extracts provided by Pioneer Overseas

satisfy the method performance requirements for this precision parameter set by the ENGL that should be below 35 % [10].

Secondly, raw data additionally needed to fulfil the requirements for repeatability: the RSD criterion was applied allowing an  $RSD_r$  and an  $RSD_{ip}$  of less than 25 % (Fig. 2). Following this approach, the mean GM content was estimated to be lower than  $mean_{all}$  for ERM-BF415e and ERM-BF425c. The RSD criterion did not have an impact on the mean for ERM-BF427c. However, allowing a RSD maximum value of 25 % as a quality criterion alone led to a drastically decreased  $s$  for ERM-BF415e and ERM-BF425c. The RSD criterion alone did not have an impact on  $s$  for ERM-BF427c (Table 4).

As a third approach, raw data from the three interlaboratory comparisons were scrutinised using method acceptance criteria from the EURL-GMFF ( $R^2 \geq 0.98$ ,  $\varepsilon_{min} = 90\%$ ,  $\varepsilon_{max} = 110\%$  and  $RSD_r \leq 25\%$  [10]). Doing so, only few data sets fulfilled the requirements: 4 out of 42 for ERM-BF415e, 23 out of 35 for ERM-BF427c and 12 out of 42 for ERM-BF425c. The mean DNA copy number ratios were lower than the  $mean_{all}$  for all three CRMs (Table 4). Also the standard deviations obtained following this approach were much lower than the ones when all data sets were accepted, corresponding to  $RSD_R$  of 29, 27 and 25 % for ERM-BF415e, ERM-BF427c and ERM-BF425c, respectively. Thus, the method performance requirement for this precision parameter set by the ENGL was met. However, the  $RSD_R$  values obtained following this approach were higher than the ones reported in the validation reports from the EURL-GMFF (Table 5). Moreover, only a few data sets could be retained, especially for ERM-BF415e and ERM-BF425c. While mean PCR efficiencies fell in the interval of 90 to 110 %, many individual calibration curves could not satisfy this requirement. As the three studies were designed so that the rejection of a single calibration curve led to the rejection of the complete data set, many data sets had to be omitted. With the exception of measurement results for the *le1* target, all mean PCR efficiencies were lower than 100 % (Table 3). Therefore, the PCR efficiency ranges were determined per GM event in a method-dependent approach as described in the “Materials and methods” section (Table 2).

As a fourth approach, the quality acceptance criteria for data sets described above were applied in a sequential manner, comparing DNA copy number ratios when no specific quality criterion was in place ( $mean_{all}$ ) to DNA copy number ratios obtained when data were successively filtered. The mean calculated DNA copy number ratios changed when raw data were subjected to scrutiny (Table 4). Interestingly, for the two maize CRMs,  $mean_{all}$  was clearly higher than  $mean_{scrutinised}$ .

For all three CRMs, the sequential application of quality criteria on qPCR raw data also showed that  $s$  was lowered

**Table 6** Quality criteria as described in Fig. 2 and the number of data sets ( $N_d$ ) that were rejected on the basis of these for measurement data on ERM-BF415e, ERM-BF427c and ERM-BF425c

	ERM-BF415e ( $N_d = 42$ )	ERM-BF427c ( $N_d = 35$ )	ERM-BF425c ( $N_d = 42$ )
Minimum requirements	5	0	3
Internal controls	4	6	10
Calibration	10	9	9
RSD	2	0	5

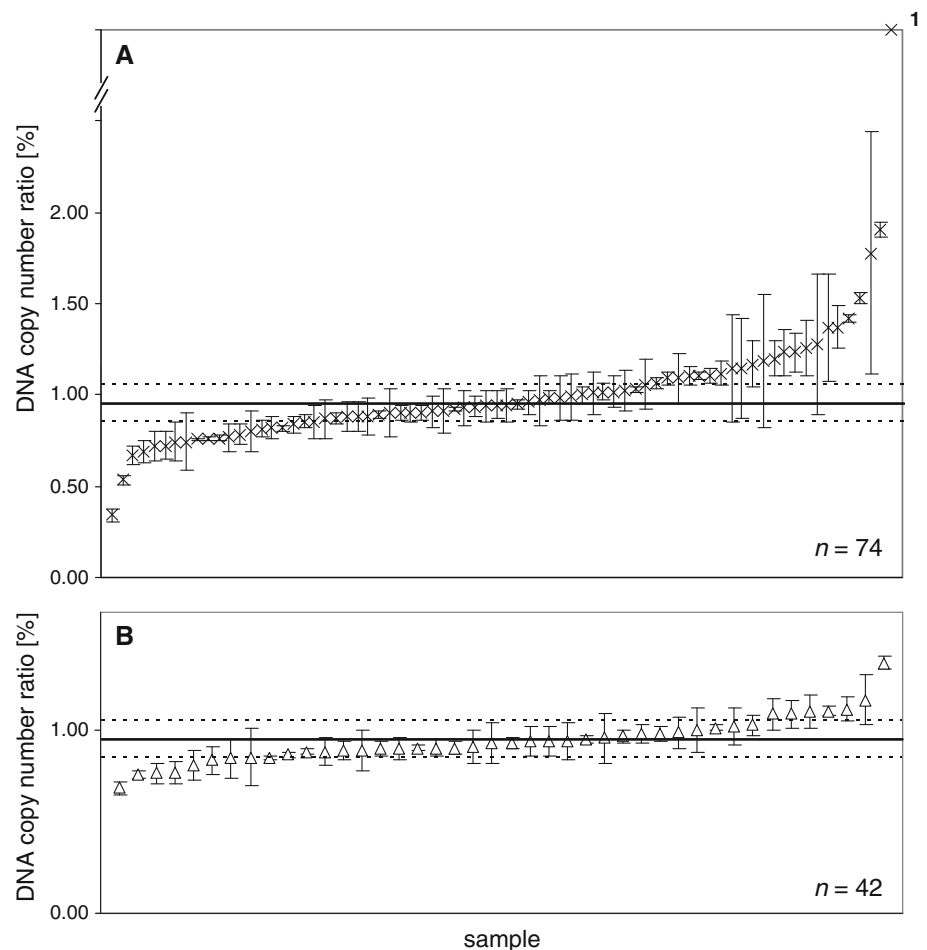
with the number of quality criteria in place (Table 4). The lowest  $s$  was obtained when all quality criteria were applied. In this comparison, the linearity of the calibration curve had the largest single impact on  $s$  for the calculated DNA copy number ratio of ERM-BF415e. For the DNA copy number ratio for ERM-BF427c,  $s$  improved most when a minimum  $\varepsilon$  was established. For ERM-BF425c, the RSD criterion had the largest single impact, as  $s$  decreased drastically when only data sets with  $RSD < 25\%$  were considered (Table 4).

The sequential application of quality criteria also allowed to identify for which reason the majority of data sets was filtered out. Table 6 summarises the number of data sets that were rejected out for particular reasons described above. For the maize CRMs ERM-BF415e and ERM-BF427c, most data sets were not accepted due to not fulfilling the set requirements for linearity and PCR efficiency of the calibration curves. For the soya CRM ERM-BF425c, most data sets were rejected because they failed internal control requirements (i.e. they showed positive results for NTCs or dilution inconsistencies) and second most for not fulfilling the established requirements for calibration curves.

The DNA copy number ratios per analysed sample and  $s$  are presented graphically for ERM-BF415e (Fig. 3), ERM-BF427c (Fig. 4) and ERM-BF425c (Fig. 5). The upper part of the graphs (A) show the results when minimum quality criteria were applied only. In these cases, also extreme values with large  $s$  were obtained. In all three cases, the DNA copy number ratios per sample followed a right-skewed distribution. In the lower part of the graphs (B), the measured DNA copy number ratios from scrutinised data sets are presented. For all three materials, extreme values were not considered for the estimation of the GM content of the samples when quality criteria were applied to the raw data. Large standard deviations were not obtained after the quality criteria were in place. Therefore, this approach was followed for the calculation of the certified value (Table 4). The  $RSD_R$  values for the certified values are 12, 14 and 15 % for ERM-BF415e, ERM-BF427c and ERM-BF425c, respectively (Table 5).



**Fig. 3** DNA copy number ratios (%) for DNA extracted from ERM-BF415e (maize NK603). Error bars represent standard deviation  $s$ . The upper graph **a** ( $\times$ ) shows all DNA copy number ratios for samples fulfilling the minimum requirements for acceptance of a data set. The lower graph **b** ( $\Delta$ ) shows the DNA copy number ratios for samples additionally fulfilling the quality criteria regarding internal controls, calibration curve and RSD as described in Fig. 2. In graph **a** and **b**, the bold line marks  $\text{mean}_{\text{scrutinised}}$  (i.e. the certified value), and the dotted line represents  $\text{mean}_{\text{scrutinised}} \pm s$ .  $n$  is the number of samples. <sup>1</sup>The DNA copy number ratio  $\pm s$  obtained for this sample was  $13.16 \pm 2.54 \%$



### Comparison of calculated DNA copy number ratios

The large number of data sets allowed to establish a reliable system for scrutiny of raw data from qPCR. The calculation and comparison of mean DNA copy number ratios revealed that the scrutiny of raw data as proposed here was advantageous for obtaining results with suitable standard deviations. In all investigated cases, mean DNA copy number ratios were higher when all data sets were accepted ( $\text{mean}_{\text{all}}$ ) than after scrutiny of raw data ( $\text{mean}_{\text{scrutinised}}$ ) (Tables 2, 7). The latter was assigned as the certified values. In the following, it was investigated which evaluation approaches are relevant in order to obtain reliable results from a smaller number of data sets.

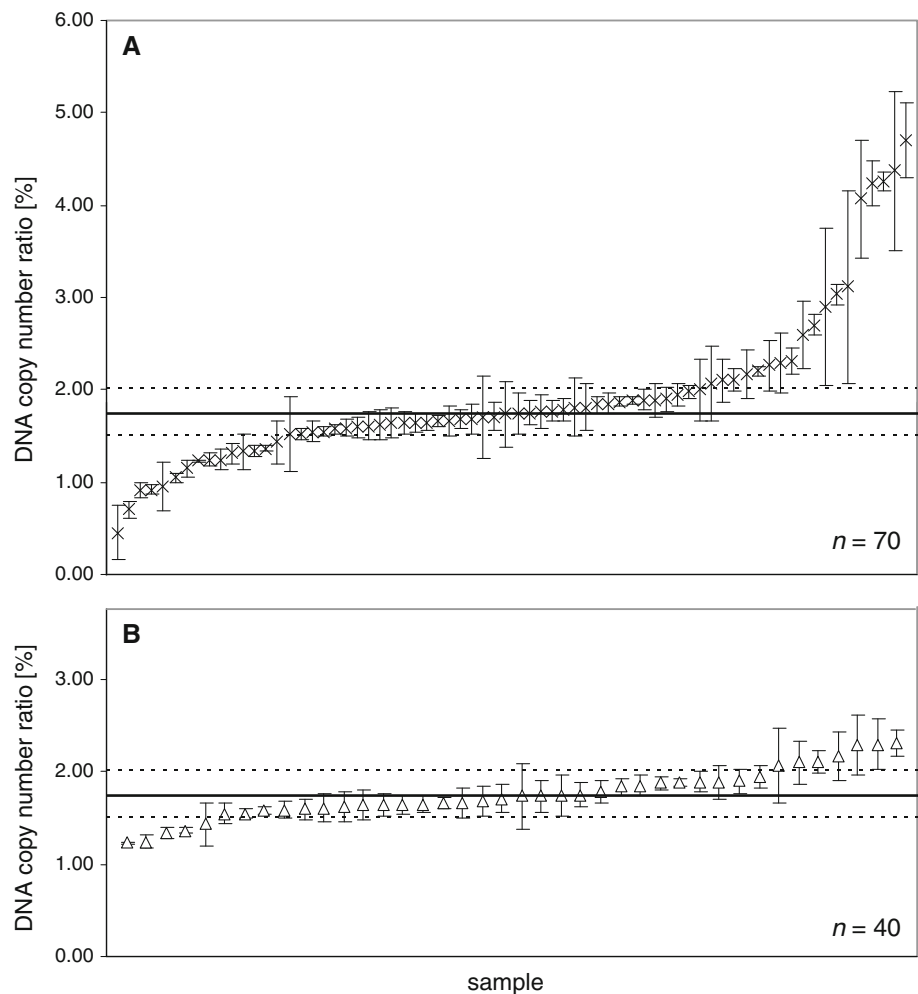
It has been already reported that results from proficiency tests of laboratories quantifying GM content by qPCR show a right-skewed distribution [24–26]. The DNA copy number ratios were transformed into the logarithm to base 10 to follow a normal distribution. For all three GM materials investigated here, the DNA copy number ratios were lower when calculated as arithmetic mean of all data sets after log-transformation ( $\text{arithmetic mean}_{\log}$ ) than by using untransformed data ( $\text{arithmetic mean}_{\text{all}}$ ) (Table 7).

For data on ERM-BF415e, the standard deviation was also smaller after log-transformation (Table 7). For data of the two maize CRMs ERM-BF415e and ERM-BF427c, DNA copy number ratios after log-transformation are still slightly higher than the certified values, while this parameter is slightly lower than the certified value for the soya CRM ERM-BF425 (Table 7).

The robust means calculated from all acceptable data sets were good estimates of the certified values. Their standard deviations were also smaller compared to those of the arithmetic means of all data sets (Table 7). Robust means and  $s$  calculated from log-transformed data were similar to robust means and  $s$  from untransformed data. In all cases investigated here, the median gave the best approximation of the certified value. The median absolute deviations (MAD) as a robust measure of variability, which itself is the median of the absolute deviations from the data's median, were much smaller compared to the standard deviations from the  $\text{mean}_{\text{all}}$  and robust mean (Table 7).

The corresponding  $\text{RSD}_R$  values are given in Table 5. They are in line with the observation above:  $\text{RSD}_R$  values are smaller for the median than for the robust mean.

**Fig. 4** DNA copy number ratios (%) for DNA extracted from ERM-BF427c (maize 98140). Error bars represent standard deviation  $s$ . The upper graph **a** ( $x$ ) shows all DNA copy number ratios for samples fulfilling the minimum requirements for acceptance of a data set. The lower graph **b** ( $\Delta$ ) shows the DNA copy number ratios for samples additionally fulfilling the quality criteria regarding internal controls, calibration curve and RSD as described in Fig. 2. In graph **a** and **b**, the bold line marks  $\text{mean}_{\text{scrutinised}}$  (i.e. the certified value), and the dotted line represents  $\text{mean}_{\text{scrutinised}}$  plus or minus  $1s$ .  $n$  is the number of samples



## Discussion

Regulation (EC) No 1830/2003 of the European Union specifies the labelling of GMOs within the food chain [1]. It demands that food and feed products consisting of or containing more than 0.9 % GMOs authorised in Europe are labelled accordingly. Consequently, analytical methods that are able to identify and quantify GMOs are required. According to Regulation (EC) No 882/2004 [27], the EURL-GMFF shall provide NRLs with such analytical methods, including validated reference methods. In the EU, the quantification of GMO in food and feed is based on one reference method per event [28], which is the method to be applied and referred to in the case of legal dispute. Therefore, it is of utmost importance that these analytical parameters are identified, which are crucial for the adequate performance of the measurement and which have to be checked to guarantee that the methods are well under control.

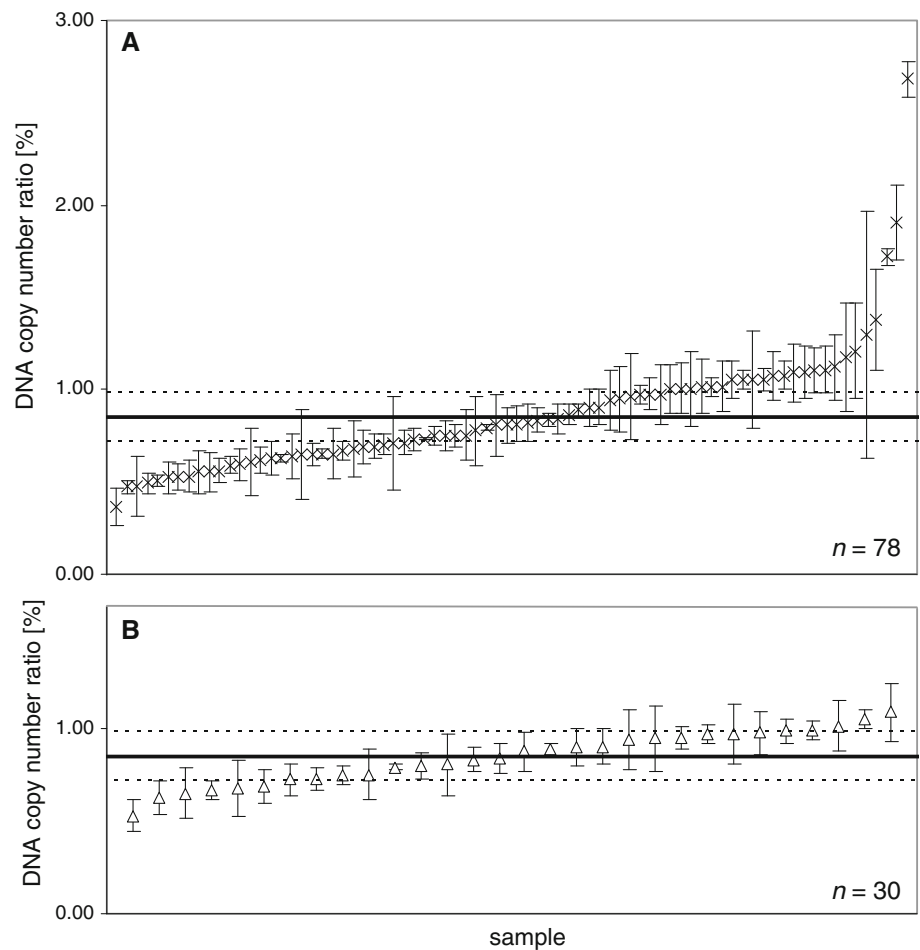
The comparison presented above is based on the evaluation of raw data obtained from the blind analysis of

samples in the frame of three independent interlaboratory studies. The participating laboratories were requested to follow strictly the protocols for DNA extraction, dilutions of calibrants and DNA extracts and the set-up of the qPCR. They had to submit their qPCR raw data and did not have the possibility to treat their data prior submission. They were also not specifically asked to apply quality checks.

By evaluating in parallel a large number of qPCR raw data of the independent interlaboratory comparisons, it was possible to identify factors that influence the reported measurement results.

Certainly, the measurement results depend on the calibration. However, even if calibration curves were not subjected to any filtering by applying any quality criteria, the large majority of calibration curves were within the limits recommended by EURL-GMFF and commonly accepted for linearity and efficiency of qPCR. Only for the PCR target NK603, two calibration curves were not accepted as no signals after amplification were detected for some dilution points (Table 3). The large standard deviation for PCR efficiencies that is observed when no selection

**Fig. 5** DNA copy number ratios (%) for DNA extracted from ERM-BF425c (soya 356043). Error bars represent standard deviation  $s$ . The upper graph **a** ( $x$ ) shows all DNA copy number ratios for samples fulfilling the minimum requirements for acceptance of a data set. The lower graph **b** ( $\Delta$ ) shows the DNA copy number ratios for samples additionally fulfilling the quality criteria regarding internal controls, calibration curve and RSD as described in Fig. 2. In graph **a** and **b**, the bold line marks  $\text{mean}_{\text{scrutinised}}$  (i.e. the certified value), and the dotted line represents  $\text{mean}_{\text{scrutinised}}$  plus or minus  $1s$ .  $n$  is the number of samples



**Table 7** Comparison of the certified value  $\pm$  standard deviation  $s$  to arithmetic  $\text{mean}_{\text{all}} \pm s$ , arithmetic  $\text{mean}_{\log} \pm s$ , robust mean  $\pm s$ , robust  $\text{mean}_{\log} \pm s$  and median  $\pm$  median absolute deviation (MAD) of DNA copy number ratios when data sets fulfil the minimum requirements only

	Certified value <sup>a</sup>	Arithmetic mean <sup>b</sup> <sub>all</sub>	Arithmetic mean <sup>b</sup> <sub>log</sub>	Robust mean <sup>b</sup>	Robust mean <sup>b</sup> <sub>log</sub>	Median <sup>b</sup>
ERM-BF415e <sup>c</sup>	0.95 $\pm$ 0.10	1.15 $\pm$ 1.01	1.03 (−0.32, +0.46)	0.99 $\pm$ 0.17	0.98 (−0.16, +0.19)	0.96 $\pm$ 0.12
ERM-BF427c <sup>d</sup>	1.75 $\pm$ 0.25	1.92 $\pm$ 0.83	1.79 (−0.54, +0.78)	1.79 $\pm$ 0.46	1.77 (−0.41, +0.54)	1.76 $\pm$ 0.26
ERM-BF425c <sup>e</sup>	0.85 $\pm$ 0.13	0.87 $\pm$ 0.27	0.83 (−0.21, +0.28)	0.85 $\pm$ 0.24	0.83 (−0.21, +0.28)	0.87 $\pm$ 0.17

$N_d$  is the number of data sets

<sup>a</sup>  $\text{Mean}_{\text{scrutinised}}$  is the certified value of the CRM

<sup>b</sup> Data sets were accepted if they fulfilled the minimum requirements, that is, (1) the calibration curve consisted of 3 or more dilution points, (2) unknowns were within the working range of the calibration curve, (3) DNA copy number ratios of the highest and one other dilution point were accepted, and (4) minimum 2 copy number ratios (from 3) per day could be calculated

<sup>c</sup> For certified value,  $N_d = 21$ ; for arithmetic  $\text{mean}_{\text{all}}$ , arithmetic  $\text{mean}_{\log}$ , robust mean, robust  $\text{mean}_{\log}$  and median,  $N_d = 37$

<sup>d</sup> For certified value,  $N_d = 20$ ; for arithmetic  $\text{mean}_{\text{all}}$ , arithmetic  $\text{mean}_{\log}$ , robust mean, robust  $\text{mean}_{\log}$  and median,  $N_d = 35$

<sup>e</sup> For certified value,  $N_d = 15$ ; for arithmetic  $\text{mean}_{\text{all}}$ , arithmetic  $\text{mean}_{\log}$ , robust mean, robust  $\text{mean}_{\log}$  and median,  $N_d = 39$

according to quality criteria is performed (Table 3) indicates that individual calibration curves may not necessarily be suitable for calibration. As a consequence, when all calibration curves regardless of their linearity or PCR efficiency were accepted, extreme values of DNA copy number ratios with large standard deviations were obtained (Table 4, Figs. 3a, 4a, 5a). Data followed a positively

skewed distribution. Also, the  $\text{mean}_{\text{all}}$  DNA copy number ratio was estimated to be higher than the certified value.

The characteristics of the calibration curves can be improved if measures are taken regarding a minimum value for  $\Delta\text{Ct}$  of a triplicate measurement in addition to the  $R^2$  and PCR efficiency criteria. In this way, standard deviations for mean PCR efficiencies would be improved.

However, when quality requirements for  $\Delta C_t$ ,  $R^2$  and PCR efficiency were in place, a number of calibration curves could not satisfy these and consequently were rejected. In the set-up of the three interlaboratory comparison studies, the rejection of a single calibration curve for either transgene or endogene triggered the exclusion of the complete data set. The rejection of calibration curves was the main reason for rejection of data sets for the two maize materials ERM-BF415e and ERM-BF427c and the second most reason in the case of the soya material ERM-BF425c (Table 6). Likewise, when the calibration curves met the quality criteria established here, mean measured DNA copy number ratios were in agreement with the certified values. This was not achieved if the RSD criterion or the technical criterion alone was used as quality criteria (Table 4).

The ENGL defined method performance criteria that need to be met before a method is considered to be ready to enter the full validation process. A method is accepted if the slope of the calibration curve is in the range of  $(-3.1 \geq \text{slope} \geq -3.6)$  corresponding to a PCR efficiency range of 90–110 % [10]. Except for the PCR target NK603, mean PCR efficiencies measured in the three interlaboratory comparisons met this requirement, regardless of the data scrutiny level (Table 3). While mean PCR efficiencies fell in the interval of 90–110 % (with the exception for the NK603 PCR target), PCR efficiencies of individual calibration curves did not necessarily fall in this interval. This can be monitored if the PCR efficiency interval of 90–110 % was applied as a quality criterion. If only data sets were considered whose calibration curves showed PCR efficiencies between 90 and 110 %, many data sets were omitted for further calculation: Only 4 out of 42 data sets for ERM-BF415e, 23 out of 35 data sets for ERM-BF427c and 11 out of 42 data sets for ERM-BF425c were accepted (Table 4). On the other hand, the standard deviations of the mean measured DNA copy number ratios were large (Table 4) resulting in higher  $RSD_R$  values than those obtained during method validation studies (Table 5). As there was a large number of data sets available that allowed us to evaluate the raw data in detail, cut-off limits for PCR efficiencies were established in a method-dependent approach for the certification of ERM-BF415e, ERM-BF427c and ERM-BF425c (Table 2). Yet, for a reliable measurement, it is more important that the PCR efficiencies demonstrated by the calibration curve and the PCR efficiency measured for the diluted test sample are in the same range [20, 29].

A tendency to overestimate GM levels by laboratories and a right-skewed distribution of data has been noted [24, 26].

The assigned value for the proficiency testing material is usually calculated using a robust statistical procedure of

log-transformed data [23, 26]. Therefore, in a part of the study presented here, DNA copy number ratios have been calculated after scrutiny of data and were compared to those calculated by robust statistics. For the three materials ERM-BF415e, ERM-BF427c and ERM-BF425c, robust means (of both, log-transformed and untransformed data) and medians are close to the certified values. The lowest standard deviations were achieved when DNA copy number ratios were calculated from technically acceptable data sets only. The data selection based on well-defined acceptance criteria describing the quality of the calibration curves (i.e.  $R^2$  of above 0.98 and PCR efficiencies within a method dependent defined interval) as well as defined minimum requirements in terms of repeatability, intermediate precision and reproducibility as described in Fig. 2 improve calculation results considerably. After scrutiny of raw data, mean DNA copy number ratios followed a normal distribution (Figs. 3b, 4b, 5b).

The data scrutiny by applying the quality criteria described here is not only important for certification campaigns for RMs but also in routine analysis. The comparison presented above shows that, if a large number of data sets are available, reliable results could also be achieved by applying robust statistical procedures. The robust mean and median of unevaluated data sets can be good estimators of the certified value (Table 7). Such large numbers of data sets are usually not available in routine analysis, and consequently, the scrutiny of data as described here is strongly advised.

Identifying and excluding technically invalid data is especially important in view of the new European legislation on low-level presence of GMOs in feed imports [30]. The requirements of this regulation are technically demanding as it sets a non-compliant limit of 0.1 % related to the mass fraction of GM material in feed. This is currently the lowest level considered by the EURL-GMFF and is referred to as minimum required performance limit (MRPL) [10]. The regulation applies to GM material in feed that is authorised in a third country and for which an authorisation procedure for the European Union is pending. It also applies to GM material for which the authorisation has expired. This means that only those GMOs for which detection and quantification methods were accepted by the EURL-GMFF and that show an  $RSD_r$  value of maximum 25 % at the level of 0.1 % related to mass fraction of the GM material are covered by this regulation. EU Member states are obliged to declare animal feed as non-compliant with Regulation (EC) 1829/2003 at the MRPL level of 0.1 % or higher taking into account the uncertainty for the measurement [30, 31].

The comparison of calculating DNA copy number ratios from qPCR raw data by different approaches demonstrates that the lowest  $RSD_R$  values are achieved only when data

were scrutinised and technically invalid data sets were not considered for the reported result. As shown here,  $RSD_R$  of 12, 14 and 15 % was obtained for the measurement of the GM content of ERM-BF415e, ERM-BF427c and ERM-BF425c, respectively (Table 5). These  $RSD$  values for reproducibility were obtained for matrix materials containing GMO mass fractions well above the MRPL of 0.1 %, but at the same time, they are also well below the maximum value for reproducibility recommended by the ENGL, which is 35 % and even 50 % for containing GMO mass fractions of below 0.2 % [10]. Considering this, the approach for scrutiny of raw data as presented here could be followed for routine analysis to lower the measurement uncertainty.

## Conclusion

Taken together, the studies demonstrated that the declared GM content is not only dependent on the ability of the method to deliver good quality data. Indeed, the importance of data evaluation and the impact of data scrutiny after the actual measurement on the reported result have been demonstrated here. For obtaining the most reliable and best possible results, the following measures should be taken: Firstly, raw data, that is, the Ct values should be verified against technical aspects. This means, for instance that NTC should not show an amplification, the variation between replicate measurements should be small and the trend in dilution should be reflected in the trend of measurement results. The latter two criteria are applicable to both the calibrant and the test samples.

Secondly, it is crucial for the measurement system that the calibration is under control which can be monitored by  $R^2$  and PCR efficiency. The comparison has shown that amplification efficiency is slightly suboptimal in most of the methods compared here. PCR efficiencies were always below 100 %. Therefore, PCR efficiency intervals were chosen in a method-dependent approach and applied as acceptance criteria in order to better reflect the real rather than the ideal PCR amplification.

Lastly, repeatability, intermediate precision and reproducibility should be under control. These parameters can easily be monitored by measuring powder CRMs such as ERM-BF415e, ERM-BF427c and ERM-BF425c. The powder CRMs are intended to be used as quality control material and have been subjected to homogeneity and stability studies leading to an uncertainty fit for this intended purpose.

We have demonstrated that technically invalid data can be identified by following this straight forward approach and can then be excluded for the calculation of the result.

In this way, the GM content can be quantified more reliably than calculating the robust mean or median from all data.

## References

1. Regulation (EC) No 1830/2003 of the European Parliament and of the Council of 22 September 2003 concerning the traceability and labelling of genetically modified organisms and the traceability of food and feed products produced from genetically modified organisms and amending Directive 2001/18/EC. Off J Eur Union L (2003) 24–28
2. Commission regulation (EC) No 641/2004 of 6 April 2004 on detailed rules for the implementation of Regulation (EC) No 1829/2003 of the European Parliament and of the Council as regards the application for the authorisation of new genetically modified food and feed, the notification of existing products and adventitious or technically unavoidable presence of genetically modified material which has benefited from a favourable risk evaluation. Off J Euro Union L (2004)
3. Trapmann S, Conneely P, Corbisier P, Gancberg D, Gioria S, Van Nyen M, Schimmel H, Emons H (2005) The certification of reference materials of dry-mixed maize powder with different mass fractions of NK603 maize certified reference materials ERM<sup>®</sup>-BF415 (ERM<sup>®</sup>-BF415-a/ERM<sup>®</sup>-BF415-b/ERM<sup>®</sup>-BF415-c/ERM<sup>®</sup>-BF415-d/ERM<sup>®</sup>-BF415-e/ERM<sup>®</sup>-BF415-f), EC certification report, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements
4. Gancberg D, Corbisier P, de Andrade Silva E, Mazoua S, Merveillie A, Tumba M-F, Trapmann S (2009) Certification of reference materials of maize seed powder with different mass fractions of genetically modified 98140 maize. Certified reference material ERM<sup>®</sup>-BF427 (ERM<sup>®</sup>-BF427a, ERM<sup>®</sup>-BF427b, ERM<sup>®</sup>-BF427c, ERM<sup>®</sup>-BF427d), EC certification report, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements p 20
5. Gancberg D, Conneely P, Contreras Lopez M, Corbisier P, Gioria S, Van Nyen M, Vincent S, Trapmann S (2007) Certification of reference materials of soya powder with different mass fractions of 356043 soya, ERM<sup>®</sup>-BF425 (ERM<sup>®</sup>-BF425a, ERM<sup>®</sup>-BF425b, ERM<sup>®</sup>-BF425c, ERM<sup>®</sup>-BF425d), EC certification report, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements p 20
6. de Andrade E, Jeynov B, Corbisier P, Kortekaas AM, Trapmann S, Vincent S, Emons H (2011) Certification of a maize NK603 reference material for its DNA copy number ratio. Certified reference material ERM<sup>®</sup>-BF415e. ISBN 978-92-79-19074-2. EC certification report EUR 24699 EN, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements
7. Caprioara-Buda M, Trapmann S, Gancberg D, Corbisier P, Merveillie A, Contreras Lopez MC, Emons H (2011) Certification of maize 98140 reference material ERM<sup>®</sup>-BF427c. ISBN 978-92-79-19279-1. EC certification report EUR 24718 EN., European Commission, Joint Research Centre, Institute for Reference Materials and Measurements
8. Meyer W, de Andrade E, Broeders S, Charels D, Corbisier P, Merveillie A, Trapmann S, Emons H (2011) Certification of a soya 356043 reference material for its DNA copy number ratio. Certified reference material ERM<sup>®</sup>-BF425c. ISBN 978-92-79-19070-4. EC certification report EUR 24695 EN, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements

9. Regulation (EC) No 882/2004 of the European parliament and of the council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules. Off J Euro Union L (2004)
10. Definition of minimum performance requirements for analytical methods of GMO testing. Available via [http://gmo-crl.jrc.ec.europa.eu/doc/Min\\_Perf\\_Requirements\\_Analytical\\_methods.pdf](http://gmo-crl.jrc.ec.europa.eu/doc/Min_Perf_Requirements_Analytical_methods.pdf), European network of GMO Laboratories (ENGL), 2008
11. Event-specific method for the quantification of maize line NK603 using real-time PCR. European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), biotechnology and GMOs Unit—community reference laboratory, 2005
12. Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucl Acids Res 8:4321–4325
13. Sampling and DNA extraction of cotton seeds. Report from the validation of the “CTAB/Genomic-tip 20” method for DNA extraction from ground cotton seeds, European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), biotechnology and GMOs Unit—community reference laboratory, 2006
14. Event-specific method for the quantification of maize 98140 using real-time PCR, EURL-GMFF: validated method maize 98140, European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), Molecular Biology and Genomics Unit—European union reference laboratory for GM food and Feed (EURL-GMFF), 2011
15. Event-specific method for the quantification of soybean event DP-356043-5 using real-time PCR. European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), biotechnology and GMOs unit—community reference laboratory, 2009
16. 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 quantitation of genetically modified maize (*Zea mays* L.). J Agric Food Chem 56:8825–8831
17. Jeynov B, de Andrade E, Broothaerts W, Corbisier P, Mazoua, Merveillie A, Trapmann S, Emons H (2011) Certification of plasmid DNA containing NK603 maize DNA fragments. Certified reference material ERM<sup>®</sup>-AD415. ISBN 978-92-79-19071-1. EC certification report EUR 24696 EN, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements
18. Caprioara-Buda M, Corbisier P, Gancberg D, Broeders S, Trapmann S, Mazoua S, Matayron G, Emons H (2011) Certification of plasmid DNA containing 98140 maize DNA fragments. Certified reference material ERM<sup>®</sup>-AD427. ISBN 978-92-79-19277-7. EC certification report EUR 24722 EN., European Commission, Joint Research Centre, Institute for Reference Materials and Measurements
19. Corbisier P, Broeders S, Meeus N, Mazoua S, Meyer W, Trapmann S, Emons H (2011) Certification of plasmid DNA containing 356043 soya DNA fragments. Certified reference material ERM<sup>®</sup>-AD425. ISBN 978-92-79-18993-7, EC certification report EUR 24687 EN, European Commission, Joint Research Centre, Institute for Reference Materials and Measurements, p 40
20. Caprioara-Buda M, Meyer W, Jeynov B, Corbisier P, Trapmann S, Emons H (2012) Evaluation of plasmid and genomic DNA calibrants used for the quantification of genetically modified organisms. Anal Bioanal Chem. doi:10.1007/s00216-012-6104-6
21. Analytical Methods Committee (2011) Technical brief No. 6. Robust statistics: a method of coping with outliers. RSC, AMC Technical Briefs 6
22. Analytical Methods Committee (1989) Robust statistics-how not to reject outliers. Part 1. Basic concepts. Analyst 114:1693–1697
23. Suslov O, Steindler DA (2005) PCR inhibition by reverse transcriptase leads to an overestimation of amplification efficiency. Nucleic Acids Res 33:e181
24. Thompson M, Ellison SL, Owen L, Mathieson K, Powell J, Key P, Wood R, Damant AP (2006) Scoring in genetically modified organism proficiency tests based on log-transformed results. J AOAC Int 89:232–239
25. Analytical Methods Committee (2004) GMO proficiency testing: interpreting z-scores derived from log-transformed data. RSC, AMC Technical Briefs 18
26. Powell J, Owen L (2002) Reliability of food measurements: the application of proficiency testing to GMO analysis. Accred Qual Assur J Qual, Comp Reliab Chem Meas 7:392–402
27. Regulation (EC) No 882/2004 of the European parliament and the council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules. Off J Euro Union L (2004)
28. Compendium of reference methods for GMO analysis, JRC reference reports, 2011
29. Mygind T, Birkelund S, Birkebaek N, Ostergaard L, Jensen J, Christiansen G (2002) Determination of PCR efficiency in chelex-100 purified clinical samples and comparison of real-time quantitative PCR and conventional PCR for detection of *Chlamydia pneumoniae*. BMC Microbiol 2:17
30. Commission regulation (EU) No 619/2011 of 24 June 2011 laying down the methods of sampling and analysis for the official control of feed as regards presence of genetically modified material for which an authorisation procedure is pending or the authorisation of which has expired. Off J Euro Union L (2011)
31. Regulation (EC) No 1829/2003 of the European parliament and of the council of 22 September 2003 on genetically modified food and feed. Off J Euro Union L (2003)
32. Event-specific method for the quantification of maize line NK603 using real-time PCR. Validation report, European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), biotechnology and GMOs unit—community reference laboratory, 2005
33. Event-specific method for the quantification of maize 98140 using real-time PCR. Validation Report, European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), molecular biology and genomics unit—European union reference laboratory for GM food and feed (EURL-GMFF), 2011
34. Event-specific Method for the quantification of soybean event DP-356043-5 using Real-time PCR. Validation Report, European Commission, Joint Research Centre, Institute for Health and Consumer Protection (IHCP), biotechnology and GMOs unit—community reference laboratory, molecular biology and genomics unit—European union reference laboratory for GM food and feed (EURL-GMFF), 2009