Seed Science Research (2005) **15**, 197–204 DOI: 10.1079/SSR2005210

Testing for adventitious presence of transgenic material in conventional seed or grain lots using quantitative laboratory methods: statistical procedures and their implementation

Jean-Louis Laffont[1], Kirk M. Remund[2]*, Deanne Wright[3], Robert D. Simpson[2] and Sylvain Grégoire[4]

[1]Pioneer Génétique, Chemin de l'Enseigure, 31840 Aussonne, France; [2]Monsanto Company, 800 North Lindbergh Blvd., St. Louis, Missouri, 63167, USA; [3]Pioneer Hi-Bred International Inc., 7300 N.W. 62nd Avenue, Johnston, Iowa, 50131-1004 USA; [4]GEVES, La Minière, 78 285 Guyancourt Cedex, France

Abstract

When the laboratory methods employed are qualitative, the statistical methodologies used in testing for the adventitious presence (AP) of transgenic material in conventional seed and grain lots are well defined. However, when the response from the method used by the laboratory is quantitative (e.g. percent transgenic DNA), the statistical methodologies developed for qualitative laboratory methods are not fully appropriate. In this paper, we present the details of procedures specific to quantitative laboratory methods. In particular we consider: (1) the assessment of quantitative laboratory method errors using linear modelling; and (2) the process of deciding whether or not a lot meets pre-specified purity standards, including the development of probability calculations needed to develop operating characteristic curves and estimate consumer and producer risks for a given lower quality limit (LQL), acceptable quality limit (AQL) and testing plan. We also describe implementation of this approach in a useful spreadsheet application.

Keywords: adventitious presence, quantitative laboratory methods, testing plans, transgenic, real time PCR, variance components, seed testing, statistical methods

*Correspondence:
Fax: +1 314 693 6673
Email: kirk.m.remund@monsanto.com

Introduction

Commercial seed lots are used to obtain grain lots for the next generation of food or feed. While the parent seed or the grain progeny may look much the same, the quality criteria to check for seed and grain are generally quite different. For example, quality characteristics for seed include germination, vigour and varietal purity; while quality criteria for grain include protein content or the processing ability of the grain for beer or bread. Testing for (low) levels of transgenic material in conventional lots is one of the few quality characteristics that is of interest for both grain and seed. The practice of testing conventional lots for the presence of transgenic material is widely referred to as adventitious presence (AP) testing, and we adopt this terminology even though the testing methods cannot differentiate between the adventitious or intentional presence of transgenic traits; in practice, appropriate documentation and procedures that are outside the scope of this paper would be required to provide evidence that any transgenic material is in fact adventitious. Throughout the remainder of this paper we use the word 'lot' to refer to a seed lot or a grain lot, and 'kernel' to refer to an individual seed or grain, as in both situations the principles and possibilities described are valid.

Numerous steps are taken to ensure that a lot is pure to pre-specified standards. Currently, one of the features that must be checked is the level of transgenic material in the lot. A lot that is claimed to carry a certain transgene must be ensured to carry it in a high percentage of the kernels. A conventional (non-transgenic) lot may also be checked and held to purity standards. The primary focus of this paper pertains to testing for AP of transgenic material in conventional lots.

As we cannot test all of the kernels in a lot because of the destructive nature of the tests, the first step of a testing regime consists of obtaining representative samples of kernels from the lot. This sampling step should be implemented carefully, as it is essential to ensure that the sample(s) is representative of the AP present in the whole lot.

Once representative samples are obtained, they are submitted to a laboratory to be assayed for AP. Depending on the method used by the laboratory, the characterization can be qualitative (i.e. absence or presence of a particular trait) or quantitative (i.e. quantification of the trait presence). It is important to understand and quantify the laboratory and assay errors that are common in the characterization of the sample(s). Qualitative laboratory methods have false-negative (i.e. a kernel or pool testing negative when in reality it is positive) and false-positive (i.e. a kernel or pool testing positive when in reality it is negative) rates. To assess false-negative rates at specified levels of AP, experiments are executed where laboratories are sent (blind) spiked samples at those levels. Similarly, laboratories can test true negative samples to evaluate their false-positive rates. For quantitative laboratory methods, there are multiple sources of variation. In this paper, we will provide guidance for obtaining estimates of these errors. Additionally, the statistical methods that we propose for developing testing plans account for the estimates of these errors.

Once a set of samples has been assayed for the AP, a decision is made regarding lot conformity to specified purity standards. This decision is supported by probability calculations, which are common in the area of statistics called statistical quality control. For qualitative laboratory methods, Remund *et al.* (2001) provided details for establishing testing plans based on the assessment of consumer and producer risks for a given lower quality limit (LQL) and acceptable quality limit (AQL), and Kobilinsky and Bertheau (2005) provided an optimization method to find minimum cost testing plans.

A number of applications exist for evaluating testing plans for qualitative laboratory methods. For example Seedcalc, developed by Remund and Simpson, is a Microsoft Excel® spreadsheet application freely available at the International Seed Testing Association (ISTA) web site (http://www.seedtest.org/en/content---1--1143.html).

Laffont and Wright (2002) introduced methods for evaluating testing plans based on quantitative laboratory methods. We provide the details of these probability calculations in this paper. We also describe how the statistical methods for quantitative laboratory methods are implemented in the newest version of Seedcalc, which is called Seedcalc6.

Materials and methods

Estimation of variance parameters

For a quantitative laboratory method [e.g. real-time polymerase chain reaction (PCR)], we propose to consider two sources of variation: flour subsampling variation (due to using only a subsample of the flour after grinding a pool of m kernels for the assay) and measurement (instrument) variation. In practice there may be other important sources of variation, such as variation over operators, machines or other factors that might change with time. We suggest to incorporate them into measurement variation by ensuring replicate measurements encompass different levels of these factors when estimating measurement variation. The following experimental design could be used to quantify flour subsampling and measurement variation when using pool sizes of 3500 kernels to test for AP levels around 0.1%:

- One pool of 3500 kernels is spiked with four homozygous positive kernels to target the 0.11% AP level. It is important that the zygosity of the transgenic kernels spiked is considered so that the desired target DNA level is achieved.
- The pool is ground into flour, five flour subsamples are taken, and DNA is extracted from each subsample.
- Each DNA extraction is measured three times using a quantitative assay.

This experimental design is proposed as a minimum requirement, especially as there are only five subsamples to be used for estimating the flour subsampling variation. It could be improved by taking more subsamples from one pool, by taking subsamples from more than one pool of 3500 kernels, or by considering various spiking levels.

The analysis of such an experiment (and of particular interest here, estimation of the primary sources of variation, e.g. the variance parameters) can be conducted using the linear mixed effects model:

$$y_{ij} = \mu + a_i + e_{ij}, \tag{1}$$

where μ is the overall mean, a_i ($i = 1,\ldots,5$) is the random effect of flour subsample i, and e_{ij} ($i = 1,\ldots,5$, $j = 1, 2, 3$) is the random effect of measurement j made on flour subsample i. We assume that the a_i effects are independently and identically distributed (iid) according to a normal distribution with mean zero and variance σ^2_{flour}, $N(0, \sigma^2_{\text{flour}})$, and the e_{ij} effects are iid $N(0, \sigma^2_{\text{measurement}})$. The variance parameters in these normal distributions, σ^2_{flour} and $\sigma^2_{\text{measurement}}$, represent the flour subsampling and measurement variation mentioned above. Estimates of these parameters are

needed for testing plan development described later in this section.

Usually, the variance parameters in such models are estimated using the restricted maximum likelihood (REML) method of estimation (Searle *et al.*, 1992). Although this method is available in most classical statistical packages, it is not available in spreadsheet applications such as Microsoft Excel®. Another method, the method of moments, provides estimates that are easy to compute and, thus, is used in our implementation of these methods in Microsoft Excel®. On the other hand, we stress that use of the method of moments must be avoided when the data are heavily unbalanced; the REML method provides estimates with better properties in this case. The variance parameter estimates are then incorporated into the development of appropriate testing plans, as described in the next subsection.

Testing plan development

Suppose that n ($n \geq 1$) pools of m ($m \geq 1$) kernels are taken from a lot, J flour subsamples from each pool are measured K times, and let y_{ijk} be the kth measurement made on flour subsample j from pool i.

We assume that y_{ijk} can be modelled as:

$$y_{ijk} = p + a_i + b_{j(i)} + e_{ijk}, \tag{2}$$

where p is the true AP probability in the lot, a_i is the random effect of pool i, $b_{j(i)}$ is the random effect of flour subsample j from pool i, and e_{ijk} is the random effect of measurement k made on flour subsample j from pool i. We assume that the a_i effects are iid $N(0, \sigma^2_{\text{sampling}})$, the $b_{j(i)}$ effects are iid $N(0, \sigma^2_{\text{flour}})$, the e_{ijk} effects are iid $N(0, \sigma^2_{\text{measurement}})$, and that the parameters $\sigma^2_{\text{sampling}}$, σ^2_{flour} and $\sigma^2_{\text{measurement}}$ are as follows:

- The variance parameter $\sigma^2_{\text{sampling}}$ is derived from the variance of a binomial distribution based on m kernels and probability p_{kernel} that a particular kernel is positive, denoted $B(m, p_{\text{kernel}})$. This gives $\sigma^2_{\text{sampling}} = \frac{p_{\text{kernel}}(1-p_{\text{kernel}})}{m}$ expressed on a kernel level. The probability p is the true presence of transgenic material in a lot expressed in %DNA. We desire to express $\sigma^2_{\text{sampling}}$ in terms of p rather than p_{kernel}. The relation between these parameters is $p_{\text{kernel}} = b \times p$, where b is a constant multiplier that converts from %DNA to %Seed units. We call this constant b the b-Factor (biological factor), since it makes unit conversions due to biological phenomena such as zygosity, ploidy and copy number. Using this relationship, the sampling variance of $\hat{p}$ is equal to the sampling variance of $\hat{p}_{\text{kernel}}/b$. The sampling variance of $\hat{p}_{\text{kernel}}/b$ is equal to $\frac{p_{\text{kernel}}(1-p_{\text{kernel}})}{b^2 m}$, which can be re-expressed as $\frac{bp(1-bp)}{b^2 m}$ using the relation $p_{\text{kernel}} = bp$. Therefore, the sampling variance expressed in %DNA units is $\sigma^2_{\text{sampling}} = \frac{p(1-bp)}{bm}$.
- The variance parameters σ^2_{flour} and $\sigma^2_{\text{measurement}}$ correspond to a specific laboratory process. To obtain estimates of these two parameters for any given laboratory or laboratory process, our recommendation is to perform experiments similar to the (minimal) experiment described in the previous subsection, or more sophisticated ones with multiple target AP levels. As the estimates to consider should reflect the current variability of the laboratory process, such experiments should be conducted on a regular basis and for each laboratory (method). At present, we recommend use of estimates provided by the analysis of the most recent experiment, realizing that these estimates can be improved by using information from previous experiments. A sufficient number of flour subsamples per AP level and measurement replicates per flour subsample will be necessary to get accurate estimates for any given laboratory process.

A fundamental objective of quantitative AP testing is to estimate the AP level in a lot (denoted by p) and determine if the lot meets pre-specified purity standards with high statistical confidence. The parameter p in model (2) is estimated by the sample mean: $\hat{p} = \frac{1}{nJK}\sum_{i,j,k} y_{ijk}$. The quantity $\hat{p}$ then gives the estimated AP in the lot. The variance of $\hat{p}$ is given by:

$$\sigma^2_{\hat{p}} = \frac{p(1-bp)}{bnm} + \frac{\sigma^2_{\text{flour}}}{nJ} + \frac{\sigma^2_{\text{measurement}}}{nJK}. \tag{3}$$

As $\sigma^2_{\text{measurement}}$ seems to depend linearly on p over a limited range of p while the measurement coefficient of variation ($\text{CV}_{\text{measurement}}$), defined as $\text{CV}_{\text{measurement}} = \frac{\sqrt{\sigma^2_{\text{measurement}}}}{p}$, is fairly constant (as an example, see Wright *et al.*, 2002, pp. 22–23), we can rewrite equation (3) as:

$$\sigma^2_{\hat{p}} = \frac{p(1-bp)}{bnm} + \frac{\sigma^2_{\text{flour}}}{nJ} + \frac{(p\text{CV}_{\text{measurement}})^2}{nJK}. \tag{4}$$

In these equations, p and $\text{CV}_{\text{measurement}}$ are expressed as proportions, that is percent/100, ranging from zero to one.

When using equation (4) one should carefully consider factors such as the zygosity, ploidy and copy number of the reference material used in PCR method calibration versus these same factors for transgenic presence in the lots to be tested. If both the reference material and the test lots have the same zygosity/ploidy/copy number, then b-Factor $= 1$ in equation (4) (i.e. %DNA is equal to %Seed). If the zygosity/ploidy/copy number of the reference material and

the test lots differ, then b-Factor $\neq 1$ in equation (4) [e.g. homozygous reference material and hemizygous test lots (b-Factor $= 2$)]. These are just two relatively simple examples. Lipp *et al.* (2005) provide a more thorough discussion of this topic.

When developing a testing plan based on a quantitative assay, we can define an acceptance limit (AL) such that if the observed AP% ($\hat{p}$) is less than or equal to the AL, the lot is 'accepted', and if not, the lot is 'rejected'. Given this AL and the variance of the average estimated AP% in equation (4), we can construct an operating characteristic (OC) curve, which is a plot of the true AP% p versus the probability of accepting the lot; such OC curves are useful for evaluating whether or not a given testing plan satisfies the testing objectives. OC curves are constructed by considering that $\hat{p}$ follows a normal distribution, such that $\hat{p} \sim N(p, \sigma^2_{\hat{p}})$. We then have

$$Pr(\hat{p} \leq \mathrm{AL}|p) = Pr\left(\frac{\hat{p}-p}{\sigma_{\hat{p}}} \leq \frac{\mathrm{AL}-p}{\sigma_{\hat{p}}}\middle| p\right) = \Phi\left(\frac{\mathrm{AL}-p}{\sigma_{\hat{p}}}\right) \quad (5)$$

where Φ is the cumulative distribution function for the standard normal distribution and $\sigma_{\hat{p}} = \sqrt{\sigma^2_{\hat{p}}}$.

Figure 1 is an example of an OC curve that can be used to investigate properties of a testing plan with the following components:

- Two pools of 3000 kernels are sampled from the lot, one flour subsample is taken per pool, and three measurements are made per flour subsample.
- The measurement CV is 15% (assumed to be fairly constant up to 1% AP) and the flour subsampling standard-deviation is $\sigma_{\mathrm{flour}} = \sqrt{\sigma^2_{\mathrm{flour}}} = 0.011\%$.
- The AL is chosen to be 0.5%.

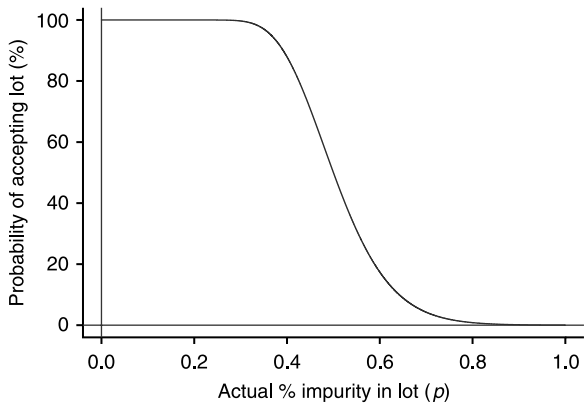

Figure 1. An example of the operating characteristic (OC) curve for a quantitative testing plan.

From this curve, we can see, for example, that with this particular testing plan and the variance parameters associated with this hypothetical assay, the probability of accepting a lot with true AP below 0.35% is very high (above 95%), while the probability of accepting a lot with true AP above 0.7% is very low (below 5%). Probability of acceptance for values between 0.35% and 0.7% AP in a lot can also be read from the y-axis, and used to determine whether or not this plan would satisfy testing objectives.

Equation (5) can also be used to estimate the consumer and the producer risks for a given lower quality limit (LQL) and a given acceptable quality limit (AQL), respectively (for a definition of consumer and producer risks, LQL and AQL, see Remund *et al.*, 2001):

$$\text{Consumer risk} = Pr(\hat{p} \leq \mathrm{AL}|\mathrm{LQL}) = \Phi\left(\frac{\mathrm{AL}-\mathrm{LQL}}{\sigma_{\hat{p}}}\right),$$

and

$$\text{Producer risk} = Pr(\hat{p} > \mathrm{AL}|\mathrm{AQL}) = 1 - \Phi\left(\frac{\mathrm{AL}-\mathrm{AQL}}{\sigma_{\hat{p}}}\right).$$

These risks represent specific values (that can also be obtained from the OC curve) in which there is a particular interest. The OC curve describes the consumer risk for accepting a lot when the true lot AP is equal to the LQL, as well as giving the probability of acceptance for AP equal to any other value shown on the x-axis; the producer risk can be obtained from the OC curve as 100 minus the y-axis value when the actual impurity is equal to the AQL.

Results

All of the methods discussed in the previous section to accommodate quantitative assays have been implemented in the newest version of the Microsoft Excel® spreadsheet Seedcalc: Seedcalc6. This new version contains the following additional worksheets: Quant Plan Design (Fig. 2), Compare Testing Plan (Fig. 3), and Quant Impurity Estimation (Fig. 4). These new worksheets are next discussed in detail, including some discussion and comparisons with the worksheets intended for qualitative plans. We do not include many details pertaining to these qualitative plans, but these details can be obtained from Remund *et al.* (2001).

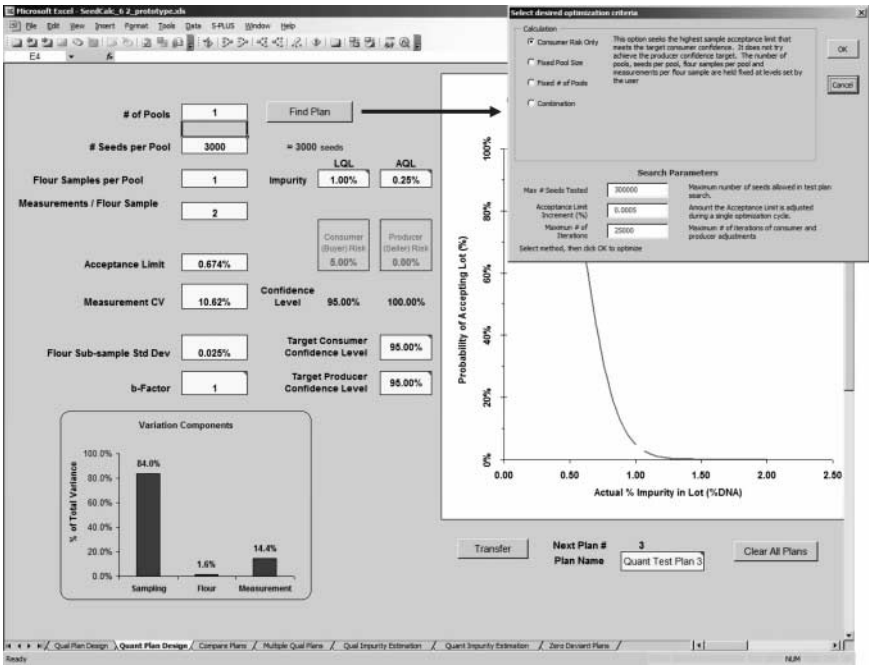

Figure 2. Seedcalc6 Quant Plan Design worksheet.

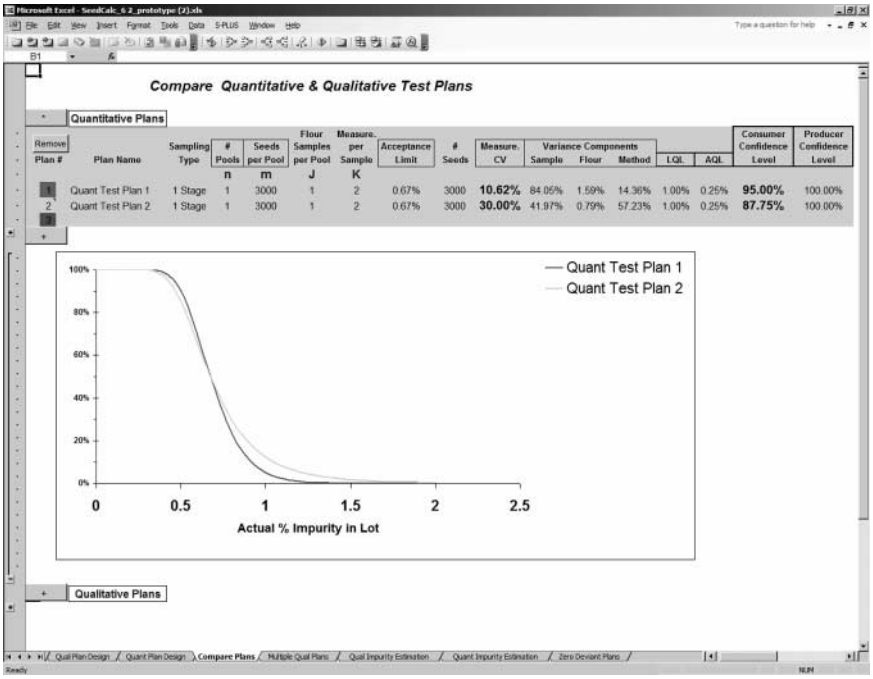

Figure 3. Seedcalc6 Compare Testing Plan worksheet.

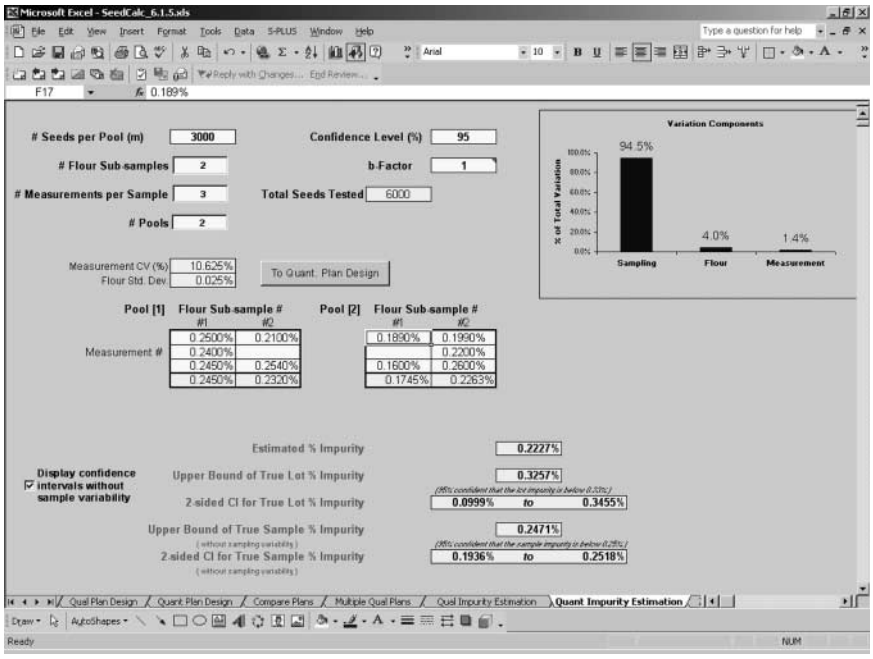

Figure 4. Seedcalc6 Quant Impurity Estimation worksheet.

Quant Plan Design worksheet (Fig. 2)

This worksheet is used to design seed/grain sampling and testing plans when a quantitative assay such as real–time PCR is used. Much of the logic and terminology is similar to the Qual Plan Design worksheet (intended for use with qualitative assays), but there are differences pertaining to the use of quantitative assays, such as the use of flour subsampling standard deviation or measurement CV, rather than false–positive and false–negative rates.

The number of pools, number of kernels per pool, number of flour subsamples per pool and number of measurements for each flour subsample can be entered by the user, based on the intended procedure for sampling and testing of the kernels. The usual LQL and AQL inputs are also provided by the user and are used to determine how good the testing plan is for the specific testing needs. The b–Factor can be used to make adjustments in the sampling variability for zygosity, ploidy, copy number or some other biological phenomenon that causes the transgenic material presence in %Seeds and %DNA units to differ. The measurement CV and flour subsample standard deviation are entered to define the performance of the assay system. These values take the place of the false–negative and false–positive rates for qualitative assays, and can be taken from past experience, or can be obtained from testing results that are transferred from the Quant Impurity Estimation worksheet (discussed later in this section). The acceptance limit (AL) is the cut–off impurity level measured in the sample for acceptance of a lot, and is also a required user input. In practice, testing results for all tested samples are averaged together, and this average is compared against the AL, as explained using formulas in the Materials and methods section; if the average sample results are equal to or less than this value, then the lot is 'accepted'; otherwise, the lot is 'rejected'. An AL equal to zero cannot be entered in the current version of Seedcalc because this value does not provide a good approximation of the binomial distribution with the normal distribution. The method limit of detection should also be considered when choosing AL values in close proximity to zero. Based upon the user inputs described in this paragraph, Seedcalc6 then gives the producer and consumer confidence levels at the AQL and LQL, respectively. It also provides graphical (OC curve) and tabular presentations of the lot acceptance probabilities at various levels of lot impurity.

An additional feature of Seedcalc6 is the 'Find Plan' tool, which can be used to help the user in searching for a testing plan. After entering the optional values, target producer and consumer levels, this tool can be used to modify the user inputs, defining the testing plan properties in the previous paragraph to achieve these levels. The 'Find Plan' tool can be selected by clicking the button above the LQL input cell. The user is then given the four options for

searching for testing plans. The descriptions for each of these four options are given in the 'Find Plan' window.

Compare Testing Plan worksheet (Fig. 3)

This worksheet can be used to visually compare the OC curves of different testing plans along with testing plan parameters. A testing plan can be designed in the Qual Plan Design or the Quant Plan Design worksheet and then transferred to this worksheet to be compared with other testing plans, or simply saved for future reference. Caution should be used when comparing testing plans based on quantitative assays with testing plans based on qualitative assays, as the impurity units may be different, and thus, the two types of testing plans may not be comparable. For example, the impurity units for a qualitative assay testing plan may be %Kernel, and units for a quantitative assay testing plan may be in terms of %DNA.

Quant Impurity Estimation worksheet (Fig. 4)

This worksheet can be used to obtain estimates of lot impurity/purity and approximate confidence limits when a quantitative assay is used in testing. Unbalanced data (e.g. missing measurements for the different flour subsamples) can be accommodated by using empty cells. The user enters data and parameters in the input cells, and the results are then reported. Notice that all individual quantitative results for each assay measurement are input into the spreadsheet so that the different components of variability can be estimated. A bar chart is displayed to visualize the contribution of each variation component to the total variation. These contributions are defined by:

- Sampling variation component: $\left(\frac{\hat{p}(1-b\hat{p})}{bnm}\right)/\hat{\sigma}^2_{\hat{p}}$
- Flour variation component: $\left(\frac{\hat{\sigma}^2_{\text{flour}}}{nJ}\right)/\hat{\sigma}^2_{\hat{p}}$ and
- Measurement variation component:

$$\left(\frac{\hat{\sigma}^2_{\text{measurement}}}{nJK}\right)/\hat{\sigma}^2_{\hat{p}}$$

where n is the number of pools, m is the number of kernels per pool, J is the number of flour subsamples from each pool, K is the number of measurements per flour subsample, b is the b-Factor, $\hat{p}$ is the mean of the quantitative results, $\hat{\sigma}^2_{\text{flour}}$ and $\hat{\sigma}^2_{\text{measurement}}$ are the estimates (determined using the method of moments) of the flour subsampling variance and measurement variance, respectively, and

$$\hat{\sigma}^2_{\hat{p}} = \frac{\hat{p}(1-b\hat{p})}{bnm} + \frac{\hat{\sigma}^2_{\text{flour}}}{nJ} + \frac{\hat{\sigma}^2_{\text{measurement}}}{nJK}.$$

The variance components estimated on this worksheet can be transferred to the Quant Plan Design worksheet to be used as inputs into the design of future testing plans when an appropriately designed preliminary experiment was used to obtain the input data.

To illustrate the use of these worksheets, an example is given (Fig. 4). Suppose that a laboratory has completed the analysis of samples to test for AP levels of a commercial transgenic trait in a conventional lot. The laboratory evaluated two pools of 3000 kernels each. Two flour subsamples were taken from each ground pool, and DNA extractions were performed. Three measurements were made using real time PCR on each extraction. These results are entered into the Quant Impurity Estimation worksheet in Fig. 4. Note that the second assay measurements for flour subsample #2 in pool #1 and subsample #1 in pool #2 are left blank to represent missing data. The worksheet reports the estimated impurity in the lot as 0.22%, and the associated 95% upper confidence limit is 0.33% for the lot; this 95% upper confidence limit gives an indication of the uncertainty of the result.

The measurement CV and flour sub-sampling variability obtained from these sample results can be transferred to the Quant Plan Design worksheet by clicking on the 'Quant. Plan Design' button shown in Fig. 4. Note again that our strong recommendation is to use specific experiments based on spiked pools with appropriate and known levels of AP to estimate the variance components, rather than to rely on estimation of variance parameters from unknown samples; when unknown samples are used, there is no way to be sure the variance parameters would be appropriate for the levels of AP appropriate for a given testing strategy. With these two variability estimates, a testing plan can be designed with the following additional criteria: LQL = 1%, AQL = 0.25%, pool size = 3000, number of flour subsamples = 1, and number of measurements per flour subsample = 2. The 'Find Plan' button can be selected along with the 'Fixed Pool Size' option and the testing plan shown in Fig. 2 is the result. Note that the consumer and producer confidence targets are met with the testing plan. If this testing plan is deemed reasonable, it can be transferred to the Compare Plans worksheet, as shown in Fig. 3, by clicking the 'Transfer' button in Fig. 2. If the testing plan designed in Fig. 2 is considered unreasonable because of the issue of inappropriate laboratory variability estimates, larger variance estimates can be considered to assess how the testing plan is impacted or changes. Note that in Fig. 3, with a measurement CV of 30%, the testing plan does not meet the target consumer confidence requirement of 95% at the LQL (consumer confidence level at LQL = 87.75%).

The confidence limits on the true lot impurity given in Fig. 4 are generally of primary interest. As

stated earlier, these confidence limits are partially a function of the sampling variability. In some instances, we may have an interest in calculating confidence limits that include lab variability, but exclude sampling variability. For example, in a laboratory proficiency test where blind samples are distributed, an estimate and confidence interval of the true value in the sample would be of interest. In this case, the sampling variability would not be included in the confidence interval if the entire sample were ground for analysis. Note that in this case, inferences can only be made to the true value in the sample. These confidence intervals, which exclude sampling variability, can be obtained in the Quant Impurity Estimation sheet (bottom of Fig. 4) by checking the box 'Displaying confidence intervals without sampling variability' on the left side of the sheet.

Discussion

The methods presented in this paper provide a useful extension to the present lot acceptance sampling methods when a quantitative assay is used. These methods can be used to design testing plans to evaluate lot purity; in particular, we focus on the application of testing for AP. Importantly, these methods are consistent with methods that have been developed for qualitative assays; the same general concepts and terminology apply for the two approaches. When quantitative assay methods are used, assessment of the variation from different sources is an important step. This assessment is done through a classical linear modelling framework; variance components are estimated and are then incorporated into the reject/accept criteria for seed/grain lots. As variation is laboratory and assay dependent, it is important to obtain estimates of variation for a given laboratory and assay prior to relying on a particular testing plan. Other factors that may affect testing plan validity through the accuracy of the variance estimates are the number of kernels in a given pool and the true AP in the lot; when developing testing plans, it is important to obtain variance estimates for appropriate ranges of these factors. Finally, we note that the methods presented here are an area of ongoing research. For example, we are investigating statistical methods for improving and incorporating the relationship between the AP level and the measurement variance ($\sigma^2_{\text{measurement}}$) directly into testing plan calculations.

Acknowledgements

The authors would like to thank Markus Lipp for his valuable advice, and Dr Henk Hilhorst and the two reviewers for their helpful comments and suggestions which have led to an improvement of the manuscript.

References

Kobilinsky, A. and Bertheau, Y. (2005) Minimum cost acceptance sampling plans for grain control, with application to GMO detection. *Chemometrics and Intelligent Laboratory Systems* **75**, 189–200.

Laffont, J.-L., Wright, D. (2002) Statistical methodologies in testing for adventitious presence. Communication at the ASTA 57th Corn & Sorghum Seed Conference, Chicago, Illinois. 13 December 2002

Lipp, M., Shillito, R., Giroux, R., Spiegelhalter, F., Charlton, S., Pinero, D. and Song, P. (2005) Polymerase chain reaction technology as analytical tool in agricultural biotechnology. *Journal of AOAC International* **88**, 136–155.

Remund, K.M., Dixon, D.A., Wright, D.L. and Holden, L.R. (2001) Statistical considerations in seed purity testing for transgenic traits. *Seed Science Research* **11**, 101–119.

Searle, S.R., Casella, G. and McCulloch, C.E. (1992) *Variance components*. New York, John Wiley & Sons.

Wright, D., Laffont, D.-L., Remund, K., Alarcon, C., Grothaus, D., Hondred, D., Kaufman, B., Lawrance, G., Meyer, S. (2002). Statistical methodologies in testing for biotech traits. *6th International Seed Testing Association Seminar in Statistics*. Available at web site http://www.seedtest.org/upload/cms/user/GMOdect.pdf (accessed 13 July 2005)

Received 30 September 2004
accepted after revision 30 May 2005