JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

# Sampling Plan and Test Protocol for the Semiquantitative Detection of Genetically Modified Canola (*Brassica napus*) Seed in Bulk Canola Seed

KERRY R. EMSLIE,* LEE WHAITES, KATE R. GRIFFITHS, AND E. JOHN MURBY

National Measurement Institute, P.O. Box 385, Pymble, New South Wales, 2073 Australia

Using a statistical approach, sampling plans for the semiquantitative detection of genetically modified (GM) canola within a bulk seed sample can be developed and tailored to meet different GM thresholds, costs, and confidence limits. This is achieved by changing the number of subsamples analyzed, the number of seeds per subsample, and the percentage of positive results allowed. These sampling plans must be devised carefully, taking into account the detection capability of the analytical assay. This is particularly important in the case of InVigor (a registered trademark of Bayer CropScience) canola, for which expression levels of the introduced protein in seed are very low. Lateral flow assays and enzyme-linked immunosorbent assays (ELISA) were both investigated for their suitability as a qualitative assay using a subsampling approach. On the basis of an ELISA, several sampling plans have been devised and validated to provide at least 99% confidence that bulk seed samples containing at least 0.9% (w/w) InVigor canola will be detected. Although the term "seed" is used throughout this paper to refer to the canola, the term "seed" is to be taken to include both seed and the canola seed (grain) that is harvested by the farmer/grower.

**KEYWORDS:** ***Brassica napus*; canola; confidence; genetically modified organisms; GMO; immunoassay; InVigor canola; seed; semiquantitative; threshold**

## INTRODUCTION

In recognition that comingling of genetically modified (GM) and non-GM seed along the supply chain is difficult to avoid completely, several countries and the European Union have implemented threshold levels for allowable GM contamination in non-GM seed lots. Such threshold levels generally apply only to GM events that have been approved for release within the relevant country. In Australia, several GM canola events have been approved for commercial release by the regulating body, the Office of the Gene Technology Regulator (OGTR). In late 2005, the Primary Industries Ministerial Council of Australia agreed to adopt a threshold level for labeling of canola grain and seed approved by the OGTR (http://www.maff.gov.au/releases/05/pimc9.html). The threshold level was set, in the first instance, at 0.9% GM seed for canola crop and 0.5% GM seed for commercial seed for sowing.

The standard process to monitor for the adventitious presence (AP) of GM canola in a seed lot involves collection of several primary samples using established sampling procedures, which take into consideration the likely heterogeneity of GM material in the sample (*1*−*3*). The primary samples are then combined and mixed to form a composite sample that is representative of the whole seed lot (*3*). The composite sample is reduced in size, through the use of mixers and dividers, such as a riffle box, into one or more working samples that are analyzed using the chosen test method. The percentage of GM seed in a working sample can be estimated using a quantitative analysis such as real-time Polymerase Chain Reaction (*4*). As there will be an error associated with the sampling process, it is unlikely that the percentage of GM seed in such a working sample will correspond exactly to that in the composite sample. Hence, the uncertainty associated with the sampling process must be considered together with the uncertainty of the analytical method when the level of confidence associated with an estimate of the percentage of GM in the composite sample is determined (*3*, *5*).

Alternatively, a statistical approach can be used which is based on the assumption that the number of GM seeds present in any working sample taken from a homogeneous composite sample follows a binomial distribution. In this case, the minimum number of seeds that must be analyzed to have a defined level of confidence that at least one of those seeds would be GM if the composite sample exceeds the desired threshold level for GM contamination is determined (*6*, *7*). A qualitative analysis can then be conducted on such a working sample, and a positive result would indicate with the defined level of confidence that the composite sample exceeds the threshold level. Using this semiquantitative approach, the size and number of working samples will depend on the capabilities and detection limit of the chosen analytical assay, the acceptable threshold level, and the level of confidence desired in the result.

* Corresponding author (telephone +61 2 9449 0141; fax +61 2 9449 1653; e-mail kerry.emslie@measurement.gov.au).

10.1021/jf070267i CCC: $37.00 Published xxxx by the American Chemical Society
Published on Web 05/08/2007 **PAGE EST: 7.7**

**Table 1.** Number of Seeds Required in Working Sample To Ensure with a Defined Degree of Confidence That at Least One GM Seed Will Be Present[a]

| | confidence level | |
|---|---|---|
| % GM seed in composite sample | 95% | 99% |
| 0.1 | 2995 | 4603 |
| 0.5 | 598 | 919 |
| 0.9 | 332 | 510 |
| 1.0 | 299 | 459 |
| 3.0 | 99 | 152 |
| 5.0 | 59 | 90 |

[a] Sample size is determined using the following formula: $n = \log(1 - CL/100))/\log[1 - (P/100)]$ where $n$ is the number of seeds required in composite sample, CL is the probability (in percent) that at least one GM seed will be present in the composite sample, and $P$ is the actual perentage of GM seed in the composite sample (*16*). Calculated values for sample size are rounded up to the next whole number to specify minimum number of discrete seeds required in the sample.

The simplest semiquantitative approach aims only to ensure that specified levels of GM material are not exceeded and does not accommodate the situation when very low levels of AP may be acceptable. Thus, a representative sample of 598 canola seeds taken from a homogeneous composite sample containing 0.5% GM seed will contain at least one GM seed 95% of the time (**Table 1**). In other words, if buyers want to be 95% confident that they will reject seed lots with adventitious GM contamination at the level of 0.5% or more in the composite sample, then the working sample must contain a minimum of 598 canola seeds. This estimate of the required number of seeds in the working sample is valid only if the analytical method has zero false-positive and false-negative rates.

There is always a finite probability of a single adventitious seed producing a positive test result by being present in the working sample even when AP is at levels much lower than the maximum allowable level. This limitation can be circumvented by considering three key factors in the statistical approach: (1) the lower quality limit (LQL) or tolerance level, which is the maximum percentage of GM seed allowable to the end user (*6*, *8*); (2) the acceptable quality limit (AQL), which is the amount of GM contamination that is considered to be acceptable by the producer (*6*, *8*); and (3) the level of confidence required at both the LQL and AQL.

Such an approach can be achieved by devising sampling plans that require the analysis of multiple working samples and permitting a defined number of positive results. Individual positive test results arising from low levels of AP are thereby permitted under such sampling plans, whereas larger numbers of positives arising from GM levels of concern are not tolerated. This approach, which reduces the probability of rejecting seed lots with low levels of AP, cannot be implemented when a single working sample is analyzed with a qualitative assay.

In the design of any sampling plan based on a semiquantitative approach, it is essential to ensure that the analytical method is capable of detecting a single GM seed among the specified number of seeds in the working sample. The number of working samples required may thus be influenced by the sensitivity of the analytical method used for the detection of GM contamination. This is particularly relevant in the case of InVigor canola because expression of the introduced phosphinothricin acetyltransferase (PAT) protein in seed tissue is very weak (*9*). In contrast, the Roundup Ready (a registered trademark of Monsanto Co.) canola construct uses a constitutive promoter so the introduced CP4 5-enolpyruvylshikimate-3-phosphate synthase (CP4 EPSPS) protein is expressed throughout the plant, including the seed (*10*, *11*).

The purpose of this study was to validate sampling plans and test protocols for rapid, cost-effective, semiquantitative detection of GM canola seed using an immunoassay. As expression of the PAT protein in canola seed is known to be weak, protocols to allow detection of this trait were taken as the model for this investigation. However, these sampling plans can be utilized together with test protocols for the detection of other traits provided that the test has an equivalent or greater level of sensitivity than the test used to detect InVigor canola seed. Although the term "seed" is used throughout this paper to refer to the canola, the term "seed" is to be taken to include both seed and the canola seed (grain) that is harvested by the farmer/grower.

Although the sampling procedure from seed lot to primary sample did not form part of this study, a recent study on the distribution of GM soybean in large shipments demonstrated that randomness in seed lots cannot be assumed (*1*). This highlights the need to consider lot heterogeneity in the design of sampling protocols for the collection of primary samples from a seed lot to ensure that the composite sample is representative of the seed lot.

## MATERIALS AND METHODS

**Materials.** Genetically modified InVigor (MS8xRF3) canola seed with a stated lot purity of >97% and conventional canola seed were provided by Bayer CropScience. Roundup Ready RT73 canola seed Certified Reference Material (AOCS 0304-B) with a lot purity above 99.19% at the 95% confidence level was obtained from the American Oil Chemists' Society (Champaign, IL). Canola seed samples, each containing approximately 3300 seeds and comprising either 0.3, 0.6, or 1.2% RoundUp Ready RT73 canola seed, were obtained from the International Seed Testing Association (ISTA) as part of the sixth ISTA proficiency test on GMO testing on canola seed.

**Methods.** *Preparation of Working Samples.* Synthetic working samples containing a single InVigor canola seed in a total of 50 or 100 canola seeds were prepared by manually counting the required number of seeds. Working samples comprising 100 canola seeds were prepared from composite samples by either using a 50-seed counting device twice or manually counting the required number of seeds. Alternatively, working samples of approximately 100 seeds were prepared from composite samples by weighing out the previously determined average mass of 100 seeds. Working samples were transferred to individual ziplock bags for storage prior to analysis.

*Preparation of Composite Samples.* Several synthetic or field-based composite samples were prepared as follows: Two synthetic homogeneous composite samples (500 g) were prepared containing either 0.1 or 0.9% (w/w) InVigor canola seed by weighing out and combining the required mass of both InVigor canola seed and conventional canola seed. Each composite sample was homogenized by passing the seed twice through a riffle divider and recombining according to the ISTA guidelines (*3*).

A synthetic stratified composite sample containing 0.5% (w/w) InVigor canola was prepared by combining 200 g portions of each of the synthetic homogeneous 0.1 and 0.9% (w/w) InVigor canola seed samples. Fourteen field-based composite samples (1−2 kg) were collected from the grain flow during out-loading from a bin containing 42 tonnes of a relatively homogeneous mix of approximately 0.6% (w/w) InVigor canola during a field trial (*12*). The synthetic stratified composite sample and the field-based composite samples were all individually mixed in accordance with ISTA guidelines (*3*) prior to splitting into working samples.

*Lateral Flow Strip Assay.* The SDI Trait√ LL test kit (part 7000043, Strategic Diagnostics, Newark, DE) detects the PAT protein produced from either the *Streptomyces viridichromogenes pat* gene or the *Streptomyces hygroscopicus bar* gene. The latter gene has been incorporated into InVigor canola. Although the manufacturer claims a

detection limit of one LibertyLink (a registered trademark of Bayer CropScience) corn kernel that expressed the PAT protein in 100 conventional corn kernels, a detection limit for InVigor canola seed has not been specified.

The SDI Trait√ LL test kit assay was performed in accordance with the manufacturer's instructions for bulk corn testing as supplied with the kit with the following exceptions. The required number of working samples each in a separate ziplock bag was allowed to equilibrate for 1 h at either 12 or 37 °C or room temperature. Working sample seeds were then crushed inside the ziplock bag using a rolling pin and mixed to form a slurry with 1.0 mL of water, which had also been pre-equilibrated at the selected temperature. The slurry mixture was allowed to settle for 10−20 min, and then 0.5 mL of the liquid was transferred to a microcentrifuge tube. Two lateral flow strip assays were conducted on each seed mixture. The strips were allowed to develop for 10 min at the appropriate temperature (12 °C, room temperature, or 37 °C), and then results were interpreted by three or four independent analysts. Once a strip had developed (within 15 min), the top and bottom absorbent pads were removed for archiving of the strip, as described in the manufacturer's application note.

Working sample negative and positive controls containing 50 conventional canola seeds or 49 conventional canola seeds plus a single InVigor canola seed, respectively, were analyzed with each batch of 10 working samples.

*ELISA Assays To Detect the PAT Protein and the CP4 EPSPS Protein.* The LibertyLink PAT/bar ELISA plate kit (catalog no. AP013, Envirologix, Portland, ME) detects the PAT protein produced from the *S. hygroscopicus bar* gene but not the PAT protein produced from the *S. viridichromogenes pat* gene. The manufacturer claims a detection limit of 0.1% Starlink corn by weight but does not specify a detection limit for InVigor canola seed.

The QualiPlate ELISA kit for Roundup Ready corn event 603 and cotton (catalog no. AP010, Envirologix) detects the CP4 EPSPS protein. The manufacturer claims a detection limit of 0.1% by weight for corn event 603 but does not specify a detection limit for Roundup Ready canola seed.

For analysis using either ELISA assay, a crude protein extract was prepared from working samples by crushing the canola seeds inside a ziplock bag with a rolling pin and then adding 1 mL of kit wash buffer to form a slurry. A portion of the slurry (0.5 mL) was transferred to a microcentrifuge tube and centrifuged at 13000 rpm in a benchtop centrifuge for 10 min. The clear supernatant from each microcentrifuge tube was then transferred to an individual well in a blank 96-well plate. The ELISA was performed on samples of clear supernatant in accordance with the manufacturer's instructions using the low-sensitivity protocol for the Roundup Ready seed and the high-sensitivity protocol for the InVigor seed.

Five negative control and at least two positive control working samples were analyzed with each ELISA run. Negative controls comprised 100 conventional canola seeds. Positive controls for the LibertyLink PAT/bar ELISA and the QuickPlate ELISA kit for Roundup Ready corn event 603 comprised 99 conventional canola seeds together with either a single InVigor canola seed or a single Roundup Ready RT73, respectively.

*Verifying Suitability of Immunoassay Limit of Detection (LOD).* The lateral flow strip assay and, for the purposes of this study, the ELISA both provide qualitative data as a positive or negative result. For a qualitative assay, the LOD is generally expressed as the smallest amount of analyte at which the analytical method will have a false-negative rate of ≤5% (*4*, *13*, *14*).

When there are only two possible outcomes to a test, the results obtained by repeated application of the test can be described by a binomial distribution. At the LOD of a test where, by definition, the probability of obtaining a false-negative result is 5%, it is possible to predict the chance of obtaining a given number of false negatives for any particular number of tests performed. In addition, the number of tests required for estimation of the LOD can also be determined. Thus, if an assay is repeated 57 times on independent synthetic working samples containing a defined percentage of GM canola and there are no false-negative results, this provides a level of confidence of 95% that the assay is operating within the LOD (**Figure 1**). On the other hand, if more than 6 false negatives are detected in a set of 57 assays, it is very probable that the assay is not operating within its LOD.

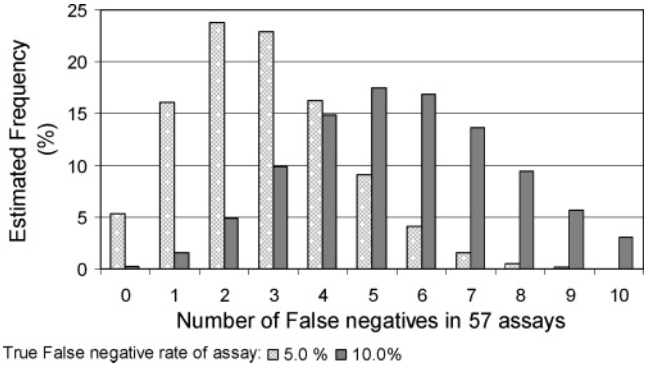


**Figure 1.** Expected number of false-negative results in a total of 57 assays when operating at a level of analyte with a true false-negative rate of either 5 or 10% assuming a binomial distribution of outcomes.

An assumption of this approach is that every synthetic working sample analyzed does contain a single InVigor canola seed. Because the batch of InVigor canola seed used in this study had a stated purity level of >97% GM seed, it is possible that a small number of working samples in a set of 57 may in fact be spiked with a non-GM seed. This would produce false false-negative results and prevent verification of the LOD by this procedure. Consequently, any assay that produced a negative test result was repeated using another lateral flow strip on the same sample extract. If the repeated analysis produced a positive result, the overall result for that sample was recorded as a true false negative. However, if the duplicate test on a given apparently false-negative assay was also negative, then the assay very probably contained no GM seed and results from this sample were disregarded. The assay was then repeated with another working sample.

*Design of Sampling and Test Plans.* Sampling and test plans were devised on the basis of probabilities predicted by the binomial distribution with the aid of SeedCalc7 software version 7.0 (http://www.seedtest.org/en/content---1--1143.html), which provides tools to enable design of sample plans based on either qualitative (*6*) or quantitative (*8*) analytical methods. Test plans were designed to work well within the LOD of the assay and, as such, no allowance was made for a false-negative rate. The percentage impurity of composite samples and the 95% confidence interval range for the true percentage impurity were estimated on the basis of the number of working samples analyzed using SeedCalc7 software (*6*).

## RESULTS AND DISCUSSION

**Verifying Suitability of Lateral Flow Strip Assay LOD.** In a previous study (*12*), consistent positive results were obtained with the SDI Trait√ LL lateral flow strips when using a ground sample containing 2% (w/w) InVigor canola seed in conventional canola seed, where the level of the PAT protein reflects the average expression level across a large number of InVigor canola seeds. However, because expression levels of the protein may vary significantly between individual seeds, the LOD may not necessarily be the same for a working sample containing a single InVigor canola seed.

To determine if a working sample comprising a single InVigor canola seed in a total of 50 canola seeds was within the LOD of the assay, 57 assays were performed in an air-conditioned environment at 20−25 °C on synthetic working samples prepared with one InVigor seed and 49 conventional canola seeds. Results were scored independently by three analysts, with two analysts reporting a single false negative and the third reporting two false negatives (data not shown). Assuming a binomial distribution of outcomes when operating at the LOD, 1 or 2 false-negative results in a total of 57 assays would be expected 16 and 24% of the time, respectively (**Figure 1**). If

the assay was operating below the LOD with a true false-negative rate of 10%, 1 or 2 false-negative results in a total of 57 assays would be expected <5% of the time. This result thus indicates that the LOD for this assay when conducted under controlled laboratory conditions at 20−25 °C is close to a single InVigor canola seed in 50 seeds. It should be noted that bands signifying a positive result for these working samples were very faint (as expected at the LOD), making result scoring very subjective at the LOD. In addition, a faint band was sometimes observed with negative controls giving the potential to score a false-positive result.

One of the advantages of the lateral flow strip format is its potential for use on site. To evaluate the robustness of the assay under simulated field conditions, synthetic working samples were analyzed at 45, 37, and 12 °C (data not shown). When the assay was conducted at elevated temperatures, false positives from conventional seed were common. For this reason, experiments to determine the LOD at this temperature were not undertaken. When the analysis was performed at 12 °C, four independent analysts reported between 5 and 7 false negatives from 30 tests. The binomial distribution predicts that at a true LOD of 1 GM seed in 50 seeds, 5 false negatives from 30 tests would be expected on approximately 1% of occasions. Six and seven false negatives are even less likely. It is therefore almost certain that the LOD of the assay is worse than 1 seed in 50 when performed at 12 °C.

Although lateral flow strips have the apparent advantages of speed and limited demands on equipment and staff resources, a critical requirement to ensure a cost-effective test plan is an assay with an adequate LOD. An assay with a poor LOD will require a larger number of tests on each composite sample. The results obtained under simulated field conditions of high and low temperatures suggest that an LOD of 1 in 50 is unlikely to be achieved reliably in routine use outside a laboratory. Because test plans should be designed to work well within the LOD of the assay, lateral flow strip assays were not considered further in this study. It should be noted, however, that this study evaluated only one commercially available lateral flow strip assay and does not necessarily reflect the suitability of other commercially available assays.

**Verifying Suitability of ELISA LOD.** To determine if a working sample comprising a single InVigor canola seed in 100 seeds was within the LOD of the LibertyLink PAT/bar ELISA, 59 assays were performed in an air-conditioned environment at 20−25 °C on synthetic working samples prepared with one InVigor seed and 99 conventional canola seeds. All 59 test samples gave a response that was significantly stronger than the 5 blank samples, verifying that the LOD of the ELISA is 1 InVigor canola seed in at least 100 conventional canola seeds (95% confidence level) (**Figure 2**). The blanks had a mean $Abs_{450nm}$ of 0.0146 and a standard deviation of 0.00055.

The 59 synthetic working samples analyzed by ELISA displayed a wide range of $Abs_{450nm}$ values with a broad distribution pattern ranging from 0.029 to 0.668 (**Figure 2**), but even the lowest of these exceeds the average of the 5 blanks by 26 times the standard deviation. This provides a very high level of confidence that the LOD of the ELISA exceeds the requirements of the sampling protocol. InVigor canola is a hybrid between the RF3 and MS8 lines, and therefore individual seeds contain either the RF3 construct alone or both the RF3 and MS8 constructs. Both the RF3 and MS8 parental lines contain the *bar* gene. However, a previous study demonstrated that expression levels of the PAT protein vary considerably between the parental lines and hybrid line, MS8xRF3, with samples of

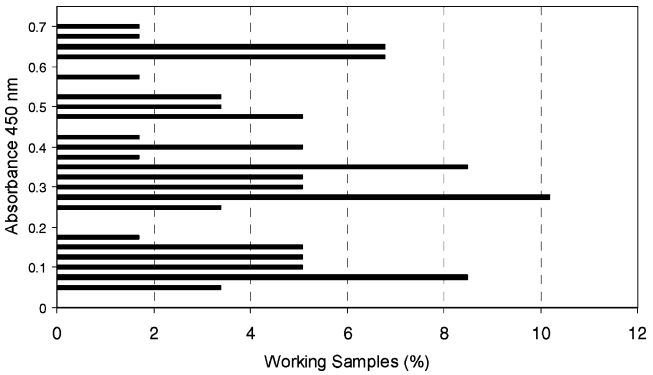


**Figure 2.** Histogram of absorbance readings at 450 nm from 59 synthetic 100-seed working samples each containing a single InVigor canola seed assayed using the Envirologix LibertyLink PAT/bar ELISA kit. Samples were binned in intervals of 0.025 absorbance reading and were counted in a particular bin if their absorbance reading was equal to the bin absorbance value (shown on the *y*-axis) or no more than 0.025 absorbance readings below that bin number. Horizontal bars indicate percentage of samples within each bin. The blank assays had a mean $Abs_{450nm}$ of 0.0146 and a standard deviation of 0.00055.

homogenized canola seed from MS8, RF3, and the MS8xRF3 lines containing 0.69, 0.07, and 0.35 $\mu$g of PAT/g of seed, respectively (*9*). The range in ELISA results from the synthetic working samples in this study is thus consistent with the range of expression levels observed in a previous study (*9*). This must be considered in designing a sampling plan because the assay must be capable of detecting an individual InVigor canola seed in the working sample even if the expression level of PAT protein in that seed is low. This implies that the detection limit when working samples containing a single InVigor canola seed in the sample are tested will be not as good as the detection limit when working samples derived from a larger, uniform, homogenized sample containing the equivalent percentage of InVigor seeds are tested because the level of PAT protein in a large, homogenized sample will reflect the average expression level over a number of InVigor canola seeds.

**Design of Test Plans.** All test plans were designed to be used in conjunction with the LibertyLink PAT/bar ELISA on working samples comprising 100 canola seeds. Because the ELISA reliably detected a single InVigor canola seed in a total of 100 seeds across 59 assays, the false-negative rate for the ELISA was assumed to be zero. On the basis of stakeholder requirements, all test plans ensured a confidence level of 99% that any composite sample at or above a designated LQL would be correctly identified. To reduce the risk to the producer that seed lots containing a very low adventitious level of GM seed would be rejected, plans also provided 95% confidence that composite samples below a designated AQL were correctly identified. These test plans are designed on the basis that the composite sample is representative of the whole seed lot.

*Single-Stage Test Plans Designed To Meet Different AQLs.* A series of single-stage test plans was designed to ensure 99% confidence of detecting composite samples containing at least 0.9% GM seed while providing varying levels of protection to the producer by changing the AQL. Operating characteristic curves, which plot the probability of accepting a lot against the true state or actual percentage of GM impurity in the lot (*7*), can be prepared and demonstrate the effect of changing the AQL from 0.0% (plan 1) to 0.3% (plan 4) GM seed on the number of working samples required (**Figure 3a**). Plan 1 comprises six 100-seed working samples, which is the minimum testing requirement, and all six samples must return a negative result

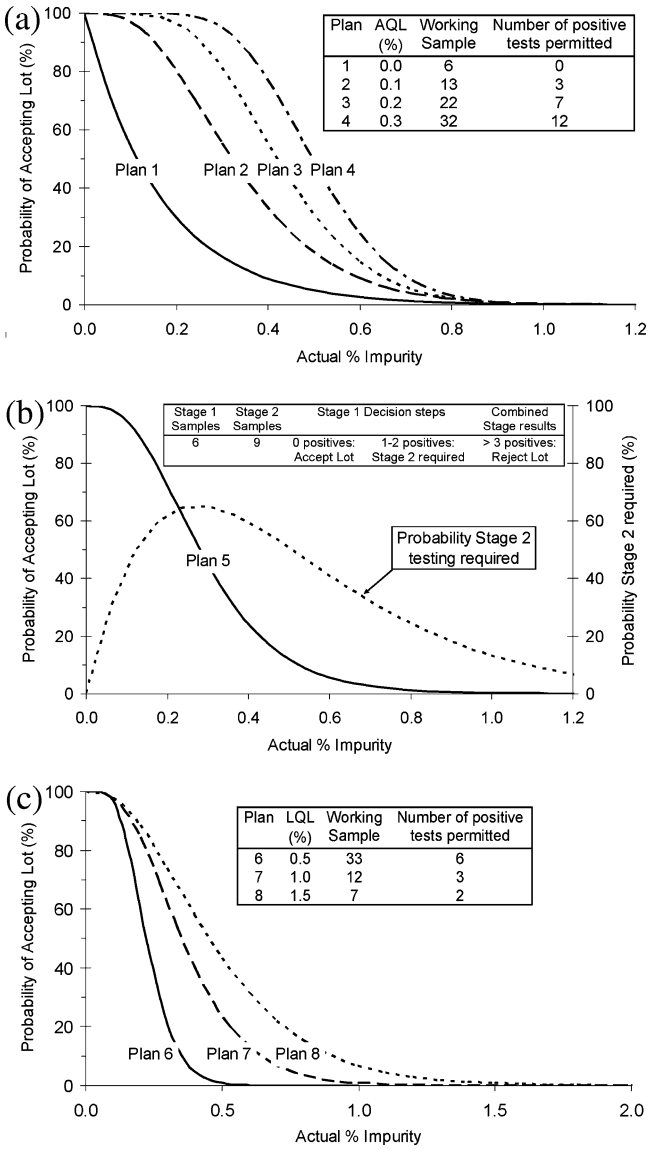


**Figure 3.** Operating characteristic curves for test plans using 100-seed working samples and designed to ensure at least 99% confidence of identifying composite samples at the designated LQL and 95% confidence at the designated AQL. Plans were derived using SeedCalc7 software: (**a**) single-stage test plans designed to meet different AQLs at a designated LQL of 0.9% GM seed; (**b**) two-stage test plan at a designated LQL of 0.9% GM seed designed to reduce costs if seed lots are expected to contain much less than the designated AQL of 0.1% GM seed [dotted line indicates probability of requiring stage two testing (secondary *y*-axis)]; (**c**) single-stage test plans designed to meet different LQLs at a designated AQL of 0.1% GM seed.

to have the required level of confidence that the composite sample contains <0.9% GM seed. Plan 1 does not specify an AQL and thus provides little protection to the producer, especially when low levels of GM seed are likely to be present in the sample. This is because there is a high probability that seed lots containing a low level of GM seed would be rejected. Plans 2, 3, and 4 are designed to provide 95% confidence to the producer that lots with an AP at the AQL of 0.1, 0.2, or 0.3%, respectively, will be accepted. This is achieved by analyzing a greater number of working samples and allowing a small number of positive samples. Hence, whereas plan 1 requires 6 working samples, plans 2, 3, and 4 require 13, 22, and 32 working samples, respectively (**Figure 3a**).

As these single-stage test plans were all designed to provide the same level of confidence at the LQL, the operating characteristic curves all show only 1% acceptance when the true concentration is 0.9%. The increased risk for the producer who adopts plan 1 compared to plan 2, 3, or 4 is clearly demonstrated through the change in shape of the operating characteristic curve where the probability of accepting a lot using test plan 1 drops dramatically as the actual percentage impurity of the lot increases from 0 to 0.5% (**Figure 3a**). The additional testing requirements for plans 2−4 lead to increased associated costs, although this cost may be balanced by the increased number of lots that are accepted. It is thus important that sampling plans are carefully designed to address the defined parameters required of both the buyer and the seller while minimizing the number of tests required.

*Two-Stage Test Plans.* Two-stage testing plans were designed to minimize the number of assays required. A small number of tests are conducted initially to identify composite samples that are clearly above or below the threshold. The second stage is required only for samples displaying an equivocal result. This can provide significant savings in time and consumables where the majority of composite samples are expected to be either well below the AQL or above the LQL.

For example, the addition of a second stage to plan 1 (**Figure 3a**), as shown in plan 5 (**Figure 3b**), retains the advantage of requiring a minimal number of tests in stage 1 while providing protection to the producer in the second stage in the case of positives resulting from low levels of AP. At an AQL of 0.1% GM, about 40% of composite samples would require the full two-stage testing protocol (**Figure 3b**). If it is anticipated that composite samples will usually contain much less than 0.1% GM seed, significant cost-savings will be achieved by adopting this two-stage sampling protocol.

*Single-Stage Test Plans Designed To Meet Different LQLs.* Plans 6, 7, and 8 are designed to ensure 99% confidence of detecting composite samples at a designated LQL of 0.5, 1.0, or 1.5%, respectively. These plans all provide 95% probability of accepting a lot when the true concentration of GM seed is 0.1% as demonstrated by the operating characteristic curves (**Figure 3c**). As the LQL decreases from 1.5 to 0.5%, the required number of working samples increases from 7 to 33.

**Validation of Sampling and Test Plans.** *Synthetic Homogeneous Composite Samples.* To validate the process of splitting composite samples into working samples and applying individual test plans with differing AQL and LQL requirements, 91 working samples were prepared from each of two synthetic composite samples containing either 0.1 or 0.9% (w/w) InVigor canola seed. Working samples were analyzed by ELISA and the resultant data used to evaluate the efficacy of the devised test plans (**Table 2**). For each test plan, results from the 91 working samples were considered consecutively to create the maximum number of independent data sets for evaluation. For example, plan 1 utilizes 6 working samples so 15 independent sets of working samples were evaluated, whereas for plan 4, 32 working samples are required, so only 2 independent data sets could be assessed.

Test plans 1−6 were designed to ensure 99% confidence of detecting composite samples at an LQL of 0.9% or below. Each of these test plans correctly classified the synthetic composite sample containing 0.9% (w/w) InVigor canola as exceeding the LQL (**Table 2**). Plans 7 and 8, which had LQLs of 1.0 and 1.5%, respectively, also consistently classified this synthetic sample as exceeding the LQL. This is not unexpected because the operating characteristic curves indicate that the probability

**Table 2.** Classification of Synthetic Composite Samples Containing 0.1, 0.5, or 0.9% (w/w) InVigor Canola Seed Using Devised Test Plans

| | test plan[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| no. of times plan evaluated[b] | 15 | 7 | 4 | 2 | *c* | 2 | 7 | 13 |
| InVigor canola sample | | | | | | | | |
| 0.1% (w/w) homogeneous | | | | | | | | |
| classified as exceeding LQL (maximum % GM seed allowable) | | | | | | | | |
| number times | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| % | 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.9% (w/w) homogeneous | | | | | | | | |
| classified as exceeding LQL (maximum % GM seed allowable) | | | | | | | | |
| number times | 15 | 7 | 4 | 2 | 15 | 2 | 7 | 13 |
| % | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.5% (w/w) stratified | | | | | | | | |
| classified as exceeding LQL (maximum % GM seed allowable) | | | | | | | | |
| number times | 14 | 6 | 2 | 1 | 5 | 2 | 3 | 7 |
| % | 93 | 86 | 50 | 50 | 83 | 100 | 43 | 54 |

[a] Test plans are as outlined in **Figure 3**. [b] Ninety-one working samples were prepared from each composite sample and analyzed using the Envirologix LibertyLink PAT/bar ELISA. [c] The 0.1, 0.9, and 0.5% (w/w) InVigor canola tests using plan 5 were evaluated 10, 15, and 6 times, respectively, and required second-stage testing on 30, 0, and 100% occasions, respectively.

of accepting a lot with a true GM contamination of 0.9% for test plans 7 and 8 is 1.7 and 10.1%, respectively (**Figure 3c**).

In the case of the synthetic composite sample containing 0.1% (w/w) InVigor canola, test plans 2−8 all correctly classified the composite sample as below the designated LQL, thus meeting the AQL for these test plans of between 0.1 and 0.3% GM canola. However, plan 1, which was evaluated using 15 independent data sets, classified this sample as exceeding the LQL of 0.9% on 4 occasions (**Table 2**). Because plan 1 has an AQL of 0.0% and has not been designed to provide protection to the producer, a frequency of 27% rejection is not unreasonable based on the operating characteristic curve for plan 1, which predicts that 45% of samples with a true GM contamination of 0.1% would be rejected.

*Synthetic Stratified Composite Sample.* To validate the sampling and testing process on a nonhomogeneous sample, a stratified composite sample comprising equal amounts of 0.1 and 0.9% (w/w) InVigor canola was prepared and mixed in accordance with ISTA guidelines (*3*). Ninety-one working samples were prepared from the mixed composite sample, and 33 of the 91 samples were positive when assayed by ELISA. On the basis of the binomial distribution, the probability of a single working sample of 100 seeds containing no GM seeds is estimated as 60.6% when the true GM contamination is 0.5% (w/w). Therefore, the probability that it would contain one or more GM seeds is 39.4%. Using this probability, 36 positive samples in a series of 91 independent working samples is the modal outcome predicted by the binomial distribution on seed with a true GM contamination of 0.5% (w/w), and hence 33 positive samples, as observed experimentally in this study, is a plausible result.

Test plans 1 and 2 classified the 0.5% (w/w) synthetic stratified composite sample as exceeding an LQL of 0.9% on 93 and 86% of occasions, respectively, whereas plans 3 and 4 both classified the sample as exceeding 0.9 on 50% of occasions (**Table 2**). These results correspond closely to those predicted from the operating characteristic curves for plans 1, 2, 3, and 4, which demonstrate probabilities of exceeding the LQL of 95, 82, 69, and 51%, respectively (**Figure 3a**). The two-stage testing plan, plan 5, was evaluated using six independent datasets

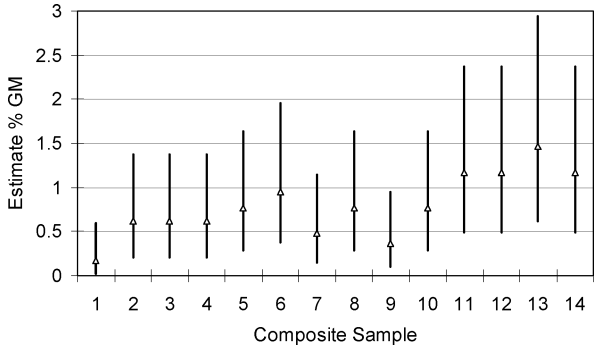


**Figure 4.** Estimated percentage of GM in composite samples collected from the grain flow during out-loading from a bin containing 42 tonnes of a relatively homogeneous mix of approximately 0.6% (w/w) InVigor canola seed. Percentage of GM (triangles) and 95% confidence intervals (vertical bars) were estimated using SeedCalc software based on analysis of 13 working samples from each composite sample.

and required progression to the second stage on each occasion (**Table 2**), suggesting that such a two-stage testing plan may not be cost-effective for samples containing contamination levels midway between the AQL and LQL. Plan 6 correctly classified the sample as meeting or exceeding an LQL of 0.5%, whereas plans 7 and 8 classified the sample as exceeding an LQL of 1.0 or 1.5% on 43 and 54% of occasions, respectively.

**Analysis of Composite Samples Collected from Grain Flow during Out-loading from a Bin.** Fourteen composite samples collected from the grain flow during out-loading from a bin containing approximately 42 tonnes of a relatively homogeneous mix of 0.6% (w/w) InVigor canola seed (*12*) were each mixed twice in their entirety using a riffle mixer; 13−15 100-seed working samples of each were then prepared by manual counting or by weighing an amount equivalent to the average weight of 100 seeds and analyzed by ELISA. The data obtained were processed using SeedCalc7 to obtain estimates of the InVigor canola content of the composite samples.

The estimated percentage impurity and 95% confidence interval range for the 14 field-based composite samples varied from 0.17% with a confidence interval range of 0.02−0.6% (composite sample 1) to 1.46% with a confidence interval range of 0.62−2.94% (composite sample 13) (**Figure 4**). For 13 of the 14 composite samples, the percentage of InVigor canola of approximately 0.6% (w/w) seed fell within the 95% confidence interval range for the true percentage impurity, and it fell just outside the range for the remaining sample, composite sample 13.

When interpreted using the devised test plans, all 14 composite samples were classified as exceeding a 0.9% LQL threshold for GM contamination using test plan 1 and the two-stage protocol, plan 5, whereas test plan 2 classified 13 of the composite samples as exceeding this threshold (data not shown). Plans 7 and 8 classified 13 or 12 composite samples as exceeding an LQL of 1.0 or 1.5%, respectively. On the basis of the operating characteristic curves, test plans 1, 2, and 5 all have a <10% probability of accepting a lot with 0.6% impurity; hence, the results obtained are consistent with predicted results.

**Analysis of ISTA Proficiency Samples Using Test Plans.** Seven blind canola seed samples containing 0.3, 0.6, or 1.2% RoundUp Ready canola seed were analyzed as part of the sixth ISTA proficiency study for GM canola seed using the Quick-Plate ELISA kit for Roundup Ready corn event 603. Each sample comprised approximately 3300 canola seeds. The results for each sample were applied to test plans 1−4, and the percentage of GM was estimated on the basis of the number of

working samples analyzed for each test plan. Although a total of 33 100-seed working samples were analyzed for each proficiency sample, test plans 1−4 required analysis of only 6, 13, 22, and 32 working samples, respectively. For each test plan, results for the working samples were considered consecutively until the required number of working samples for the test plan had been considered.

Test plans 1−4 all correctly classified two canola seed samples containing 1.2% RoundUp Ready canola seed as exceeding an LQL of 0.9% (data not shown). These plans also classified three proficiency samples containing 0.6% RoundUp Ready canola seed as exceeding an LQL of 0.9% (data not shown). However, this is not unexpected because test plans 1−4 all have a <25% probability of accepting a lot with 0.6% impurity (**Figure 3a**). Test plans 3 and 4 correctly classified two proficiency samples containing 0.3% RoundUp Ready canola seed as below an LQL of 0.9% (data not shown). However, one or both of the proficiency samples containing 0.3% RoundUp Ready canola seed were classified as exceeding an LQL of 0.9% using test plans 2 and 1, respectively (data not shown). This difference in classification of the samples containing 0.3% RoundUp Ready canola seed reflects the differences in AQLs among plans 1−4 and demonstrates the potential advantage for the producer when using either test plan 3 or 4.

The percentage of GM together with 95% confidence intervals was estimated for each proficiency sample, on the basis of the required number of working samples for each test plan. All four test plans correctly estimated the percentage of GM in each proficiency sample, taking into consideration the 95% confidence intervals for the estimated percentage of GM (**Figure 5**). Plan 1 requires the minimal number of six working samples and, consequently, the 95% confidence intervals for this plan are quite large. For example, using test plan 1 proficiency samples S8 and S10, which both comprised 0.3% RoundUp Ready canola seed, had 95% confidence intervals for the estimated percentages of GM of 0.13−2.11 and 0.04−1.49, respectively (**Figure 5a**). By increasing the number of working samples to 13 and 22, as for test plans 2 and 3, respectively, the 95% confidence interval range was reduced substantially. However, further increasing the number of working samples to 32 samples using test plan 4 had only a marginal impact on the confidence interval range, suggesting that the increased costs associated with test plan 4 may not be justified. A cost−benefit analysis, which evaluates the requirements of both the buyer and the producer and the costs associated with each test plan, would be required to determine the most appropriate plan for a specific purpose.

**Overall Uncertainty of the Process of Sampling and Analyzing a Seed Lot.** The uncertainty associated with determining GM content in a seed lot comprises the uncertainty in the collection of the composite sample together with that of the analytical procedure, which includes both reduction of the composite sample to one or more working samples and analysis of these samples. If a quantitative assay such as quantitative PCR is used, it would be possible to combine the uncertainties as the square root of the sum of the squared relative standard uncertainties in the usual way (*15*). This is not the case when a qualitative ELISA assay is used in the semiquantitative approach used in this study. It is, however, possible to obtain an indication of the overall uncertainty by a sensitivity analysis using the operating characteristic curve of the selected sampling plan and the maximum uncertainty in the preparation of the composite sample (at the 95% confidence level). These sampling

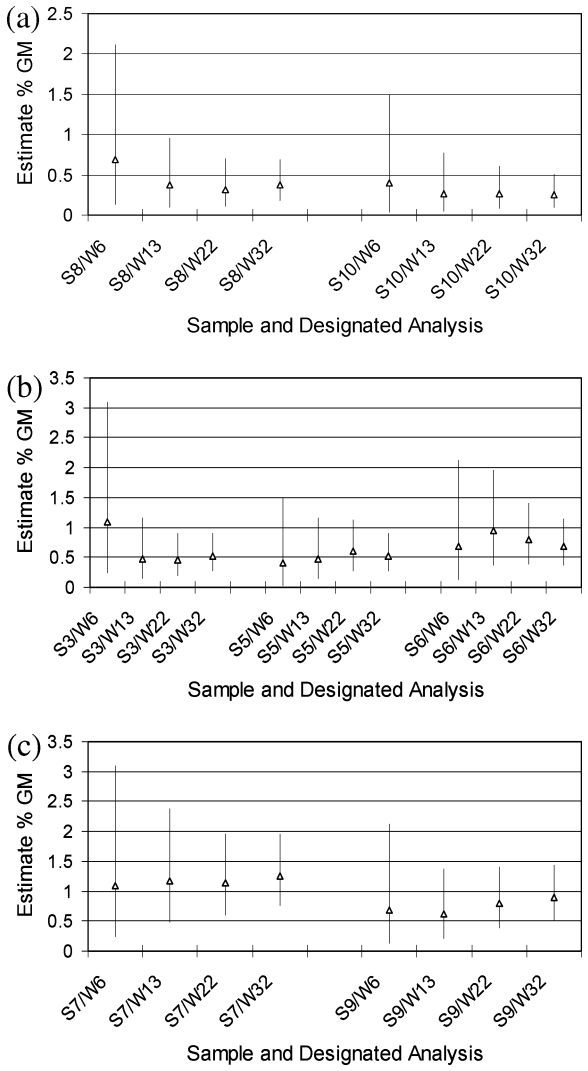


**Figure 5.** Estimated percentage of GM in ISTA proficiency samples containing (**a**) 0.3%, (**b**) 0.6%, and (**c**) 1.2% RoundUp Ready RT73 canola. W6, W13, W22, and W32 indicate analysis of 6, 13, 22, and 32 working samples, respectively, which correspond to the numbers of working samples required for test plans 1–4, respectively. S3, S5, S6, S7, S8, S9, and S10 denote ISTA proficiency samples. Percentage of GM (triangles) and 95% confidence intervals (vertical bars) were estimated using SeedCalc software based on analysis of the required number of working samples for the corresponding test plan. See Figure 3a for details on test plans 1–4.

plans are predicated on reliable detection of a single seed in the sample size stated, and no allowance for false negatives has been made. It is thus important to ensure that the assay on which the sampling plan is based will be required to detect only GM concentrations above that of its LOD.

**Application of Sampling Plans.** This study has investigated and validated sampling plans and test protocols for rapid, cost-effective, semiquantitative detection of GM canola seed in composite samples. A critical requirement to ensure a cost-effective test plan is an assay with an adequate LOD. This is particularly relevant to immunoassays, which rely on adequate expression of the introduced protein in seed tissue. In the case of InVigor canola, expression of the introduced PAT protein in seeds is weak and the lateral flow strip assay was not sensitive enough for practical implementation into a sampling plan. The ELISA demonstrated greater reliability and a lower LOD, and both single- and two-stage sampling plans that provided at least

99% confidence of identifying composite samples at the designated LQL and 95% confidence at the AQL were validated using this assay. The devised sampling plans must be used in conjunction with an established procedure for collection of primary samples from the bulk sample and preparation of the composite sample to ensure the composite sample is representative of the seed lot.

## ABBREVIATIONS USED

$Abs_{450nm}$, absorbance reading at 450 nm; AQL, acceptable quality limit; AP, adventitious presence; CP4 EPSPS, CP4 5-enolpyruvylshikimate-3-phosphate synthase; ELISA, enzyme-linked immunosorbent assay; GM, genetically modified; GMO, genetically modified organism; ISTA, International Seed Testing Association; LOD, limit of detection; LQL, lower quality limit; OGTR, Office of the Gene Technology Regulator; PAT, phosphinothricin acetyltransferase.

## ACKNOWLEDGMENT

The technical assistance of Gursharan Bains and Zena Kassir is gratefully acknowledged. The authors thank Dr. Wendy Lawson for her detailed comments on the manuscript.

## LITERATURE CITED

(1) Paoletti, C.; Heissenberger, A.; Mazzara, M.; Larcher, S.; Grazioli, E.; Corbisier, P.; Hess, N.; Berben, G.; Lubeck, P. S.; De Loose, M.; Moran, G.; Henry, C.; Brera, C.; Folch, I.; Ovesna, J.; Van den Eede, G. Kernel lot distribution assessment (*KeLDA*): a study on the distribution of GMO in large soybean shipments. *Eur. Food Res. Technol.* **2006**, *224*, 129−139.

(2) Macarthur, R.; Murray, A. W.; Allnutt, T. R.; Deppe, C.; Hird, H. J.; Kerins, G. M.; Blackburn, J.; Brown, J.; Stones, R.; Hugo, S. Model for tuning GMO detection in seed and grain. *Nat. Biotechnol.* **2007**, *25* (2), 169−170.

(3) Kruse, M. *ISTA Handbook on Seed Sampling*, 2nd ed.; International Seed Testing Association: Bassersdorf, Switzerland, 2004.

(4) Lipp, M.; Shillito, R.; Giroux, R.; Spiegelhalter, F.; Charlton, S.; Pinero, D.; Song, P. Polymerase Chain Reaction technology as analytical tool in agricultural biotechnology. *J. AOAC Int.* **2005**, *88* (1), 136−155.

(5) Begg, G.; Cullen, D.; Iannetta, P.; Squire, G. Sources of uncertainty in the quantification of genetically modified oilseed rape contamination in seed lots. *Transgenic Res.* **2007**, *16* (1), 51−63.

(6) Remund, K. M.; Dixon, D. A.; Wright, D. L.; Holden, L. R. Statistical considerations in seed purity testing for transgenic traits. *Seed Sci. Res.* **2001**, *11*, 101−119.

(7) Whitaker, T. B.; Freese, L.; Giesbrecht, F. G.; Slate, A. B. Sampling grain shipments to detect genetically modified seed. *J. AOAC Int.* **2001**, *84* (6), 1941−1946.

(8) Laffont, J.-L.; Remund, K. M.; Wright, D.; Simpson, R. D.; Grégoire, S. Testing for adventitious presence of transgenic material in conventional seed or grain lots using quantitative laboratory methods: statistical procedures and their implementation. *Seed Sci. Res.* **2005**, *15* (3), 197−204(8).

(9) OGTR. *Risk assessment and risk management plan. Application for licence for dealings involving an intentional release into the environment. DIR021/2002. Commercial release of genetically modified (InVigor® hybrid) canola*; Australian Government Office of the Gene Technology Regulator: Canberra, Australia, 2003.

(10) Maiti, I. B.; Gowda, S.; Kiernan, J.; Ghosh, S. K.; Shepherd, R. J. Promoter/leader deletion analysis and plant expression vectors with the figwort mosaic virus (FMV) full length transcript (FLt) promoter containing single or double enhancer domains. *Transgenic Res.* **1997**, *6* (2), 143−156.

(11) Sanger, M.; Daubert, S.; Goodman, R. M. Characteristics of a strong promoter from figwort mosaic virus: comparison with the analogous 35S promoter from cauliflower mosaic virus and the regulated mannopine synthase promoter. *Plant Mol. Biol.* **1990**, *14* (3), 433−443.

(12) Viljoen, J.; Griffiths, K.; Murphy, B.; Robinson, G.; Lwin, T.; Clamp, P.; Wilson, P. *Segregating GM and non-GM Grain in the Australian Grain Storage System*; Commonwealth of Australia: Canberra, Australia, 2005.

(13) Codex Alimentarius Commission. *Procedural Manual, Principles for the Establishment of Codex Methods of Analysis, Guidelines on the Application of the Criteria Approach*, 12th ed.; 2001.

(14) ISO. *ISO 24276:2006(E) Foodstuffs—methods of analysis for the detection of genetically modified organisms and derived products—general requirements and definitions*; International Standardization Organization: Geneva, Switzerland, 2006.

(15) ISO. *Guide to the Expression of Uncertainty in Measurement*; International Standardization Organization: Geneva, Switzerland, 1995.

(16) GIPSA. Sampling for the detection of biotech grains; http://archive.gipsa.usda.gov/biotech/sample2.htm.

**Received for review January 30, 2007. Revised manuscript received March 15, 2007. Accepted March 22, 2007. This project was funded in part by the Australian Department of Agriculture, Fisheries and Forestry.**

JF070267I