

# Probability models for detecting transgenic plants

Carlos M. Hernández-Suárez<sup>1</sup>, Osval A. Montesinos-López<sup>2</sup>, Graham McLaren<sup>3</sup>  
and José Crossa<sup>4\*</sup>

<sup>1</sup>Facultad de Ciencias, Universidad de Colima, Bernal Díaz del Castillo No. 340 Col. Villas San Sebastián, C.P. 28045 Colima, Colima, México; <sup>2</sup>Facultad de Telemática, Universidad de Colima, Bernal Díaz del Castillo No. 340 Col. Villas San Sebastián, C.P. 28045 Colima, Colima, México; <sup>3</sup>Biometrics and Bioinformatics Unit, Crop Informatics Laboratory (CRIL), International Rice Research Institute (IRRI), DAPO Box 7777, Manila, Philippines; <sup>4</sup>Biometrics and Statistics Unit, Crop Informatics Laboratory (CRIL), International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, México DF, México

## Abstract

When detecting the adventitious presence of transgenic plants (AP), it is important to use an appropriate testing method in the laboratory. Dorfman's group testing method is effective for reducing the number of laboratory analyses, but does not consider the case where AP is diluted below the sensitivity of the analyses, which causes the rate of false negatives to increase. The objective of this study is to propose binomial and negative binomial probabilistic models for determining the required sample size ( $n$ ), number of pools ( $g$ ), and size of the pool ( $k$ ) for detecting individuals possessing AP with a probability  $\geq (1 - \alpha)$  (for a small  $\alpha$ ) given: (1) pool size ( $k$ ); (2) estimated proportion of individuals with AP in the population ( $p$ ); (3) concentration of the trait of interest (AP) in individual seeds ( $w$ ); and (4) detection limit of the test ( $c$ ) (AP concentration in a pool below which it cannot be detected). The proposed models consider the different rates of false positives ( $\delta$ ) and false negatives ( $\lambda$ ), and the assessment of consumer and producer risks. Results have shown that when using the negative binomial, a required sample size  $n$  can be determined that guarantees a high probability that  $m$  individuals or  $g$  pools containing AP will be found. The pools formed have an optimum size, such that one element with AP will be detected at a low cost. The negative binomial distribution should be used when it is known that the proportion of individuals with AP in the population is  $p < 0.1$ ; thus, it is guaranteed that  $m$  individuals or  $g$  pools of individuals with AP will be detected with high probability.

**Keywords:** adventitious presence of transgenic plants, binomial distribution, negative binomial distribution, pool number, pool size, sample size

## Introduction

The presence of genetically modified plants, hereafter named the adventitious presence of unwanted transgenic plants (AP), is becoming common in modern crop production systems. This reality has created concerns regarding possible gene flow through outcrossing between AP crops and their landraces and wild relatives. This is especially important in a country such as Mexico, a centre of diversity for maize, where the effects of AP maize outcrossing with traditional maize landraces and wild relatives, such as tripsacum and teosinte, are unknown. Recently, different authors have reported contrasting results in terms of detecting AP maize in Mexico. Quist and Chapela (2001, 2002) were the first to report AP landraces collected in the Sierra Juarez region of the Mexican State of Oaxaca; they specifically identified genes from *Bacillus thuringiensis* (Bt), a soil bacterium gene used to create maize that is resistant to some insects. In contrast, 4 years later, Ortiz-García *et al.* (2005a, b) sampled maize landraces in the same region of Oaxaca State and failed to detect AP.

When testing for AP, two distinct activities should be emphasized. The first is determining the optimal sample size ( $n$ ) and sampling strategy to be used when taking seeds at random from a seed lot (Cleveland *et al.*, 2005; Ortiz-García *et al.*, 2005c); the second is determining the sample preparation and testing method to be used in the laboratory (Remund *et al.*, 2001). The sensitivity of the analyses and specificity of the tests are important factors that may affect the rates of false-negative and false-positive results (Remund *et al.*, 2001). Usually, it

\*Correspondence  
Email: j.crossa@cgiar.org

will be necessary to collect a large number of seeds from a reference population located in the region of interest, since the frequency of AP is likely to be very low.

Because laboratory tests are expensive, it is not feasible to analyse all  $n$  individual seeds collected from a lot. There are several testing plans for reducing the number of samples to be analysed (Montgomery, 1997). One plan consists of testing pooled seed samples (Remund *et al.*, 2001). Conditions listed by Federer (1991) for pooled samples are that: (1) the trait is discrete, i.e. it can be measured as presence or absence, or some countable quantity; (2) the proportion of positives (e.g. AP) is relatively small; and (3) pooling the samples does not alter the characteristics of individual samples.

The group testing method of Dorfman (1943) is effective for reducing the number of laboratory analyses and can result in up to 80% savings in the number of laboratory analyses (Federer, 1991). This method consists of dividing  $n$  individual samples (e.g. seeds) into  $g$  groups (or pools), each of size  $k$ . If a group tests positive, then at least one individual in the pool is positive; the author gives an approximate solution for the optimal value of  $k$ . A formula for determining the sample size ( $n$ ) required for detecting AP can be derived from the Dorfman method. However, a major disadvantage of the Dorfman testing plan is that it is insensitive to the dilution that arises when the group is formed. This is particularly true for large group sizes where the number of AP kernels in the pool can be diluted below the sensitivity of the analyses, which causes the rate of false negatives to increase.

All quality laboratory methods have false positives ( $\delta$  = probability of falsely detecting a seed with impurity) and false negatives ( $\lambda$  = probability of failing to detect seed with impurity). Furthermore, these two types of errors, which commonly occur in any testing plan, can be integrated in an overall consumer and producer risk assessment. Remund *et al.* (2001) proposed testing plans that integrate a given lower quality limit (LQL) and acceptable quality limit (AQL) for the consumer and producer risks, respectively.

For sampling seeds from a seed lot (USDA/GIPSA, 2000a, b) and testing seeds in the laboratory (Kay and Van den Eede, 2001; Kay and Paoletti, 2002), a uniform distribution of the number of individuals with AP in the seed lot is assumed. Therefore, when the size of the seed sample ( $n$ ) is small in relation to the size of the reference population ( $N$ ), the acceptance sampling method and testing plan generally use a binomial probability distribution with parameter  $p$  (frequency of AP in the population). However, when AP is rare, i.e.  $p \leq 0.1$ , using a binomial distribution may not provide an unbiased and precise estimate of  $p$  (Haldane, 1945; Cochran, 1980); in this case, using a negative binomial distribution is suggested (Gerrard and Cook, 1972; Kalton and Anderson, 1986).

The main objective of this research is to propose probabilistic models for determining the required (1) sample size,  $n$ ; (2) number of pools,  $g$ ; and (3) size of the pool,  $k$ , that will detect individuals containing AP with a probability  $\geq (1 - \alpha)$  (for small  $\alpha$ ). The proposed models were developed within the framework of the Dorfman model, but considering: (1) the dilution effect when forming groups (pools) of seeds to be tested; (2) the detection limit of the laboratory test; (3) the different rates of false positives and false negatives; and (4) the assessment of consumer and producer risks. The probability distributions used in this study were binomial and negative binomial distributions for: (1) pool size,  $k$ ; (2) estimated proportion of individuals with AP in the population,  $p$ ; (3) known concentration of the trait of interest (AP) in individual seeds,  $w$ ; and (4) the detection limit of the test,  $c$  (AP concentration in a pool below which it cannot be detected). These models can be used for detecting the presence/absence of AP or any other trait of interest.

### The Dorfman model

The procedure proposed by Dorfman (1943) consists of dividing  $n$  individuals into  $g$  groups or pools, each of size  $k$ . Each group is tested; if a group has the AP, then at least one individual has the AP. All  $k$  individuals in that pool must then be examined to identify individuals with AP or to estimate the proportion of individuals with AP. This may not be necessary if the objective is simply to know if any individuals have AP. The probabilistic model in this procedure is useful for determining sample size,  $n$ , in order to detect individuals possessing AP with acceptable probability, and for determining the optimum pool size,  $k$ , and the number of pools,  $g$  (note that the sample size  $n = gk$ ).

Assume a population of size  $N$  in which a fraction  $p$  has AP [say type (+)]. We consider the problem of determining the optimum values of  $n$  and  $k$  such that the probability of detecting at least one individual with AP is greater than  $(1 - \alpha)$  (for a given  $\alpha$ ). For sample size  $n$  and group size  $k$ ,  $g = n/k$  pools can be formed. If  $X$  is the number of + individuals in a pool, then  $P(X = j)$  ( $j = 1, 2, \dots, k$ ) follows a binomial distribution  $X \sim \text{Bin}(k, p)$ . The probability that a group is (+) is one minus the probability that  $k$  randomly selected individuals are negative

$$P(X > 0) = 1 - (1 - p)^k$$

The probability of a pool testing negative (-) is  $P(X = 0) = (1 - p)^k$ . Because there are  $g = n/k$  pools, the probability of detecting only (-) groups, given that the proportion of (+) individuals in the

population is  $p$ , is

$$[(1 - p)^k]^{n/k} = (1 - p)^n \tag{1}$$

If a small probability,  $\alpha$ , of detecting only (-) individuals is required, given that there is a proportion  $p$  of (+) individuals in the population, then equation (1) can be written as

$$(1 - p)^n < \alpha \tag{2}$$

It should be pointed out that the Dorfman model was not developed with the objective of determining the sample size  $n$ , but rather for determining the required number of pools,  $g$ , and the size of the pools,  $k$ , that will minimize the number of laboratory tests,  $T$ . Under these premises, the expected value of  $T$ ,  $E(T) = g + kgp'$ , is a function of the number of pools ( $g = n/k$ ), plus the number of individuals in the positive pools that need to be analysed, where  $p' = 1 - (1 - p)^k$  is the probability of a pool being detected positive. Therefore, the ratio between the expected number of laboratory tests required ( $T$ ) and the required sample size for each lab method is a measure of its expected relative cost  $E(T)/n = (g + kgp')/n = (1/k) + p'$  (Dorfman, 1943). Thus, minimizing the number of lab tests is equivalent to finding the minimum relative cost. However, Dorfman's model assumes that when  $k$  individuals in a pool are mixed, AP concentration would not be diluted. Therefore, under these assumptions, the value of  $n$  that satisfies equation (2) can be obtained as

$$n = \frac{\log(\alpha)}{\log(1 - p)} \tag{3}$$

The expression given in equation (3) is used by the United States Department of Agriculture (USDA/GIPSA, 2000a, b) to determine sample sizes for detecting AP seeds. It is mentioned that detecting AP is not different from detecting seeds with other discrete traits. This analysis would suggest a single group; however, in practice AP cannot always be detected when the proportion of AP seeds in the group is very small because the analytical methods used may not be sensitive enough. It has been suggested that for a pool size of  $k = 400$  grains, standard analytical procedures in the laboratory should be able to detect the presence of one AP grain. Equation (3) does not give any guidelines as to how the concentration of the trait of interest (AP) (impurity) in individual seeds ( $w$ ) could affect the pool size  $k$ , or how the dilution effect could make AP undetectable by standard analytical procedures in the laboratory ( $c$ ).

**Binomial sampling with the dilution effect**

When  $k$  individuals that form a pool are mixed or homogenized, the AP will be diluted; this dilution

effect increases with the size of the pool, and may decrease the AP concentration in the pool below the test's detection limit ( $c$ ), thereby increasing the number of false negatives [i.e. seed(s) with AP not detected when, in fact, it is present in the group].

We propose a model that considers the dilution effect as well as the laboratory's detection limit for a pool sampling method based on the Dorfman model (1943). We assume a reference population of size  $N$ , with a proportion  $p$  of individuals with AP [or type (+)]. We also assume that the concentration of AP per individual,  $w$ , is known (i.e. transgenic DNA as % of the total DNA in the seed). When  $g$  pools are formed from a total of  $n$  individuals, the AP concentration in a single (+) individual in a pool is reduced to  $wg/n = w/k$ . If  $c$  is the laboratory detection limit, it is required that  $(w/k) \geq c$ , in which case the probability of detecting AP in a pool with at least one (+) individual is 1, and zero otherwise. If a pool has  $X$  (+) individuals, then we require  $[(wX)/k] \geq c$ . Note that in this study, the units of AP concentration,  $w$ , can be given in % DNA, whereas the units of  $w$  for other traits of interest, such as unwanted diseases in the grain, may be given in % kernel (Laffont *et al.*, 2005).

The question is: what is the required sample size  $n$  and pool size  $k$  such that the probability of detecting individuals of type (+) in the population is equal to or greater than  $(1 - \alpha)$ ? Variable  $X =$  number of (+) individuals in the pool of size  $k$  ( $X = 0, 1, 2, \dots, k$ ) is a binomial variable with parameters  $k$  and  $p$ , that is,  $X \sim \text{Bin}(k, p)$ . Hence, the probability that a group will be detected (+) is

$$P\left(X \geq \frac{ck}{w}\right) = \sum_{j=ck/w}^k P(X = j) = \sum_{j=ck/w}^k \binom{k}{j} p^j (1 - p)^{k-j} \tag{4}$$

To compute more precise probability values from equation (4) and avoid rounding errors when calculating  $ck/w$ , and because the binomial distribution computes probability for discrete values between 0 and  $n$ , we will use the relationship between the binomial and beta distributions given by

$$\sum_{x=a}^k \binom{k}{x} p^x (1 - p)^{k-x} = \frac{1}{B(a, b)} \int_0^p x^{a-1} (1 - x)^{k-a} dx = \frac{\Gamma(k + 1)}{\Gamma(a)\Gamma(k - a + 1)} \int_0^p x^{a-1} (1 - x)^{k-a} dx,$$

where the beta function is related to the gamma function by

$$\frac{1}{B(a, b)} = \frac{1}{\Gamma(a)\Gamma(b)} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} = \frac{\Gamma(k+1)}{\Gamma(a)\Gamma(k-a+1)}$$

for  $a > 0$  and  $b = (k - a + 1) > 0$ .

Thus,  $P(X \geq a) = P(Y \leq p)$ , where  $X \sim \text{Bin}(k, p)$  and  $Y \sim \text{Beta}(x|a, b = k - a + 1)$ . This has the advantage that, for the beta distribution,  $a > 0$  and  $b > 0$ , which is not possible with the binomial distribution. Therefore, equation (4) can be rewritten as

$$P\left(X \geq \frac{ck}{w}\right) = \sum_{j=ck/w}^k P(X = j) = \frac{\Gamma(k+1)}{\Gamma(ck/w)\Gamma(k-ck/w+1)} \int_0^p j^{ck/w-1} (1-j)^{k-ck/w} dj \tag{5}$$

For the detection of AP, two types of error rates are important. One is the proportion of false positives,  $\delta$ , which is the probability that one individual (or group) is detected as (+) even though it is (-) ( $1 - \delta$  is the test specificity); the other is the rate of false negatives,  $\lambda$ , which is the probability of an individual or pool testing (-) even though it is (+) ( $1 - \lambda$  is the test sensitivity) (Remund *et al.*, 2001). Therefore, the adjusted probability of detecting AP in a group is given by  $p_a = P[(+)|(+)]P[(+)] + P[(+)|(-)]P[(-)]$  where  $P[(+)|(+)] = 1 - \lambda$ ,  $P[(+)] = p_b$ ,  $P[(-)] = 1 - p_b$ , and  $P[(+)|(-)] = \delta$  where

$$p_b = \frac{\Gamma(k+1)}{\Gamma(ck/w)\Gamma(k-ck/w+1)} \int_0^p j^{ck/w-1} (1-j)^{k-ck/w} dj \tag{6}$$

Therefore,

$$p_a = (1 - \lambda)p_b + \delta(1 - p_b) \tag{7}$$

When the rate of false positives ( $\delta$ ) and false negatives ( $\lambda$ ) are introduced into these equations, and since the pools are formed independently, then the number of pools testing (+) can be considered a random variable ( $Y$ ) with a binomial distribution  $Y \sim \text{Bin}(g, p_a)$ . Therefore, the probability of finding at least one (+) pool is

$$P(Y \geq 1) = 1 - \binom{g}{0} p_a^0 (1 - p_a)^g = 1 - (1 - p_a)^g \tag{8}$$

We want this probability to be  $\geq (1 - \alpha)$ ; alternatively, we want the probability that none is detected (+) to be  $< \alpha$

$$P(Y = 0) = \binom{g}{0} p_a^0 (1 - p_a)^g < \alpha \tag{9}$$

Note that equation (9) determines the required sample size  $n$  with  $k = \lceil w/c \rceil - 1$  and with  $p_a = (1 - \lambda)p_b + \delta(1 - p_b)$ . We used  $k = \lceil w/c \rceil - 1$ , because it is the maximum possible value of the pool size; however, with this value of  $n$  and  $k$ , the minimum number of laboratory tests  $[E(T) = g + kgp']$  and the minimum relative cost of the tests being  $(E(T)/n = (g + kgp')/n = (1/k) + p_a)$  are not achieved (Dorfman, 1943). The strategy for minimizing the number of laboratory tests is to find, for a given  $n$ , a value of  $k$  between 1 and  $\min(n, k1 = \lceil w/c \rceil - 1)$  that satisfies equation (9) with the minimum relative cost.

### Negative binomial sampling with the dilution effect considering false positives and false negatives

Haldane (1945) proposed the inverse sampling method (or negative binomial sampling or inverse binomial sampling) for cases where  $p$  is small (i.e.  $p \leq 0.1$ ). In this method, sampling continues until  $m$  individuals with AP are obtained. Assume a finite population  $U$  of size  $N$  is divided into two disjoint and complementary sets:  $C_i$  ( $i = 1, 2$ ) of size  $N_i$  ( $U = C_1 \cup C_2, C_1 \cap C_2 = \emptyset$ ) and  $N = \sum_{i=1}^2 N_i$ , where class  $C_1$  contains individuals with AP and class  $C_2$  individuals without AP. Then,  $p = N_1/N$  and  $q = N_2/N = (1 - p)$  such that  $N = N_1 + N_2 = Np + Nq$ .

To estimate  $p$ , individuals are sequentially sampled until the  $m$ th individual of set  $C_1$  is obtained. The total sample size,  $n$ , is a random variable with a probability distribution given by

$$P(n) = P(E)P(F|E)$$

where  $E$  is the event that in a sample of size  $n - 1$ , exactly  $m - 1$  individuals belong to set  $C_1$ , and  $F$  is the event where the last individual belongs to set  $C_1$  (Guenther, 1969). This probability distribution is

$$P(n) = \begin{cases} \frac{\binom{Np}{m-1} \binom{Nq}{n-m}}{\binom{N}{n-1}} \left[ \frac{Np - (m-1)}{N - (n-1)} \right] & (n = m, m+1, \dots, m+Nq) \\ 0 & \text{otherwise} \end{cases}$$

which is a negative hypergeometric distribution with expectation  $E(n) = ((N+1)m)/(Np+1)$  and variance  $\text{Var}(n) = ((N+1)(Np-m+1)(N-Np))/((Np+1)^2(Np+2))$ . However, when trying to detect AP plants, the size of the reference population  $N$  is unknown, but very large. In this case, it is reasonable to assume that

$N \rightarrow \infty$ . For this case, it is well known (Kotz *et al.*, 1988) that the random variable  $n$  has a negative binomial distribution

$$P(n) = \begin{cases} \binom{n-1}{m-1} p^m q^{n-m}; n=m, m+1, \dots & (10) \\ 0 & \text{otherwise} \end{cases}$$

with  $E(n) = m/p$  and  $Var(n) = m(1-p)/p^2$ . An unbiased estimate of  $p$  if  $m > 1$  is  $\hat{p} = (m-1)/(n-1)$  and an unbiased estimate of the variance of  $\hat{p}$  is  $Var(\hat{p}) = \hat{p}(1-\hat{p})/(n-2)$ . Furthermore, the required probabilities from the negative binomial can be computed directly from the binomial distribution due to the following equality (Patil, 1960; Bartko, 1962; Morris, 1963):

$$\sum_{i=m}^{n^*} \binom{n^*}{i} p^i q^{n^*-i} = \sum_{n=m}^{n^*} \binom{n-1}{m-1} p^m q^{n-m}$$

The inverse sampling method suggested by Haldane (1945) is more precise than binomial sampling because when  $m > 1$ , the coefficient of variation of  $p$  decreases. Therefore, when the dilution effect is considered, and  $g = n/k$  independent pools are formed from the sample of  $n$  individuals, it is necessary to consider the number of pools with at least one element carrying AP as a random variable with a binomial distribution  $Y \sim Bin(g, p_a)$ , where  $p_a = (1-\lambda)p_b + \delta(1-p_b)$  [equation (7)] is the probability that at least one positive element is in a pool adjusted by the rate of false positives and false negatives. Therefore, the probability of finding at least  $m$  positive pools out of  $g$  ( $g = n/k$ ) is

$$P(Y \geq m) = \sum_{j=m}^g \binom{g}{j} p_a^j (1-p_a)^{g-j} \tag{11}$$

$$= 1 - \sum_{j=0}^{m-1} \binom{g}{j} p_a^j (1-p_a)^{g-j}$$

where  $g = n/k$  and  $k = [w/c] - 1$ .

We wish to find  $n$  such that  $P(Y \geq m) \geq 1 - \alpha$ . Using equation (11), this is equivalent to

$$\sum_{j=0}^{m-1} \binom{n/k}{j} p_a^j (1-p_a)^{n/k-j} < \alpha \tag{12}$$

Given  $m$ ,  $p_a$ , and  $\alpha$ , equation (12) can be solved numerically for  $n$ . Equation (12) is used to compute the required  $n$  with  $k = [w/c] - 1$ ,  $g = n/k$ , and  $p_a = (1-\lambda)p_b + \delta(1-p_b)$ . Similar to the binomial case, for this value of  $n$ , values of  $k$  between 1 and

$\min(n, k1 = [w/c] - 1)$  that satisfy equation (12) should be found that will have a minimum relative cost ( $E(T)/n = (g + kgp')/n = (1/k) + p_a$ ) (Dorfman, 1943). Note that for  $m = 1$ , equation (12) reduces to

$$P(Y \geq 1) = 1 - \binom{g}{0} p_a^0 (1-p_a)^g < \alpha \tag{13}$$

Thus, the binomial sampling method with the dilution effect proposed in equations (13) and (8) is a particular case of the inverse sampling method shown in equation (12). Also, it is interesting to show that for  $\delta = \lambda > 0$ ;  $p_a$  will increase and thus  $n$  will decrease, compared to the case where  $\delta = \lambda = 0$ . However,  $p_a$  considers all individuals that will be detected as (+), even if they are not. However, for determining  $n$ , only the true probability of the real (+) positives should be considered, as was the case with  $\delta = 0$ , that is, when  $p_a = (1-\lambda)p_b$ .

**Testing seed plans with the dilution effect considering false positives, false negatives, lower quality limit (LQL), and a given acceptable quality limit (AQL)**

The AP testing plan previously outlined has the aim of computing the required sample size ( $n$ ), pool size ( $k$ ), and number of pools ( $g$ ) to guarantee, with probability  $\geq (1 - \alpha)$ , that at least one AP plant (or AP pool) will be in the sample. These types of testing plans have zero tolerance because they focus only on the limiting quality level ( $LQL = p$ ), and do not consider the acceptable quality level (AQL) (i.e.  $AQL = 0$ ), which refers to different seed production levels under normal conditions. Zero-tolerance testing plans generally have a high producer risk (Remund *et al.*, 2001). In practice, tolerance testing plans have two main parameters: (1) the number of individual seeds (or seed pools); and (2) the maximum number of unacceptable seeds (or seed pools) that can be tolerated in the sample before the seed lot is rejected. Thus, in practice, testing plans should consider the consumer's and the producer's interests and risks by assessing the LQL and AQL. More details on the definition and description of LQL and AQL for consumer and producer risks can be found in Remund *et al.* (2001).

When  $n$  individuals (or  $g$  seed pools) are chosen from a seed lot, and it is decided to reject the lot if more than  $m$  AP plants are observed, it is necessary to construct an operating characteristic (OC) curve that plots the true AP proportion ( $p$ ) versus the probability of accepting the lot. OC curves are useful for evaluating whether or not a given testing plan satisfies the testing objectives. In this case, the probability that the lot will be accepted, given its

true AP proportion ( $p$ ),  $\lambda$ ,  $\delta$ , and considering  $w$  and  $c$  is given by

$$P(\text{Accept Lot}|p, n, k, c, w) = P(Y \leq m) \\ = \sum_{j=0}^m \binom{n/k}{j} p_a^j (1 - p_a)^{n/k-j} \quad (14)$$

where  $p_a = (1 - \lambda)p_b + \delta(1 - p_b)$ . Equation (14) can be used to estimate the consumer and producer risks for a given lower quality limit (LQL) and a given acceptable quality limit (AQL), such as

$$\text{Consumer Risk} = P(Y \leq m|p = \text{LQL}, n, k, c, w) \\ = \sum_{j=0}^m \binom{n/k}{j} p_a^j (1 - p_a)^{n/k-j} \quad (15)$$

and

$$\text{Producer Risk} = P(Y > m|p = \text{AQL}, n, k, c, w) \\ = \sum_{j=m+1}^{n/k} \binom{n/k}{j} p_a^j (1 - p_a)^{n/k-j} \quad (16)$$

Equations (15) and (16) are adapted from Remund *et al.* (2001), including the dilution effect and the testing limit  $c$ . If the value of  $k$  from equations (14), (15) and (16) is  $> [w/c] - 1$ , the program will search for a  $k$  between 1 and  $\min(n, k1 = [w/c] - 1)$  that will guarantee that the AP concentration ( $w$ ) in the pool is larger than the detection limit ( $c$ ) and that the size of the pool has the minimum relative cost.

## Results

### **The binomial sampling method with the dilution effect**

For different values of  $p$  and  $w$ , Table 1 shows the sample size,  $n$ , number of pools,  $g$ , and pool size,  $k$ , required to achieve 95% and 99% probability of detecting AP in the sample using the modified Dorfman method with the dilution effect. Clearly, as  $p$  increases and  $w$  increases, the required  $n$  decreases. When  $w = 0.0002$ , all individuals must be tested,  $g = n$  and  $k = 1$ ; as  $w$  increases, a smaller number of groups ( $g$ ) of larger sizes ( $k$ ) must be tested.

Sample sizes,  $n$ , for different values of  $p$  and  $w$ , are very different for the standard Dorfman model, as compared with those obtained from the

modified Dorfman model with the dilution effect. For example, consider that the laboratory detection limit is  $c = 0.0001$ ; then for  $p = 0.01$  and  $w = 0.0002$ , the required sample size to guarantee, with 0.95 probability, the detection of at least one (+) individual is 299 individuals using the standard Dorfman method of equation (3). However, the modified Dorfman method with the dilution effect required testing 7036 individuals [equation (8)] (Table 1). For  $p = 0.01$  and  $w = 0.01$ , the required sample size from equation (3) is 299 individuals, while using the modified Dorfman method of equation (8), the required sample size is 892 individuals (Table 1); however, while the modified Dorfman method with the dilution effect recommends performing 60 laboratory tests with 15 grains each ( $g = 60, k = 15$ ), the traditional Dorfman method without the dilution effect does not say how many groups ( $g$ ) (i.e.  $g$ , laboratory tests) and group sizes ( $k$ ) are required.

Therefore, the two methods give rise to different sample sizes,  $n$ , but the modified Dorfman method with the dilution effect has the advantage that it gives a precise value for  $g$  and  $k$ . The Dorfman model, which emphasizes pool size, was proposed with the objective of minimizing the number of laboratory tests. However, it disregards the dilution effect and the laboratory detection limit, thus increasing the probability of false negatives (i.e. detecting no AP grains in a sample when, in fact, there are some present). The modified Dorfman model adjusted for the dilution effect considers pool size as well as the laboratory detection limit. The problem is that when pool size increases, the dilution of AP increases; that is, AP concentration,  $w$ , decreases and may become smaller than the laboratory detection limit,  $c$  (i.e.  $w < c$ ). It is important to point out that  $w$  is considered a fixed quantity for a given grain, but this concentration decreases when grains in the pool are mixed because only a fraction,  $p$ , contain AP.

When  $p$  increases,  $n$ ,  $k$  and  $g$  decrease, but for a given value of  $p$ , when  $w$  increases,  $k$  may increase or decrease (Table 1). For  $c = 0.0001$ ,  $p = 0.01$ , and  $w = 0.0002$ , the sample size for  $\alpha = 0.05$  was  $n = 7036$ , and the pool size was  $k = 1$ . This result indicates that each individual must be tested separately, otherwise  $w < 0.0001$  (below the detection limit). For the same case, but assuming  $w = 0.01$  with  $\alpha = 0.05$ , then  $n = 892, k = 15$ . The increase in AP concentration should reduce the total cost of laboratory testing because now we have an optimum group size that minimizes the number of laboratory tests without diluting the AP below the laboratory detection limit. Even if only one element of the 15 had AP, the AP concentration in the group would be 0.000666, which is still detectable.

**Table 1.** Sample size ( $n$ ), group size ( $k$ ), and number of groups ( $g$ ) for various values of  $p$  and  $w$  required for achieving a 95% and 99% probability of detecting at least one individual with AP using the binomial distribution with  $\lambda = 0$  (rate of false negatives),  $\delta = 0$  (rate of false positives) and the dilution effect ( $c = 0.0001$ )

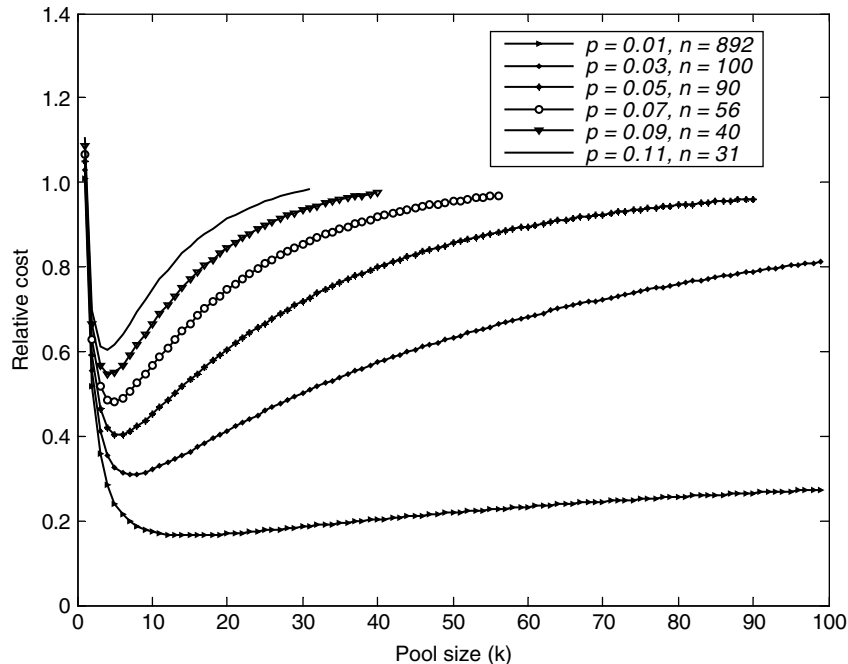
$p$	$w = 0.0002$			$w = 0.0008$			$w = 0.002$			$w = 0.006$			$w = 0.010$			$w = 0.014$		
	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$
$(1 - \alpha) = 95\%$																		
0.005	19934	1	19934	23192	7	3314	11876	19	626	4662	59	80	3070	34	91	2364	22	108
0.010	7036	1	7036	6420	7	918	3231	19	171	1358	26	53	892	15	60	696	13	54
0.020	2480	1	2480	1800	7	258	913	19	49	414	11	38	298	9	34	279	9	31
0.030	1345	1	1345	862	7	124	438	19	24	178	8	23	100	9	12	140	7	20
0.040	871	1	871	519	7	75	267	19	15	119	7	17	100	6	17	108	6	18
0.050	621	1	621	344	7	50	172	11	16	60	5	12	90	6	15	79	5	16
0.060	471	1	471	253	7	37	134	8	17	60	5	12	69	5	14	62	5	13
0.070	372	1	372	190	7	28	96	7	14	60	5	12	56	5	12	51	5	11
0.080	303	1	303	155	7	23	77	6	13	57	5	12	47	4	12	43	4	11
0.090	253	1	253	127	7	19	58	6	10	47	4	12	40	4	10	38	4	10
0.100	215	1	215	106	7	16	58	5	12	41	4	11	35	4	9	33	4	9
0.110	186	1	186	85	7	13	39	5	8	36	4	9	31	4	8	30	4	8
0.120	163	1	163	78	7	12	39	5	8	32	4	8	28	4	7	27	4	7
$(1 - \alpha) = 99\%$																		
0.005	30642	1	30642	35652	7	5094	18260	19	962	7199	59	123	4753	34	140	3754	22	171
0.010	10816	1	10816	9871	7	1411	4960	19	262	2066	26	80	1486	15	100	1113	13	86
0.020	3811	1	3811	2766	7	396	1407	19	75	650	11	60	496	9	56	418	9	47
0.030	2067	1	2067	1331	7	191	685	19	37	296	8	37	199	7	29	140	6	24
0.040	1338	1	1338	792	7	114	419	19	23	178	7	26	100	6	17	140	6	24
0.050	954	1	954	533	7	77	286	11	26	119	6	20	100	6	17	127	5	26
0.060	723	1	723	386	7	56	210	8	27	119	5	24	100	5	20	99	5	20
0.070	572	1	572	295	7	43	153	7	22	60	4	15	90	5	18	81	5	17
0.080	466	1	466	232	7	34	134	6	23	60	4	15	75	4	19	68	4	17
0.090	389	1	389	190	7	28	96	6	16	60	4	15	64	4	16	59	4	15
0.100	331	1	331	162	7	24	96	5	20	60	4	15	56	4	14	52	4	13
0.110	286	1	286	134	7	20	77	5	16	57	4	15	49	4	13	46	4	12
0.120	250	1	250	120	7	18	58	5	12	50	4	13	44	4	11	42	4	11

Figure 1 plots the relative cost of laboratory tests versus the group size ( $k$ ) under different values of  $n$  and  $p$ ; it shows, under the binomial distribution with dilution, that the optimum pool size is between 1 and  $\min(n, k1 = [w/c] - 1)$ . For example, for  $p = 0.01$ ,  $n = 892$ , the optimum group size that minimizes the relative cost is  $k = 15$ ; for  $p = 0.03$ ,  $n = 100$ , the optimum pool size for minimizing the relative cost is  $k = 9$ ; for  $p = 0.05$ ,  $n = 90$ , the best pool size is  $k = 6$ ; for  $p = 0.07$ ,  $n = 56$ , the best size is  $k = 5$ ; and for  $p = 0.09$  and  $0.11$ ,  $n = 40$  and  $31$ , respectively, the best size for both is  $k = 4$ .

**The negative binomial sampling method with the dilution effect considering false positives and false negatives**

Table 2 shows, for different values of  $m$  and  $p$ , the sample size,  $n$ , group size,  $k$ , and number of pools,  $g$ , needed to be able to detect the AP with a 95% probability obtained using equation (11) or (12); this assumes that the rate of false negatives is  $\lambda = 0.02$ , and

the proportion of false positives is  $\delta = 0$ . Results show that as  $p$  increases, sample size ( $n$ ) decreases. When more AP individuals must be detected ( $m$ ), the sample size increases for all levels of the other factors. As already shown in the theoretical sections, an increase in the number of individuals to be detected increases the precision of the estimate of  $p$ . For example, for  $p = 0.01$ ,  $w = 0.0008$ ,  $c = 0.0001$ , and  $m = 1$ , the required  $n$  to detect an individual containing AP with a 95% probability is 6553, whereas for  $p = 0.01$ ,  $w = 0.0008$ ,  $c = 0.0001$ , and  $m = 3$  and  $m = 11$ , the required sample sizes are 13,777 and 37,143, respectively, and the pool size is 7. However, for the same value of  $p$ , but  $w = 0.01$ , the  $n$  required for  $m = 3$  and  $m = 11$  are  $n = 2179$  and  $n = 6139$ , respectively, with a pool size of  $k = 15$ . On the other hand, assuming  $p = 0.03$ ,  $w = 0.0008$  in order to detect  $m = 1, 3, 5, 7, 9$ , and  $11$  AP individuals with a 95% probability, the required sample sizes are  $n = 883, 1863, 2717, 3515, 4285$ , and  $5041$ , respectively, with  $k = 7$  for each pool. However, if  $w = 0.006$ , the required sample sizes are  $n = 237, 532, 827, 1122, 1358$ , and  $1653$ , with a required pool size of  $k = 8$ .



**Figure 1.** The relative costs of laboratory tests as a function of optimum group size ( $k$ ) and sample size ( $n$ ) for different values of  $p$  [proportion of the adventitious presence of transgenic plants (AP) in the population] under the binomial distribution with the dilution effect for  $c = 0.0001$ ,  $\lambda = 0$ ,  $\delta = 0$ ,  $w = 0.01$  and  $(1 - \alpha) = 0.95$ .

When  $p$  is modelled by the negative binomial, its coefficient of variation does not change much with different values of  $p$  (Fig. 2). However, it decreases considerably when more individuals with AP must be detected. With  $m \geq 20$ , the coefficient of variation of  $p$  decreases at least 30% compared to the case where  $m = 1$ ; however, this will significantly increase the sample size and, therefore, the total cost of sampling and testing.

Results using the negative binomial distribution indicate that an important increase in detection precision (decrease in the coefficient of variation of  $p$ ) can be achieved in the testing plan by increasing the number of individuals with the AP that need to be detected,  $m$ . However, this implies an important increase in the total cost of sampling and testing. When  $\lambda > 0$  and  $\delta = 0$ , the sample sizes tended to increase, whereas when both error rates ( $\lambda$  and  $\delta$ ) increase, sample size tends to decrease, but group size remains constant (data not shown).

**Testing seed plans with the dilution effect considering false positives, false negatives, lower quality limit (LQL), and a given acceptable quality limit (AQL)**

Figure 3 depicts the operating characteristics (OC) curves for four different laboratory testing plans with

different values of  $g$  and  $k$  ( $g = 80, k = 40$ ;  $g = 54, k = 60$ ;  $g = 40, k = 80$  and  $g = 32, k = 100$  with  $n = 3200, m = 8, c = 0.0001$  and  $w = 0.014$ ) for an LQL of 1.0% and an AQL = 0.5% using equations (14), (15) and (16). The consumer risks for the four plans are 0.0060, 0.0248, 0.0759 and 0.1526, whereas the producer risks are 0.3787, 0.1863, 0.0764 and 0.0317. Results show that the best testing plan is when  $g = 40$  and  $k = 80$ , because it represents the lowest consumer (7.59%) and producer (7.64%) risks; that is, when  $g = 40$  and  $k = 80$ , seed lots with 1.0% of AP plants will be accepted 7.59% of the time, and seed lots with 0.5% AP plants will be accepted 7.64% of the time. Figure 3 shows the influence of  $m, c,$  and  $w$  on producer and consumer risks. Specific testing plans guarantee that if there is at least one pool with AP, this will be detected. However, these specific testing plans do not guarantee that it will have low consumer and producer risks.

Figure 4 shows the effect of the detection limit ( $c$ ) on the OC curves. The first three OC curves (from right to left) have detecting limit values of  $c = 0.00014$ ,  $c = 0.00012$  and  $c = 0.00010$ , with  $n = 1820, k = 15, m = 6,$  and  $w = 0.0088$ . The OC farthest to the left shows the standard testing plan (for details, see Remund *et al.*, 2001), which does not include the dilution effect; no values of  $c$  and  $w$  are used. If the LQL and AQL are set at 1.1% and 0.4% and  $c = 0.00014, 0.00012, 0.0001$  and without considering

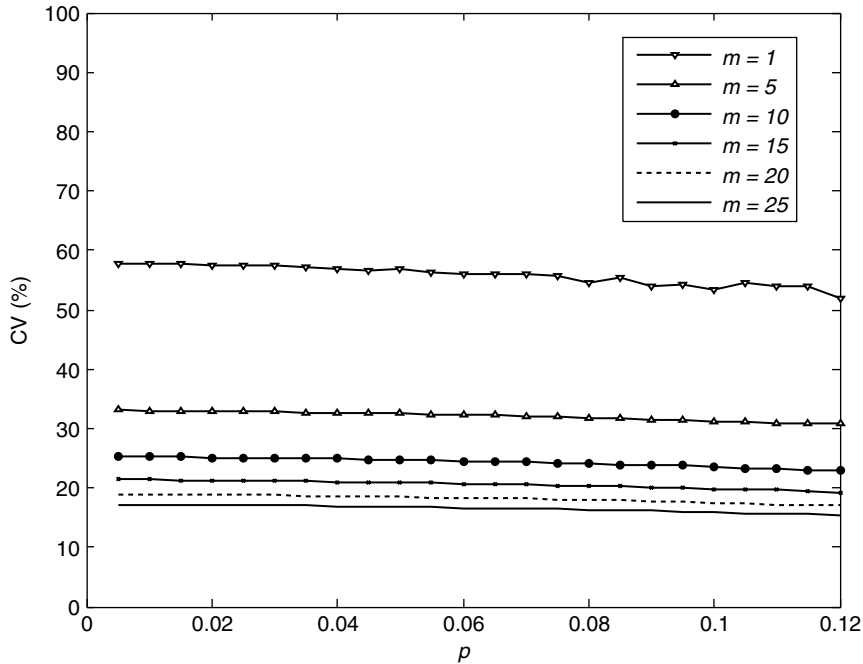


**Table 2.** Sample size ( $n$ ), group size ( $k$ ), and number of groups ( $g$ ) for various values of  $p$  and  $w$  required for achieving a probability of  $(1 - \alpha) = 95\%$  of detecting different numbers of individuals ( $m$ ) with the trait of interest using the negative binomial distribution with the dilution effect and a probability of  $\lambda = 0.002$  of false negatives and  $\delta = 0.00$  of false positives

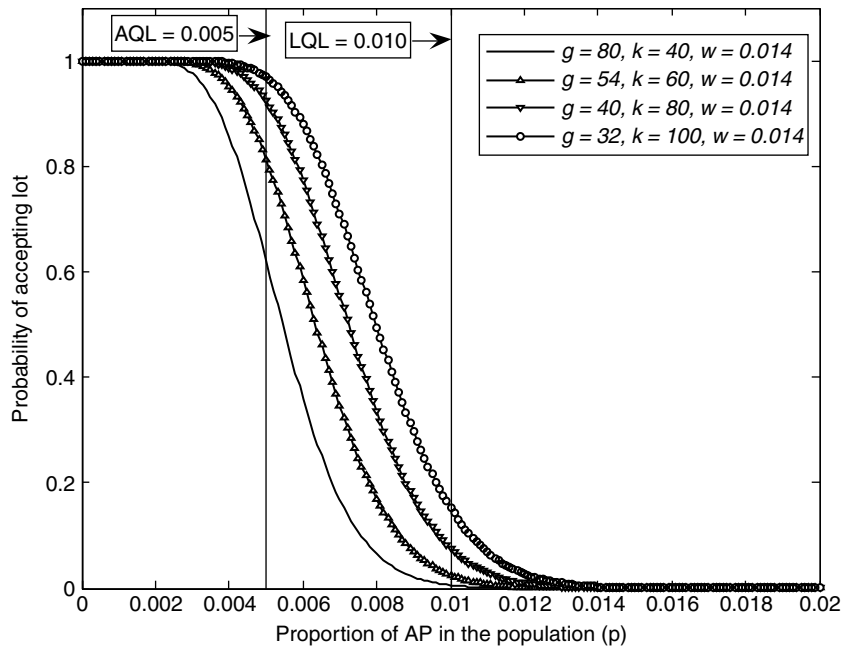
$m$	$w = 0.0002$			$w = 0.0008$			$w = 0.002$			$w = 0.006$			$w = 0.01$		
	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$	$n$	$k$	$g$
$p = 0.01$															
1	7180	1	7180	6553	7	937	3288	19	174	1358	27	51	991	15	67
3	15090	1	15090	13777	7	1969	6936	19	366	2951	27	110	2179	15	146
5	21939	1	21939	20035	7	2863	10109	19	533	4367	27	162	3169	15	212
7	28384	1	28384	25929	7	3705	13092	19	690	5665	27	210	4159	15	278
9	34598	1	34598	31606	7	4516	15961	19	841	6963	27	258	5149	15	344
11	40656	1	40656	37143	7	5307	18773	19	989	8202	27	304	6139	15	410
13	46602	1	46602	42575	7	6083	21528	19	1134	9441	27	350	7030	15	469
15	52460	1	52460	47930	7	6848	24226	19	1276	10621	27	394	7921	15	529
17	58248	1	58248	53215	7	7603	26905	19	1417	11801	27	438	8812	15	588
19	63978	1	63978	58458	7	8352	29565	19	1557	12981	27	481	9703	15	647
21	69659	1	69659	63645	7	9093	32187	19	1695	14161	27	525	10594	15	707
23	75299	1	75299	68804	7	9830	34790	19	1832	15341	27	569	11485	15	766
25	80902	1	80902	73921	7	10561	37393	19	1969	16462	27	610	12376	15	826
27	86473	1	86473	79017	7	11289	39977	19	2105	17642	27	654	13267	15	885
$p = 0.02$															
1	2530	1	2530	1835	7	263	932	19	50	414	11	38	298	9	34
3	5318	1	5318	3872	7	554	1977	19	105	945	11	86	793	9	89
5	7732	1	7732	5636	7	806	2889	19	153	1476	11	135	1189	9	133
7	10003	1	10003	7295	7	1043	3744	19	198	1889	11	172	1585	9	177
9	12193	1	12193	8891	7	1271	4561	19	241	2361	11	215	2080	9	232
11	14329	1	14329	10452	7	1494	5378	19	284	2774	11	253	2476	9	276
13	16424	1	16424	11985	7	1713	6176	19	326	3246	11	296	2872	9	320
15	18489	1	18489	13490	7	1928	6955	19	367	3659	11	333	3169	9	353
17	20529	1	20529	14981	7	2141	7715	19	407	4072	11	371	3565	9	397
19	22548	1	22548	16458	7	2352	8494	19	448	4485	11	408	3961	9	441
21	24551	1	24551	17921	7	2561	9235	19	487	4898	11	446	4357	9	485
23	26539	1	26539	19370	7	2768	9995	19	527	5311	11	483	4753	9	529
25	28514	1	28514	20812	7	2974	10736	19	566	5724	11	521	5149	9	573
27	30477	1	30477	22247	7	3179	11496	19	606	6137	11	558	5446	9	606
$p = 0.03$															
1	1373	1	1373	883	7	127	457	19	25	237	8	30	100	95	2
3	2885	1	2885	1863	7	267	970	19	52	532	8	67	496	7	71
5	4195	1	4195	2717	7	389	1426	19	76	827	8	104	793	7	114
7	5427	1	5427	3515	7	503	1863	19	99	1122	8	141	1090	7	156
9	6616	1	6616	4285	7	613	2281	19	121	1358	8	170	1387	7	199
11	7774	1	7774	5041	7	721	2680	19	142	1653	8	207	1684	7	241
13	8911	1	8911	5776	7	826	3079	19	163	1889	8	237	1882	7	269
15	10032	1	10032	6511	7	931	3478	19	184	2184	8	273	2179	7	312
17	11139	1	11139	7225	7	1033	3858	19	204	2420	8	303	2476	7	354
19	12235	1	12235	7939	7	1135	4257	19	225	2656	8	332	2773	7	397
21	13321	1	13321	8646	7	1236	4637	19	245	2892	8	362	2971	7	425
23	14400	1	14400	9346	7	1336	5017	19	265	3187	8	399	3268	7	467
25	15471	1	15471	10046	7	1436	5397	19	285	3423	8	428	3565	7	510
27	16537	1	16537	10739	7	1535	5758	19	304	3659	8	458	3862	7	552

$c$  (standard testing plan), consumer risks are 0.0868, 0.0514, 0.0281 and 0.0003, whereas producer risks are 0.0419, 0.0692, 0.1102 and 0.5727. The best testing plan is when  $c = 0.00012$ , because it gives the lowest consumer (5.14%) and producer (6.92%) risks. It is interesting to note that, in the standard testing plan,

the consumer risk is 0.03%, and producer risk is 57.27%. Results show the differences among the four testing plans and the important role that the detection limit ( $c$ ) plays for designing balanced testing plans, because different values of  $c$  and  $w$  greatly affect the laboratory testing plans.



**Figure 2.** Coefficient of variation (CV%) of  $p$  under the negative binomial distribution with the dilution effect for different values of  $m$  and  $p$  for obtaining sample sizes with  $c = 0.0001$ ,  $\lambda = 0$ ,  $\delta = 0$ ,  $w = 0.0008$  with 95% probability of detecting individuals with the adventitious presence of transgenic plants (AP).

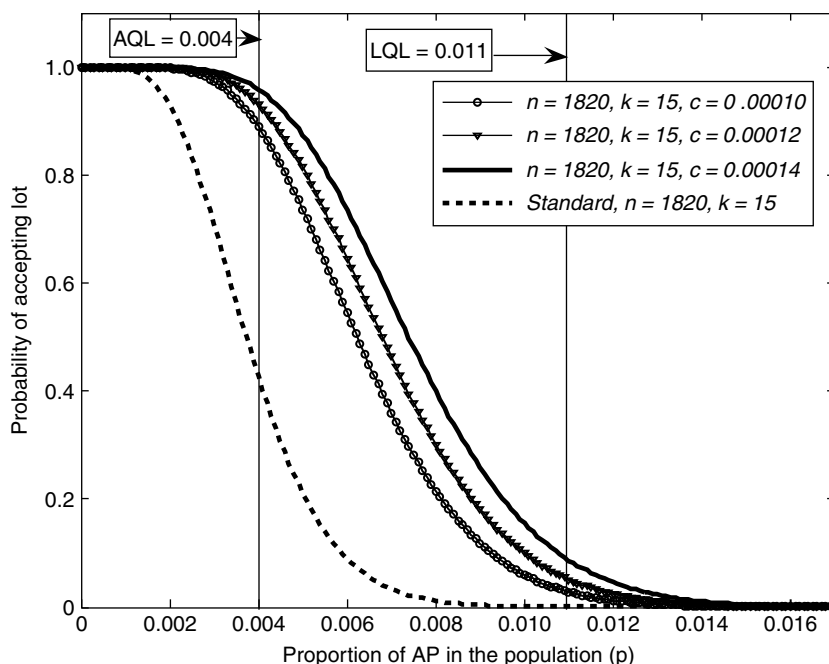


**Figure 3.** Operating characteristic curves for four different seed lot testing plans (with  $n = 3200$ ,  $m = 8$ ,  $\lambda = 0$ ,  $\delta = 0$ ,  $c = 0.0001$  and  $w = 0.014$ ) with consumer risk (lower quality limit, LQL) = 1.0% and producer risk (acceptable quality limit, AQL) = 0.5%.

**Conclusions**

A sample size that guarantees a high probability  $(1 - \alpha)$  that at least one individual with AP will be

detected can be obtained by applying the modified Dorfman method using the binomial distribution and considering the dilution effect and the laboratory detection limit. When using the negative binomial and



**Figure 4.** Operating characteristic curves for four different seed lot testing plans with detection limits of  $c = 0.00010$ ,  $0.00012$ ,  $0.00014$ ,  $m = 6$ ,  $\lambda = 0$ ,  $\delta = 0$  and  $w = 0.0088$  (with  $n = 1820$ ,  $k = 15$ ) and for the standard testing plan with consumer risk (lower quality limit, LQL) = 1.1% and producer risk (acceptable quality limit, AQL) = 0.4%.

performing inverse sampling, a required sample size  $n$  can be determined that guarantees a high probability  $(1 - \alpha)$  that  $m$  individuals or groups containing AP will be found. The method offers a strategy for forming groups (pools) from the sample that will be subjected to laboratory tests, and for determining an optimal number of pools that guarantees that if there is at least one group with AP, there is a high probability that it will be detected. The groups formed have an optimum size, such that one element with AP will be detected at a low cost.

The binomial distribution should be used (modified Dorfman method) when it is known that the proportion of AP in the population is large,  $p > 0.1$ ; otherwise, the inverse sampling method is recommended, because it guarantees that  $m$  individuals or pools of individuals with AP will be detected with high probability. It is important to point out that the precision with which  $p$  is estimated for detecting the AP is related to the value of  $m$ . This research shows that good precision is achieved for estimates of  $p$  with  $m > 11$ , since this leads to a coefficient of variation  $< 25\%$  for any value of  $p \leq 0.1$ . However, the increase in precision is accompanied by a significant increase in the cost of laboratory testing. Larger values of concentration of the AP in the seeds,  $w$ , will require fewer laboratory tests and, therefore, the overall cost of testing will decrease. Performing laboratory tests pool by pool is recommended until AP is detected in

the  $g$ th pool, and the probability of false positives is equal to zero. This is sufficient to conclude that there are individuals with AP in the lot.

The approach used in this study considers the proportion of false negatives, which is never zero in practice. Assuming a low proportion of false negatives, the models proposed in this study facilitate computing a more precise sample size for detecting AP that is larger than that obtained when it is assumed that the proportion of false negatives is zero. Furthermore, for designing sampling testing plans, the authors propose incorporating the dilution effect and considering the rates of false positives, false negatives, the lower quality limit (LQL), and a given acceptable quality limit (AQL). The LQL and AQL give the OC curves and the producer and consumer risks, which facilitate making decisions on important practical matters.

Possible disadvantages of the method are that (1) it does not provide a closed solution for sample size, pool size, and total cost (however, a computer program in MatLab is available); and (2) the value of  $w$  may be difficult to obtain (but an average value given by the % DNA per grain may be used). In the case of genetically modified plants, the weight of the enzyme or other protein formed by that specific DNA sequence may be estimated as a proportion of the total weight of the grain or as a percentage of DNA.

### Program in MatLab

The second author of this paper has developed a program in MatLab for computing the optimal sample size,  $n$ , number of pools,  $g$ , and pool size,  $k$ , for different cases, and for calculating the consumer and producer risks. The software can be downloaded from <http://docente.ucol.mx/oamontes1> and has two windows, sample size (window 1) and OC curves (window 2).

The sample size window computes the optimal sample size ( $n$ ), number of groups ( $g$ ), and size of the groups ( $k$ ), for different cases. The following information must be given: an estimate of the proportion of individuals with AP in the population ( $p$ ); an estimate of AP concentration in the seed ( $w$ ); the laboratory detection limit ( $c$ ); number of pools or individuals with AP that should be detected ( $m$ ); rate of false negatives ( $\lambda$ ); and the value of alpha ( $\alpha$ ). The program will then provide  $n$ ,  $k$ , and  $g$ . Note that when  $m = 1$ , sampling is based on the binomial distribution (since this is a particular case of the inverse sampling method); when  $m > 1$ , sampling is based on the negative binomial. When the rate of false negatives,  $\lambda$ , is unknown, a value of 0 should be given.

Window 2 draws the OC curves and calculates the producer and the consumer risks. The computer software that generates the proposed testing plans requires the following information: the value of the sample size ( $n$ ), the pool size ( $k$ ), the limit of AP seeds ( $m$ ) that will be accepted, the value of LQL (proportion of individuals with AP in the population), the value of AQL; an estimate of AP concentration in the seed ( $w$ ); the laboratory detection limit ( $c$ ); the rate of false negatives ( $\lambda$ ); the rate of false positives ( $\delta$ ); and the value of alpha ( $\alpha$ ). The values of  $\lambda$  and  $\delta$  could be  $\geq 0$ . The program will then generate the OC curves with consumer and producer risks that are useful for evaluating whether or not a given testing plan satisfies the testing objectives. If the value of  $k$  proposed by the consumer and producer is  $\leq \min(n, [w/c] - 1)$ , the program automatically draws the curves and calculates the consumer and producer risks; otherwise, the values used will be between one and  $\min(n, k1 = [w/c] - 1)$  with the aim of guaranteeing that AP concentration ( $w$ ) in the pool is larger than the detection limit ( $c$ ), and that the size of the pool represents a minimum relative cost.

### References

- Bartko, J.J. (1962) A note on the negative binomial distribution. *Technometrics* **4**, 609–610.  
 Cleveland, D.A., Soleri, D., Aragón-Cuevas, F., Crossa, J. and Gepts, P. (2005) Detecting (trans)gene flow to

- landraces in centers of crop origin: lessons from the case of maize in Mexico. *Environmental Biosafety Research* **4**, 197–208.  
 Cochran, W.G. (1980) *Técnicas de Muestreo*. México, Editorial Continental.  
 Dorfman, R. (1943) The detection of defective members of large populations. *Annals of Mathematical Statistics* **14**, 436–440.  
 Federer, W.T. (1991) *Statistics and society. Data collection and interpretation*. New York, Marcel Dekker.  
 Gerrard, D.J. and Cook, R.D. (1972) Inverse binomial sampling as a basis for estimating negative binomial population densities. *Biometrics* **28**, 971–980.  
 Guenther, W.C. (1969) Modified sampling, binomial and hypergeometric cases. *Technometrics* **11**, 639–647.  
 Haldane, J.B.S. (1945) On a method of estimating frequencies. *Biometrika* **33**, 222–225.  
 Kalton, G. and Anderson, D.W. (1986) Sampling rare populations. *Journal of the Royal Statistical Society. Series A (General)* **149**, 65–82.  
 Kay, S. and Paoletti, C. (2002) *Sampling strategies for GMO detection and/or quantification*. European Commission Report, Code EUR20239EN, Joint Research Centre Publication Office. Available online at <http://bgmo.jrc.ec.europa.eu/home/docs.htm#articles2002> (accessed 10 October 2007).  
 Kay, S. and Van den Eede, G. (2001) The limits of GMO detection. *Nature Biotechnology* **19**, 405.  
 Kotz, S., Johnson, N.L. and Read, C.B. (1988) *Encyclopedia of statistical sciences*, Vol. 3 (1st edition). Toronto, Wiley.  
 Laffont, J.-L., Remund, K.M., Wright, D., Simpson, R.D. and Grégoire, S. (2005) Testing for adventitious presence of transgenic material in conventional seed or grain lots using quantitative laboratory methods: statistical procedures and their implementation. *Seed Science Research* **15**, 197–204.  
 Montgomery, D.C. (1997) *Introduction to statistical quality control* (3rd edition). New York, Wiley.  
 Morris, K.W. (1963) A note on direct and inverse binomial sampling. *Biometrika* **50**, 544–545.  
 Ortiz-García, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberón, J. and Snow, A.A. (2005a) Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003–2004). *Proceedings of the National Academy of Sciences, USA* **102**, 12338–12343.  
 Ortiz-García, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberón, J. and Snow, A.A. (2005b) Correction. *Proceedings of the National Academy of Sciences, USA* **102**, 18242.  
 Ortiz-García, S., Ezcurra, E., Schoel, B., Acevedo, F., Soberón, J. and Snow, A.A. (2005c) Reply to Cleveland *et al.*'s 'Detecting (trans)gene flow to landraces in centers of crop origin: lessons from the case of maize in Mexico'. *Environmental Biosafety Research* **4**, 209–215.  
 Patil, G.P. (1960c) On the evaluation of the negative binomial distribution with examples. *Technometrics* **2**, 501–505.  
 Quist, D. and Chapela, I.H. (2001) Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* **414**, 541–543.  
 Quist, D. and Chapela, I.H. (2002) Biodiversity (Communications arising (reply)): suspect evidence of transgenic contamination. Maize transgene results in Mexico are artefacts. *Nature* **416**, 602.

- Remund, K.M., Dixon, D.A., Wright, D.L. and Holden, R.L.**  
(2001) Statistical considerations in seed purity testing for transgenic traits. *Seed Science Research* **11**, 101–119.
- USDA/GIPSA (2000a)** Sampling for the detection of biotech grains. Washington, DC, United States Department of Agriculture. Available at website <http://archive.gipsa.usda.gov/biotech/sample1.htm> (accessed 20 November 2007).
- USDA/GIPSA (2000b)** Practical application of sampling for the detection of biotech grains. Washington, DC, United

States Department of Agriculture. Available at website <http://archive.gipsa.usda.gov/biotech/sample1.htm> (accessed 20 November 2007).

Received 5 December 2006  
accepted after revision 29 February 2008  
© 2008 Cambridge University Press