

## **Copyright Notice**

This electronic reprint is provided by the author(s) to be consulted by fellow scientists. It is not to be used for any purpose other than private study, scholarship, or research.

Further reproduction or distribution of this reprint is restricted by copyright laws. If in doubt about fair use of reprints for research purposes, the user should review the copyright notice contained in the original journal from which this electronic reprint was made.

## Roundtable

# 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”

Sol ORTIZ-GARCÍA<sup>1</sup>, Exequiel EZCURRA<sup>2</sup>, Bernd SCHOEL<sup>3</sup>, Francisca ACEVEDO<sup>4</sup>, Jorge SOBERÓN<sup>5</sup>  
and Allison A. SNOW<sup>6\*</sup>

<sup>1</sup> Instituto Nacional de Ecología, SEMARNAT, Av. Periférico Sur 5000, Colonia Insurgentes Cuicuilco, Delegación Coyoacán, 04530 - México, D.F., Mexico

<sup>2</sup> Biodiversity Research Center of the Californias, San Diego Natural History Museum, 1788 El Prado, San Diego, CA 92101, USA

<sup>3</sup> Genetic ID North America, Inc., Fairfield, IA 52556, USA

<sup>4</sup> Comisión Nacional para el Conocimiento y Uso de la Biodiversidad en México, Avenida Liga Periférico–Insurgentes Sur 4903, Colonia Parques del Pedregal, Delegación Tlalpan, 14010 - México, D.F., México

<sup>5</sup> Biodiversity Institute, University of Kansas, Lawrence, KS, 66047, USA

<sup>6</sup> Department of Evolution, Ecology, and Organismal Biology, Ohio State University, Columbus, OH 43210, USA

Cleveland et al. (2005, *Environ. Biosafety Res.* **4**: 197–208) offer useful suggestions for monitoring transgenes in landraces of maize, but we disagree with their statement that the scientific conclusions of our paper (Ortiz-García et al., 2005, *Proc. Natl. Acad. Sci. USA* **102**: 12338–12343) are not justified. First, contrary to their perception, our survey was not designed to evaluate transgenes in the whole State of Oaxaca, but rather to monitor a specific portion of the District of Ixtlán de Juárez where the presence of transgenes had been reported previously by Quist and Chapela (2001, *Nature* **414**: 541–543). Second, our paper described two methods for estimating frequencies of undetected transgenic seeds, while Cleveland et al. recommend a third approach that explicitly estimates effective population size. They argue that the effective population size of our seed samples is smaller than we assumed, leading to false claims about our detection accuracy. However, we employed a robust statistical approach to compensate for possible bias by using numbers of maternal plants, in addition to numbers of seeds, to provide a conservative estimate of the minimum number of independent samples. When we re-analyzed our 2004 data using effective population sizes, our conclusion that transgenic seeds were “absent or extremely rare” did not change, nor did the general range of possible frequencies of undetected transgenic seeds. Unlike Cleveland et al., we advocate using combined probability tests to analyze data across localities. Third, our critics argue that we accepted the null hypothesis that transgenes were absent. Actually, we assumed that transgenes were present in local landraces, and we used parameter estimation methods to calculate the probability of failing to detect transgenic individuals at a range of frequencies. In agreement with Cleveland et al., we reiterate that there is a clear need for additional surveys with rigorous sampling methods to provide estimates of transgene frequencies over broad geographic areas in Mexico.

**Keywords:** maize / landrace / transgenic / gene flow / introgression / effective population size

Cleveland et al. (2005, *Environ. Biosafety Res.* **4**: 197–208) ofrecen sugerencias útiles para el monitoreo de transgenes en variedades de maíz, pero no concordamos con su argumento que las conclusiones de nuestro artículo (Ortiz-García et al., 2005, *Proc. Natl. Acad. Sci. USA* **102**: 12338–12343) son injustificadas. Primero, nuestro inventario no fue diseñado para evaluar transgenes en todo el Estado de Oaxaca, sino para monitorear una porción específica del Distrito de Ixtlán de Juárez donde la presencia de transgenes había sido reportada previamente por Quist and Chapela (2001, *Nature* **414**: 541–543). Segundo, nuestro estudio utilizó dos métodos

\* Corresponding author: snow.1@osu.edu

para estimar posibles frecuencias transgénicas; Cleveland et al. recomiendan un método para estimar el tamaño efectivo de la población. Argumentan que el tamaño efectivo de nuestra población de semillas es menor que lo supuesto, lo cual lleva a falsas conclusiones sobre la exactitud de nuestra detección. Sin embargo, si utilizamos un enfoque estadístico robusto para compensar un posible sesgo, al emplear, además de las semillas, el número de plantas maternas como un estimador conservador del número de unidades. Al re-analizar nuestros datos de 2004 usando tamaño efectivo, nuestra conclusión básica que las semillas transgénicas estaban “ausentes” o eran “extremadamente raras” no cambió, como tampoco el rango general de posibles frecuencias de semillas transgénicas que pudieran haber evadido la detección. En desacuerdo con Cleveland et al., proponemos el uso de pruebas de probabilidades combinadas para analizar conjuntamente las localidades. Tercero, nuestros críticos argumentan que tomamos como hipótesis nula la ausencia de transgenes. En realidad, asumimos que los transgenes se encontraban presentes en las variedades tradicionales locales, y usamos métodos de estimación de parámetros para calcular la probabilidad de fallar en la detección a distintas frecuencias. En acuerdo con Cleveland et al., reiteramos que hay necesidad de inventarios adicionales con métodos rigurosos de muestreo para proveer estimadores de posible frecuencia de transgénicos a lo largo de grandes regiones de México.

## INTRODUCTION

Cleveland et al. (2005) offer many useful suggestions for monitoring the frequency of transgenes in landraces of maize (*Zea mays* L.). For example, we agree that sampling methods for detecting transgenic seeds should be designed to maximize variance effective population size,  $N_{e(v)}$ , by taking a small number of seeds from each of many, ideally unrelated, plants. Furthermore, efforts to monitor transgene frequencies should include sites that are representative of the entire region that is under consideration. In addition, like Cleveland et al., we are concerned by the fact that some scientists, government agencies, news journalists, and others oversimplified and therefore misinterpreted the conclusions of our paper titled “Absence of detectable transgenes in local landraces in Oaxaca, Mexico (2003, 2004)” (Ortiz-García et al., 2005). However, we object to Cleveland et al.’s statement that the conclusions presented in our paper are “not scientifically justified”. Here, we explain why their key criticisms of our paper are unwarranted.

## REPRESENTATIVENESS OF THE LOCALITIES IN OUR STUDY

Contrary to statements in Cleveland et al., our survey was not designed to monitor the frequencies of transgenes in the state of Oaxaca or the entire region of the Sierra de Juárez. Rather, we focused on a specific portion of the district of Ixtlán de Juárez where Quist and Chapela (2001) reported the presence of transgenes in local landraces grown in 2000, as well as on nearby localities where small farmers were interested in collaborating with us. Quist and Chapela’s study was based on seeds from only six ears of maize, which did not allow them to calculate transgene frequencies. Nonetheless, they concluded that “there is a

high level of gene flow from industrially produced maize towards populations of progenitor landraces.” In a reply to published criticisms of their paper, they stated that further DNA hybridization analysis “confirms our original detection of transgenic DNA integrated into the genomes of local landraces in Oaxaca” (Quist and Chapela, 2002). Subsequent studies in this same area by the Mexican government, including members of our group, also reported that transgenes were present in 2000 and 2001 (Álvarez Morales, 2002; Ezcurra et al., 2002; note: these reports were not peer-reviewed). Therefore, we focused our sampling efforts on the same local area where transgenes had been reported, so we could evaluate the persistence of transgenes and inform local farmers about our findings. Our total sample size was 153 746 seeds from 870 maternal plants and 125 fields at 18 localities in Oaxaca (2003 and 2004). We also reported results from analyses of 19316 seeds from 98 maternal plants and 37 fields in the states of Guerrero and Michoacán in 2002 and 2003 (Ortiz-García et al., 2005; supporting information).

Cleveland et al. argue that our sample “was a very small proportion of fields and locations within the reference metapopulation”, and that “the area sampled was not representative of Oaxaca.” In fact, we never defined a “reference metapopulation”, but we clearly indicated that our study was intended to serve as a systematic survey of transgenic elements in Mexican landraces in the same area of Ixtlán de Juárez that had been examined by Quist and Chapela and three government agencies. By comparing our sampled area with that of the whole State of Oaxaca, Cleveland et al. suggest that our sample was inadequate and does not represent the State, the region, or even the local district. However, we emphasized that “our results should not be extrapolated to

other regions of Mexico, or to previous years in the same region.” Cleveland et al. also argue that the Ixtlán de Juárez District is not the best area for detecting transgene presence in Oaxacan landraces, as transgene flow to landraces is likely to be lower here compared with large, contiguous areas of maize that include some less stressful environments like the central valleys of Oaxaca. These arguments have merit, but we reiterate that the objective of our study, clearly stated in our paper, was not to monitor for transgenes in general, but to provide much-needed data from areas where the presence of transgenic landraces had been reported previously.

### **CORRECTION FOR EFFECTIVE POPULATION SIZE**

The second criticism brought forward by Cleveland et al. is that we did not take into account the fact that the effective population size of our samples could be substantially smaller than the numbers of seeds that were analyzed, leading to false claims about the lowest frequencies at which transgenic seeds could be detected in our study. Ideally, sampling designs for detecting transgenes should account for the fact that seeds collected from the same field often share common parents and ancestors. Farmers’ practices of repeatedly saving seeds from a few maternal plants in each generation to use for the next year’s crop (e.g., Soleri et al., 2000) could lead to genetic bottlenecks and drift (Crossa and Venkovsky, 1994). Although maize outcrosses within and among adjacent fields (e.g., Luna et al., 2001), and despite the fact that farmers periodically exchange or purchase new seeds (e.g., Louette and Smale, 2000; Pressoir and Berthaud, 2004), genetic bottlenecks that continue for several seasons are likely to lead to biparental inbreeding. In this case, rare alleles such as transgenes could become very common or, more likely, become lost from the population due to random selection of a small number of seed parents by farmers.

Therefore, Cleveland et al. note that the census population size of seeds sampled is inappropriate for estimating probabilities of detecting transgenic alleles because this measure of population size  $n$  is likely to be much larger than the true effective population size of the sample,  $N_e$ . Conceptually, we agree completely with this point. In our analyses, we tried to compensate for the fact that seeds from the same maternal plant (ear) are related by using two statistical analyses. In our first analysis, seeds taken from the same ear were considered to be independent samples, assuming that they could have received pollen from many different paternal plants.

Recognizing that this assumption could artificially enlarge our effective population size and introduce bias in our analysis, we did a second analysis considering all seeds from the same ear to be non-independent, given that their paternity was not known, and using maternal plants as the unit of observation. In short, although we did not calculate effective population size as suggested by Cleveland et al., we did use a robust statistical approach to compensate for possible bias in the estimation of ears. As we noted in our paper, although this second approach is likely to be too conservative, it provided a robust estimate of the minimum number of independent observations in the study.

A major contribution of Cleveland et al. (2005) is their suggestion to use a third index of population size, known as the variance effective population size ( $N_{e(v)}$ ), as a measure of the genetic representativeness of a given sample of seeds. To account for relatedness among seeds from the same ear, they recommend using a formula for variance effective population size derived from Hernández and Crossa (1993) and Vencovsky and Crossa (1999):  $N_{e(v)} = n / [((n-1)/4m) + 1]$ , where  $N_{e(v)}$  indicates the effective population size based on allelic variance,  $n$  = census population size (number of seeds), and  $m$  = number of parent plants contributing seeds or plants sampled. We have re-analyzed our data following their suggestion, and we will show that our basic conclusions still hold after correcting for effective population size in this way.

First, the basic binomial equation used by Cleveland et al. predicts, for a single locality  $i$ , the probability  $p_i$  of failing to detect a transgene if present at a given frequency  $f$ :  $p_i = (1 - f)^{N_e}$ , where  $N_e$  is the effective population size. Using our data from 2004, we calculated  $N_{e(v)}$  for each of 16 localities and used binomial probabilities to estimate the minimum detection frequency at each locality with 95% certainty ( $P < 0.05$ , Tab. 1), as in Cleveland et al. 2005. Thus, we are 95% confident that we would have detected transgenic seeds if they were present at a frequency of greater ~1–2% at each locality except one (Santiago Comaltepec), where only 21 putative maternal plants were used for calculations and the detection limit was 3.6% (Tab. 1). However, for this locality, as stated in our paper, 21 was the *minimum* number of maternal plants sampled because in some fields we collected seeds that were already detached from the cob and it is very likely that the true number of maternal plants was much higher than 21.

Cleveland et al. use the results in our Table 1 (and their Tab. 2) to argue that “transgenes could be present in maize landraces at frequencies of ~1–4%” in the fields that were

**Table 1.** Sample sizes,  $N_{e(v)}$ , and statistical tests for the sixteen localities sampled in 2004 in Oaxaca, Mexico (data from Ortiz-García et al., 2005). We used  $N_{e(v)}$  to calculate the minimum detectable frequency ( $f$ ) of transgenic seeds at  $P < 0.05$  for each locality;  $m$  indicates number of maternal plants (= ears) and  $n$  indicates numbers of seeds sampled from each locality. The binomial probability ( $p_i$ ) of failing to detect a transgene in the sixteen localities as a function of different hypothetical transgene frequencies ( $f$ ) was calculated using  $N_{e(v)}$ . The simple combined probability based  $N_{e(v)}$  is shown. The Chi-squared significance values for Fisher's combined probability test using  $N_{e(v)}$  for each locality are indicated by asterisks (\*\*\*:  $\alpha < 0.001$ ; \*\*:  $\alpha < 0.01$ ; ns: not significant). For comparison, the simple combined probabilities based on  $m$  and  $n$ , also calculated in our original paper, are listed in the lower lines.

Locality	$m$	$n$	$N_{e(v)}$	Detectable frequency ( $f$ ) based on $N_{e(v)}$ and $P < 0.05$	Binomial probability ( $p_i$ ) for calculated $N_{e(v)}$ of failing to detect transgenic seeds if present at a frequency ( $f$ ) of:				
					0.1	0.01	0.001	0.0001	0.00001
Ixtlán	52	6100	201	0.015	<0.0001	0.1326	0.8178	0.9801	0.9980
San Andrés Yatuni	35	4450	136	0.022	<0.0001	0.2549	0.8728	0.9865	0.9986
Santiago Comaltepec <sup>‡</sup>	21 <sup>‡</sup>	4320	82	0.036	0.0002	0.4386	0.9212	0.9918	0.9992
San Pablo Macuiltianguis	47	7530	183	0.016	<0.0001	0.1589	0.8327	0.9819	0.9982
San Juan Luvina	52	6340	201	0.015	<0.0001	0.1326	0.8178	0.9801	0.9980
San Juan Bautista Atepec	56	7600	218	0.014	<0.0001	0.1118	0.8040	0.9784	0.9978
Trinidad	41	5240	159	0.019	<0.0001	0.2023	0.8529	0.9842	0.9984
San Francisco La Reforma	80	11260	311	0.010	<0.0001	0.0439	0.7326	0.9694	0.9969
San Juan Ev. Analco	51	7350	199	0.015	<0.0001	0.1353	0.8195	0.9803	0.9980
Santa María Jaltianguis	53	6400	205	0.015	<0.0001	0.1274	0.8146	0.9797	0.9980
Capulalpan de Méndez	51	5900	197	0.015	<0.0001	0.1381	0.8211	0.9805	0.9980
Santiago Xiacui	42	5620	163	0.018	<0.0001	0.1943	0.8495	0.9838	0.9984
Santa María Yahuiche	38 <sup>‡</sup>	5580	148	0.020	<0.0001	0.2259	0.8624	0.9853	0.9985
San Miguel Amatlán	38	4450	147	0.020	<0.0001	0.2282	0.8632	0.9854	0.9985
San Juan Chicomezuchitl	51	7600	199	0.015	<0.0001	0.1353	0.8195	0.9803	0.9980
Santa Catarina Lachatao	57	7500	221	0.013	<0.0001	0.1085	0.8016	0.9781	0.9978
Combined probability for $N_{e(v)}$ ( $P$ )					<0.0001	<0.0001	0.0512	0.7430	0.9707
Chi-squared significance (Fisher's combined probability test)					***	**	ns	ns	ns
Combined probability for seeds ( $P_{(n)}$ )					<0.0001	<0.0001	<0.0001	<0.0001	0.3561
Combined probability for ears ( $P_{(m)}$ )					<0.0001	0.0005	0.4652	0.9263	0.9924

<sup>‡</sup> These are the minimum number of maternal plants sampled. Farmers had already harvested and separated the kernels from the cobs in some of the fields sampled in these localities.

**Table 2.** Minimum frequencies ( $f$ ) at which transgenic seeds can be detected with 95% confidence in the whole set of localities for the 2004 data, using the same binomial model as Cleveland et al. in their Table 2 and combining probabilities using Fisher's method (see text). The minimum detectable frequency ( $f$ ) was calculated following two alternative procedures: (a) finding a value of  $f$  that would yield a combined probability ( $P$ ) of 0.05, or (b) finding a value of  $f$  that would yield a  $\chi^2$  value significant at the 5% level ( $\alpha = 0.05$  for the log-likelihood test). Note that even in the conservative case of correcting for effective population size and using Fisher's log-likelihood test to account for possible non-independence in the locality-level samples, our detection precision is lower than 1% (0.0775%).

Unit of analysis	Minimum detectable frequency ( $f$ )	
	a. Using simple combined probability	b. Using Fisher's log-likelihood criterion
Seeds ( $n$ )	0.00003	0.00022
Effective population ( $N_{e(v)}$ )	0.00101	0.00775
Maternal plants ( $m$ )	0.00391	0.02974

sampled in 2004 in Oaxaca. We assume that they are referring to frequencies of “transgenic seeds” rather than “transgenes” in this context. This interpretation of our data is based on the fact that they chose not to analyze our data set as a whole. However, we argue that it is useful to analyze the entire data set to calculate the likelihood of repeatedly failing to detect transgenic seeds at frequencies lower than 1% (or other thresholds) across the 16 localities that were sampled in 2004, as discussed below.

Cleveland et al. state that it is inappropriate to combine data from multiple localities because the unknown genetic structure of metapopulations of maize makes it impossible to accurately estimate  $N_{e(v)}$  across localities. According to Vencovsky and Crossa (2003: 1915) “it is impossible to measure  $N_e$  appropriately if the structure of the metapopulation is unknown or, in other words, if the number of the component subpopulations ( $S^*$ ) and the interpopulation diversity ( $F_{ST}$ ) under real conditions is unknown.” Therefore, Cleveland et al. conclude that “ $N_e$  must be calculated for each population within the metapopulation and these  $N_e$  cannot be summed for a metapopulation estimate.” Following this line of reasoning, Cleveland et al. suggest that data from each locality should be considered separately. We suggest that this conclusion is too conservative for the following reasons.

First, Cleveland et al. argue that population numbers cannot be summed for a metapopulation estimate, and that hence an overall probability cannot be estimated across all localities. What we did in our original paper, however, was not to add the population numbers of each locality to get an overall estimate. Rather, we multiplied the probabilities from each locality to get an overall probability of repeatedly failing to detect transgenic seeds at frequencies lower than a given threshold. We did this based on the assumption that the probability calculated for each locality was methodologically independent of other localities, and all were part of the same experimental design. Second, if we accept their argument that the use of a joint statistical analysis is not appropriate for separate estimates of  $N_{e(v)}$ , other statistical methods – known as “meta-analyses” – are available for combining the results of many separate tests of the same research question (e.g., DuMouchel, 1990; Eddy et al., 1992; Schulze, 2004; Sokal and Rohlf, 1995). Among these methods, one of the most commonly used is Fisher’s combined probability test (Fisher, 1932; Folks, 1984; Rice, 1990), which is used to estimate the significance of an overall likelihood when the separate tests to be combined are all testing the same hypothesis but a joint statistic is not considered to be appropriate (Sokal and Rohlf, 1995). In this reply we present the

results from this test to address the concern over the validity of calculating a joint probability based on  $N_{e(v)}$  for all localities.

We used two statistical methods – the simple combined probability test in which localities are assumed to represent independent subsamples of the same sample, as in Ortiz-García et al., 2005, and the more conservative Fisher’s combined probability test, in which they are assumed to derive from different independent tests (Tabs. 1 and 2). For a series of  $l$  localities, the overall combined likelihood  $P$  of repeatedly failing to detect a transgene that is regionally present at a frequency,  $f$ , can be estimated as the product of all the different within-locality probabilities:  $P = \prod_i p_i$ . Assuming that the individual probabilities come from independent trials within the same sampling design, this overall likelihood can be used as an estimate of the combined probability that a stated hypothesis is true. Second, we used Fisher’s combined probability test, which is derived from the log-transform of the overall likelihood ( $L = -2 \ln(P)$ ), and, like all log-likelihood estimates, is distributed as  $\chi^2$  with degrees of freedom equal to two times the number of localities (Sokal and Rohlf, 1995).

For illustration, let us now follow the calculations proposed by Cleveland et al. in their Table 2. Following their analysis, we have applied their calculation based on  $N_{e(v)}$  to our 2004 data for different hypothetical values of  $f$  and for all localities (Tab. 1). We then calculated the simple combined probability ( $P$ ) for all localities. These combined probabilities show that if transgenic seeds had been present in the region at frequencies greater than 1 in a thousand ( $f > 0.001$ ), they would have been detected with a ~5% combined probability of failure (Tab. 1). However, if we assume non-independence of localities and apply the Fisher’s log-likelihood test, the minimum detection level for transgenic seeds across localities is closer to 1 in one hundred (0.0077, Tab. 2). We do not have sufficient knowledge of the genetic relatedness of samples from different localities in our study area to judge whether this more conservative test is warranted.

To summarize, the estimated detection probabilities obtained using effective population size show a higher probability of failing to detect transgenes than the values we obtained taking individual seeds as our unit of observation, but a lower probability of failure than the thresholds we reported for maternal plants, our most conservative estimate (Tab. 2). The values obtained using Fisher’s combined probability test are in agreement with the conclusion in our original paper stating that “the maximum frequency of transgenic seeds in our study should fall somewhere between these seed-based and

plant-based estimates". However, in our original paper, we also concluded that "this value is likely to be closer to 0.0001 (0.01%) because we expect that many seeds from the same cob were sired by different paternal plants." When we re-analyzed our data using  $N_e$ , our estimate of the minimum detectable frequency was 0.001 using the simple combined probability test, but the value from the more conservative Fisher's combined probability test was 0.0077 (Tab. 2). Other tests for combined probabilities could be applied (*e.g.*, Olkin, 1995; Whitlock, 2005), but it is doubtful that our conclusions would change. Furthermore, we expect that similar results would be obtained using Bayesian methods to analyze frequency data from multiple years and sites (D. Andow, personal communication to A. Snow).

Referring to our data from 2004 in Oaxaca, Cleveland et al. state that "because Ortiz-García et al. sampled a large number of seeds ( $n = 103\,620$ ) from a small number of ears (765), the  $N_e$  for seeds (2972) for their sample is much smaller than  $n$ ." We agree with them; so much so, that for that same reason we did the second analysis using maternal plants (ears) as the unit of observation rather than seeds (this part of our paper is not acknowledged by Cleveland et al.). Even using this statistically conservative approach, we estimated in our original paper with 95% certainty that the frequency of ears with at least one transgenic seed was lower than 0.004 in 2004 and less than 0.003 for both 2003 and 2004 combined. We are grateful to Cleveland et al. for highlighting the utility of using effective population size in statistical calculations, and for leading us to understand better the techniques of meta-analysis (including Fisher's combined probability test), but these corrections do not change our basic conclusions.

Finally, it goes without saying that even if the frequency of transgenic maize plants is 1 in 10000 or less, thousands if not millions of transgenic plants could be present in the region without being detected (Cleveland et al., 2005). This is a self-evident statistical truth that was not disputed nor discussed in our paper. Because it is impossible to prove that transgenes are absent in a given region, discussions about the *consequences* of undetected transgenic plants should acknowledge that even extremely low frequencies could result in biological and/or socioeconomic effects, depending on the transgenes in question and how they are viewed by local farmers.

## NULL HYPOTHESES

Lastly, Cleveland et al. contend that we accepted "the null hypothesis of absence of transgenes at detectable frequencies" instead of "accepting the alternative

hypothesis of presence of transgenes, as Quist and Chapela did," leading to a greater likelihood of Type II error in our study. This argument is confusing, because our study involved parameter estimation rather than hypothesis testing. While there is a clear mathematical relationship between these two methods, it is important to clarify the purpose of a study and focus the statistical analyses on either hypothesis testing or parameter estimation. Our goal was to estimate the frequency of transgenes (which we assumed were present) in populations of maize plants and, finding none, to estimate minimum detectable frequencies of transgenic constructs based on our sample sizes.

Thus, although we used the phrase "absence of detectable transgenes" in the paper, this was not our null hypothesis. As explained, we assumed that transgenic constructs were present in the area, and we calculated the frequency ( $f$ ) at which we would have included at least one transgenic individual in our sample with 95% certainty. After failing to detect them, we did allow for the possibility that transgenes could still be present in the region, although at very low frequencies and below the detection capacity of our 2003–2004 sampling design. In a carefully crafted phrase, we concluded that "evidence that transgenes are rare or absent in the sampled area should not be extrapolated to other regions of Mexico without quantitative data, nor is the current situation likely to remain static." We stand by our conclusion.

## FINAL COMMENTS

In reality, the results we present in our study are quite clear-cut, although the interpretation may be complex. We referred to our study as a "preliminary baseline" with several expectations in mind. First, we hope that many more surveys of this type will be published in the near future, making it possible to generalize about the prevalence of transgenes in landraces of maize over a much broader geographic area. Second, we expect that other investigators will continue to suggest improved methods for this type of monitoring. Scientific progress is built upon critical analyses of previous work, and Cleveland et al. provide many useful recommendations for further studies. However, it may not be practical for investigators to follow all of their suggested guidelines, and even the most comprehensive studies may fail to detect the presence of very rare alleles. We expect that the sampling methods, laboratory analyses, and statistical approaches for monitoring transgene frequencies will improve over time as we continue to study traditional landraces. Meanwhile, we encourage more investigators

to conduct similar surveys and publish their findings in peer-reviewed journals so these results can be evaluated and interpreted by others. Ultimately, more field data and further monitoring will help to understand the movement of transgenes in non-target organisms and unravel some of the remaining questions about transgene persistence in the Ixtlán maize crops.

## ACKNOWLEDGEMENTS

We thank Lesley Campbell, David Andow, Thomas Waite, Kristin Mercer, Daniel Piñero, David Gernandt, and Michael Reagon for helpful comments and discussion.

Received March 13, 2006; accepted March 22, 2006.

## REFERENCES

- Álvarez-Morales A (2002) Transgenes in maize landraces in Oaxaca: Official report on the extent and implications. Available from Proceedings of the 7th International Symposium on the Biosafety of Genetically Modified Organisms Beijing, October 10-16, 2002, p 65. <http://www.bba.de/gentech/isbgmo.pdf> (verified 20 Mar 2006)
- Cleveland DA, Soleri D, Aragón Cuevas F, Crossa J, Gepts P (2005) Detecting (trans)gene flow to landraces in centers of crop origin: lessons from the case of maize in Mexico. *Environ. Biosafety Res.* **4**: 197–208
- Crossa J, Vencovsky R (1994) Implication of the variance in effective population size on the genetic conservation of monoecious species. *Theor. Appl. Genet.* **89**: 936–942
- DuMouchel WH (1990) Bayesian meta-analysis. In Berry DA, ed, *Statistical methodology in the pharmaceutical sciences*. Dekker, New York pp 509–529
- Eddy DM, Hasselblad V, Shachter R, eds (1992) *Meta-Analysis by the Confidence Profile Method*, Academic Press, New York
- Ezcurra E, Ortiz-García S, Soberón J (2002) Evidence of gene flow from transgenic maize to local varieties in Mexico. In *LMOs and the Environment: Proceedings of an International Conference*, Roseland CR ed, OECD, pp 289–295. [http://www.oecd.org/document/18/0,2340,en\\_2649\\_34387\\_2509330\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/18/0,2340,en_2649_34387_2509330_1_1_1_1,00.html) (verified 20 Mar 2006)
- Hernández SCM, Crossa J (1993) A program to estimate the optimum sample size for germplasm conservation. *J. Hered.* **84**: 85–86
- Fisher RA (1932) *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh
- Folks JL (1984) Combination of Independent Tests. In Krishnaiah PK, Sen PK, eds, *Handbook of Statistics, 4, Nonparametric Methods*, North-Holland, New York
- Louette D, Smale M (2000) Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. *Euphytica* **113**: 25–41
- Luna VS, Figueroa MJ, Baltazar MB, Gomez LR, Townsend R, Schoper JB (2001) Maize pollen longevity and distance isolation requirements for effective pollen control. *Crop Sci.* **41**: 1551–1557
- Olkin I (1995) Meta-analysis: reconciling the results of independent studies. *Stat. Med.* **14**: 457–472
- Ortiz-García S, Ezcurra R, Schoel B, Acevedo F, Soberón J, Snow AA (2005) Absence of detectable transgenes in local landraces of maize in Oaxaca, Mexico (2003-2004). *Proc. Natl. Acad. Sci. USA* **102**: 12338–12343
- Pressoir G, Berthaud J (2004) Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity* **92**: 88–94
- Quist D, Chapela IH (2001) Transgenic DNA introgressed into traditional maize landraces in Oaxaca, Mexico. *Nature* **414**: 541–543
- Quist D, Chapela IH (2002) Quist and Chapela reply. *Nature* **416**: 602
- Rice WR (1990) A consensus combined *P*-value test and the family-wide significance of component tests. *Biometrics* **46**: 303–308
- Schulze R (2004) *Meta-Analysis: A Comparison of Approaches*, Hogrefe and Huber Publishing, New York
- Sokal RR, Rohlf FJ (1995) *Biometry*, W.H. Freeman, New York
- Soleri D, Smith SE, Cleveland DA (2000) Evaluating the potential for farmer and plant breeder collaboration: a case study of farmer maize selection in Oaxaca, Mexico. *Euphytica* **116**: 41–57
- Vencovsky R, Crossa J (1999) Variance effective population size under mixed self and random mating with applications to genetic conservation of species. *Crop Sci.* **39**: 1282–1294
- Vencovsky R, Crossa J (2003) Measurements of representativeness used in genetic resources conservation and plant breeding. *Crop Sci.* **43**: 1912–1921
- Whitlock MC (2005) Combining probability from independent tests: the weighted *Z*-method is superior to Fisher's approach. *J. Evolution. Biol.* **18**: 1368