

Testing the Robustness of Validated Methods for Quantitative Detection of GMOs Across qPCR Instruments

E. Luque-Perez · M. Mazzara · T. P. Weber · N. Foti ·
E. Grazioli · B. Munaro · G. Pinski · G. Bellocchi ·
G. Van den Eede · C. Savini

Received: 14 November 2011 / Accepted: 16 May 2012 / Published online: 12 June 2012
© The Author(s) 2012. This article is published with open access at SpringerLink.com

Abstract The number of real-time polymerase chain reaction (qPCR) instruments available has greatly increased over recent years. However, little information is available on the performance of validated quantitative methods when tested in different instruments. A study has been designed to evaluate the robustness of three validated methods for genetically modified organisms (GMO) quantification across six real-time platforms from four different suppliers. The performance of three validated methods for event-specific detection of Bt11 maize, DAS 59122 maize and MON 89788 soybean was evaluated on six qPCR platforms (ABI 7900 HT, ABI Prism® 7700 and ABI 7500 from Applied Biosystems; LightCycler® 480 from Roche; Mx3005P® from Stratagene; and iQ™5 from Bio-Rad). Method performance criteria were compared against European Network of GMO Laboratories method performance requirements for analytical methods of GMO testing. The latter comparison indicates that the criteria are fulfilled for most of the platforms and levels of GMO concentrations, though with minor exceptions. The analysis of variance (one-way) indicated that the quantification of GMOs was affected by the platforms, which did not respond consistently across GM levels. A two-way analysis of variance confirmed that there were significant interactions between platforms and GM levels. The pairwise comparison of the platforms'

performance indicated that some deviate significantly from others, though differences between methods were observed.

Keywords Comparison PCR equipments · European Union · EU-RL GMFF · Food · GMO · Method validation · qPCR instrument · Robustness · Maize 59122 · Maize Bt11 · Soy MON89788 · Traceability · Labelling · JRC

Introduction

In the European Union (EU), labelling of any food and feed product containing or consisting of genetically modified organisms (GMOs) is mandatory (Regulation No 1830; Regulation No 1829). In order to comply with existing legislation, analytical methods for GMO quantification and detection must be available. Currently, the polymerase chain reaction (PCR) and, particularly, the quantitative real-time PCR for the determination of a target analyte in a sample (qPCR) has become the technique of choice for GMO determination.

The accurate and reliable identification and quantification of a target DNA in a sample is a complex operation. It entails adequate knowledge and sufficient control of several factors, such as the quality of the nucleic acids, appropriate choice of reference materials, comparable amplification efficiencies between calibrants and samples, stability of the event- and reference systems, and the specificity levels of the assays (Shokere et al. 2009; Berdal et al. 2008; Corbisier et al. 2007; Cankar et al. 2006; Broothaerts et al. 2008; ISO 21570 European Committee for Standardization/International Organization for Standardization 2005; ISO 24276 European Committee for Standardization/International Organization for Standardization 2006). Accreditation under an international standard (ISO 17025) is a pre-requisite and,

E. Luque-Perez · M. Mazzara · T. P. Weber · N. Foti · E. Grazioli ·
B. Munaro · G. Pinski · G. Bellocchi · G. Van den Eede ·
C. Savini (✉)

Molecular Biology and Genomics Unit, Institute for Health
and Consumer Protection, Joint Research Centre, European
Commission,
via E. Fermi 2749,
21027 Ispra, Varese, Italy
e-mail: cristian.savini@jrc.ec.europa.eu

importantly, control laboratories should make use of validated methods for quantitative detection.

International bodies (ISO 5725 International Organization for Standardization 1994; IUPAC 1988) provide comprehensive standards describing procedures to assess analytical method performance. The Codex Alimentarius Commission recently adopted into the international food code the ‘Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods’ (Codex Alimentarius Commission 2010). In synergy with the international standards and guidelines, the European Network of GMO Laboratories (ENGL) provided recommendations (European Network of GMO Laboratories 2008) on how methods for GMO analysis should be evaluated and validated by the European Union Reference Laboratory for Genetically Modified Food and Feed (EU-RL GMFF), which was set up to validate methods for detection of GMOs as its main task (Regulation No 1829). So far, the EU-RL GMFF has validated 40 event-specific methods for single GMO events, through inter-laboratory collaboration with members of the ENGL, in accordance with internationally accepted standards (ISO 5725 International Organization for Standardization 1994; Horwitz 1995) and has verified the performance of the methods on 21 stacked events. Additionally, the EU-RL GMFF is responsible for providing harmonisation and communication of scientific data among laboratories, for monitoring the quality levels of the GMO control laboratories and for contributing to level capacities through training and workshops (EC Regulation No 882/2004 Off. J. Eur. Union, L165, 1 2004).

ENGL guidelines on method evaluation take ‘robustness’ into account among the method acceptance criteria and state that it is ‘a measure of the method capacity to remain unaffected by small and deliberate variations from the experimental conditions described in the procedure.’ In the latest revision of the document on method acceptance criteria (2008), the ENGL has further elaborated on the issue of robustness and—in total consistency with the aforementioned CODEX document (Annex II: validation of a quantitative PCR method, clause 24)—provided practical guidance on the elements for consideration and recommended testing the method with different thermal cycler models. In recent years, in fact, the number of available qPCR instruments has greatly increased, with several companies commercialising a wide variety of models. However, the scientific community has little information on the performance of methods when transferred to different instruments. A study of the repeatability of three qPCR platforms using plasmid DNA was completed by Donald et al. (Donald et al. 2005). Also two qPCR platforms were compared to develop a quantitative assay for the detection of a human virus DNA

(Nitsche et al. 1999). However, in the field of GMO determination, a study of the cross-platform mobility of a method has not been performed.

In the GMO testing field, international guidelines do not introduce the distinction between ruggedness and robustness in force in analytical chemistry (Dejaegher and Vander Heyden 2007; Vander et al. 2001) about the source of the influence on test results, whether external or internal to the experimental procedure; only the term robustness is applied. Accordingly, in the framework of this study, the latter term is used. In this context, the EU-RL GMFF designed a study to evaluate the cross-platform mobility of different validated methods (robustness) for identification and quantification of certain GMOs (namely Bt11 maize, DAS 59122 maize and MON 89788 soybean) in single-laboratory conditions. These three single-line GMOs have been variously inbred with other GMOs to give rise to many ‘stacked’ products, characterised by the fact of harbouring more than one GM trait per organism, thus mainly tackling insect pests and carrying tolerance to herbicides. Methods were chosen based on the different sample matrices (species), Cq vs. Δ Cq calibration procedure, passive reference dye (Rox or sulforhodamine) and the practicability of the reaction setup (i.e. number of PCR mix components) which can make the system more prone to deviations. Six different models of qPCR equipment (detailed in ‘Materials and Methods’) were also selected for the study ABI 7900 HT Fast Real-Time PCR system, ABI Prism® 7700 Sequence Detection System and ABI 7500 Real-Time PCR system, from Applied Biosystems; LightCycler® 480 Real-Time PCR System, from Roche; Mx3005P® PCR system, from Stratagene; and iQ™5 Real-Time PCR Detection System, from Bio-Rad. Selection was carried out in order to meet some conditions: to match the limited availability of qPCR equipments at EU-RL GMFF with the frequency of qPCR models occurring in the ENGL community and, subsequently, to consider technical characteristics so as to attain the maximum possible variability among instruments, in terms of excitation source, detector, thermocycling system, excitation spectrum and detection channels (Biocompare 2011; Biochemica 2005; Logan et al. 2009).

Materials and Methods

Samples, DNA Extraction and Quantification

Samples used in this study were Bt11 and DAS-59122-7 maize seeds, heterozygous for the hosted modification and conventional maize seeds, MON 89788 soybean homozygous seeds and conventional soybean seeds. Each seed sample was ground independently in a Grindomix model GM200 from Retsch GmbH to obtain a homogeneous

powder. The extraction of genomic DNA from seeds was carried out on samples of 6 g of the powder previously obtained, using two different methods: (1) for maize events, a modified CTAB method, well established in the GMO testing community and verified by the EU-RL GMFF in the context of method for maize line MON 88017 (European Union Reference Laboratory 2999a) and (2) for the soybean event, a method also verified by the EU-RL GMFF in the context of soybean line MON 89788 (European Union Reference Laboratory 2999b).

The fragmentation state of the DNA was checked by agarose gel electrophoresis, using a Power Pac 300 electrophoresis power supply from Bio-Rad. The purity of the extracted DNA was tested for the absence of PCR inhibitors (Žel et al. 2008) and for contamination of the samples with other GMOs. The latter test was performed using the ready-to-use multitarget analytical system, developed in the Molecular Biology and Genomics unit at the Institute for Health and Consumer Protection of the EC-Joint Research Centre (Querci et al. 2009). High molecular-weight DNA that was proven to be not inhibited and not contaminated was used during the study.

Concentration of extracted DNA solutions was determined by fluorimetry, using the PicoGreen dsDNA quantitation kit from Invitrogen, Molecular Probes Inc. and the fluorescence measured in a VersaFluor fluorometer from Bio-Rad.

Sample Preparation

Samples containing different GMO levels of Bt11 maize, DAS 59122 maize and MON 89788 soybean, respectively, were prepared by mixing DNA extracted from each of the genetically modified (GM) events (100 %) with genomic DNA of its conventional counterpart. Table 1 shows the five GM levels used in the study.

The set of DNA standards and test samples at different GM levels was prepared for each method and thoroughly homogenised. Each GM level standard and test sample was aliquoted into different tubes to produce several sets consisting of the different GM levels. The amount of each

Table 1 Bt11 maize, DAS 59122 maize and MON 89788 soybean GM contents of the unknown samples used

Bt11 GM% (mass/mass)	DAS 59122 GM% (mass/mass)	MON 89788 GM% (mass/mass)
0.09	0.1	0.1
0.4	0.4	0.4
0.9	0.9	0.9
5	2.0	4.0
8	4.5	8.0

aliquot covered for the needs of testing a method on a certain qPCR machine plus an excess factor. All the aliquots had been frozen at the same time and kept at -20°C up to their use, so to minimise any degrading effect of the DNA. The DNA standards covered the validated dynamic range and the concentration of the test samples. All the GM levels samples and the calibration standards were built in DNA mass/mass ratio.

Real-Time PCR Instruments

Six instruments of four different brands were used in this study: ABI 7900 HT Fast Real-Time PCR system, ABI Prism[®] 7700 Sequence Detection System and ABI 7500 Real-Time PCR system, from Applied Biosystems; Light-Cycler[®] 480 Real-Time PCR System, from Roche; Mx3005P[®] PCR system, from Stratagene; and iQTM5 Real-Time PCR Detection System, from Bio-Rad.¹ Relevant differences in terms of technical characteristics of the real-time PCR instruments are: excitation source (argon laser in ABI 7900HT and ABI 7700, tungsten lamps in ABI 7500 and iQ5, quartz tungsten lamp in Mx3005P and xenon lamp in LC480); thermal cycling system (all used Peltier blocks with slight variations among them); detector (most used CCD camera: ABI 7500, ABI 7700, LC480 and iQ5; ABI 7900 used spectrograph and CCD camera and Mx3005P used photomultiplier tube) and a variety of wavelengths in excitation spectrum and detection channels.

All the instruments were maintained fully operational in accordance with the provisions of ISO 17025. In particular, the instruments were calibrated and regularly checked under instrument performance verification schemes run by the suppliers. In addition, tests to show absence of background noise in the blocks were applied on a routine basis, RNase P 96-Well Instrument Verification Plate was applied to ABI instruments. Only plates recommended by the manufacturers were applied on the respective qPCR platforms.

Real-Time PCR Methods

Three event-specific methods for the quantification of three different GMO events, already validated by the EU-RL GMFF, were chosen for this study: maize event Bt11 (European Union Reference Laboratory 2999c), maize event DAS 59122 (European Union Reference Laboratory 2999d) and soybean event MON 89788 (European Union Reference Laboratory 2999e).

The method for maize event Bt11 uses sulforhodamine as passive reference dye in PCR and the quantification is based on a delta Cq method (ΔCq). On the other hand, the methods

¹ The platforms in this paper will be abbreviated as: 7900, 7700, 7500, LC480, 3005 and iQ5 respectively.

for maize event DAS 59122 and soybean event MON 89788 use Rox as passive reference dye and the quantification is based on a C_q method, with two calibration curves: one for the GM- and another for the reference system (taxon-specific).

Experimental Design

Each GM level was tested in 16 replicates per each method on each PCR platform. Each replicate sample resulted from the average of three repetitions pipetted in three adjacent wells according to a format elsewhere discussed (Žel et al. 2008). For method Bt11 and for each platform, 16 replicates of the GM levels 8.00, 5.00, 0.40 and 0.09 % were assayed in eight qPCR runs over 4 days (two replicates/run per day); 16 replicates of GM level 0.90 % were assayed in two qPCR runs (eight replicates/run) in 1 day; for method DAS59122 and for each platform, 16 replicates of the GM levels 4.50, 2.00, 0.40 and 0.10 % were assayed in eight qPCR runs over 4 days (two replicates/run per day) and 16 replicates of GM level 0.90 % were assayed in two qPCR runs (eight replicates/run) in 1 day; for method MON89788 and for each platform, 16 replicates of the GM levels 8.00, 4.00 and 0.40 % were assayed in eight qPCR runs over 4 days (two replicates/run) and 16 replicates of GM levels 0.90 and 0.10 % were assayed each in two qPCR runs (eight replicates/run) in 1 day. No template controls were added to each amplification run.

The samples were consecutively loaded on the 96-wells plate, starting from the left upper corner of the plate: first, calibration samples, then the tested samples at the various GM concentrations, followed by controls in all qPCR runs.

Each method was tested by the same operator on the six PCR platforms under study in repeatability conditions, i.e. the same GM levels were tested by the same operator with the same method on the same equipment, with the same reagents. ISO 5725-1:1994 requires that under repeatability conditions the tests are performed within a short interval of time.

To assess whether the factor ‘time’ was not affecting the outcome of the experiments, the uncertainty of measurement was estimated via an analysis of variance (ANOVA) where the four-level factor ‘number of days’—over which the experiments were run—was tested to verify whether the between-day variability was not significant. The ANOVA results provided variability measures (i.e. standard deviations) for quantifying both intermediate precision relative standard deviation (RSD_i, in percent) reported in Tables 2, 3 and 4.

The procedures for raw data analysis of the runs were standardised, i.e. settings of threshold and baseline. Threshold and baseline were set manually according to the EU-RL

GMFF accredited procedures for method validation and verification.

Data Processing

Performance Criteria

For each GMO event-specific method, platform and GMO level, two accuracy measurements were computed according to the standards of International Organization for Standardization (ISO 24276 European Committee for Standardization/International Organization for Standardization 2006): trueness and precision. Trueness assesses the agreement between measured values and an accepted reference value and is expressed in terms of average bias (*B*) in percent (difference between the average value of the test results and an accepted reference value).

Precision indicates the agreement among measured values. In this study, where test results with the same method are obtained in one laboratory under repeatability conditions, precision is expressed as RSD_r in percent.

According to ENGL minimum performance criteria for analytical methods of GMO testing (Codex Alimentarius Commission 2010), *B* should be within ±25 % and RSD_r ≤ 25 %, over the whole dynamic range investigated. Two performance parameters were also considered in association to the qPCR runs: amplification efficiency (*E*) in percent and linearity (expressed as the coefficient of determination *R*²). Amplification efficiency rates for both reference and GM target systems (*E*₁ and *E*₂, respectively) were obtained from the average slopes of the standard curves using the formula

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

Linearity is measured for both amplification systems (*R*₁², reference system; *R*₂², GM system) as determination coefficients of the standard calibration curves obtained by linear regression analysis (except for Bt11 where only a single calibration ΔC_q curve is built). Also for these PCR performance parameters, ENGL acceptance thresholds are used as reference: 90 % ≤ *E* ≤ 110 %; *R*² ≥ 0.98.

Statistical Analysis

All statistical analyses were carried out using Statistica 9.1® (Statsoft Inc., Tulsa, OK, USA). Unless stated otherwise, the criterion for significance was *p* < 0.05 for all comparisons. Levene’s test was used to assess departures from homogeneous variances. The normality of the data was tested using the Kolmogorov–Smirnov test and the Liliefors test. Outliers were identified by using Grubbs’ test. Values lying beyond 99 % of the range of the characterised distribution

Table 2 GMO levels and platforms used for Bt11 maize determination, average values of % GM DNA, qPCR performance measurements (*E* amplification efficiency in percent and R^2 linearity of standard curve) and method accuracy measurements (*B* bias in percent and RSD_r relative repeatability standard deviation in percent)

Platform	GMO level	Average GMO content (%)	Performance parameters (average values)			
			qPCR		Accuracy	
			<i>E</i> (± 95 % CI)	R^2	<i>B</i> (%)	RSD_r (%)
iQ5	A (8.00 %)	7.42	90.7 (± 2.8)	0.969	-7.2	7.9
	B (5.00 %)	3.57			-28.6	9.6
	D (0.40 %)	0.48			19.2	14.5
	E (0.09 %)	0.07			-19.1	21.5
	C (0.90 %)	1.37	109.1 ^a 98.0 ^a	0.998 0.982	52.3	22.1
3005	A (8.00 %)	7.66	95.9 (± 4.7)	0.988	-4.2	10.3
	B (5.00 %)	5.05			0.9	15.5
	D (0.40 %)	0.42			5.1	14.0
	E (0.09 %)	0.09			-2.8	17.4
	C (0.90 %)	0.83	98.8 ^a 97.1 ^a	0.998 0.997	-8.13	17.0
7900	A (8.00 %)	8.15	95.5 (± 4.2)	0.998	1.9	8.3
	B (5.00 %)	4.85			-3.1	8.7
	D (0.40 %)	0.42			3.8	17.1
	E (0.09 %)	0.09			-4.8	15.1
	C (0.90 %)	0.9	101.6 ^a 96.5 ^a	0.997 0.999	-0.5	8.1
7700	A (8.00 %)	8.11	97.9 (± 2.5)	0.998	1.3	4.5
	B (5.00 %)	5.31			6.2	7.5
	D (0.40 %)	0.43			6.8	9.3
	E (0.09 %)	0.1			8.2	14.5
	C (0.90 %)	0.96	98.5 ^a 92.6 ^a	0.996 0.996	6.9	6.7
7500	A (8.00 %)	7.62	94 (± 2.3)	0.998	-4.8	4.2
	B (5.00 %)	4.71			-5.8	3.0
	D (0.40 %)	0.4			1.1	8.8
	E (0.09 %)	0.09			-4.1	14.6
	C (0.90 %)	0.87	93.5 ^a 90.5 ^a	1.000 0.998	-3.8	6.6
480	A (8.00 %)	8.29	98.7 (± 3.4)	0.998	3.6	9.7
	B (5.00 %)	5.31			6.2	9.2
	D (0.40 %)	0.43			6.5	8.5
	E (0.09 %)	0.09			0.6	17.3
	C (0.90 %)	0.91	95.2 ^a 99.1 ^a	0.997 0.9982	1.1	5.9

GM levels A, B, D and E were assayed in eight qPCR runs (two replicates per run); GM level C was tested in two distinct qPCR runs (eight replicates per run)

^a95 % CI not shown

(i.e. $p < 0.01$) were removed from the analysis. In order to establish the effect of PCR platform and percentage of GM content on quantification results, factorial ANOVAs with platform type and GM level as factors were carried out. Tukey's honestly significant difference test was used to identify homogeneous groups across GM levels.

The option for post hoc power analyses (Faul et al. 2007) provided by G*Power3 (Cohen 1988) was run to assess whether or not the ANOVA tests performed in this study

in fact had a fair chance of rejecting an incorrect null hypothesis. In all cases, the variance explained by main effects and interaction was high, as reflected by high values of Cohen's f effect size ($f > 3$). Under these conditions, the calculated power approached 1, which means that the probability of falsely retaining an incorrect null hypothesis is near 0.

We employed pairwise comparisons between platforms to establish whether the differences between two platforms

Table 3 GMO levels and platforms used for DAS 59122 maize determination, average values of % GM DNA, qPCR performance measurements (E_1 , E_2 reference and GM target gene efficiencies in percent; R_1^2 , R_2^2 linearity of standard curves of reference gene and target gene, respectively), accuracy measurements (B bias in percent and RSD_r relative repeatability standard deviation in percent)

	GMO level	Average GMO content (%)	Performance measurements (average values)					
			qPCR				Accuracy	
			E_1 (±95 % CI)	E_2 (±95 % CI)	R_1^2	R_2^2	B (%)	RSD_r (%)
iQ5	A (4.50 %)	3.36	101 (±2.6)	103.1 (±3.8)	0.994	0.986	-25.4	8.6
	B (2.00 %)	1.39					-30.7	7.3
	D (0.40 %)	0.25					-37.4	16.3
	E (0.10 %)	0.06					-37.1	36.1
	C (0.90 %)	0.72	100.4 ^a 94.5 ^a	101.5 ^a 98.6 ^a	0.987 0.987	0.988 0.990	-20.1	7.1
3005	A (4.50 %)	4.36	105.4 (±4.2)	95.6 (±7.7)	0.996	0.980	-3.0	10.0
	B (2.00 %)	1.77					-11.7	11.3
	D (0.40 %)	0.44					10.6	24.6
	E (0.10 %)	0.1					2.0	27.4
	C (0.90 %)	0.82	109.4 ^a 106.0 ^a	103.2 ^a 101.1 ^a	0.996 0.997	0.974 0.982	-8.2	17.0
7900	A (4.50 %)	4.26	106.8 (±2.8)	101.8 (±4.1)	0.998	0.989	-5.3	5.8
	B (2.00 %)	1.67					-16.4	6.0
	D (0.40 %)	0.36					-9.3	14.8
	E (0.10 %)	0.08					-21.5	16.1
	C (0.90 %)	0.86	105.6 ^a 104.7 ^a	111.0 ^a 106.4 ^a	0.999 0.999	0.986 0.994	-4.9	10.1
7700	A (4.50 %)	4.33	101.8 (±1.8)	102.5 (±5.8)	0.998	0.991	-3.7	7.8
	B (2.00 %)	1.87					-6.4	9.2
	D (0.40 %)	0.36					-10.4	8.3
	E (0.10 %)	0.09					-12.8	15
	C (0.90 %)	0.80	101.0 ^a 102.6 ^a	118.4 ^a 94.9 ^a	0.999 0.995	0.998 0.995	-10.9	14.5
7500	A (4.50 %)	4.83	101.9 (±1.1)	103.2 (±2.9)	0.998	0.990	7.4	4.6
	B (2.00 %)	2.08					4.1	5.7
	D (0.40 %)	0.4					0.6	11.1
	E (0.10 %)	0.1					-1.3	18.4
	C (0.90 %)	0.84	104.0 ^a 103.6 ^a	99.3 ^a 105.1 ^a	0.999 0.998	0.995 0.975	-6.8	6.7
LC480	A (4.50 %)	4.48	97.5 (±1.5)	110.7 (±2.6)	0.997	0.993	-0.4	12.0
	B (2.00 %)	1.97					-1.5	6.6
	D (0.40 %)	0.28					-30.4	13.3
	E (0.10 %)	0.07					-34.4	22.5
	C (0.90 %)	0.8	98.1 ^a 98.2 ^a	104.5 ^a 95.1 ^a	1.000 0.999	0.989 0.979	-11.6	5.0

GM levels A, B, D and E were assayed together in eight qPCR runs (two replicates per run); GM level C was tested in two distinct qPCR runs (eight replicates per run)

^a95 % CI not shown

were relevant. There are a number of regression techniques to compare the agreement between two measurement methods. Ordinary least square (OLS) regression can only be

used if the independent variable is known without error. In pairwise comparisons of measurement results, this assumption does not apply, and type II regression analysis such as

Table 4 GMO levels and platforms used for MON 89788 soybean determination, average values of % GM DNA, qPCR performance measurements (E_1 , E_2 reference and GM target gene efficiencies in percent; R_1^2 , R_2^2 linearity of standard curves of reference gene and target gene, respectively), accuracy measurements (B bias in percent and RSD_r relative repeatability standard deviation in percent)

Platform	GMO level	Average GMO content (%)	Performance measurements (average values)						
			qPCR				Accuracy		
			E_1 (± 95 % CI)	E_2 (± 95 % CI)	R_1^2	R_2^2	B (%)	RSD_r (%)	
iQ5	A (8.00 %)	9.77	109.8 (± 1.3)	107.4 (± 1.7)	0.996	0.990	22.1	8.6	
	B (4.00 %)	4.69					17.2	8.4	
	D (0.40 %)	0.37					-7.2	7.9	
	C (0.90 %)	1.04	102.5 ^a	113.9 ^a	0.992	0.962	15.5	16.7	
			103.8 ^a	105.7 ^a	0.995	0.998			
E (0.10 %)	0.1	102.5 ^a	106.3 ^a	0.997	0.998	-2.5	8.6		
		106.6 ^a	107.8 ^a	0.996	0.992				
3005	A (8.00 %)	7.25	107.5 (± 3.0)	106.3 (± 2.9)	0.994	0.993	-9.4	9.0	
	B (4.00 %)	3.5					-12.6	9.5	
	D (0.40 %)	0.28					-30.8	20.0	
	C (0.90 %)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
	E (0.10 %)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	
7900	A (8.00 %)	8.44	108.5 (± 1.0)	102.9 (± 1.2)	0.996	0.996	5.4	5.4	
	B (4.00 %)	3.59					-10.3	6.3	
	D (0.40 %)	0.35					-12.4	7.8	
	C (0.90 %)	0.88	101.7 ^a	99.4 ^a	0.995	0.994	-2.4	6.5	
			102.5 ^a	100.9 ^a	0.997	0.997			
7700	A (8.00 %)	8.25	107.5 (± 2.9)	103.6 (± 2.9)	0.992	0.996	3.1	8.0	
	B (4.00 %)	3.86					-3.6	14.0	
	D (0.40 %)	0.31					-22.3	19.5	
	C (0.90 %)	0.79	101.9 ^a	98.8 ^a	0.991	0.993	-11.8	12.8	
			102.4 ^a	101.8 ^a	0.997	0.997			
7500	A (8.00 %)	8.57	104.9 (± 2.0)	100.8 (± 1.2)	0.996	0.996	7.1	7.3	
	B (4.00 %)	3.64					-8.9	8.0	
	D (0.40 %)	0.38					-5.1	8.9	
	C (0.90 %)	0.93	103.7 ^a	99.6 ^a	0.997	0.996	3.1	10.8	
			102.7 ^a	100.0 ^a	0.999	0.996			
LC480	A (8.00 %)	7.61	106.9 (± 1.4)	109.2 (± 1.8)	0.993	0.993	-4.9	7.6	
	B (4.00 %)	3.4					-14.9	9.3	
	D (0.40 %)	0.28					-30.8	6.6	
	C (0.90 %)	0.78	106.9 ^a	104.9 ^a	0.996	0.994	-13.2	6.4	
			103.3 ^a	108.0 ^a	0.997	0.995			
E (0.10 %)	0.07	101.4 ^a	105.4 ^a	0.998	0.992	-31.4	7.2		
		101.3 ^a	104.8 ^a	0.996	0.993				

GM levels A, B and D were assayed together in eight qPCR runs (two replicates per run); GM levels C and E were tested each in two distinct qPCR runs (eight replicates per run)

n.a. data for levels C and E in platform 3005 are not available due to a technical failure of the instrument

^a 95 % CI not shown

major axis (MA) or reduced major axis (RMA) regression should be employed (Brace 1977; Ludbrock 1997, 2010). These methods differ in the way errors from the fitted line are measured. However, provided that the coefficient of determination R^2 exceeds 0.95, OLS, MA and RMA regression give nearly equivalent numerical values for the regression coefficients (Niklas 2004): for example, slope (RMA) \approx slope (OLS)/ R . As $R^2 > 0.95$ in all our data sets and because the estimation of confidence intervals for type II regression analysis is complex (Jolicoeur 1990), while OLS regression allows the straightforward calculation of confidence intervals of the estimated coefficients, we have chosen to apply OLS regression in our analysis. Recommendations of MIQE guidelines were taken in consideration when applicable to the present work (Bustin et al. 2009).

Results

Selection of GMO Detection Methods

Three methods for identification and quantification of GMOs were selected in this study. They concern maize events Bt11, DAS59122 and soybean event MON89788. The three single lines had been variously inbred with other GM lines, so originating stacked GMO products where multiple traits are harboured in the same organism (GM-crop database 2999). An online register is available to search for the EU authorisation status of GMO products (EU register of genetically modified food and feed Available at http://ec.europa.eu/food/dyna/gm_register/index_en.cfm 2999).

The three methods can be considered good examples of two general approaches to GMO quantification, based on ‘two-standard curves’—method (DAS59122, MON89788) or on the ‘ ΔC_q ’—method (Bt11). The reaction setup differed across the three methods: universal master-mix (Applied Biosystems) to assemble the MON89788 reaction, the Jumpstart Readymix (Sigma) for Bt11. The simplified approach of using pre-made master-mixes is now in use to streamline the optimisation of the reaction and to make the system less prone to operator’s errors. The opposite approach is the one in use for DAS 59122, where up to 15 different ingredients were separately added as required by the validated protocol.

Therefore, the two factors taken into consideration for the selection of the validated methods used in this study were the calibration method and complexity of the chemistry of the PCR reactions.

Selection of Platforms

The platforms were selected after reviewing those available among control laboratory members of the ENGL network.

The data correspond to 81 laboratories with 164 instruments available. Figure 1 represents the distribution of instruments per company and model.

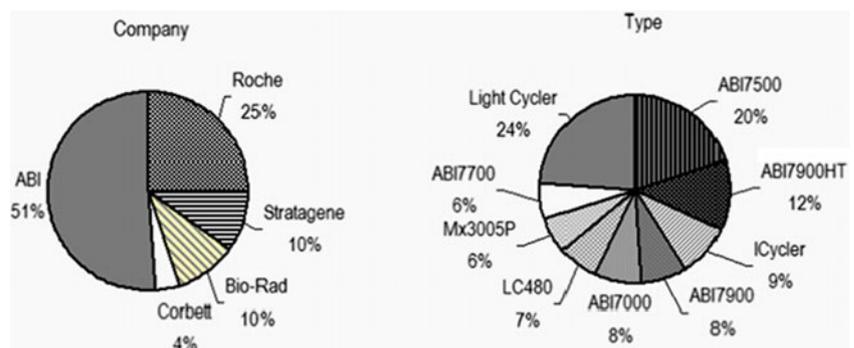
Method Performance Parameters

Tables 2, 3 and 4 summarises the measured GMO content when applying the three GM event-specific methods (Bt11 maize (Table 2), DAS 59122 maize (Table 3) and MON 89788 soybean (Table 4)) in six different qPCR platforms (iQ5, 3005, 7900, 7700, 7500 and LC480) for five different GM levels and shows the accuracy of the analysed GMO levels, expressed as bias (B , in percent) and relative repeatability standard deviation (RSD_r , in percent). Performance measurements of qPCR runs, such as amplification efficiency (E) and linearity (R^2) parameters, are also indicated.

The Bt11 event-specific method (Table 2) meets the ENGL performance criteria for accuracy parameters of trueness (B) and precision (RSD_r) (within ± 25 and ≤ 25 %, respectively), in all platforms except in iQ5. This instrument showed a bias of -28.6 % at the 5 % GM level and 52.3 % for the 0.9 % GM level and very high variation throughout the dynamic range, with no clear pattern of underestimation or overestimation of the GM content. Additional method performance measurements, i.e. amplification efficiency (E) and R^2 coefficient, also comply with the ENGL criteria ($90 \% \leq E \leq 110$ % and $R^2 \geq 0.98$) in most cases. The performance parameters for the event-specific DAS 59122 maize method (Table 3) generally comply with the ENGL performance criteria when using most of the platforms but with some exceptions. In platform iQ5, the bias values are above the ENGL criteria over the whole linearity range, except for the 0.9 % GM level, and the precision for the lowest concentration (0.1 % GM) exceeds the acceptance limit of 25 %. In the LC480 platform, the biases are poor at low GM levels (-34.4 % at 0.1 % GM and -30.4 % at 0.4 % GM). Finally, in platform 3005, the lowest GM level (0.1 %) shows a precision of 27.4 %, outside the 25 % ENGL criterion. Data in Table 4 show that most of the performance parameters for the event-specific MON 89788 soybean method met the ENGL criteria. However, the MON 89788 method shows a higher variation of bias values independently of the platform used in most of the platforms. It was not possible to obtain data in platform 3005 for levels 0.9 and 0.10 % due to a technical failure in the course of the experiments.

The Bt11 and the MON 89788 methods show acceptable precisions independently of the platform used, with RSD_r values below 25 % and acceptable variability throughout the whole dynamic range. The DAS 59122 method generally shows precisions compliant with the ENGL criteria, with the exception of platforms 3005 and iQ5 that present lower

Fig. 1 Distribution of most commonly disseminated qPCR instruments within ENGL laboratories classified by the company and the most frequently used models



precisions at the lowest GM levels (27 and 36 %, respectively).

Effect of the Platform and GM Level on GMO Quantification Results

Tables 5, 6 and 7 summarise the mean GM content measured by each platform and at each GM level for all three events. Outliers were identified using the Grubbs' test (indicated in the table) and removed as described before ('Materials and Methods') before the data were summarised and analysed. In a first step, at each GM level, a one-factor ANOVA with platform type as the main factor was carried out in order to derive groups of platforms that provide similar measurement results at a given GM level. These groupings are summarised in the last column of Tables 5, 6 and 7, in which platforms are ordered from the lowest to the highest estimate. Means followed by the same number are not significantly different. Means marked with two numbers are not significantly different from means marked with either number alone.

For Bt11, levels A (8.0 %), B (5.0 %) and E (0.09 %) show similar patterns, with iQ5 producing low measurement results and LC480 giving high estimates; however, for levels C (0.9 %) and D (0.4 %), iQ5 measures high values, while LC480 measures low values. Additionally, 7700 also measures low values at all GM levels. For DAS 59122, iQ5 shows low measurement results at all levels; LC480 shows a similar pattern of low levels, but overestimates the two highest levels (A and B); and 7500 measures high values at all GM levels. For MON 89788, iQ5 overestimates and LC480 underestimates at all GM levels.

The pattern displayed by iQ5 and LC480 are thus the most relevant. The iQ5 often estimates GM contents deviating significantly from other platforms. The LC480 platform shows some deviations, but these are generally less pronounced than for iQ5.

A higher number of groupings shows higher variability between platforms (Fig. 2). The DAS 59122 and MON 89788 methods show the same number of groups, though distributed differently in the several levels of concentration.

This indicates that their performance in the different platforms shows similar variability. The Bt11 method presents a lower number of groups, ranging from two to three. This points to good consistency of the method's performance among instruments, though the reduced grouping number can also reflect the lesser variability into play when the measurement exercise is resolved by interpolation against one standard curve (ΔCq) of the difference between the Cq values for the GM and reference system of each sample tested, instead of interpolating the two variables separately against the respective calibration curves as it occurs in a 'two-standard curve' method.

This analysis demonstrates that platforms do not respond consistently across GM levels, i.e. some platforms overestimate GM content at some levels and underestimate it at others. This means that there are interactions between platform types and GM levels. Thus, a two-factor ANOVA was employed to examine the effect of the main factors, platform and GM level and of platform-GM level interactions on the measurement result. Significant interactions indicate that results cannot be interpreted by considering each main factor in isolation, i.e., it is not possible to predict the measurement result as the sum of the effects of platform, GM level and error. If there is a significant interaction indicated in the ANOVA table, each combination of platform type i (six levels) and GM level j (five levels) requires an additional term to predict the expected measurement result. This can be written in a simplified equation as: $\text{measurement}_{ij} = K_0 + K_1 \times \text{platform}_i + K_2 \times \text{GM level}_j + K_3 \times \text{platform}_i \times \text{GM level}_j + \text{error}$, where K_0 – K_3 are fitted constants. There are thus $6 \times 5 = 30$ interaction terms; out of which 20 terms are free to vary (the rows and columns of the interaction matrix $\text{platform}_i \times \text{GM level}_j$ must always sum to 0). In our analysis, interactions including platform LC480 and GM level C were fixed by the statistical software and therefore do not show up in the analysis.

Table 8 illustrates that the measurement results for GM levels are, as expected, significantly different from each other, but also that platforms show significantly different responses. Furthermore, there is a significant interaction between platform type and GM level.

Table 5 Grubbs' test for outliers, one-factor (platform) ANOVA on % GM DNA quantification and means grouping (Tukey HSD) with five GM levels for Bt11 maize event-specific method

GMO level	Grubbs' outliers ($p < 0.01$)	p ANOVA Platform	Means grouping (Tukey HSD, $p < 0.05$)		
			Platform	Mean GM content (%)	Group
Bt11					
A (8.00 %)	No	0.013	iQ5	7.42	1
			7500	7.62	1–2
			3005	7.66	1–2
			7700	8.11	1–2
			7900	8.15	1–2
B (5.00 %)	No	<0.0001	LC480	8.29	2
			iQ5	3.57	1
			7500	4.71	2
			7900	4.85	2–3
			3005	5.05	2–3
C (0.90 %)	No	<0.0001	7700	5.31	3
			LC480	5.31	3
			3005	0.83	1
			7500	0.87	1
			7900	0.90	1
D (0.40 %)	No	<0.0001	LC480	0.91	1
			7700	0.96	1
			iQ5	1.37	2
			7500	0.40	1
			7900	0.42	1
E (0.09 %)	1 ^a	0.002	3005	0.42	1
			LC480	0.43	1–2
			7700	0.43	1–2
			iQ5	0.48	2
			iQ5	0.07	1
			7900	0.09	1–2
			7500	0.09	1–2
			3005	0.09	1–2
			LC480	0.09	1–2
			7700	0.10	2

^aOutlier value=0.14 %

Visual inspection shows that in plots of measurement results against platform, the lines for the different GM levels are not parallel (Fig. 3a), which is an indication of interaction effects. A detailed inspection of all the individual interaction terms shows that 8 out of the 20 interactions are significant: iQ5 shows significant interactions with all GM levels: A ($p=0.0211$), B ($p<0.0001$), D ($p=0.0007$) and E ($p=0.0073$); Mx3005P with GM levels A ($p=0.032$) and B ($p=0.0021$); ABI 7900 with GM level A ($p=0.0093$) and ABI 7700 with GM level B ($p<0.0001$). Overall, the interaction is mainly caused by platform iQ5, as it shows significant interactions with all GM levels that are free to vary.

The ANOVA for DAS 59122 (Table 9) shows significant differences for both platform type and GM level and significant interaction (Fig. 3b). Nine out of the 20 interaction

terms are significant: iQ5 shows significant interactions with all GM levels: A ($p<0.0001$), B ($p=0.0123$), D ($p<0.0001$) and E ($p<0.0001$); ABI 7900 with GM level B ($p=0.0096$) and ABI 7500 with all GM levels: A ($p<0.0001$), B ($p=0.0121$), D ($p=0.0007$) and E ($p<0.0001$). In this case, the interaction is mainly caused by iQ5 and ABI 7500.

Also in the case of MON89788, the ANOVA (Table 10) shows significant differences for both platform type and GM level and significant interaction (Fig. 3c). Seven interaction terms out of 16 are significant: iQ5 shows significant interactions with GM levels A, B and E ($p<0.0001$ for each interaction); ABI 7900 with GM level B ($p=0.032$); ABI 7700 with GM level A ($p=0.0074$) and ABI 7500 with GM level B ($p=0.033$). This analysis confirms that the measurement results from platform iQ5 are mainly responsible for

Table 6 Grubbs’ test for outliers, one-factor (platform) ANOVA on % GM DNA quantification and means grouping (Tukey HSD) with five GM levels for DAS 59122 maize event-specific method

GMO level	Grubbs’ outliers ($p < 0.01$)	p ANOVA Platform	Means grouping (Tukey HSD, $p < 0.05$)		
			Platform	Mean GM content (%)	Group
DAS 59122					
A (4.50 %)	No	<0.0001	iQ5	3.36	1
			7900	4.26	2
			7700	4.33	2
			3005	4.36	2
			LC480	4.48	2–3
B (2.00 %)	No	<0.0001	iQ5	1.39	1
			7900	1.67	2
			3005	1.77	2–3
			7700	1.87	3–4
			LC480	1.97	4–5
C (0.90 %)	No	<0.0001	iQ5	0.72	1
			LC480	0.80	1–2
			3005	0.82	2
			7700	0.82	2
			7500	0.84	2
D (0.40 %)	No	<0.0001	iQ5	0.25	1
			LC480	0.28	1
			7700	0.36	2
			7900	0.36	2
			7500	0.40	2–3
E (0.10 %)	1 ^a	<0.0001	LC480	0.06	1
			iQ5	0.06	1
			7900	0.08	1–2
			7700	0.09	2–3
			7500	0.10	2–3
			3005	0.10	3

^aOutlier value=0.11

the complicated pattern emerging from the analysis across platforms and GM levels.

Measurement Agreement Between Platforms

Pairwise comparison charts (regression lines) were produced to better interpret the relative performance of one platform against another. Intercept equal to 0 and slope equal to 1 indicate identical performance of the instruments being compared. Above 0 intercepts indicate that at low GMO level, the platform on y -axis tends to overestimate the GMO content compared with the platform on x -axis. The opposite applies to intercept values below 0. Slope values below 1 characterise the tendency for the platform on y -axis to underestimate the GMO content (as the content increases)

with respect to the platform on x -axis. The opposite applies to slopes above 1.

Tables 11, 12 and 13 show the intercepts (left-hand lower corner) and the slopes (right-hand upper corner) of the pairwise OLS regressions between the mean measurement results at each GM level for the three event-specific GMO methods. The values in brackets are the upper and lower 95 % confidence limits of the estimated parameter values. The regressions were done using the measurement results from the platforms in rows as x values and the measurement results from the platforms in columns as y values for calculating the slopes and the intercepts. If there is measurement agreement between two platforms, then the confidence interval for the intercept should include 0 and the confidence interval for the slope should include 1. For slopes, the values

Table 7 Grubbs’ test for outliers, one-factor (platform) ANOVA on % GM DNA quantification and means grouping (Tukey HSD) with five GM levels for MON 89788 soybean event-specific method

GMO level	Grubbs’ outliers ($p < 0.01$)	p ANOVA Platform	Means grouping (Tukey HSD, $p < 0.05$)		
			Platform	Mean GM content (%)	Group
MON 89788					
A (8.00 %)	No	<0.0001	3005	7.25	1
			LC480	7.61	1–2
			7700	8.25	2–3
			7900	8.44	3–4
			7500	8.57	4
B (4.00 %)	No	<0.0001	iQ5	9.77	5
			LC480	3.40	1
			3005	3.50	1–2
			7900	3.59	1–2
			7500	3.64	1–2
C (0.90 %)	No	<0.0001	7700	3.86	2
			LC480	0.78	1
			7900	0.79	1
			7500	0.88	1–2
			iQ5	0.93	2
D (0.40 %)	No	<0.0001	iQ5	1.04	3
			LC480	0.28	1
			3005	0.28	1
			7700	0.30	1
			7900	0.35	2
E (0.10 %)	No	<0.0001	iQ5	0.37	2
			7500	0.38	2
			LC480	0.07	1
			7900	0.08	2
			7500	0.08	2
			7700	0.09	2
			iQ5	0.10	3

in italics do not include 1, but in many of these cases, either the lower or upper confidence limits are close to 1. The confidence limits calculated for the parameter estimates of the OLS regression will not exactly coincide with the confidence limits estimated from type II regression methods,

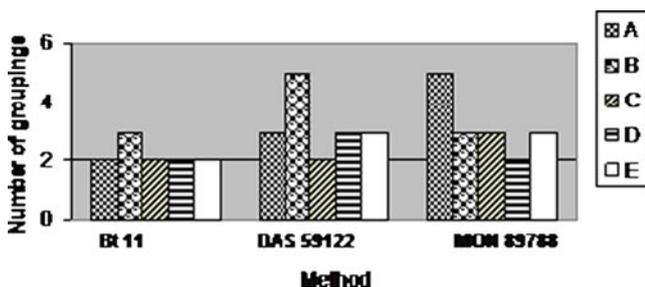


Fig. 2 Number of groups resulting from the Tukey HSD test by method and GMO level

which in general give highly conservative estimates. Furthermore, as $\text{slope}_{\text{RMA}} \approx \text{slope}_{\text{OLS}}/R$, the OLS regression gives somewhat lower estimates of the slope than RMA regression. Therefore, strong conclusions should not be drawn concerning cases where the estimate of the slope and/or one of the two confidence limits is within 1 ± 0.03 .

For Bt11 (Table 11), the intercepts of all pairs of platforms include 0 in its confidence intervals. There are five pairwise comparisons for which the confidence interval of the slope does not include 1; however, in four out of these five cases, one confidence limit is close to 1, only 3005 vs. LC480 clearly does not include 1. Similarly, LC480 tends to overestimate the GM content in comparison with platform 7500. All comparisons including platform iQ5 stand out by having very broad confidence intervals. This reflects the low precision of this platform (see Tables 2, 3 and 4). Additionally, platform iQ5 consistently

Table 8 Two-way ANOVA analysis for Bt11 results

	Effect	Univariate test of significance for Bt11				
		SS	<i>df</i>	MS	<i>F</i>	<i>p</i>
<i>SS</i> sum of squares, <i>df</i> degrees of freedom, <i>MS</i> mean squares, <i>F</i> <i>F</i> ratio, <i>p</i> probability of obtaining a specific result, given by the null hypothesis	Platform	10.042	5	2.008	10.94	<i>p</i> <0.001
	GM level	4415.952	4	1103.988	6013.80	<i>p</i> <0.001
	Platform × GM level	36.737	20	1.837	10.01	<i>p</i> <0.001
	Error	82.426	449	0.184		

gives slopes above 1, indicating that it underestimates the GM content compared to the others in the Bt11 method.

For event DAS 59122 (Table 12), the intercepts of the regressions for all platform comparisons include 0 in their confidence intervals. For slopes, ten pairwise comparisons do not include 1 in their confidence intervals, of which two (480 vs. 7700 and vs. 7500) are very close to 1. On the other hand, platform 7500 deviates from 7900, 7700 and 3005, and platform iQ5, despite showing broad confidence intervals, significantly deviates from all the others, with values consistently above 1 that indicate underestimation of the GM content.

For MON 89788 (Table 13), the analysis is compromised by the fact that for platform 3005, one GM level is missing and, thus, the regressions including this platform are based on just four pairs of values. All intercepts include 0 in their confidence limits, although these limits are very broad for all 3005 comparisons, except iQ5. For slopes, eight pairwise comparisons have confidence intervals that do not include 1, of which one (7500 vs. 7900) has one of the limits very close to 1. For the rest, LC480 slightly deviates from 7900, 7700 and 7500, and platform iQ5 deviates from all other platforms. In particular, LC480 and 3005 tend to underestimate, while platform iQ5 strongly overestimates the GM content in comparison to all other platforms. The behaviour of iQ5 is the opposite of that shown for the Bt11 and DAS 59122 methods, where it overestimated.

In general, iQ5 shows deviations from all other platforms for DAS 59122 and MON 89788 methods, and Bt11, although within the limits of confidence, also tends to deviate. In the case of Bt11 and DAS 59122, the tendency is to overestimate and for MON 89788 to underestimate. Platform 7500 also deviates from other platforms in all three methods, although some of these deviations are very slight, with confidence limits very close to 1; also, there is no clear tendency to underestimation or overestimation, showing both behaviours. Platform LC480 deviates from some platforms mainly in the MON 89788 method, although it can also be observed in the DAS 59122 with slight deviations. Finally, platform 3005 shows slight deviations with some platforms in the Bt11 and DAS 59122 methods, with some confidence limits very close to 1.

Discussion

Transferability of a method to different laboratories requires careful evaluation of the influences on the test results, such as those that may be introduced by the application of the method on different qPCR instruments. A qPCR method is shown to be robust when the whole validation process has been properly carried out. This demands proper choice of the target sequences and correct selection of the matching oligonucleotides (method development). It also requires appropriate balance of reagents and reaction conditions for an efficient, accurate and specific amplification of the target (method optimisation); next, the method parameters are corroborated by in-house method verification and full validation through a multicentric collaborative study where the fitness-for-purpose of the method is assessed against method performance criteria (Žel et al. 2008).

In PCR technology, robustness has long been confined to testing how much a method is affected when small and deliberate changes in concentrations of some reagents or in the annealing temperature are introduced in respect of standard protocols. However, it should be recognised that small changes in volume or in temperature or in reagent concentrations should not be frequent in the GMO testing field, where the use of validated methods and the requirements for being accredited according to demanding quality standards (i.e. ISO 17025) should prevent the risk of deliberate or accidental protocol modifications. Rather, a wider picture about the understanding of robustness as well as about the developing scenario of GMO detection suggests that attention should be also devoted to understanding how a method performs on different qPCR equipment whose fast-evolving technology has already introduced a constellation of brands and models to the market.

The robustness of analytical methods is a prerequisite for large method acceptance and laboratory implementation. Therefore, cross-platform mobility is usually investigated in the phase of method development and ring-trials for testing method performance do envisage, though not necessarily prescribe among the enrolment criteria, the inclusion of different analytical equipments.

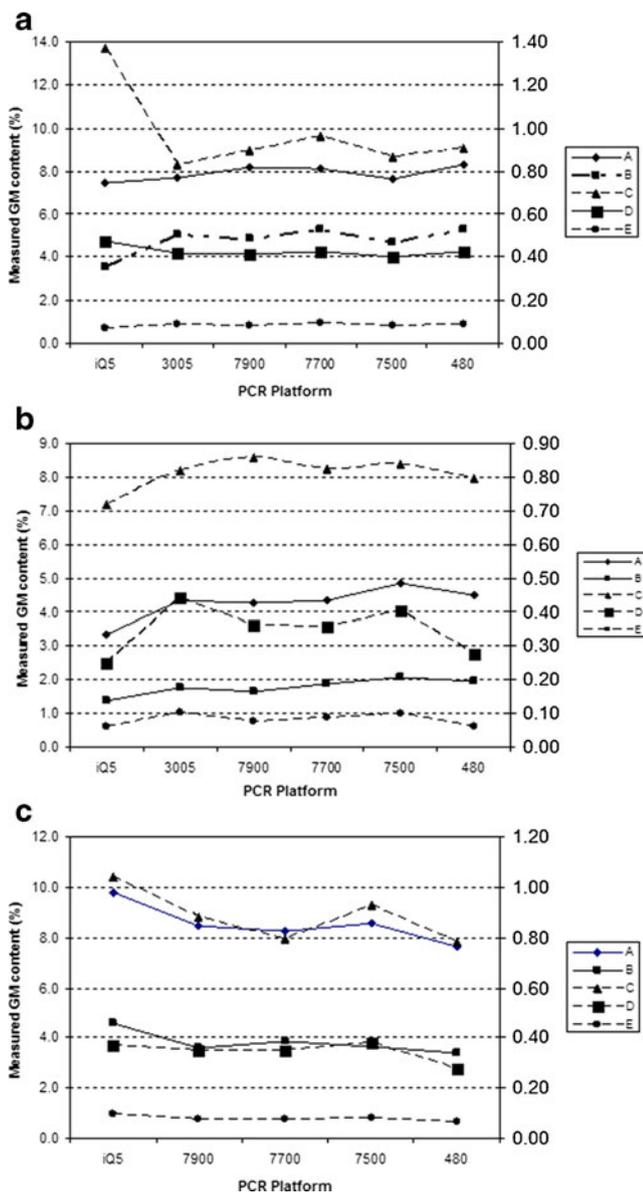


Fig. 3 **a** Measured GM content for each platform and each GM level for Bt11 maize event (A, B, C, D and E GM levels analysed). The *left axis* shows the scale corresponding to the measured GM contents (in percent) of samples A and B and the *right axis* the scale for measured GM contents (in percent) of samples C, D and E. **b** Measured GM content for each platform and each GM level for or DAS 59122 maize event (A, B, C, D and E GM levels analysed). The *left axis* shows the scale corresponding to the measured GM contents (in percent) of samples A and B and the *right axis* the scale for measured GM contents (in percent) of samples C, D and E. **c** Measured GM content for each platform and each GM level for or MON 89788 maize event (A, B, C, D and E GM levels analysed). The *left axis* shows the scale corresponding to the measured GM contents (in percent) of samples A and B and the *right axis* the scale for measured GM contents (in percent) of samples C, D and E

In certain cases, experimental evidence was attained indicating that method performances were dependent on the qPCR apparatus in use. A qPCR assay targeting an envelope

protein gene for detection of an orthopox virus monkeypox (MPXV) during the 2003 US outbreak was reported to perform optimally only in the iCycler IQ (Bio-Rad) and in the SmartCycler (Cepheid) platforms (Li et al. 2010). Interestingly, in a study conducted to evaluate the performance of commercial real-time reverse transcription PCR master-mix kits for detection of Ebola virus on four qPCR platforms (CFX96, SmartCyclerII, LightCycler480 II, and Mx3000P), the Bioterrorism Rapid Response and Advanced Technology Laboratory observed that, other things being equal, PCR efficiencies varied from one platform to another: for instance, the amplification efficiency on SmartCycler II was the best performing (99 %) with the Invitrogen master-mix the second best performing (103 %) with ABI master-mix and the worst performing (120 %) with Qiagen master-mix (Stephens et al. 2010).

In the GMO testing field, limited scientific data are available. A comparative cross-platform evaluation of qPCR quantitative detection of Roundup Ready soya was undertaken using the ABI 7700 and Roche Lightcycler detection systems in combination with different detection chemistries and showed that TaqMan probes on ABI 7700 and Scorpion primers used on Lightcycler exhibited inaccuracies of quantification of Roundup Ready soybean content in two baked biscuit samples (Terry et al. 2002). A 1 % RRS content was quantified as $1.4\% \pm 0.5$ and $0.6\% \pm 0.38$, respectively with ABI/Taqman probe and LightCycler/Scorpion primers. In another research, various chemistries were applied to the detection of MON810 target with ABI7300 and, in the fast-mode, with ABI7500, and all assays exhibited satisfactory performance (La Paz et al. 2007). These comparisons, though exploring a novel field, were not designed to look into the performance of a validated method across a range of qPCR equipments and to evaluate the accuracy results over a dynamic range against internationally accepted method performance criteria.

Noteworthy, an evaluation of the performance of different qPCR machines used to quantify MON810 maize in either a simplex (separate reactions for the GM- and reference-target) or duplex approach (concurrent reactions in the same well) was carried out recently by Allnutt and others in an inter-laboratory study with some participating laboratories (Allnutt et al. 2010). In that study, due to the fact that laboratories do not ordinarily own several and identical qPCR equipments, it was not possible to test for the significant effect on quantification by specific machines; however, the authors showed that significant differences among instruments could be linked to the method in use (for the simplex format), though the effect was not larger than run-to-run variation.

Hence, to investigate the robustness of methods for GMO detection, we undertook a study where three methods—validated in ring-trial conditions—were assayed by the same

Table 9 Two-way ANOVA analysis for DAS 59122 results

	Effect	Univariate test of significance for DAS 59122				
		SS	df	MS	F	p
SS sum of squares, <i>df</i> degrees of freedom, <i>MS</i> mean squares, <i>F</i> <i>F</i> ratio, <i>p</i> probability of obtaining a specific result, given by the null hypothesis	Platform	10.859	5	2.172	56.61	<i>p</i> <0.001
	GM level	1107.793	4	276.948	7219.05	<i>p</i> <0.001
	Platform × GM level	13.940	20	0.697	18.17	<i>p</i> <0.001
	Error	17.148	447	0.038		

qualified operator in single-laboratory conditions, in different qPCR platforms on identical test items. It is also recognised that in order to understand to what extent the findings of this study would be transferable, the same robustness study should be replicated in more reproducible conditions, preferentially involving several testing sites.

The single GM lines—maize events Bt11, DAS59122 and soybean event MON89788—were used by companies for production of ‘stacks’—due to wide insect resistance and herbicide tolerance provided by the harboured genes—and resulted in a number of ‘multiple traits’ GM products, most of which are presently under the EU authorisation procedure. In perspective, the frequency of identification of these events in food and feed can be predicted to increase, thus triggering the use of the respective quantitative methods.

Six different platforms were chosen based on their frequency of occurrence in the GMO testing community and on EU-RL GMFF availability. Our findings showed that the three methods meet the ENGL acceptance criteria for amplification efficiency and linearity of the calibration curves, when tested in all the qPCR platforms under study, with the sole exception of the linearity of the Bt11 method on iQ5. The three methods also complied with the ENGL method acceptance criteria for the accuracy of quantification measured as bias (%*B*) and relative repeatability standard deviation at most GM levels and across most of the platforms (Tables 2, 3 and 4). However, some deviations to this general rule were identified. In general, both in DAS 59122 and Bt11 methods, the trueness at the various GM levels is more dispersed around the assigned value with the iQ5; method for DAS 59122 showed departures from the expected performance in platform LC480 at the lowest GM levels. A pairwise comparison of the performance of the

platforms against each other confirmed that iQ5 significantly deviates from all other platforms in the DAS 59122 and MON 89788 methods, with tendencies.

RSD_r to be larger at the lower extreme of the dynamic range is a general trend for all calibration systems and it is also reflected in the findings of this study with a few minor exceptions. The RSD_r values of the samples at lower GM concentration are overall in line with those resulting from the figures reported in full validation studies. The relative standard deviation of repeatability (ISO 5725 standards) for the ‘near-LOQ’ GM levels was 16 % for the MON89788 method, 17 % for the Bt11 method and 18 % for DAS 59122.

The analysis of variance (one-way) of the three methods across platforms revealed that the quantified GMO contents were significantly affected by the platforms (Tables 5, 6 and 7); however, it also indicated that platforms did not respond consistently across GM levels, with some overestimation at some levels and underestimation at others.

A two-way ANOVA (Tables 8, 9 and 10) demonstrated that there are significant interactions between the platform type and the GMO level: iQ5 platform being mostly responsible for the interactions in the three methods and 7500 also participating in the interaction in the DAS 59122 method. Hence, results cannot be interpreted by considering each main factor in isolation, i.e. it is not possible to predict the result as the sum of the effects of platform and GM level.

This study was not planned to address the question of what factor(s) at stake is(are) the most influential in the outcome of qPCR analysis. However, some considerations can be proposed. Real-time PCR platforms basically consist of three components: a thermal cycler which shuttles the polymerase chain reaction between denaturation (95 °C) and

Table 10 Two-way ANOVA analysis for MON 89788 results (excluding platform Mx3005P)

	Effect	Univariate test of significance for MON 89788				
		SS	df	MS	F	p
SS sum of squares, <i>df</i> degrees of freedom, <i>MS</i> mean squares, <i>F</i> <i>F</i> ratio, <i>p</i> probability of obtaining a specific result, given by the null hypothesis	Platform	23.375	4	5.844	40.72	<i>p</i> <0.001
	GM level	4065.208	4	1016.302	7082.58	<i>p</i> <0.001
	Platform × GM level	29.971	16	1.873	13.05	<i>p</i> <0.001
	Error	53.523	373	0.143		

Table 11 Intercept and slopes values of the linear regression (ordinary least squares, OLS) obtained comparing GM contents obtained by different platforms in pairs for event Bt11 maize

Bt11						
	7900	7700	7500	LC480	3005	iQ5
7900	–	0.987 (0.884, 1.090)	<i>1.064</i> <i>(1.021, 1.107)</i>	0.967 (0.887, 1.046)	1.039 (0.922, 1.157)	1.149 (0.824, 1.473)
7700	–0.063 (–0.513, 0.386)	–	<i>1.0757</i> <i>(1.007, 1.144)</i>	0.979 (0.954, 1.004)	<i>1.053</i> <i>(1.03, 1.077)</i>	1.152 (0.720, 1.584)
7500	–0.033 (–0.207, 0.141)	0.036 (–0.240, 0.313)	–	<i>0.909</i> <i>(0.871, 0.948)</i>	0.978 (0.907, 1.049)	1.076 (0.735, 1.418)
LC480	–0.028 (–0.380, 0.324)	0.038 (–0.072, 0.148)	0.003 (–0.167, 0.173)	–	<i>1.075</i> <i>(1.042, 1.109)</i>	1.178 (0.755, 1.602)
3005	–0.041 (–0.524, 0.442)	0.022 (–0.075, 0.119)	–0.01 (–0.304, 0.284)	–0.015 (–0.153, 0.122)	–	1.092 (0.665, 1.518)
iQ5	–0.089 (–1.302, 1.125)	0.005 (–1.611, 1.623)	–0.043 (–1.321, 1.235)	–0.038 (–1.621, 1.546)	–0.011 (–1.607, 1.585)	–

Intercepts are presented in the shadowed boxes. For slopes, confidence intervals not including 1 are marked in italic

annealing/extension temperature (usually at 60 °C), an optical system which emits light to excite fluorophores linked to oligonucleotides in reaction and which detects the fluorescence produced and a software to interface with the operator on the one side and to implement the algorithms necessary to analyse the data on the other side. To serve these functions, manufacturers generated a bunch of different technical solutions. In our study qPCR platforms were equipped with Peltier-based heating and cooling systems. In a recent modelling of PCR instruments (Lee 2010), a parameter diagram was developed to classify all the variables and determine the influential factors, based on the outcome of experimental results. The inter-assay coefficient of

variation (CV) was taken to evaluate the reliability of qPCR machines. The simulation model was used to predict the CV curves for various instruments: the ABI 7000 series, Roche LightCycler and Bio-Rad iCycler. Out of five factors considered, it is suggested that only two play a role on instrument performance: the noise in fluorescence measurements, mainly due to moving parts in the optics, and the temperature uniformity. Variations of temperature uniformity can lead to variations in reaction kinetics and can make identical samples crossing the threshold at different cycle fractions. Therefore, it is not trivial to ensure that anomalies in the heating elements or the temperature gradient from the plate centre to the edge, likely occurring in Peltier-based thermal

Table 12 Intercept and slopes values of the linear regression (ordinary least squares, OLS) obtained comparing the GM contents obtained by different platforms in pairs for event DAS 59122 maize

DAS 59122						
	7900	7700	7500	LC480	3005	iQ5
7900	–	0.973 (0.882, 1.065)	<i>0.871</i> <i>(0.778, 0.964)</i>	0.927 (0.811, 1.043)	0.98 (0.929, 1.031)	<i>1.264</i> <i>(1.194, 1.335)</i>
7700	–0.005 (–0.202, 0.192)	–	<i>0.895</i> <i>(0.874, 0.915)</i>	<i>0.953</i> <i>(0.915, 0.991)</i>	1.005 (0.932, 1.079)	<i>1.296</i> <i>(1.189, 1.404)</i>
7500	0.008 (–0.213, 0.230)	0.013 (–0.035, 0.062)	–	<i>1.065</i> <i>(1.014, 1.115)</i>	<i>1.123</i> <i>(1.038, 1.208)</i>	<i>1.448</i> <i>(1.298, 1.598)</i>
LC480	0.039 (–0.219, 0.296)	0.044 (–0.041, 0.128)	0.034 (–0.078, 0.147)	–	1.053 (0.936, 1.171)	<i>1.359</i> <i>(1.224, 1.495)</i>
3005	–0.024 (–0.133, 0.086)	–0.017 (–0.175, 0.142)	–0.033 (–0.216, 0.149)	–0.061 (–0.314, 0.192)	–	<i>1.288</i> <i>(1.174, 1.403)</i>
iQ5	–0.014 (–0.131, 0.103)	–0.007 (–0.185, 0.172)	–0.021 (–0.27, 0.227)	–0.052 (–0.276, 0.173)	0.011 (–0.178, 0.201)	–

Intercepts are presented in the shadowed boxes. For slopes, confidence intervals not including 1 are marked in italic

Table 13 Intercept and slopes values of the linear regression (ordinary least squares, OLS) obtained comparing GM contents obtained by different platforms in pairs for event MON 89788 soybean

MON 89788						
	7900	7700	7500	LC480	3005	iQ5
7900	–	1.01 (0.922, 1.098)	0.987 (0.980, 0.994)	1.101 (1.051, 1.152)	1.163 ^a (0.129, 2.198)	0.853 (0.763, 0.944)
7700	–0.02 (–0.378, 0.339)	–	0.975 (0.886, 1.064)	1.089 (1.043, 1.136)	1.139 ^a (0.892, 1.386)	0.845 (0.822, 0.868)
7500	–0.018 (–0.049, 0.012)	0.007 (–0.365, 0.380)	–	1.116 (1.060, 1.171)	1.178 ^a (0.097, 2.259)	0.864 (0.77, 0.959)
LC480	–0.007 (–0.196, 0.182)	0.0152 (–0.159, 0.190)	0.011 (–0.197, 0.220)	–	1.053 ^a (0.514, 1.592)	0.775 (0.728, 0.822)
3005	–0.148 ^a (–4.958, 4.662)	–0.047 ^a (–1.194, 1.101)	–0.131 ^a (–5.157, 4.895)	–0.107 ^a (–2.612, 2.398)	–	0.742 ^a (0.719, 0.765)
iQ5	–0.057 (–0.497, 0.383)	–0.038 (–0.151, 0.074)	–0.039 (–0.499, 0.421)	–0.047 (–0.275, 0.181)	0.007 (–0.137, 0.151)	–

Intercepts are presented in the shadowed boxes. For slopes, confidence intervals not including 1 are marked in italic

^aRegressions based on four points

blocks, do not affect the amplification reactions occurring in the overlying wells of the plate. An indirect way to assess that is to test for Cq onset of replicate samples over the 96-wells of the plate. In our laboratory, verification of instrument performance was regularly performed under the quality assurance scheme in force; RNase P test was conducted and background emission of blocks was implemented as well.

Moreover, in our experimental format (Žel et al. 2008), each replicate sample is tested in three adjacent wells and the resulting values of copy number are then averaged. Therefore, it is unlikely that gross variation in temperature uniformity could relate the accuracy of GMO quantification here reported. As far as the optical systems is concerned, there is great variety of light sources: from broad spectrum light emission typical of tungsten halogen lamp like in the 7500 and in the iQ5 to Argon-ion laser like in the 7900. Detectors can also vary. All the equipments here investigated made use of CCD but one of a photomultiplier (Mx3005) (Logan et al. 2009). Furthermore, CCD cameras can vary for instance in terms of colour depth, resolution, sensitivity and other characteristics not always disclosed by the manufacturer. However, the software (e.g. processing algorithms) is believed to vary more significantly from instrument to instrument (PCR troubleshooting and optimisation. The essential guide. S. Kennedy 2999). In this context, advanced techniques such as partial least squares regression (Wold et al. 2001) might in principle be able to identify combinations of factors that influence the measurement outcome. However, potentially important technical and software specifications (i.e. deconvolution and curve-fitting algorithms) of the qPCR platforms are not readily shared by the manufacturers

and a more causal analysis of the patterns described here thus remains difficult.

Transferability of this pattern of data to other laboratory settings requires corroboration in more reproducible conditions; the results of this study are particularly relevant: a sound experimental design with setting of repeatability conditions, the thorough testing of three validated methods and the large collection of data generated and analysed could show that complex patterns of interactions exist between equipment and GM levels in single-laboratory conditions and, notably, that these interactions change when the same GM level is analysed on the same set of qPCR equipment in combination with another method. Thus, to stipulate broad predictions of constant over- or underestimation of the true GM content of sample by given equipment is difficult. The requirements laid down in the ISO 17025 accreditation scheme provide that validated methods for analysis are subject to laboratory verification prior to implementation in the routine monitoring plans. Guidelines for correct implementation have been recently released by the ENGL (Guidance document from the European Network of GMO laboratories 2999). Importantly, they maintain that principle methods should be implemented as validated and that, in such case, robustness does not need to be re-evaluated in a verification study. Therefore, it becomes increasingly important that networks carrying out method validations incorporate the robustness requirement with particular focus on the interchangeability of qPCR platforms and corresponding limitations in order to spread knowledge about robustness of qPCR methods.

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

References

- Allnutt TR, Ayadi M, Berben G, Broadmann P, Lee DJ (2010) AOAC Intl 93:1243
- Berdal KG, Bøydler C, Tengs T, Holst-Jensen A (2008) Eur Food Res Tech 227:1149
- Biochemica (2005), Available at http://www.roche-applied-science.com/PROD_INF/BIOCHEM/no4_05/pdf/p04.pdf
- Biocompare (2011). Available at <http://www.biocompare.com/ProductListings/2838/Real-Time-PCR-Thermal-Cyclers-Thermocyclers.html>
- Brace RA (1977) Am J Physiol 233:R94
- Broothaerts W, Corbisier P, Schimmel H, Trapmann S, Vincent S, Emons HJ (2008) Agr Food Chem 56:8825
- Bustin SD, Benes V, Garson JA, Hellems J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer T (2009) Clin Chem 55:4
- Cankar K, Štebih D, Dreo T, Žel J, Gruden K (2006) BMC Biotechnol 6:37
- Codex Alimentarius Commission (2010) Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods
- Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Erlbaum, Hillsdale
- Corbisier P, Broothaerts W, Gioria S, Schimmel H, Burns M, Baoutina A, Emslie KR, Furu S, Kurosawa Y, Holden MJ, Kim JH, Lee Y, Kawaharasaki M, Sin D, Wang JJ (2007) Agr Food Chem 55:3249
- Dejaegher B, Vander Heyden Y (2007) Ruggedness and robustness testing. J Chromatogr A 1158:138
- Donald CE, Qureshi F, Burns MJ, Holden MJ, Blasic JR Jr, Woolford AJ (2005) BMC Biotechnology 5:15
- EC Regulation No 1829/2003 Off. J. Eur. Union, L268, 1 (2003)
- EC Regulation No 1830/2003 Off. J. Eur. Union, L 268, 24 (2003)
- EC Regulation No 882/2004 Off. J. Eur. Union, L165, 1 (2004)
- EU register of genetically modified food and feed (2009) Available at http://ec.europa.eu/food/dyna/gm_register/index_en.cfm
- European Network of GMO Laboratories (2008). Available at <http://gmo-crl.jrc.ec.europa.eu/guidancedocs.htm>
- European Union Reference Laboratory, Validation report, Available at http://gmo-crl.jrc.ec.europa.eu/summaries/MON88017_DNAExtr_report.pdf
- European Union Reference Laboratory, Validation report, Available at http://gmo-crl.jrc.ec.europa.eu/summaries/MON89788_Soya_DNAExtrSampl_report.pdf
- European Union Reference Laboratory, Validation report, Available at http://gmo-crl.jrc.ec.europa.eu/summaries/Bt11_CRLVL1007_Val_Report.pdf
- European Union Reference Laboratory, Validation report, Available at <http://gmo-crl.jrc.ec.europa.eu/summaries/59122-ProtocolValidation.pdf>
- European Union Reference Laboratory, Validation report, Available at http://gmo-crl.jrc.ec.europa.eu/summaries/MON89788_validated_Method.pdf
- Faul F, Erdfelder E, Lang A-G, Buchner A (2007) Behav Res Methods 39:175
- GM-crop database (2009) Available at http://www.cera-gmc.org/?action=gm_crop_database
- Guidance document from the European Network of GMO laboratories (ENGL) (2009) Available at <http://gmo-crl.jrc.ec.europa.eu/doc/ENGL%20MV%20WG%20Report%20July%202011.pdf>
- Horwitz W (1995) Pure Appl Chem 67:331
- ISO 21570 European Committee for Standardization/International Organization for Standardization, Genève, Switzerland (2005)
- ISO 24276 European Committee for Standardization/International Organization for Standardization, Genève, Switzerland (2006)
- ISO 5725 International Organization for Standardization, Geneva, Switzerland (1994)
- IUPAC (1988) Pure Appl Chem 60:856
- Jolicoeur PJ (1990) Theor Biol 144:275
- La Paz J, Esteve T, Pla MJ (2007) Agr Food Chem 55:4312
- Lee DS (2010) Sensor 10:697
- Li Y, Zhao H, Wilkins K, Hughes C, Damon IK (2010) J Virol Methods 169:223
- Logan J, Edwards K, Saunders N (2009) Real-time PCR: current technology and applications. Caister Academic, Norfolk
- Ludbrock J (1997) Clin Exp Pharmacol Physiol 24:193
- Ludbrock J (2010) Clin Exp Pharmacol Physiol 37:692
- Niklas KJ (2004) Biol Rev 79:871
- Nitsche A, Steuer N, Schmidt CA, Landt O, Siebert W (1999) Clin Chem 45:1932
- PCR troubleshooting and optimisation (2009) The essential guide. S. Kennedy, N. Oswald Editors. Caister Academic, Norfolk
- Querci M, Foti N, Bogni A, Kluga L, Broll H, Van den Eede G (2009) Food Anal Method 2:325
- Shokere LA, Holden MJ, Jenkins GR (2009) Food Contr 20:391
- Stephens KW, Hutchins RJ, Dauphin LA (2010) Mol Cell Probes 24:370
- Terry CF, Shanahan DJ, Ballam LD, Harris N, McDowell DG, Parkes HCJ (2002) AOAC Intl 85:938
- Vander HY, Nijhuis A, Smeyers-Verbeke J, Vandeginste BGM, Massart DL (2001) J Pharm Biomed Anal 24:723
- Wold S, Sjöström M, Eriksson L (2001) Chemometr Intell Lab Syst 58:109
- Žel J, Mazzara M, Savini C, Cordeil S, Camloh M, Štebih D, Cankar K, Gruden K, Morisset D, Van den Eede G (2008) Food Anal Method 1:61