

EROSIÓN GENÉTICA EN LAS CRUZAS SIMPLES PROGENITORAS DE UNA VARIEDAD SINTÉTICA

José Luis Escalante-González; Jaime Sahagún-Castellanos*;
Juan Enrique Rodríguez-Pérez; Aureliano Peña-Lomelí

Universidad Autónoma Chapingo. Departamento de Fitotecnia
km 38.5 Carretera México-Texcoco. Chapingo, Estado de México. MÉXICO. C.P. 56230
Correo-e: jsahagunc@yahoo.com.mx Tel.: 9521500 Ext. 6185 (*Autor para correspondencia)

RESUMEN

El azar, el número finito de individuos y la variabilidad genética entre las plantas que representan cada progenitor de un sintético de cultivos como la cebolla (*Allium cepa* L.) y el maíz (*Zea mays* L.) pueden ocasionar erosión genética. Esto puede ocurrir en el desarrollo de una variedad sintética cuyos progenitores son $L/2$ cruza simples (CSs) formadas con L líneas. Si el coeficiente de endogamia de éstas (F_L) es menor a uno, y el número de plantas de cada CS (m) es pequeño, el número de genes no idénticos por descendencia (NIPD) de cada CS se convierte en una variable aleatoria (Xm). Los objetivos de este trabajo fueron determinar: 1) la media de Xm [$E(Xm)$], 2) la varianza de Xm [$Var(Xm)$], y 3) la pérdida promedio de genes NIPD que ocurre en la transmisión de genes de dos líneas a su CS. Se encontró que $E(Xm) = 2 + [2 - (1/2)^{m-2}](1 - F_L)$. Asimismo, se encontró que la pérdida promedio de estos genes es $(1/2)^{m-2}(1 - F_L)$ y que $Var(Xm) = [2^{m+2} - 8](1 - F_L)^2/4^m$. Estos dos resultados implican que con líneas puras ($F_L = 1$) la pérdida y la inestabilidad del número de genes NIPD se reducen a su mínimo (cero), y se aproximan a éste a partir de $m = 8$, más rápidamente cuando F_L es más grande.

PALABRAS CLAVE ADICIONALES: *Allium cepa* L., *Zea mays* L., variedad sintética, genes no idénticos por descendencia, apareamiento aleatorio.

GENETIC EROSION IN SINGLE-CROSS PROGENITORS OF A SYNTHETIC VARIETY

ABSTRACT

Randomness, finite number of individuals and genetic variability among those that represent each parent of a synthetic of crop species such as onion (*Allium cepa* L.) and maize (*Zea mays* L.) can cause genetic erosion during the development of a synthetic variety. This may happen when the parents are $L/2$ single crosses (SCs) derived from L lines. If the inbreeding coefficient of the lines (F_L) is less than 1, and the number (m) of plants that represent each SC is small, the number of non-identical by descent (NIBD) genes of each SC is a random variable (Xm). The objectives of this study were to determine: 1) the mean of Xm [$E(Xm)$], 2) the variance of Xm [$Var(Xm)$], and 3) the average loss of NIBD genes which are contributed by two lines to their SC. It was found that $E(Xm) = 2 + [2 - (1/2)^{m-2}](1 - F_L)$. In addition, the average loss of NIBD genes was $(1/2)^{m-2}(1 - F_L)$ and $Var(Xm) = [2^{m+2} - 8](1 - F_L)^2/4^m$. These two results imply that when lines are pure ($F_L=1$), the loss of NIBD genes and the instability of the number of these genes reach their minimum (zero), and approach to it from $m = 8$ on, more rapidly as F_L becomes larger.

ADDITIONAL KEYWORDS: *Allium cepa* L., *Zea mays* L., synthetic variety, non-identical by descent genes, random mating.

INTRODUCCIÓN

Las variedades sintéticas (VSs) de especies cultivadas como la cebolla (*Allium cepa* L.), el maíz (*Zea mays* L.), etc., tienen como origen el apareamiento aleatorio de varios progenitores, usualmente entre 6 y 12 (Márquez-Sánchez, 1992; Kutka y Smith, 2007; Sahagún, 2011). Si el coeficiente de endogamia de las líneas es uno y cada una se representa por m plantas, de acuerdo con el modelo de un locus de una especie diploide, los $2m$ genes de las m plantas de cada línea deben ser idénticos por descendencia (Falconer y Mackay, 2007). Sin embargo, si las líneas no son puras, cada una podrá contener genes que entre sí son idénticos por descendencia, así como genes que no lo son.

Por otra parte, la formación de VSs tiene variantes que no han sido estudiadas en forma completa. En particular, varios autores han propuesto su formación mediante el apareamiento aleatorio de varios híbridos (Villanueva *et al.*, 1994; Sahagún-Castellanos *et al.*, 2005; Márquez-Sánchez, 2008); sin embargo, este procedimiento todavía tiene temas interesantes inéditos. Al respecto, una característica que se supone inseparable de una variedad sintética (VS) es la estabilidad de su arreglo genotípico. De acuerdo con este supuesto, la VS debe reproducirse a sí misma año tras año, lo que significa que el agricultor puede producir la semilla de su VS en su propio campo de producción. Desafortunadamente, hay casos en que la ocurrencia de esta característica puede cuestionarse aun cuando la liberación de la variedad no ha ocurrido. Por ejemplo, supóngase el caso de una VS generada por el apareamiento aleatorio de $L/2$ cruza simples (CSs), formadas con L líneas (Sahagún-Castellanos y Villanueva, 1997); cada gameto que emita una planta que representa una CS para formar la VS (Sin_{CS}) puede portar un gen que es copia del gen que le transmitió una u otra de sus dos líneas progenitoras. Si la progenie de esta CS es pequeña, el número de genes no idénticos por descendencia (NIPD) que aporta al Sin_{CS} por azar puede ser menor que lo esperado en términos de las frecuencias teóricas. Más aún, sólo si las líneas son puras, la aportación de cada una de las dos a su cruce simple (CS) será de m genes idénticos por descendencia. Cuando las líneas no son puras, pueden aportar un número mayor que uno de genes NIPD, lo que a su vez generaría variabilidad aleatoria que puede afectar la estabilidad genética de la variedad.

Una medida de la estabilidad y heterogeneidad genética de una VS puede expresarse en términos del número de genes NIPD que contiene. Como ya se mencionó, este número, puede comportarse como una variable aleatoria de un ciclo a otro; y la magnitud de la variabilidad que muestre debe reflejar el grado de incertidumbre del proceso de formación de la VS en términos de los genes que debe contener según la constitución genética de las líneas progenitoras iniciales. Evidentemente, la inestabilidad en

INTRODUCTION

Synthetic varieties (SVs) of crop species as onion (*Allium cepa* L.) and maize (*Zea mays* L.), etc., derive from random mating of several parents, usually between 6 and 12 (Márquez-Sánchez, 1992; Kutka and Smith, 2007; Sahagún, 2011). If the inbreeding coefficient of lines is one and each is represented by m plants, according to the model of a locus of a diploid species, the $2m$ genes of m plants of each line must be identical by descent (Falconer and Mackay, 2007). However, if the lines are not pure, each one may contain genes that are identical by descent to each other, and non-identical by descent genes.

Moreover, the derivation of SVs has variants that have not been fully studied. Particularly, various authors have proposed the derivation by means of random mating of several hybrids (Villanueva *et al.*, 1994; Sahagún-Castellanos *et al.*, 2005; Márquez-Sánchez, 2008); nonetheless, this procedure still has unpublished interesting topics. In this regard, a feature that is supposed to be inseparable from a synthetic variety (SV) is the stability of the genotypic array. According to this scenario, the SV must reproduce itself, cycle after cycle; this means that the farmer can produce the seeds of the SV in his own production field. Unfortunately, there are cases in which the occurrence of this feature can be questioned even if the release of the variety has not happened. For example, assume the case of a SV derived by random mating of $L/2$ single crosses (SCs) derived from L lines (Sahagún-Castellanos and Villanueva, 1997); each gamete emitted by a plant representing a SC to form the SV (Sin_{CS}) carries a gene, which is a copy of a gene received from one of the two parental lines. If the progeny of this SC is small, the number of non-identical by descent (NIBD) genes provided randomly to the Sin_{CS} can be less than that expected in terms of the theoretical frequencies. Moreover, only if the lines are pure, the contribution of each one to the single cross (SC) is m identical by descent genes. When the lines are not pure, they can provide a greater number than one of NIBD genes, which in turn would generate random variability that can affect the genetic stability of the variety.

A measure of genetic heterogeneity and stability of SV can be expressed in terms of the number of NIBD genes contained. As already mentioned, this number can behave as a random variable from one cycle to another, and the magnitude of the variability showed should reflect the degree of uncertainty of the process of derivation of the SV in terms of the genes that should contain according to the genetic constitution of the initial parental lines. Instability in this case means gene loss. This is of particular importance, first to the breeder and secondly to the farmers.

The instability in the number of NIBP genes affects the genotypic composition, inbreeding coefficient, and

este caso significa pérdida de genes. Esto es de importancia singular para el fitomejorador en primera instancia, y para los agricultores posteriormente.

La inestabilidad en el número de genes NIPD repercute en la composición genotípica, el coeficiente de endogamia, y en las medias genotípica y fenotípica de la VS. Generar información relativa a la magnitud de la inestabilidad y de la pérdida de los genes NIPD que ocurren en el proceso de desarrollo de una VS debe contribuir al mejor diseño de las VS. El objetivo de este estudio fue determinar la media, la varianza y la reducción del número de genes no idénticos por descendencia en las cruza simples que generarán un sintético.

MÉTODOS Y MARCO TEÓRICO

En este estudio se utilizaron dos conceptos para una población en que el gen A_i y el genotipo $A_i A_j$ tienen frecuencias p_i y p_{ij} , respectivamente; los arreglos gamético (AGA) y genotípico (AGE) se definen en la forma:

$$AGA = \sum_i p_i A_i \text{ y } AGE = \sum_i \sum_j p_{ij} A_i A_j$$

Además, si los arreglos gaméticos de las poblaciones P_1 y P_2 son $\sum_i p_i A_i$ y $\sum_i q_i A_i$, la cruce de estas dos poblaciones produce una progenie cuyo arreglo genotípico $[AGE(P_1 \times P_2)]$ es el producto de sus arreglos gaméticos; es decir,

$$AGE(P_1 \times P_2) = \sum_i \sum_j p_i q_j A_i A_j$$

En particular, si como sucede en los sintéticos, la reproducción de una población que tiene el arreglo gamético $\sum_i p_i A_i$ es por apareamiento aleatorio, el arreglo genotípico de su progenie es (Sahagún-Castellanos, 1994):

$$\sum_i \sum_j p_i p_j A_i A_j$$

Por otra parte, respecto a las L líneas iniciales, para la formación de una variedad sintética se consideró que no tenían parentesco, y la línea i ($i = 1, 2, \dots, L$) se visualizó como una población formada por individuos que se clasificaron en dos grupos: los portadores de genotipos formados por dos genes idénticos por descendencia ($A_i A_i$) y los formados por dos genes no idénticos por descendencia ($A_i A_{ci}$, $i \neq ci$) con frecuencias F_L y $1-F_L$, respectivamente. Además, se consideró que los genes no idénticos por descendencia de una línea son: 2, cuando $F_L < 1$; 1, cuando $F_L = 1$; y, en general, si $0 \leq F_L \leq 1$ el número de genes no idénticos por individuo en promedio debe ser $1(F_L) + 2(1-F_L) = 2-F_L$.

genotypic and phenotypic means of a SV. To generate information concerning the size of the instability and loss of NIBP genes that occur in the process of developing a SV should contribute to make a better design of a SV. The objective of this study was to determine the mean, variance, and the reduction of the number of non-identical by descent genes in single-cross progenitors of a synthetic variety.

METHODS AND THEORETICAL FRAMEWORK

In this study, we used two concepts for a population in which the gene A_i and the genotype $A_i A_j$ have frequencies p_i and p_{ij} , respectively; the gametic (GAA) and genotypic (GEA) arrays are defined as follows:

$$GAA = \sum_i p_i A_i \text{ y } GEA = \sum_i \sum_j p_{ij} A_i A_j$$

Moreover, if the gametic arrays of the populations P_1 and P_2 are $\sum_i p_i A_i$ and $\sum_i q_i A_i$, the cross of these two populations produce a progeny whose genotypic array $[AGE(P_1 \times P_2)]$ is the product of its gametic arrays; i.e.,

$$AGE(P_1 \times P_2) = \sum_i \sum_j p_i q_j A_i A_j$$

As in the case of synthetics, if the reproduction of a population with the gametic array $\sum_i p_i A_i$ is by random mating, the genotypic array of its progeny is (Sahagún-Castellanos, 1994):

$$\sum_i \sum_j p_i p_j A_i A_j$$

Moreover, with respect to the L initial lines used to form a synthetic variety, they were considered unrelated, and the line i ($i = 1, 2, \dots, L$) was seen as a population conformed by individuals classified into two groups: carriers of genotypes composed of two identical by descent genes ($A_i A_i$) and those formed by two non-identical by descent genes ($A_i A_{ci}$, $i \neq ci$) with frequencies F_L and $1-F_L$, respectively. In addition, it was considered that non-identical by descent genes of one line are: 2, when $F_L < 1$; 1, when $F_L = 1$; and, in general, if $0 \leq F_L \leq 1$ the number of non-identical by descent genes per individual on average should be $1(F_L) + 2(1-F_L) = 2-F_L$.

DERIVATION OF RESULTS

Number of non-identical by descent genes

If $B_{ip1} B_{ip2}$ is the genotype of the p -th individual ($p = 1, 2, \dots, m$) of the line i ($i = 1, 2, \dots, L$; L is an even number) the frequency of the gamete B_{ipk} ($k = 1, 2$) is $(2mL)^{-1}$ and

DERIVACIÓN DE RESULTADOS

Número de genes no idénticos por descendencia

Si $B_{ip1} B_{ip2}$ es el genotipo del p -ésimo individuo ($p = 1, 2, \dots, m$) de la línea i ($i = 1, 2, \dots, L$; L es un número par) la frecuencia del gameto B_{ipk} ($k = 1, 2$) es $(2mL)^{-1}$ y el arreglo genotípico del sintético formado por el apareamiento aleatorio de los mL individuos ($AGES_L$) es:

$$AGES_L = (2mL)^{-2} \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^2 \sum_{l=1}^2 B_{ipk} B_{jqk} \quad (1)$$

Para investigar qué es $AGES_L$ (Ecuación 1) en términos del sintético que genera el apareamiento aleatorio de $L/2$ cruza simples formadas con la participación de cada una de las L líneas, se consideró que la cruce simple h ($h = 2, 4, 6, \dots, L$), representada por m plantas, está formada por las líneas h y $h-1$. El arreglo genotípico de la población que producen las $L/2$ cruza ($AGES_{CS}$) es de la forma:

$$AGES_{CS} = [2m(L/2)]^{-2} \sum_{p=1}^m \sum_{q=1}^m \sum_{h=2}^L \sum_{h'=2}^L \sum_{k=1}^2 \sum_{l=1}^2 A_{hpk} A_{h'ql} \quad (2)$$

$h, h' = 2, 4, 6, \dots, L$

Evidentemente, si las L líneas son puras ($F_L = 1$) no habrá pérdida de genes no idénticos por descendencia (NIPD) en la formación de los m genotipos que resultarán del apareamiento de las dos líneas de cada una de las $L/2$ cruza simples. Sin embargo, esto puede ser diferente cuando las L líneas no son puras ($F_L < 1$) y m es pequeño. Por ejemplo, si $m = 1$ y $F_L < 1$, como cada línea inicial, en promedio, es portadora de $2 - F_L$ genes NIPD, las 2 líneas de una cruce simple portan en promedio $4 - 2F_L$ genes NIPD, y a una planta ($m = 1$) que produzca su cruce le aportarán exactamente 2 de estos genes porque las líneas no están emparentadas. Esto implica una pérdida promedio de $2 - 2F_L$ genes NIPD por cada uno de los $L/2$ individuos que representan sendas cruza simples.

Para $m = 2$, supóngase que los genotipos de las 2 líneas progenitoras de una cruce simple se representan como $A_1 A_2$ y $B_1 B_2$. Las probabilidades (P) de que los genes de cada genotipo sean idénticos por descendencia (\equiv) son $P(A_1 \equiv A_2) = P(B_1 \equiv B_2) = F_L$ y de que no lo sean (\neq) es $P(A_1 \neq A_2) = P(B_1 \neq B_2) = 1 - F_L$. Además, el arreglo genotípico de la cruce entre las dos líneas ($AGEC_L$) es:

$$AGEC_L = (1/4)A_1 B_1 + (1/4)A_1 B_2 + (1/4)A_2 B_1 + (1/4)A_2 B_2$$

Las dos plantas de la cruce que participarán en el Sin_{CS} de este ejemplo ($m = 2$) se pueden visualizar como el resultado de un muestreo aleatorio de tamaño 2 con reemplazo del conjunto de los 4 genotipos diferentes del $AGEC_L$. Los tipos de resultados posibles son:

the genotypic array of the synthetic variety derived from the random mating of the mL individuals ($AGES_L$) is:

$$AGES_L = (2mL)^{-2} \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^2 \sum_{l=1}^2 B_{ipk} B_{jqk} \quad (1)$$

To know what $AGES_L$ (Equation 1) means in terms of the synthetic variety derived by randomly mating the $L/2$ single crosses formed with the participation of each one of the L lines, it was considered that the single cross h ($h = 2, 4, 6, \dots, L$), represented by m plants, is formed by lines h and $h-1$. The genotypic array of the population produced by $L/2$ crosses ($AGES_{CS}$) is:

$$AGES_{CS} = [2m(L/2)]^{-2} \sum_{p=1}^m \sum_{q=1}^m \sum_{h=2}^L \sum_{h'=2}^L \sum_{k=1}^2 \sum_{l=1}^2 A_{hpk} A_{h'ql} \quad (2)$$

$h, h' = 2, 4, 6, \dots, L$

If the L lines were pure ($F_L = 1$) there would not be loss of non-identical by descent (NIBD) genes in the derivation of m genotypes, produced by crossing the two lines of each one of the $L/2$ single crosses. However, this could be different when the L lines are not pure ($F_L < 1$) and m is small. For example, if $m = 1$ and $F_L < 1$, as every initial line, on average, is carrier of $2 - F_L$ NIBD genes, the two lines of a single cross carry on average $4 - 2F_L$ NIBD genes, and a plant ($m = 1$) produced by its cross will receive exactly 2 of those genes because the two lines are not related. This implies an average loss of $2 - 2F_L$ NIBD genes per each of the $L/2$ individuals representing each a single cross.

For $m = 2$, let us assume that the genotypes of the 2 parental lines of a single cross are represented by $A_1 A_2$ and $B_1 B_2$. The probabilities (P) that the genes of each genotype are identical by descent (\equiv) are $P(A_1 \equiv A_2) = P(B_1 \equiv B_2) = F_L$ and non-identical (\neq) is $P(A_1 \neq A_2) = P(B_1 \neq B_2) = 1 - F_L$. Moreover, the genotypic array of the cross between the two lines ($AGEC_L$) is:

$$AGEC_L = (1/4)A_1 B_1 + (1/4)A_1 B_2 + (1/4)A_2 B_1 + (1/4)A_2 B_2$$

The two plants of the cross participating in the Sin_{CS} of this example ($m = 2$) can be seen as the result of the random sampling of size 2 with replacement of all four different genotypes of $AGEC_L$. The types of possible outcomes are:

- Both times the same genotype. For example, $A_1 B_1$ and $A_1 B_1$ (2 NIBD genes);
- Two genotypes like: $A_1 B_1$ and $A_1 B_2$ ($3 - F_L$ NIBD genes), and
- Two genotypes like: $A_1 B_1$ and $A_2 B_2$ ($4 - 2F_L$ NIBD genes).

- a) el mismo genotipo las dos veces. Por ejemplo, A_1B_1 y A_1B_1 (2 genes NIPD);
- b) dos genotipos diferentes de la forma: A_1B_1 y A_1B_2 ($3 - F_L$ genes NIPD), y
- c) dos genotipos diferentes de la forma: A_1B_1 y A_2B_2 ($4 - 2F_L$ genes NIPD).

Las probabilidades de ocurrencia de estos tres tipos de resultados son, en el mismo orden: 4/16, 8/16 y 4/16. Y el número de genes NIPD que se espera que aporte esta cruce simple al arreglo genotípico del Sin_{CS} en promedio $[E(X_2)]$ se deriva a continuación:

$$E(X_2) = (2)(1/4) + (3 - F_L)(1/2) + (4 - 2F_L)(1/4) = 3 - F_L \quad (3)$$

Según la Ecuación 3, con $m = 2$ la cantidad esperada de genes NIPD en una cruce simple representada por dos plantas cuando las líneas son puras ($F_L = 1$) es igual a 2; en este caso con un individuo por cada línea progenitora se aportarían estos 2 genes. Sin embargo, cuando $F_L < 1$, cada individuo de una línea aportaría, en promedio, $2 - F_L$ genes NIPD; y los dos individuos aportarían, en promedio, $4 - 2F_L$ genes NIPD. Esto significa que con $m = 2$ en la cruce simple y $m = 1$ en sus dos líneas progenitoras, a pesar de la compensación del número de individuos (2 en la cruce simple y 1 en cada línea progenitora), cada cruce simple pierde $1 - F_L$ genes NIPD ya que sólo aporta $3 - F_L$ de estos $4 - 2F_L$ genes NIPD (Ecuación 3).

Un caso más para entender mejor los detalles de la derivación general se hará a continuación. Si $m = 3$ y la variable aleatoria X_3 representa el número de genes no idénticos por descendencia (NIPD) que se aporta a una cruce simple, $E(X_3)$ se puede derivar según la expresión

$$E(X_3) = \sum_{x_3} x_3 P(X_3 = x_3) \quad (4)$$

En la Ecuación 4, x_3 representa los números de genes NIPD que puede aportar una cruce simple al arreglo genotípico del Sin_{CS} , y $P(X_3 = x_3)$ es la probabilidad de que ocurra el valor x_3 cuando se efectúa el fenómeno visualizable como la toma de una muestra aleatoria con reemplazo de tamaño 3 del conjunto que forman los 4 genotipos que produce la cruce simple $A_1A_2 \times B_1B_2$. El muestreo de tamaño 3 puede producir los tipos de resultados siguientes:

- a) el mismo genotipo las 3 veces (por ejemplo A_1B_1 , A_1B_1 y A_1B_1). Esto puede ocurrir de 4 formas (los cuatro genotipos posibles); y $x_3 = 2$.
- b) 2 genotipos iguales y uno diferente: i) ejemplo 1: A_1B_1 , A_1B_1 y A_1B_2 ; ii) ejemplo 2: A_1B_1 , A_1B_1 y A_2B_2 . En los casos i) y ii), $x_3 = 3 - F_L$ y $x_3 = 4 - 2F_L$, respectivamente. Además, los números de casos del tipo i) son 24 [4 combinaciones de 2 genotipos, 2

The probabilities of occurrence of these three results are, in the same order: 4/16, 8/16 and 4/16. And the expected number of NIBD genes provided by this single cross to the genotypic array of the SV, on average $[E(X_2)]$ is derived as follows:

$$E(X_2) = (2)(1/4) + (3 - F_L)(1/2) + (4 - 2F_L)(1/4) = 3 - F_L \quad (3)$$

According to Equation 3, with $m = 2$ the expected number of NIBD genes in a single cross represented by two plants when the lines are pure ($F_L = 1$) is 2; in this case these 2 genes would be provided each by an individual of each parent line. However, when $F_L < 1$, each individual of a line would contribute, on average, $2 - F_L$ NIBD genes; and the two individuals would provide, on average, $4 - 2F_L$ NIBD genes. This means that with $m = 2$ in the single cross and $m = 1$ in the two parental lines, despite the offset of the number of individuals (2 in the single cross and 1 in each parental line), each single cross loses $1 - F_L$ NIBD genes, because it only provides $3 - F_L$ out of these $4 - 2F_L$ NIBD genes (Equation 3).

Another case will be shown to better understand the details of the general derivation. If $m = 3$ and the random variable X_3 represents the number of non-identical by descent (NIBD) genes provided to a single cross, $E(X_3)$ can be derived according to the expression

$$E(X_3) = \sum_{x_3} x_3 P(X_3 = x_3) \quad (4)$$

In the Equation 4, x_3 represents the numbers of NIBD genes that a single cross can provide to the genotypic array of Sin_{CS} , and $P(X_3 = x_3)$ is the probability of occurrence of the x_3 value when performing the phenomenon viewable as taking a random sample with replacement of size 3 from the set formed by 4 genotypes produced by the single cross $A_1A_2 \times B_1B_2$. Sampling of size 3 can generate the following types of results:

- a) All three times the same genotype (for example A_1B_1 , A_1B_1 , and A_1B_1). This can occur in 4 ways (the four possible genotypes); and $x_3 = 2$.
- b) 2 identical and one different genotype: i) example 1: A_1B_1 , A_1B_1 , and A_1B_2 ; ii) example 2: A_1B_1 , A_1B_1 , and A_2B_2 . In cases i) and ii), $x_3 = 3 - F_L$ and $x_3 = 4 - 2F_L$, respectively. Furthermore, the number of cases of type i) are 24 [4 combinations of 2 genotypes, 2 possibilities with respect to that that occurs 2 times, and 3 orders of occurrence in each case, making a total of $4 \times 2 \times 3 = 24$], and type ii) are 12 [2 combinations of 2 genotypes, 2 possibilities with respect to that that occurs 2 times, and 3 orders of occurrence in each case].
- c) 3 different genotypes. For example: A_1B_1 , A_1B_2 and A_2B_1 . In this case, $x_3 = 4 - 2F_L$. The number of cases

posibilidades respecto al que ocurre 2 veces, y 3 órdenes de ocurrencia en cada caso, que hacen un total de $4 \times 2 \times 3 = 24$] y del tipo ii) son 12 [2 combinaciones de 2 genotipos, 2 posibilidades respecto al que ocurre 2 veces, y 3 órdenes de ocurrencia en cada caso].

- c) 3 genotipos diferentes. Por ejemplo: A_1B_1 , A_1B_2 y A_2B_1 . En estos casos, $x_3 = 4 - 2F_L$. El número de casos es 24 [4 combinaciones de 3 genotipos, y $3! = 6$ órdenes de ocurrencia posibles en cada caso].

De acuerdo con la información generada para $m = 3$ en a), b) y c) y la consideración de que el muestreo produce 64 (4^3) resultados posibles:

$$\begin{aligned} E(X_3) &= \sum_{x_3} x_3 P(X_3 = x_3) \\ &= (2)(4/64) + (3 - F_L)(24/64) \\ &\quad + (4 - 2F_L)(12/64) + (4 - 2F_L)(24/64) \\ &= 3.5 - 1.5F_L \end{aligned} \quad (5)$$

En general, cuando cada cruce simple se representa por m plantas, los datos requeridos para calcular $E(Xm)$ se consignan en el Cuadro 1. De acuerdo con éstos:

$$\begin{aligned} E(Xm) &= (2)(4/4^m) + (3 - F_L)[4(2^m - 2)]/4^m \\ &\quad + (4 - 2F_L)[4^m - 4 - 4(2^m - 2)]/4^m \\ &= \{8 + 12(2^m - 2) + 4^{m+1} - 4^2 - 16(2^m - 2) \\ &\quad - 2[2(2^m - 2) + (4^m - 4) - 4(2^m - 2)]F_L\}/4^m \\ &= \{4[4^m - 2] - 4(2^m - 2) \\ &\quad - 2[4^m - 2^{m+1}]F_L\}/4^m \\ &= [2^{2m+2} - 2^{m+2}]/2^{2m} \\ &\quad - 2[1 - (1/2)^{m-1}]F_L \\ &= 4 - 1/2^{m-2} - [2 - 1/2^{m-2}]F_L \\ &= 2 + [2 - (1/2)^{m-2}](1 - F_L) \end{aligned} \quad (6)$$

Según la Ecuación 6, el número total de genes NIPD que se espera que contengan todas las cruces simples (NTG_{NIPD}) debe ser $(L/2)[E(Xm)]$; es decir:

$$NTG_{NIPD} = L + L[1 - (1/2)^{m-1}](1 - F_L) \quad (7)$$

is 24 [4 combinations of 3 genotypes, and $3! = 6$ orders of occurrence in each case].

According to the information generated for $m = 3$ in a), b) and c) and the assumption that the sampling produces 64 (4^3) possible results:

$$\begin{aligned} E(X_3) &= \sum_{x_3} x_3 P(X_3 = x_3) \\ &= (2)(4/64) + (3 - F_L)(24/64) \\ &\quad + (4 - 2F_L)(12/64) + (4 - 2F_L)(24/64) \\ &= 3.5 - 1.5F_L \end{aligned} \quad (5)$$

In general, when every single cross is represented by m plants, the data required to calculate $E(Xm)$ are listed in Table 1. According to these:

$$\begin{aligned} E(Xm) &= (2)(4/4^m) + (3 - F_L)[4(2^m - 2)]/4^m \\ &\quad + (4 - 2F_L)[4^m - 4 - 4(2^m - 2)]/4^m \\ &= \{8 + 12(2^m - 2) + 4^{m+1} - 4^2 - 16(2^m - 2) \\ &\quad - 2[2(2^m - 2) + (4^m - 4) - 4(2^m - 2)]F_L\}/4^m \\ &= \{4[4^m - 2] - 4(2^m - 2) \\ &\quad - 2[4^m - 2^{m+1}]F_L\}/4^m \\ &= [2^{2m+2} - 2^{m+2}]/2^{2m} \\ &\quad - 2[1 - (1/2)^{m-1}]F_L \\ &= 4 - 1/2^{m-2} - [2 - 1/2^{m-2}]F_L \\ &= 2 + [2 - (1/2)^{m-2}](1 - F_L) \end{aligned} \quad (6)$$

According to Equation 6, the total number of NIBD genes expected to have all single crosses (NTG_{NIBD}) should be $(L/2)[E(Xm)]$; i.e.:

$$NTG_{NIBD} = L + L[1 - (1/2)^{m-1}](1 - F_L) \quad (7)$$

Variability of the number of NIBD genes

The variance (Var) of the number (Xm) of non-identical by descent (NIBD) genes for each single cross, using the Equation 6 and Table 1 can be derived based on the following initial expression:

$$\begin{aligned} Var(Xm) &= E[Xm]^2 - [E(Xm)]^2 \\ &= \frac{4}{4^m} (2^2) + \frac{4(2^{m-2})}{4^m} (3 - F_L)^2 + [4^m - 4 - 4(2^m - 2) (4 - 2F_L)]^2 \\ &\quad - \{2 + [2 - (1/2)^{m-2}](1 - F_L)\}^2 \end{aligned} \quad (8)$$

CUADRO 1. Número de genes no idénticos por descendencia (Xm) en los genotipos de m plantas que produce la cruce simple entre dos líneas, número de formas de ocurrencia de cada valor de Xm ($NFVXm$), y probabilidad de ocurrencia de cada valor de Xm [$P(Xm = xm)$].

TABLE 1. Number of non-identical by descent genes (Xm) in the genotypes of m plants produced by a single cross between two lines, number of ways of occurrence of each value of Xm ($NFVXm$), and probability of occurrence of each value of Xm [$P(Xm = xm)$].

Genotipos resultantes / Resulting genotypes	xm^z	$NFVXm$	$P(Xm = xm)$
El mismo m veces/ The same m times	2	4	$4/4^m$
Uno n veces y otro $m-n$ veces/One n times and the other $m-n$ times			
Ejemplo: A_1B_1 y A_1B_2	$3 - F_L$	$4(2^m - 2)$	$4(2^m - 2)/4^m$
Example: A_1B_1 and A_1B_2			
Uno n veces y otro $m-n$ veces/One n times and the other $m-n$ times			
Ejemplo: A_1B_1 y A_2B_2	$4 - 2F_L$	$2(2^m - 2)$	$2(2^m - 2)/4^m$
Example: A_1B_1 and A_2B_2			
El resto de resultados/ The rest of the results	$4 - 2F_L$	$4^m - 4 - 6(2^m - 2)$	$[4^m - 4 - 6(2^m - 2)]/4^m$

^z F_L es el coeficiente de endogamia de las líneas.

^z F_L is the inbreeding coefficient of the lines.

Variabilidad del número de genes NIPD

La varianza (Var) del número de genes (Xm) no idénticos por descendencia (NIPD) para cada cruce simple, con ayuda de la Ecuación 6 y del Cuadro 1 puede derivarse con base en la expresión inicial siguiente:

$$\begin{aligned}
 Var(Xm) &= E[Xm]^2 - [E(Xm)]^2 \\
 &= \frac{4}{4^m} (2^2) + \frac{4(2^m-2)}{4^m} (3 - F_L)^2 + [4^m - 4 - 4(2^m - 2) (4 - 2F_L)^2 \\
 &\quad - \{2 + [2 - (1/2)^{m-2}](1-F_L)\}^2
 \end{aligned}
 \tag{8}$$

El desarrollo de la Ecuación 8, se hará agrupando términos de acuerdo con la inclusión o ausencia de F_L y F_L^2 . La contribución a $Var(Xm)$ de los términos que no contienen el coeficiente de endogamia F_L ni F_L^2 (NCE) es:

$$\begin{aligned}
 NCE &= [16 + (36 - 64)(2^m - 2) + 16(4^m) - 64]/4^m - \{4 + 4[2 - (1/2)^{m-2} \\
 &\quad + 4 - 4(1/2)^{m-2} + (1/2)^{2m-4}]\} \\
 &= [16 - 28(2^m) + 56 - 64]/4^m + 16 - [16 - (1/2)^{m-4} - (1/2)^{m-4} + (1/2)^{2m-4}] \\
 &= [8 - 28(2^m) + (1/2)^{m-5-2m} - (1/2)^{2m-4-2m}]/4^m \\
 &= [8 - 28(2^m) + (1/2)^{m-5} - (1/2)^{-4}]/4^m \\
 &= [8 - 28(2^m) + 32(2^m) - 16]/4^m \\
 &= [4(2^m) - 8]/4^m
 \end{aligned}
 \tag{9}$$

The development of the Equation 8 will be done grouping terms according to the inclusion or absence of F_L and F_L^2 . The contribution to $Var(Xm)$ of terms that do not have the inbreeding coefficient F_L or F_L^2 (NCE) is:

$$\begin{aligned}
 NCE &= [16 + (36 - 64)(2^m - 2) + 16(4^m) - 64]/4^m - \{4 + 4[2 - (1/2)^{m-2} \\
 &\quad + 4 - 4(1/2)^{m-2} + (1/2)^{2m-4}]\} \\
 &= [16 - 28(2^m) + 56 - 64]/4^m + 16 - [16 - (1/2)^{m-4} - (1/2)^{m-4} + (1/2)^{2m-4}] \\
 &= [8 - 28(2^m) + (1/2)^{m-5-2m} - (1/2)^{2m-4-2m}]/4^m \\
 &= [8 - 28(2^m) + (1/2)^{m-5} - (1/2)^{-4}]/4^m \\
 &= [8 - 28(2^m) + 32(2^m) - 16]/4^m \\
 &= [4(2^m) - 8]/4^m
 \end{aligned}
 \tag{9}$$

Similarly, the contributions of the terms containing F_L (CCE) and F_L^2 [$C(CE)^2$] to $Var(Xm)$ are:

$$CCE = [16 - 8(2^m)]F_L/4^m \tag{10}$$

and

$$C(CE)^2 = [4(2^m) - 8]F_L^2/4^m \tag{11}$$

According to the Equations 8 to 11, the expression of the variance of the number of NIBD genes received by a single cross [$Var(Xm)$] is:

Similarmente, las contribuciones de los términos que contienen F_L (CCE) y F_L^2 [$C(CE)^2$] a $Var(Xm)$ corresponden a las expresiones siguientes:

$$CCE = [16 - 8(2^m)]F_L/4^m \quad (10)$$

y

$$C(CE)^2 = [4(2^m) - 8]F_L^2/4^m \quad (11)$$

De acuerdo con las Ecuaciones 8 a 11, la expresión de la varianza del número de genes NIPD que recibe una cruce simple [$Var(Xm)$] es:

$$\begin{aligned} Var(Xm) &= [4(2^m) - 8][1 + F_L^2]/4^m \\ &+ [16 - 8(2^m)]F_L/4^m \\ &= [2^{m+2} - 2^3](1 - F_L)^2/4^m \end{aligned} \quad (12)$$

La Ecuación 12 implica que la varianza del número de genes NIPD es mayor cuando el coeficiente de endogamia de las líneas iniciales (F_L) y el número de individuos que representan cada línea (m) son menores. De hecho, cuando las líneas progenitoras iniciales son puras ($F_L=1$), independientemente del tamaño de m , $Var(Xm)$ es igual a cero; y la contribución esperada a cada cruce simple es de 2 genes NIPD (Ecuación 6). Notablemente, las Ecuaciones 6 y 12 implican que los incrementos de m redundan en la reducción de la pérdida de genes NIPD y en la estabilidad del arreglo genotípico del Sin_{cs} (Ecuación 2). En este contexto es de particular interés la selección recurrente en que la prueba temprana de aptitud combinatoria general se aplica a líneas de una sola autofecundación ($F_L=1/2$) y la formación de sintéticos con estas líneas. En ambos casos, con $m \geq 8$ la pérdida esperada y la varianza de genes NIPD de las cruces simples se acercan considerablemente a cero (Cuadro 2, Ecuaciones 6 y 12).

DISCUSIÓN

Si el coeficiente de endogamia de las L líneas con que se genera las $L/2$ cruces simples (CS) es menor que 1 ($F_L < 1$), y cada CS se representa por más de 1 planta ($m > 1$), en los m genotipos de las m plantas de cada cruce simple debe haber de 2 a 4 genes no idénticos por descendencia (NIPD). Según la Ecuación 6, el número esperado de genes NIPD [$E(Xm)$] que tiene una cruce simple es $2 + [2 - (1/2)^{m-2}](1 - F_L)$. Esta expresión implica que $E(Xm)$ aumenta cuando m es mayor. Esto se debe a que cada gen adicional que contribuye a la formación de un genotipo de una cruce simple es una oportunidad de que sea un gen NIPD que no se hubiera contribuido previamente. Sin embargo, de acuerdo con la Ecuación 6, estos incrementos del número de genes NIPD aportados son cada vez más pequeños. El Cuadro 2 refleja claramente esta tendencia. En el extremo opuesto, cuando

$$\begin{aligned} Var(Xm) &= [4(2^m) - 8][1 + F_L^2]/4^m \\ &+ [16 - 8(2^m)]F_L/4^m \\ &= [2^{m+2} - 2^3](1 - F_L)^2/4^m \end{aligned} \quad (12)$$

Equation 12 implies that the variance of the number of NIBD genes is greater when the inbreeding coefficient of the initial lines (F_L) and the number of individuals representing each line (m) are small. In fact, when the initial parental lines are pure ($F_L=1$), regardless of the size of m , $Var(Xm)$ is equal to zero; and the expected contribution to each single cross is 2 NIBD genes (Equation 6). The Equations 6 and 12 imply that the increases of m result in reducing loss of NIBD genes and the stability of the genotypic array of Sin_{cs} (Equation 2). In this context it is important to note that the early generation test of general combining ability is applied to S_1 lines ($F_L=1/2$) and the derivation of synthetics to these lines. In both cases, with $m \geq 8$ the expected loss and variance of NIBD genes of single crosses considerably approach to zero (Table 2, Equations 6 and 12).

DISCUSSION

If the coefficient of inbreeding of the L lines that generates the $L/2$ single crosses (SC) is smaller than 1 ($F_L < 1$), and each SC is represented by more than 1 plant ($m > 1$), in the m genotypes of these m plants of each single cross there must be from 2 to 4 non-identical by descent (NIBD) genes. According to Equation 6, the expected number of NIBD genes [$E(Xm)$] that a single cross has is $2 + [2 - (1/2)^{m-2}](1 - F_L)$. This expression implies that $E(Xm)$ increases when m is greater. This is because each additional gene contributing to the derivation of a genotype of a single cross is a chance of a non-previously contributed NIBD gene. However, according to the Equation 6, these increases in the number of NIBD genes provided are progressively smaller. Table 2 clearly reflects this trend. On the other hand, when every single cross is represented by a plant ($m = 1$) its genotype always has to be formed by two NIBD genes because lines are unrelated; and the mean and variance of the number of NIBD genes must be 2 (Equation 6) and 0 (Equation 12), respectively, for any value of F_L (Table 2).

Equation 6 also shows that $E(Xm)$ and F_L have an inverse relationship, which was expected because F_L is a measure of the degree of identity by descent of the two genes that make up each genotype of the initial lines. Furthermore, the relationship between $E(Xm)$ and F_L is linear, that is, every time F_L increases (or decreases) an equal amount, $E(Xm)$ also decreases (or increases) a constant amount.

With respect to the variance of Xm , the differences in magnitude due to changes in F_L and changes in m are predictable according to Equation 12. In contrast, the

cada cruza simple se representa por una planta ($m = 1$) su genotipo siempre tiene que estar formado por 2 genes NIPD debido a que las líneas no están emparentadas; y la media y la varianza del número de genes NIPD deben ser 2 (Ecuación 6) y 0 (Ecuación 12), respectivamente, para cualquier valor de F_L (Cuadro 2).

La Ecuación 6 también refleja que $E(Xm)$ y F_L tienen una relación inversa, lo que era de esperarse porque F_L es una medida del grado de identidad por descendencia de los dos genes que integran cada genotipo de las líneas iniciales. Además, la relación entre $E(Xm)$ y F_L es lineal; es decir, cada que F_L aumenta (o disminuye) una misma cantidad, $E(Xm)$ disminuye (o aumenta) una cantidad constante también.

Respecto a la varianza de Xm , las diferencias en sus magnitudes debidas a cambios en F_L y a cambios en m son predecibles según la Ecuación 12. En cambio, la variación que se consigna para cada combinación de valores de m y F_L en el Cuadro 2 es variación que se debe a la naturaleza aleatoria del mecanismo genético. Esta variabilidad se intensifica cuando las L líneas iniciales no son puras ($F_L < 1$) y m es pequeña pero mayor que 1. Esta variación aleatoria es importante porque impacta la constitución genotípica de las cruza simples en primera instancia, y posteriormente la del sintético que se forme con esas $L/2$ cruza sim-

variation recorded for each combination of values of m and F_L in Table 2 is a variation due to the random nature of the genetic mechanism. This variability is intensified when the initial L lines are not pure ($F_L < 1$) and m is small but greater than 1. This random variation is important because, in the first instance, it affects the genotypic derivation of the single crosses, and subsequently the synthetic derivation consisting of these $L/2$ single crosses. Equation 12 implies that if $m > 1$, the magnitude of $Var(Xm)$ is inversely related to F_L and m . When the initial lines are pure ($F_L = 1$), the single crosses are genotypically uniform and carry only 2 NIBD genes, regardless of the size of m , which is consistent with the Equation 12.

Due to its characteristics (coefficient of inbreeding F_L and no relationship between them), the L initial lines contain an expected number of $L(2-F_L)$ NIBD genes, whatever the magnitude of m . Furthermore, according to Equation 6, it is expected that the $L/2$ single crosses be carriers of $L + L[1 - (1/2)^{m-1}](1 - F_L)$ NIBD genes. The difference between the first and the second result, $(1/2)^{m-1}(1 - F_L)$, is the average number of NIBD genes lost in the derivation of the $L/2$ single crosses. Clearly, the gene loss is larger as F_L and m are smaller. This loss affects the genotypic array of Sin_{cs} (Equation 2) being different from Sin_L (Equation 1). However, when we start from pure lines ($F_L = 1$) no NIBD gene is lost until the derivation of the single crosses.

CUADRO 2. Media [$E(Xm)$] y varianza [$Var(Xm)$] del número de genes (Xm) no idénticos por descendencia que recibe una cruza simple que será un progenitor de un sintético formado por $L/2$ cruza simples (CSs). F_L es la endogamia de las L líneas progenitoras de las CSs y m es el número de individuos que representan cada CS.

TABLE 2. Mean [$E(Xm)$] and variance [$Var(Xm)$] of the number (Xm) of non-identical by descent genes received by a single cross that will be a parent of the synthetic formed by $L/2$ single crosses (SCs). F_L is the inbreeding of the L lines progenitors of the SCs and m is the number of individuals representing each SC.

m	$F_L = 0.25$		$F_L = 0.50$		$F_L = 0.75$		$F_L = 1.0$	
	$E(Xm)$	$Var(Xm)$	$E(Xm)$	$Var(Xm)$	$E(Xm)$	$Var(Xm)$	$E(Xm)$	$Var(Xm)$
1	2.000	0.000	2.000	0.000	2.000	0.000	2.000	0.000
2	2.750	0.281	2.500	0.125	2.250	0.031	2.000	0.000
3	3.125	0.211	2.750	0.094	2.375	0.023	2.000	0.000
4	3.313	0.123	2.875	0.055	2.438	0.014	2.000	0.000
5	3.410	0.070	2.938	0.029	2.469	0.007	2.000	0.000
8	3.488	0.009	2.992	0.004	2.496	0.001	2.000	0.000
10	3.497	0.002	2.998	0.001	2.499	0.000	2.000	0.000
15	3.500	0.000	3.000	0.000	2.500	0.000	2.000	0.000
25	3.500	0.000	3.000	0.000	2.500	0.000	2.000	0.000
50	3.500	0.000	3.000	0.000	2.500	0.000	2.000	0.000
100	3.500	0.000	3.000	0.000	2.500	0.000	2.000	0.000

ples. La Ecuación 12 implica que si $m > 1$ la magnitud de $Var(Xm)$ se relaciona inversamente con F_L y con m . En un extremo, cuando las líneas iniciales son puras ($F_L = 1$), las cruza simples son genotípicamente uniformes y portan sólo 2 genes NIPD, independientemente del tamaño de m , en congruencia con la Ecuación 12.

Por sus características (coeficiente de endogamia F_L y ausencia de parentesco entre ellas), las L líneas iniciales contienen un número esperado de $L(2-F_L)$ genes NIPD, sea cual fuere la magnitud de m . Además, según la Ecuación 6, se espera que las $L/2$ cruza simples sean portadoras de $L + L[1 - (1/2)^{m-1}](1 - F_L)$ genes NIPD. La diferencia entre el primero y el segundo de estos resultados, $(1/2)^{m-1}[1 - F_L]$, es el número promedio de genes NIPD que se pierde en la formación de las $L/2$ cruza simples. Evidentemente, la pérdida de genes es mayor a medida que F_L y m son más pequeños. Y esta pérdida influye en que el arreglo genotípico del Sin_{CS} (Ecuación 2) sea diferente del Sin_L (Ecuación 1). En cambio, cuando se parte de líneas puras ($F_L = 1$) no se pierde ningún gen NIPD hasta la formación de las cruza simples.

La reducción del número de genes NIPD debe repercutir en las frecuencias génicas y genotípicas y, consecuentemente, fenotípicas. Por esta razón, si m y/o F_L son pequeñas la representación del Sin_L mediante el Sin_{CS} no será fidedigna. Sin embargo, si las líneas no son puras, según los datos del Cuadro 2 y la Ecuación 6, a partir de $m = 8$, aproximadamente, la pérdida de genes se reduce prácticamente a 0, y el número de genes NIPD se estabiliza en $2(2 - F_L)$ en cada cruza simple.

En el apareamiento aleatorio de las CSs para producir el Sin_{CS} , sin embargo, nuevamente el azar y el tamaño de m actúan como factores que pueden causar una pérdida adicional de genes NIPD, y, en consecuencia, que el $AGES_{CS}$ (Ecuación 2) difiera del $AGES_L$ (Ecuación 1).

CONCLUSIONES

Con la participación de cada una de L líneas no emparentadas y un coeficiente de endogamia F_L se consideró la formación de las $L/2$ cruza simples (CSs) representadas por sendos conjuntos de m plantas. La media y la varianza del número de genes no idénticos por descendencia (NIPD) que transmiten cada dos líneas progenitoras a su CS fueron $2 + [2 - (1/2)^{m-2}](1-F_L)$ y $[2^{m+2}-8](1-F_L)^2/4^m$, respectivamente. Como cada par de líneas progenitoras de una CS porta en promedio $2(2 - F_L)$ genes NIPD, por cada CS se pierden en promedio $(1/2)^{m-2}(1 - F_L)$. Estos resultados sugieren la conveniencia del uso de líneas puras ($F_L = 1$), para evitar pérdida de genes NIPD y aumentar la estabilidad genotípica de los progenitores del sintético.

Reducing the number of NIBD genes should affect the gene and genotypic frequencies and hence the phenotypic frequencies. Therefore, if m and/or F_L are small the representation of Sin_L by Sin_{CS} will not be reliable. However, if the lines are not pure, according to the data of Table 2 and Equation 6, from $m = 8$ on approximately, the gene loss is practically reduced to 0, and the number of NIBD genes is stabilized at $2(2 - F_L)$ in each single cross.

In random mating of the single crosses to produce Sin_{CS} , randomness and size of m act as factors that can cause an additional loss of NIBD genes, and consequently, that $AGES_{CS}$ (Equation 2) differs from $AGES_L$ (Equation 1).

CONCLUSIONS

The derivation of the $L/2$ single crosses (SCs), represented each by a set of m plants was done with the participation of each of L unrelated lines whose inbreeding coefficient was F_L . Mean and variance of the number of non-identical by descent (NIBD) genes transmitted by two parental lines to its SC were $2 + [2 - (1/2)^{m-2}](1-F_L)$ and $[2^{m+2}-8](1-F_L)^2/4^m$, respectively. As each pair of parental lines of a SC carries on average $2(2 - F_L)$ NIBD genes, $(1/2)^{m-2}(1 - F_L)$ are lost on average for each SC. These results suggest the benefits of using pure lines ($F_L = 1$), to avoid loss of NIBD genes and increase the genotypic stability of parents of synthetic varieties. Furthermore, even with $F_L < 1$ from $m = 8$ on, the loss of NIBD genes is practically reduced to 0 and $E(Xm)$ is stabilized at $2(2 - F_L)$.

End of English Version

Además, aún con $F_L < 1$ a partir de $m = 8$, la pérdida de genes NIPD prácticamente se reduce a 0 y $E(Xm)$ se estabiliza en $2(2 - F_L)$.

LITERATURA CITADA

- FALCONER, D. S.; MACKAY, T. F. C. 2007. Introducción a la Genética Cuantitativa. 4ª Edición. Editorial Acribia. S. A. Zaragoza, España. 469 p.
- KUTKA, F.J.; SMITH, M. E. 2007. How many parents give the highest yield in predicted synthetic and composite populations of maize? *Crop Sci* 47: 1905-1913. doi: 10.2135/cropsci2006.12.0802sc
- MÁRQUEZ-SÁNCHEZ, F. 1992. Inbreeding and yield prediction in synthetic maize cultivars made with parental lines: Basic methods. *Crop Sci*. 32: 345-349. doi: 10.2135/cropsci1992.0011183X003200020013x
- MÁRQUEZ-SÁNCHEZ, F. 2008. Endogamia y predicción de sintéticos de maíz de cruza dobles. *Revista Fitotecnia Mexi-*

- cana 31(Especial 3): 1-4. <http://www.revistafitotecniamexicana.org/documentos/31-1%20Especial%203/1a.pdf>
- SAHAGÚN-CASTELLANOS, J. 1994. Sobre el cálculo del coeficiente de endogamia de variedades sintéticas. *Agrociencia Serie Fitociencia* 5: 67-78.
- SAHAGÚN-CASTELLANOS, J., RODRÍGUEZ-PÉREZ, J. E.; PEÑA-LOMELÍ, A. 2005. Predicting yield of synthetics derived from double crosses. *Maydica* 50: 129-136. http://www.maydica.org/articles/50_129.pdf
- SAHAGÚN C., J. 2011. Inbreeding and yield of synthetic varieties derived from single and double cross hybrids. *Maydica* 56: 265-271. http://www.maydica.org/articles/56_265.pdf
- SAHAGÚN C., J.; VILLANUEVA V., C. 1997. Teoría de las variedades sintéticas formadas con híbridos de cruza simple. *Revista Fitotecnia Mexicana* 20: 69-72. <http://www.redalyc.org/articulo.oa?id=61020106>
- VILLANUEVA V., C.; CASTILLO G., F.; MOLINA G., J. D. 1994. Aprovechamiento de cruzamientos dialélicos entre híbridos comerciales de maíz: análisis de progenitores y cruas. *Revista Fitotecnia Mexicana* 17: 175-185.