# G03025 Final Report: Quantitation using informative zeros (QUIZ): proof of concept study

**David Lee, Maurizio La Mura, Andrew Greenland and Ian Mackay**
National Institute of Agricultural Botany, Huntingdon Road, Cambridge CB3 0LE

## Introduction

Quantitation using informative zeros (QUIZ) is a statistics-based approach to quantifying the number of molecules in a sample. It relies on the ability to differentiate between the presence and absence of selected markers, down to single molecules. However, the statistics behind QUIZ (most probable number) are well-established and used by microbiologists to measure the number of bacteria in samples, as well as by seed analysts for determining seed contamination.
In short, samples are diluted to a point where the (DNA) molecules of interest may be present or absent in aliquots of the diluted samples. In our system, the detection of DNA target is by PCR amplification. Since it is not possible to differentiate between samples containing one or more targets, the absence of target is the only unambiguous result i.e. the 'informative zero'.
In this proof of concept study we proposed to apply QUIZ to measure the GMO content of RoundUp Ready™ soya samples constructed by mixing two differing GMO-containing samples in different proportions. RoundUp Ready™ soya is an ideal system to use because the transgene has been well characterised [(1)], providing sequence data for primer design, and the availability of certified reference materials (CRMs) for the event in the form of ground soya meal (Fluka Biochemika, Sigma-Aldrich Company Ltd, Dorset). Furthermore, though absolute number of GMO DNA molecules can be measured within each sample using QUIZ, determination of proportion of GMO content also requires measurement of the number of a taxon-specific (reference) target: in soya the lectin gene can be used as reference [(2)].

## Materials and Methods

Primers (Sigma-Genosys, Haverhill) were designed from the sequences for the transgene junction for RoundUp Ready™ soya (GI: 14270370) and for the lectin gene target (GI: 170005) as the taxon-specific reference (**Table 1**).

Soya Roundup Ready™ GMO Reference Material was purchased from Fluka Biochemika: Lot&Filling code 1129490, 43404065 (Individual Sample Number 6937). Five samples, containing different GMO content, were constructed by mixing varying proportions of 0% and 5% CRMs (**Table 2**). DNA was extracted from each using a DNeasy kit (QIAGEN, Crawley) following the manufacturer's instructions. An aliquot of each (2 $\mu$L) was run out on an agarose gel to determine quality, and quantified by comparison with known amounts of lambda DNA (New England Biolabs, Hitchins) (not shown) and estimated to be ~20 $\eta g \mu L^{-1}$.
The haploid genome size of soya is approximately 1 pg [(1)]; so, 20 $\eta g \mu L^{-1}$ equates to 20000 genomes in one $\mu$L. A dilution of $10^{-3}$ would therefore yield 20 copies per $\mu$L of a single copy gene. For the lectin marker, 5 $\mu$L of this dilution was added to each tube containing one mL of mastermix for all samples apart from the 0% GMO

sample. Corresponding dilutions were made to each sample to generate an equal number of GM markers. For the 0% sample, $10^{-4}$ dilutions were made of the template for lectin gene detection and 25 μL of the DNA sample was added to each μL of reagents to try and equalize the numbers of each marker (**Table 2**).

Amber 1.5 mL microcentrifuge tubes (Star Lab, Milton Keynes - catalogue No. I 1415 1007) were siliconised by submerging in 'Silanization Solution II' (2% dimethylchlorosilane in 1, 1, 1–trichloroethane; Fluka Biochemika) for 10 mins. The tubes were rinsed 10 times with water, purified by reverse osmosis using a PRO\RO unit (Labconco Corporation, Missouri, USA) and air dried before using to make PCR mastermixes.

**PCR Conditions**

PCR mix containing 1 x PCR buffer, 2.5 mM $MgCl_2$, 0.25 μM each primer, 200 μM each dNTPs, 0.05 $U\mu L^{-1}$ Fastart Taq (Roche, Lewes), λ DNA to 1 $\eta g\mu L^{-1}$ with appropriate amounts of template added prior to dispensing into 384-well microtitre plates. One mL of reaction mix was made for 96 samples; four mL were required for each plate. Ten μL of reaction mix were dispensed into each well using an epMotion 5070 workstation (Eppendorf, Histon).

Cycling parameters: 5 mins 94°C initial denaturation, followed by 50 cycles of 92°C 30s, 55°C 30s, 72°C 30s; with final extension of 10 mins 72°C. After cycling, each lectin amplification reaction is mixed with a reaction for the GMO target. One μL of the mix is added to 10 μL of Hi-Di (Applied Biosystems) with Rox-labelled size standards (81, 106, 146, 173, 358 and 408 nucleotides) and loaded onto a 3100 Genetic Analyzer (Applied Biosystems).

The alignment of the tracks using the size standards was performed using Genemapper 3.7 software (Applied Biosystems) though each track was manually checked to ensure correct alignment. The presence, or absence of the two markers (see **Figure 1**) were recorded and their numbers used to calculate the percentage GMO content in each sample. In cases (~20%) where the peaks were not oversized (fluorescence height of >8000), peak heights greater than 500 were accepted as positives and everything else considered negatives.

The average number of molecules in each sample was calculated as:

$$\mu = -\log(\text{proportion of zeros}) \times \text{dilution factor}$$

The GMO content was determined as: $\mu_{gm} / \mu_{reference} \times 100\%$

Variance of percentage was calculated from:

$$\text{Var}_{(\text{ratio})} \approx \left( \left( \frac{p_l q_l}{N_1} \right) \cdot \left( \frac{1}{p_l} \right)^2 \cdot \left( \frac{\mu_g}{\mu_l} \right)^2 + \left( \frac{p_g q_g}{N_2} \right) \cdot \left( \frac{1}{p_g} \right)^2 \right) \cdot \left( \frac{1}{\mu_l} \right)^2$$

where $p$ is the proportion of zeros i.e. the proportion of negative amplifications, $q$ is the frequency of positive (1-p) amplifications for the lectin ($_l$) and GMO ($_g$) markers; $N_1$ and $N_2$ represent the total numbers of lectin and GMO reactions, respectively; $\mu_g$ and $\mu_l$ are the estimated numbers of molecules.
The formula for the variance of the ratio of (no. of GM molecules)/(no. of lectin molecules) was derived using the delta method, taking the distribution of p as binomial.

**Quantitation using real-time PCR**

A sample containing 40 μL of each of the five admixtures, labeled 1, 2, 3, 4 and 5, were sent to iDna Genetics Ltd (Norwich) to be tested for their RoundUp Ready™ content. A copy of their report is included as appendix 1. The results are summarised in **Table 2**.

**Results and Discussion**

The outline of the QUIZ testing procedure is shown in **Figure 2**. We have exploited the nominal GMO contents to dilute the samples such that the expected numbers of reference and GMO targets are approximately equal. Though the number of reference targets can always be manipulated from the DNA concentrations this is not possible for GMO target where the GMO content is unknown. However, there are empirical methods to determine the dilutions that are required: for example, testing a dilution series of template. In fact these data may be sufficient to determine whether further testing is required if performed with rigour. Replicates of dilution series of template could be sufficient to determine a sample to be >>5% or <0.1%.
The numbers of positive reactions for both lectin and GMO targets using extended numbers of PCR cycles (50) and FAM-labelled primers, for detection in automated genotypers, show successful amplification from single molecules. In fact, most peaks observed in the ABI3100 (Applied Biosystems) were oversized, even though only 0.5 μL of the reactions were loaded. Given the DNA concentrations of the samples, it was statistically unlikely for all the reactions to contain more than one target amplicon.

Quantitation obtained using QUIZ compares favourably with those of the RT-PCR assay: the correlation between the measured and nominal values are 0.9901 and 0.9878, respectively (**Figure 3**). However, the dilutions used for all the samples, except the 0% CRM, should contain approximately one amplicon per reaction. At this concentration, we expect two-thirds of the reactions to be positive. The actual percentage positives varied between 21 and 64 (with a mean of 44). The starting DNA concentrations, approximately 20 $\eta g \mu L^{-1}$, were corroborated by the $C_t$ values of the RT-PCR reactions for the reference gene (see iDNA Genetics report). The lower than expected numbers may reflect the impact of organellar DNA in the samples, which would lead to an overestimation of the expected numbers of amplicon, and loss of DNA during the preparative stages. Two steps in the procedure were included to reduce the loss of template DNA: the silanization of the tubes to reduce DNA

binding and the addition of bacteriophage λ DNA (to 1 ηgμL$^{-1}$) in the reaction mixes as competitor to DNA binding sites on the tubes and pipette tips.

## Future perspectives

These results show that QUIZ may be a realistic alternative for GMO quantitation. Clearly the drawback is the numbers of reactions that need to be performed for accurate determination. The generation of this data set required 5 x 2 x 384 (3840) PCR reactions and 1920 gel runs. However, as already stated above, QUIZ allows testing at different levels. Replicates of a dilution series may provide sufficient confidence to condemn or accept samples with a given threshold/tolerance GMO level. Furthermore, there are technical improvements that can be made to the procedure: for example, the use of an ABI 3730 with 96 capillaries would reduce the time taken to test the samples; PCR reactions could be performed in 1 μL reactions under oil (as oppose to 10 μL in current study) thus reducing reaction costs tenfold.

Improvements to the application of QUIZ will be investigated within the next stage of the project (G03030). These include optimisation of the testing procedures by modelling combined results from different dilutions and dilution series. If stage 2 is successful in delivering a sensitive and robust alternative to real time PCR, further development of a sensitive non-PCR amplification and detection method (isothermal amplification) would further enhance the application of QUIZ as a viable cost-effective method for the quantitation of GMOs.

The strongest arguments for the development of QUIZ are:

- **GMO quantitation is independent of CRMs, which are not available for many events;**
- **The threshold of detection (and quantitation) is down to single molecules and may be the only feasible quantitative test for highly processed foods.**

**Acknowledgements**
Huw Jones and Richard Horsnell are thanked for their help and patience demonstrating the ABI 3100.

**References**
**(1)** Windels, P, Taverniers, A, Depicker, A, Van Bockstaele, E and De Loose, M. (2001). Characterisation of the Roundup Ready soybean insert. Eur. Food Res. Tech. 213: 107-112.
**(2)** Germini A, Zanetti A, Salati C, Rossi S, Forre C, Schmid S, Marchelli R and Fogher C. (2004). Development of a seven-target multiplex PCR for the simultaneous detection of transgenic soybean and maize in feeds and foods. J. Agric. Food Chem. 52:3275-80. Erratum in: J. Agric. Food Chem. 52:4350.
**(3)** Genome size taken from the Plant DNA C-values Database, Royal Botanic Gardens, Kew, UK (http://www.rbgkew.org.uk/cva1/database1.html).

**Table 1. Primer sequences and product length**

| Primer | Sequence | Amplicon length (bp) |
|---|---|---|
| RRS GMF | *-ACTGCTTCTCCCAGAATGATC | 285 |
| RRS GMR | TCGAGCTTCTTCACGAACTT | |
| LecF | *-ATGGGCTTGCCTTCTTTCT | 157 |
| LecR | CCGATGTGTGGATTTGGTG | |

* primer labeled with FAM

**Table 2. Sample composition and measured GMO content**

| Sample | % (weight) 5% CRM | Nominal GMO Content (%) | Number of Lectin Amplicons (*Total) | Number of GM Amplicons (*Total) | Dilution Factor | QUIZ Determined GMO Content (%) | Average RT-PCR Determined GMO Content (%) |
|---|---|---|---|---|---|---|---|
| A | 0 | 0 | 12 (384) | 13 (384) | $5 \times 10^4$ | 0.0022 ± 0.0009 | 0.01 |
| B | 9.1 | 0.45 | 213 (384) | 134 (368) | 200 | 0.28 ± 0.03 | 0.130 |
| C | 16.7 | 0.83 | 171 (384) | 84 (384) | 100 | 0.42 ± 0.06 | 1.095 |
| D | 64.7 | 3.24 | 130 (384) | 121 (368) | 33.3 | 2.89 ± 0.37 | 3.089 |
| E | 100 | 5 | 242 (383) | 246 (382) | 20 | 5.17 ± 0.49 | 4.601 |

* Though 384 reactions were performed for each marker/dilution, some reactions were lost to evaporation.

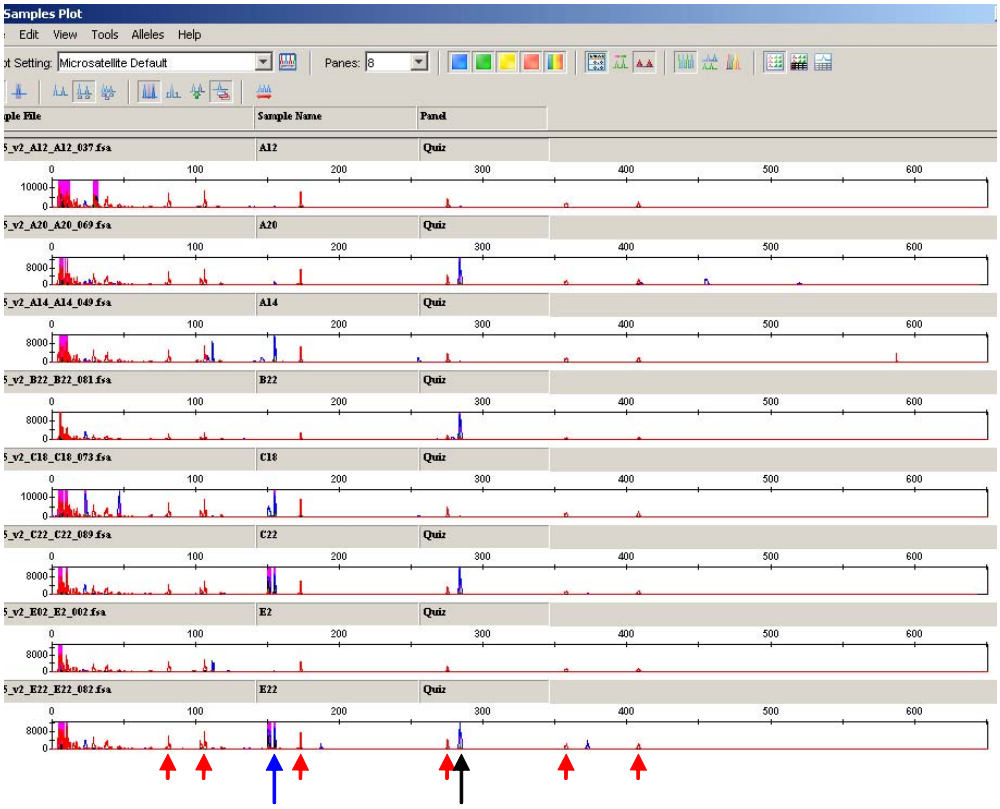



**Figure 1. Chromatograms showing the amplification of the lectin (blue arrow) and GMO-specific amplicons (black arrows) in post-PCR mixes. Red arrows signify the internal Rox-labelled standards used to align tracks and numbers represent sizes in nucleotides.**

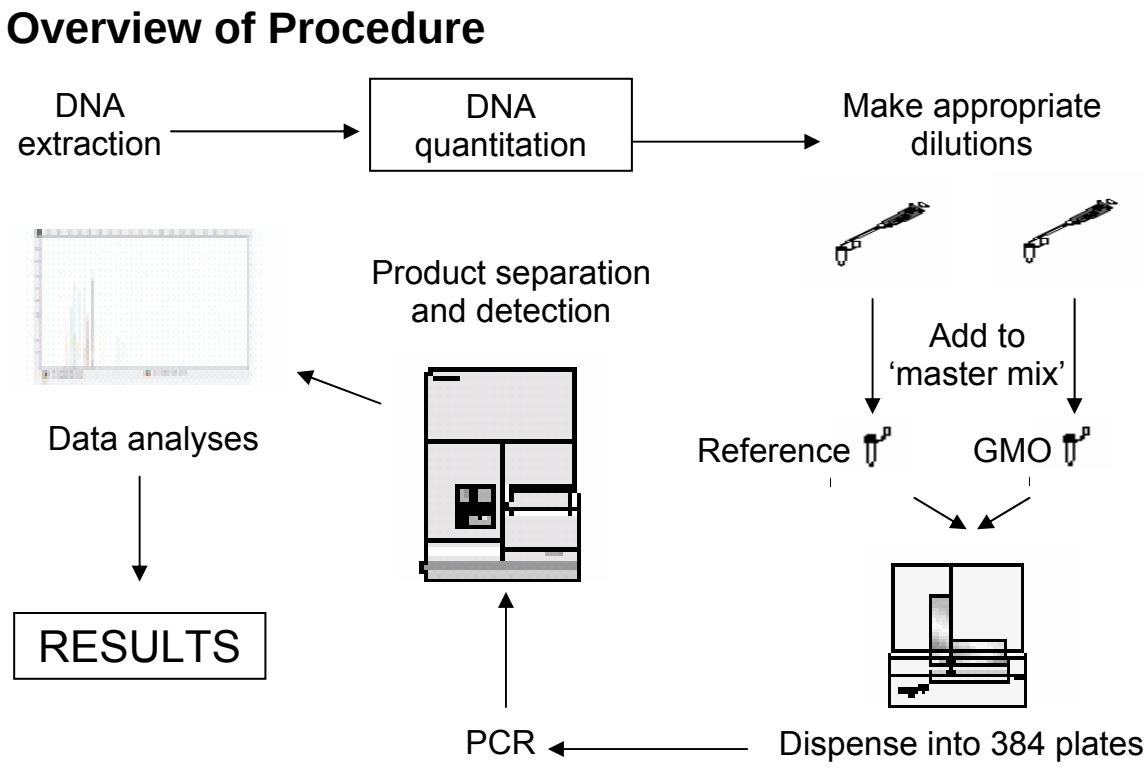



**Figure 2. Outline of the QUIZ procedure used in this study.**

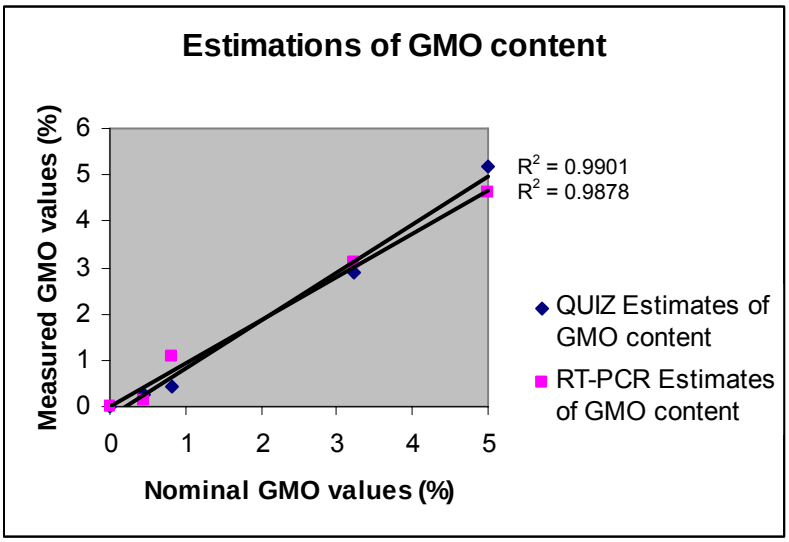



**Figure 3. The estimated GMO contents of the five samples, as determined using QUIZ and RT-PCR (iDNA Genetics Ltd) are plotted against the nominal GMO content. The best fitting line is drawn for each method and the correlation of fit $R^2$ for each curve is shown.**

**Appendix 1**

| | |
|---|---|
| **Project Title** | **Quantification of Roundup ready GMO Soya** |
| **Project Number** | 06M0097 |
| **Report** | 25th January 2007 |


**Client**

David Lee
NIAB
Huntingdon Road
Cambridge
CB3 0LE
Tel 01223 342200
Email david.lee@niab.com

**Report by**

Peter Isaac
CEO
IDna Genetics Ltd
Norwich Bioincubator
Norwich Research Park
Norwich
NR4 7UH
Tel 01603 450941
Email peter.Isaac@idnagenetics.com


7 Pages


**Summary**

*Five DNA samples supplied by the Client, derived from soya flour containing varying amounts of RoundUp Ready certified reference material, were analysed by realtime PCR relative to standards. All of the samples were positive for the transgene sequence, although one of the samples was at a level known to be present in some 0% certified reference materials.*

## Introduction

The Client has 5 samples of SNA derived from Soybean (Glycine max) flour. Each sample contains a known percentage of Roundup Ready GMO in non-GMO flour. The flour mixtures were formulated from Soya certified reference material (RoundUp Ready, Fluka). The client wishes to have an independent estimate of the percentage GMO in the samples. IDna Genetics has undertaken to test the DNA with realtime PCR and to estimate the percentage GMO with reference to certified reference materials.

## Materials and Methods

*Client samples*

Supplied by the client – 5 DNA samples, 40ul per sample, quantified to approximately 20ug/ul of DNA in each tube.

*Preparation of standards*

The standards used were Fluka Soybean Powder Set Certified reference Material IRMM Nr4105 , lot and filling code 1129490 / 43404065, samples all numbered Sample 6932, covering the range 0.1% to 2%. The 0% control from this set was not used (see Discussion). A 0% control was made from ground soybeans that had been repeatedly tested and found to be negative for GMO content.

The standards were extracted using SOP QT0113 (SOP for DNA extraction from Fluka standards for GMO detection). The DNA concentrations were measured by fluorometry using the Sigma DNA quantification kit Fluorescence Assay (lot 103K4099 in a LabSystems Fluoroskan II instrument (SOP QT0103 SOP for Quantification of DNA using BisBenzimide ). DNA concentrations were adjusted to 40ng/ul with water. A 0.01% standard was made by mixing 1 part 0.1% GMO soybean standard with 9 parts 0% GMO standard.

## PCR conditions

Real time PCR assays were carried out in an AppliedBiosystems Prism 7900 HT Sequence Detection System instrument equipped with a 384 place block. The reaction plate was Applied Biosystems MicroAmp Optical 384 well reaction plate, p/n 4309849, bearing the barcode S300819K.

The transgene amplicon was a 121 base fragment spanning the CTP4-Agrobacterium CP4 EPSPS gene junction, the Internal Positive Control (IPC) was a 118 base fragment not known to be variant in soybean. The probes carried 6FAM reporter and TAMRA quencher.

The PCR reactions were carried out as SOP QT1190 (SOP for Detection and Quantification of RoundUp Ready Soybean ). The reaction volumes were 20ul and used AbGene AbSolute QPCR mix lot 1810/13. The reactions for the GMO target and the IPC were carried out in simplex, each sample and standard being assayed in duplicate, except the 0.01% GMO samples which were tested four times each for GMO target and IPC. Duplicate samples containing water, but no DNA, were also assayed for both the transgene and the IPC (No Template Controls (NTC)). 10 x Probe and primer mixes contained each primer at 2uM, and the probe at 1uM. Master mixes were made containing AbSolute mix, probe and primer mix and water. 15ul of master mix was added to each reaction well and 5ul of DNA (100ng for Client samples, 200ng for standards) added to each well. Each reaction contained:-

| | ul Per rxn |
|---|---|
| 2 x AbSolute QPCR | 10 |
| 10 xProbe & primer mix | 2 |
| water | 3 |
| DNA | 5 |

| | |
|---|---|
| Total volume | 20 |

The plate was sealed with Applied Biosystems optical adhesive film p/n 4311971 lot 200607-142. The plate was centrifuged at 3.2kg for two minutes, and then placed in the drawer of the ABI7900 instrument.

The PCR conditions were

2 minutes at 50C
15 minutes at 95C (enzyme activation)
15 seconds 95C, 60 seconds 60C for 45 cycles

The fluorescence data was recorded at 60C only. 9600 emulation was switched off. The results were collected in the SDS file RRS_250107.

At the end of the PCR the samples were analysed with the threshold line set to 0.1 (the middle of the logarithmic amplification phase when displayed on a log scale), using the default setting for background subtraction (cycles 3 to 15). The Ct values were exported from the SDS software, and reimported into Excel for further analysis. DeltaCt values (CtTransgene minus CtIPC) were calculated for each sample and standard.

/*rest of page intentionally blank*/

**Results**

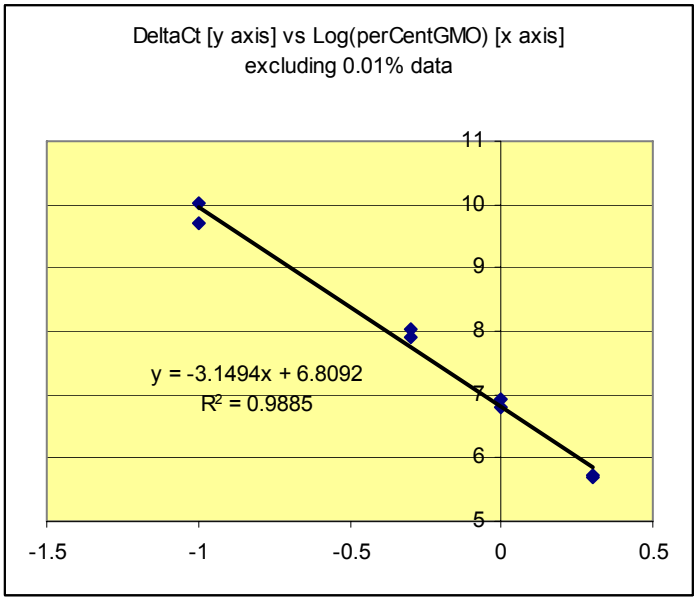
No amplification, for either the transgene, nor the IPC was observed for the NTCs

No amplification for the transgene was observed for the 0% GMO control sample, though the IPC amplified as expected.

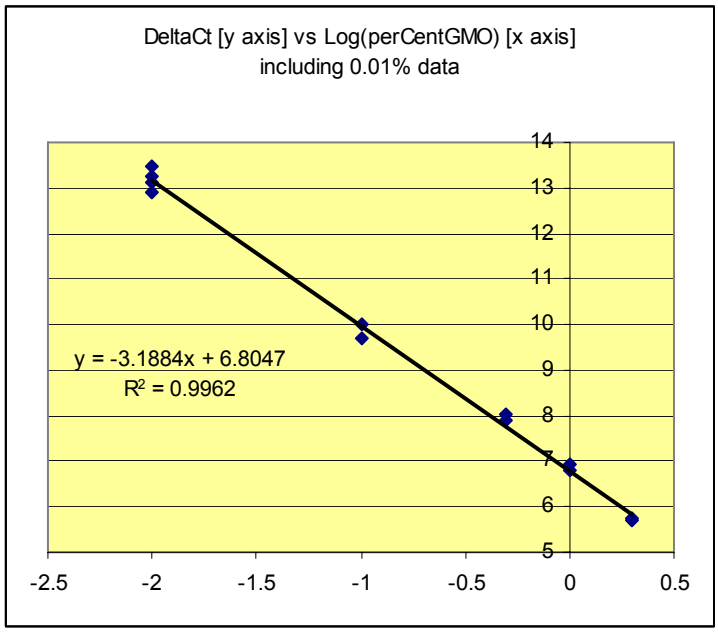
All other samples and standards amplified for both the transgene and the IPC. The Ct and DeltaCt values for the standards were as follows

| PercentGMO | Ct_GMO | Ct_IPC | DeltaCt |
|---|---|---|---|
| 2.00 | 27.715 | 22.024 | 5.691 |
| 2.00 | 27.776 | 22.038 | 5.738 |
| 1.00 | 28.892 | 21.968 | 6.924 |
| 1.00 | 28.821 | 22.021 | 6.800 |
| 0.50 | 29.411 | 21.391 | 8.020 |
| 0.50 | 29.126 | 21.235 | 7.891 |
| 0.10 | 31.563 | 21.552 | 10.012 |
| 0.10 | 31.299 | 21.601 | 9.698 |
| 0.01 | 34.923 | 21.800 | 13.124 |
| 0.01 | 34.868 | 21.947 | 12.921 |
| 0.01 | 35.010 | 21.775 | 13.235 |
| 0.01 | 35.293 | 21.805 | 13.488 |
| 0.00 | Undetermined | 21.760 | |
| 0.00 | Undetermined | 21.829 | |
| NTC | Undetermined | Undetermined | |
| NTC | Undetermined | Undetermined | |

A linear regression of the DeltaCt values against the log(percentGMO) should produce a straight line graph with R^2 >0.98. From the equation of the line, the amount of GMO in the test samples can be estimated. In a full GMO procedure, normally only the standards in the range 0.1% to 2% are used, as the sampling errors throughout the procedure preclude quantification below 0.1%:-




However, in this experiment, the flour mixtures were made directly, and not from sampling seeds, and so regression down to 0.01% may be acceptable:-




The Ct and Delta Ct values for the Client samples were as follows:-

| Sample_ID | Ct_GMO | Ct_IPC | DeltaCt |
|---|---|---|---|
| Sample_1 | 29.479 | 22.987 | 6.492 |
| Sample_1 | 29.876 | 22.982 | 6.894 |
| Sample_2 | 27.946 | 22.698 | 5.248 |
| Sample_2 | 27.986 | 22.748 | 5.238 |
| Sample_3 | 27.397 | 22.604 | 4.793 |
| Sample_3 | 27.216 | 22.620 | 4.596 |
| Sample_4 | 32.045 | 22.602 | 9.443 |
| Sample_4 | 32.232 | 22.390 | 9.842 |
| Sample_5 | 36.100 | 22.786 | 13.314 |
| Sample_5 | 35.766 | 22.799 | 12.966 |

The amount of GMO in each sample can be estimated from the delta Ct, the slope and the intercept of the regression line using the formula:-

PercentGMO=10^((DeltaCt – intercept)/(slope))

The estimations, with or without the use of the 0.01% standards, are shown below, together with the average estimate from each pair of readings:-

*/rest of page intentionally blank/*

| Sample_ID | Calc_Exc_0.01 | Av_Exc_0.01 | Calc_Inc_0.01 | Av_Inc_0.01 |
|---|---|---|---|---|
| Sample_1 | 1.261 | 1.100 | 1.253 | 1.095 |
| Sample_1 | 0.940 | | 0.938 | |
| Sample_2 | 3.132 | 3.143 | 3.078 | 3.089 |
| Sample_2 | 3.155 | | 3.100 | |
| Sample_3 | 4.367 | 4.705 | 4.275 | 4.601 |
| Sample_3 | 5.043 | | 4.928 | |
| Sample_4 | 0.146 | 0.127 | 0.149 | 0.130 |
| Sample_4 | 0.109 | | 0.112 | |
| Sample_5 | 0.009 | 0.010 | 0.009 | 0.010 |
| Sample_5 | 0.011 | | 0.012 | |

**Discussion**

The Ct value for all standards and samples was <24 and therefore passes acceptance standard. No amplification was observed in the NTCs and only the IPC amplified in the 0% GMO soya control. All four 0.01% GMO control samples had amplification of the GMO target. The linear regression coefficient R^2 for the plot of the deltaCt against log(PercentGMO) was greater than 0.98 Therefore all standards and samples pass the acceptance criteria.

The assays of the standards contained 200ng DNA (IDna fluorometry measurement), and the Client samples contained 100ng DNA (agarose gel densitometry relative to a lambda standard). The average Ct value for the IPC of the standards was 21.77, whilst that of the samples was 22.72, corresponding nearly exactly to a difference of 1 Ct unit, expected from halving the amount of DNA in the assay.

100ng correspond to around 83000 haploid soybean genomes in each assay. Consequently, a concentration of 0.01% GMO is represented by the equivalent of 8 genomes in each PCR assay.

No matter whether the 0.01% GMO standard is used for regression or not, the estimates for GMO content are very similar, being as follows:-

| Sample_ID | EstimatePercentGMO |
|---|---|
| Sample_1 | 1.10 |
| Sample_2 | 3.10 |
| Sample_3 | 4.70 |
| Sample_4 | 0.13 |
| Sample_5 | 0.01 |

*Special note concerning sample 5*. We note that sample 5 has a GMO concentration of around 0.01%. We have noted in the past that batches of Fluka certified reference material of 0% GMO do contain appreciable levels of GMO, particularly RoundUp Ready. Other laboratories involved in GMO testing have noted similar problems. For instance, the 0% GMO standard from Fluka lot and filling code 1129490 / 43404065, sample 6932 is known to contain RoundUp Ready soy at around 0.01%, hence the reason why IDna has verified an independent source of non-GMO soybeans to use as a standard. If the Client suspects that sample 5

should be 0% GMO, it may be that the certified reference standard that was used to make the samples is not 0% GMO as claimed, but 0.01%.

The estimated GMO concentration in one sample is above the range of the standards tested (sample 3, estimated concentration 4.7%). As only half the amount of DNA was used in the assays for the samples as the standards, the Ct value for the GMO target is very close to the range covered by the standards, hence the error in extrapolating this concentration is unlikely to be large. In a practical sense, if seed or grain lots are dramatically above 0.9%, the precision of the estimate is rarely of interest to the Client.

Peter G. Isaac
25 January 2007

/end of document/