

EROSIÓN GENÉTICA DE LOS HÍBRIDOS TRILINEALES DE MAÍZ PROGENITOR-RES DE UNA VARIEDAD SINTÉTICA

GENETIC EROSION OF THE THREE-WAY LINE MAIZE HYBRID PARENTS OF A SYNTHETIC VARIETY

J. Enrique Rodríguez-Pérez, Jaime Sahagún-Castellanos*, Aureliano Peña-Lomelí,
Lucas Hernández-Ibañez, J. Luis Escalante-González

Fitotecnia. Universidad Autónoma Chapingo. 56230. Chapingo, Estado de México. (jsahagunc@yahoo.com.mx)

RESUMEN

Las generaciones avanzadas de híbridos trilineales (HTs) de maíz (*Zea mays* L.) podrían considerarse como variedades sintéticas (VSs) formadas por apareamiento aleatorio de sus líneas (Sin_L), pero se ha mostrado que hay diferencias dependientes del tipo de híbrido y del coeficiente de endogamia de las líneas que forman los híbridos (F), entre otros factores. Para investigar a los HTs en este contexto se determinó la pérdida de genes no idénticos por descendencia (NIPD) que ocurre durante la formación de la muestra de los m individuos que representan a cada uno como progenitor de una VS. La hipótesis de estudio fue que sí ocurre erosión genética. Se consideró el uso de líneas no emparentadas y que cada una constituye una población virtual con $2-F$ genes NIPD. El número promedio de genes NIPD que pierde un HT por muestreo (GNIPDP) se expresó como $3(2-F)$ menos el número de genes NIPD que sí llegan a la muestra. Se encontró que GNIPDP se relaciona inversamente con m y F . Si $F=1$, para que GNIPDP sea 0.15 o menos, m debe ser igual o mayor a 4, o 15 o más cuando $F=0$. Estos números indican que no debe ocurrir erosión con muestras de varios centenares como ocurre en la producción de semilla. Cuando $F>0$ el desbalance de las frecuencias génicas de las líneas en cada HT debe producir un coeficiente de endogamia en el Sin_t mayor al del Sin_L .

Palabras clave: Genes no idénticos por descendencia, coeficiente de endogamia, *Zea mays* L.

INTRODUCCIÓN

En México, el uso de variedades sintéticas (VSs) de maíz (*Zea mays* L.) formadas con líneas puras de rendimiento alto no es común debido

* Autor responsable ♦ Author for correspondence.

Recibido: octubre, 2015. Aprobado: agosto, 2016.

Publicado como ARTÍCULO en Agrociencia 50: 1081-1090. 2016.

ABSTRACT

The advanced generations of three-way line maize (*Zea mays* L.) hybrids (THs) could be considered synthetic varieties (SVs) formed by random mating of their lines (Sin_L). However, it was shown that there are differences depending on the type of hybrid and the inbreeding coefficient of the lines that form the hybrids (F), among other factors. To investigate the THs in this context, the loss of genes that are non-identical by descent (NIBD) was determined, which takes place during the formation of the sample of the m individuals that represent each one as parent of an SV. The hypothesis of the study was that genetic erosion does occur. The use of non-related lines was considered and that each one comprises a virtual population with $2-F$ NIBD genes. The average number of NIBD genes lost by a TH per sampling (GNIBDP) is expressed as $3(2-F)$ minus the number of NIBD genes that reach the sample. It was found that GNIBDP is inversely related with m and F . If $F=1$, in order for GNIBDP to be 0.15 or less, m must be equal to or greater than 4, or 15 or more when $F=0$. These numbers indicate that erosion should not occur with samples of several hundred, as occurs in seed production. When $F>0$ the imbalance of the gene frequencies of the lines in each TH should produce an inbreeding coefficient in the Sin_t higher than that of Sin_L .

Key words: Genes non-identical by descent, inbreeding coefficient, *Zea mays* L.

INTRODUCTION

In Mexico, the use of synthetic varieties (SVs) of maize (*Zea mays* L.) formed with pure high yield lines is not common, due to the scarcity of this type of parents (Espinosa *et al.*, 2002; Luna *et al.*, 2012) and the limited availability of inexpensive seed of outstanding single cross hybrids. Pérez-López *et al.*

a la escasez de este tipo de progenitores (Espinosa *et al.*, 2002; Luna *et al.*, 2012) y a la reducida oferta de semilla barata de híbridos de cruza simple sobresalientes. Pérez-López *et al.* (2014) mencionan que por esto se recurre a la producción de híbridos trilineales (HTs) o dobles, pero su siembra por un mismo agricultor en forma recurrente es poco frecuente. Algunos agricultores explotan sus generaciones avanzadas en lugar de comprar semilla original nuevamente. Las poblaciones resultantes se podrían considerar como las VSs que se formarían con los progenitores de los híbridos cultivados, pero puede no ser así (Márquez-Sánchez, 2010; Sahagún-Castellanos y Villanueva-Verduzco, 2012). De las diferencias entre estos dos tipos de poblaciones destaca la debida a la participación balanceada del material genético de las líneas en la VS que éstas formarían, lo cual no ocurre cuando los progenitores son HTs formados con las mismas líneas, por ejemplo. En este caso, cada línea de una cruza simple contribuye con la mitad del material genético en relación a lo que aporta la tercera línea de la cruza trilineal. Esta diferencia importante se manifiesta en el coeficiente de endogamia que reporta Márquez-Sánchez (2010) para una VS formada con HTs contruidos con líneas puras. Las diferencias también pueden surgir cuando hay erosión genética. Escalante-González *et al.* (2013) consideran que el azar del mecanismo genético, el número finito de individuos que representan cada progenitor de una VS, y la condición heterocigótica de sus genotipos posibilitan la pérdida de genes durante su desarrollo. Ellos estudiaron la ocurrencia de erosión genética durante la formación de VSs cuyos progenitores son cruza simples y encontraron que la pérdida de genes no idénticos por descendencia (*NIPD*) se relaciona inversamente con el coeficiente de endogamia de sus progenitores. Además, por diferencias en las frecuencias génicas con que las líneas participan en las VSs debe haber pérdida de genes *NIPD* en cuantía diferente entre aquellas formadas por sólo líneas y las generadas por HTs cuyos progenitores son las mismas líneas, cualquiera que sea su coeficiente de endogamia.

Respecto a la formación de VSs a partir de HTs, el objetivo de esta investigación fue determinar la magnitud de la erosión genética que ocurriría durante la formación de la muestra de individuos que representan a cada progenitor.

(2014) mention this as the reason for recurring to the production of three-way (THs) or double-cross hybrids, but they are seldom sown recurrently by the same farmer. Some farmers exploit their advanced generations instead of buying original seed again. The resulting populations could be considered like the SVs that would form with the parents of the cultivated hybrids, but this may not be the case (Márquez-Sánchez, 2010; Sahagún-Castellanos and Villanueva-Verduzco, 2012). Of the differences between these two types of populations, that due to the balanced participation of the genetic material of the lines in the SV that would be formed stands out, given that this does not occur when the parents are THs formed with the same lines, for example. In this case, each line of a single cross contributes half of the genetic material with respect to what is supplied by the third line of the three-way cross. This important difference is manifested in the inbreeding coefficient reported by Márquez-Sánchez (2010) for an SV formed with THs constructed with pure lines. The differences may also appear when there is genetic erosion. Escalante-González *et al.* (2013) consider that the randomness of the genetic mechanism, the finite number of individuals represented by each parent of an SV, and the heterozygous condition of its genotypes enable the loss of genes during its development. These authors studied the occurrence of genetic erosion during the formation of SVs whose parents are single crosses, and found that the loss of genes that are non-identical by descent (*NIBD*) is inversely related to the inbreeding coefficient of their parents. Furthermore, due to differences in the gene frequencies with which the lines participate in the SVs, there must be a loss of genes *NIBD* in different quantities among those formed only by lines and those generated by THs whose parents are the same lines, regardless of their inbreeding coefficient.

With respect to the formation of SVs from THs, the objective of the present investigation was to determine the magnitude of genetic erosion that would occur during the formation of the sample of individuals that represent each parent.

MATERIALS AND METHODS

The model of a *locus* in a diploid species reproduced by random mating was considered in this investigation. To this

MATERIALES Y MÉTODOS

El modelo de un *locus* en una especie diploide reproducida por apareamiento aleatorio se consideró en esta investigación. Al respecto se supuso que las L líneas iniciales con las que se forman los HTs no tienen parentesco, que su coeficiente de endogamia es F ($0 \leq F \leq 1$), y que L es múltiplo de tres. De acuerdo con las características de las líneas, independientemente de las frecuencias génicas, se supuso que cada una de ellas contiene $2-F$ genes *NIPD*.

El número de genes *NIPD* que en promedio pierde un HT $[(NIPD)_{pt}]$ se expresó como el número inicial de los que tienen sus tres progenitores $[3(2-F)]$, menos el número promedio de genes *NIPD* que llegan a sus m representantes. La variable que representó este número se definió como Y_m , y su promedio como $E(Y_m)$, o sea, $(NIPD)_{pt} = 3(2-F) - E(Y_m)$.

Se asumió que los progenitores de un HT son A_1A_2 , B_1B_2 y C_1C_2 , y que, sin pérdida de generalidad, éste se representa como $(A_1A_2 \times B_1B_2) \times C_1C_2$, con arreglo genotípico $[(AGE)_{CT}]$:

$$(AGE)_{CT} = (1/8)[A_1C_1 + A_1C_2 + A_2C_1 + A_2C_2 + B_1C_1 + B_1C_2 + B_2C_1 + B_2C_2] \quad (1)$$

Además, sus m representantes son los elementos de una muestra aleatoria tomada con reemplazo de la población formada por sus ocho genotipos (Ecuación 1).

Los resultados posibles del muestreo con reemplazo de la progenie de un HT (Ecuación 1) incluyen desde el caso en que sus m elementos son el mismo genotipo m veces, hasta cuando la muestra contenga los ocho genotipos de la progenie. Para este último caso se requiere que $m \geq 8$.

Cuando m es grande hay un gran número de posibilidades de frecuencias y órdenes de ocurrencia de los genotipos que integran cada muestra y en ésta se determinó el número de genes *NIPD* diferentes (sin importar su frecuencia), y el número de formas (permutaciones) en que este número de genes puede ocurrir. Al calcular el número de eventos igualmente posibles y mutuamente excluyentes que puede producir el muestreo (8^m), se calcularon $E(Y_m)$ y el número de genes *NIPD* que se pierden en el paso de la población virtual de un HT (arreglo genotípico de la Ecuación 1) a la muestra que lo representa.

RESULTADOS Y DISCUSIÓN

Para derivar una fórmula para $E(Y_m)$ se consideró que la muestra de tamaño m con reemplazo extrae g genotipos de los indicados en la Ecuación 1 ($g=1, 2, \dots, 8$) y se definió:

respect, it was assumed that the initial L lines with which the THs are formed are not related, that their inbreeding coefficient is F ($0 \leq F \leq 1$), and that L is a multiple of three. According to the characteristics of the lines, independently of the gene frequencies, it was assumed that each one of these contains $2 - F$ *NIBD* genes.

The number of *NIBD* genes that on the average are lost by a TH $[(NIBD)_{pt}]$ was expressed as the initial number of the genes that have their three parents $[3(2 - F)]$, minus the average number of *NIBD* genes that reach their m representatives. The variable represented by this number was defined as Y_m , and its average as $E(Y_m)$, or that is, $(NIBD)_{pt} = 3(2 - F) - E(Y_m)$.

It was assumed that the parents of a TH are A_1A_2 , B_1B_2 and C_1C_2 , and without loss of generality, this is represented as $(A_1A_2 \times B_1B_2) \times C_1C_2$, with genotypic array $[(AGE)_{CT}]$:

$$(AGE)_{CT} = (1/8)[A_1C_1 + A_1C_2 + A_2C_1 + A_2C_2 + B_1C_1 + B_1C_2 + B_2C_1 + B_2C_2] \quad (1)$$

Furthermore, their m representatives are the elements of a random sample taken with replacement of the population formed by its eight genotypes (Equation 1).

The possible results of the sampling with replacement of the offspring of a TH (Equation 1) include the extreme cases in which their m elements are the same genotype m times, and when the sample contains the eight genotypes of the progeny. For this latter case it is required that $m \geq 8$.

When m is large, there are a great number of possibilities of frequencies and orders of occurrence of the genotypes that integrate each sample, and in this case the number of different *NIBD* genes was determined (regardless of their frequency), and the number of forms (permutations) in which this number of genes may occur. When calculating the number of events that are equally possible and mutually exclusive that can be produced by the sampling (8^m), calculations were made of $E(Y_m)$ and the number of *NIBD* genes that are lost in the passing of the virtual population of a TH (genotypic array of Equation 1) to the sample that represents it.

RESULTS AND DISCUSSION

To derive a formula for $E(Y_m)$, it was considered that the sample of m size with replacement extracts g genotypes of those indicated in Equation 1 ($g=1, 2, \dots, 8$) and was defined as:

- 1) ${}^g f_i^m$: frequency of genotype i ($i=1, 2, \dots, g$) in the k -th ($k=1, 2, \dots, Ug$) possible set of g frequencies of occurrence, each for one of the g genotypes that may be in the sample.

- 1) ${}^g_k f_i^m$: frecuencia del genotipo i ($i = 1, 2, \dots, g$) en el k -ésimo ($k = 1, 2, \dots, Ug$) juego posible de g frecuencias de ocurrencia de sendos genotipos que pueden estar en la muestra.
- 2) ${}^g_k p_l^m$: número de veces que aparece la l -ésima frecuencia más pequeña del k -ésimo juego posible de g frecuencias de ocurrencia ($l = 1, 2, \dots, l_g$) de sendos genotipos que pueden estar en la muestra.

Con ambos términos se construyeron dos cocientes que permiten calcular:

- 1) El número de formas en que pueden ocurrir en la muestra g genotipos cualesquiera con el k -ésimo juego de g frecuencias del tipo ${}^g_k f_i^m \left[{}^g_k (NF)^m \right]$, definido en una distribución multinomial como:

$${}^g_k (NF)^m = m! / \left[\left({}^g_k f_1^m \right)! \left({}^g_k f_2^m \right)! \dots \left({}^g_k f_g^m \right)! \right] \quad (2)$$

El número de permutaciones para asignar los g elementos del k -ésimo juego de frecuencias a sendos genotipos de la muestra $\left[{}^g_k (NP)^m \right]$:

$${}^g_k (NP)^m = g! / \left[\left({}^g_k p_1^m \right)! \left({}^g_k p_2^m \right)! \dots \left({}^g_k p_{l_g}^m \right)! \right] \quad (3)$$

Cualquiera que sea la forma de asociación entre frecuencias y genotipos, el número de genes $NIPD$ que éstos aportan es el mismo.

El producto $\left[{}^g_k (NF)^m \right] \left[{}^g_k (NP)^m \right]$ (Ecuaciones 2 y 3) es el total de formas en que pueden ocurrir g genotipos, cualesquiera que éstos sean, en una muestra aleatoria de tamaño m . La muestra con reemplazo se toma de los ocho genotipos que se indican en la Ecuación 1 con el k -ésimo juego de frecuencias.

Si del r -ésimo grupo de ${}^g_k S_r^m$ conjuntos diferentes de g genotipos cada conjunto contiene ${}^g_k N_r^m$ genes $NIPD$ ($r=1, 2, \dots, n_g$), el muestreo que captura g genotipos con el k -ésimo juego de frecuencias debe contribuir con un total de genes $NIPD \left\{ \left[{}^g_k CG \right]^m \right\}$ expresable como:

$$\left[{}^g_k CG \right]^m = \left[{}^g_k (NF)^m \right] \left[{}^g_k (NP)^m \right] \left[\sum_{r=1}^{n_g} \left({}^g_k S_r^m \right) \left({}^g_k N_r^m \right) \right] \quad (4)$$

- 2) ${}^g_k p_l^m$: frequency of the l -th smallest frequency of the k -th possible set of the frequencies of occurrence ($l = 1, 2, \dots, l_g$) each for each one of the g genotypes that may be in the sample.

Two quotients were constructed with both terms with which the following can be calculated:

- 1) The number of forms in which any g genotypes can occur in the sample with the k -th set of g frequencies of the type ${}^g_k f_i^m \left[{}^g_k (NF)^m \right]$, defined in a multinomial distribution as:

$${}^g_k (NF)^m = m! / \left[\left({}^g_k f_1^m \right)! \left({}^g_k f_2^m \right)! \dots \left({}^g_k f_g^m \right)! \right] \quad (2)$$

The number of permutations to assign the g elements of the k -th set of frequencies to two separate genotypes of the sample $\left[{}^g_k (NP)^m \right]$:

$${}^g_k (NP)^m = g! / \left[\left({}^g_k p_1^m \right)! \left({}^g_k p_2^m \right)! \dots \left({}^g_k p_{l_g}^m \right)! \right] \quad (3)$$

Whatever the form of association between frequencies and genotypes, the number of $NIBD$ genes they supply is the same.

The product $\left[{}^g_k (NF)^m \right] \left[{}^g_k (NP)^m \right]$ (Equations 2 and 3) is the total of forms in which g genotypes can occur, regardless of what they are, in a random sample of m size. The sample with replacement is taken from the eight genotypes indicated in Equation 1 with the k -th set of frequencies.

If from the r -th group of ${}^g_k S_r^m$ different sets of g genotypes each set contains ${}^g_k N_r^m$ $NIBD$ genes ($r = 1, 2, \dots, n_g$), the sampling that captures g genotypes with the k -th set of frequencies should contribute with a total of $NIBD$ genes $NIPD \left\{ \left[{}^g_k CG \right]^m \right\}$ expressed as:

$$\left[{}^g_k CG \right]^m = \left[{}^g_k (NF)^m \right] \left[{}^g_k (NP)^m \right] \left[\sum_{r=1}^{n_g} \left({}^g_k S_r^m \right) \left({}^g_k N_r^m \right) \right] \quad (4)$$

En una muestra de tamaño m el número de genotipos diferentes (g) puede ser 1, 2, 3, ..., 8 (cuando $m \geq 8$). Además, el número de juegos de frecuencias depende de g ($k=1, 2, 3, \dots, Ug$). Con ambas consideraciones y por la Ecuación 4, el número esperado (promedio) de genes *NIPD* en los m representantes de un HT [$E(Y_m)$] es:

$$E(Y_m) = \sum_{g=1}^8 \sum_{k=1}^{Ug} \left[CG \right] / 8^m \quad (5)$$

Como el número de genes *NIPD* que contiene el arreglo genotípico de un HT es $3(2-F)$, el promedio de genes *NIPD* que pierde cada híbrido trilineal [$(NIPD)_{pt}$] se expresó como:

$$(NIPD)_{pt} = 3(2-F) - E(Y_m) \quad (6)$$

Como ejemplo, cuando $m=12$ y $g=6$, digamos $G_1, G_2, \dots, G_6$, hay 11 juegos de frecuencias posibles en que estos últimos pueden ocurrir ($k = 1, 2, \dots, 11$; Cuadro 1). Si $k = 1$ (juego de frecuencias para seis genotipos cualesquiera en que cada uno de cinco de ellos ocurre una vez y el sexto lo hace siete veces), las frecuencias pueden asociarse a los seis genotipos en seis formas diferentes [$6!/(5! 1!) = 6$]. Además, el número de formas en que pueden ocurrir estos seis genotipos con $m=12$ es $12!/(1!1!1!1!1!7!) = 95\,040$ (Cuadro 1).

En este ejemplo, para cada uno de los 28 conjuntos diferentes de seis genotipos los juegos posibles de frecuencias son 11. Complementariamente, se muestra los números de formas de asignar las frecuencias de cada juego $\left[{}^6_k(NP)^{12} \right]$, los números correspondientes de las formas de ocurrencia de cada conjunto de seis genotipos $\left[{}^6_k(NF)^{12} \right]$ y su producto (Cuadro 1).

Si los seis genotipos que aparecen en la muestra fueran $A_1C_1, A_1C_2, A_2C_1, A_2C_2, B_1C_1$ y B_1C_2 (Ecuación 1), éstos aportarían 5-2F genes *NIPD*. Con base en esto, por la información en el Cuadro 1 y aplicando las Ecuaciones 1 a 4, su contribución cuando

$$k=1 \left\{ \left[{}^6_1 CG \right]^{12} \right\} a E(Y_m) \text{ es:}$$

In a sample of m size the number of different genotypes (g) may be 1, 2, 3, ..., 8 (when $m \geq 8$). Furthermore, the number of sets of frequencies depends on g ($k = 1, 2, 3, \dots, Ug$). With both considerations and by Equation 4, the expected number (average) of *NIPD* genes in the m representatives of a TH [$E(Y_m)$] is:

$$E(Y_m) = \sum_{g=1}^8 \sum_{k=1}^{Ug} \left[CG \right] / 8^m \quad (5)$$

As the number of *NIPD* genes that contain the genotypic arrangement of a TH is $3(2-F)$, the average of *NIPD* lost by each three-way hybrid [$(NIPD)_{pt}$] was expressed as follows:

$$(NIPD)_{pt} = 3(2-F) - E(Y_m) \quad (6)$$

As an example, when $m = 12$ and $g = 6$, let's say $G_1, G_2, \dots, G_6$, there are 11 sets of possible frequencies in which the six genotypes may occur ($k = 1, 2, \dots, 11$; Table 1). If $k = 1$ (set of frequencies for any six genotypes in which each one of five of these occurs once and the sixth one occurs seven times), the frequencies can be associated to the six genotypes in six different forms [$6!/(5! 1!) = 6$]. Furthermore, the number of forms in which these six genotypes can occur with $m = 12$ is $12!/(1!1!1!1!1!7!) = 95,040$ [Table 1].

In this example, for each one of the 28 different sets of six genotypes the possible sets of frequencies are 11. Complementarily, the numbers of forms of assigning the frequencies of each set are shown $\left[{}^6_k(NP)^{12} \right]$, the corresponding numbers of the forms of occurrence of each set of six genotypes $\left[{}^6_k(NF)^{12} \right]$ and their product (Table 1).

If the six genotypes that appear in the sample were $A_1C_1, A_1C_2, A_2C_1, A_2C_2, B_1C_1$ and B_1C_2 (Equation 1), they would supply 5-2F genes *NIPD*. Based on the above, the information in Table 1 and applying Equations 1 to 4, their contribution when

$$k=1 \left\{ \left[{}^6_1 CG \right]^{12} \right\} a E(Y_m) \text{ to } E(Y_m) \text{ is as follows:}$$

Cuadro 1. Tres características del muestreo aleatorio con reemplazo cuando $m=12$ y $g=6$ del arreglo genotípico que produce $(A_1A_2 \times B_1B_2) \times C_1C_2$: 1) los 11 juegos posibles de frecuencias genotípicas (k), 2) las permutaciones para asociar las frecuencias de cada juego a los seis genotipos $\left[{}^6_k(NP)^{12} \right]$ y, 3) las formas de ocurrencia de seis genotipos con una asignación de las frecuencias del juego k {Ecuación 2, $\left[{}^6_k(NF)^{12} \right]}$.

Table 1. Three characteristics of the random sampling with replacement when $m = 12$ and $g = 6$ of the genotypic array produced by $(A_1A_2 \times B_1B_2) \times C_1C_2$: 1) the 11 possible sets of genotypic frequencies (k), 2) the permutations for associating the frequencies of each set to the six genotypes $\left[{}^6_k(NP)^{12} \right]$ and 3) the forms of occurrence of six genotypes with an assignment of the frequencies of set k {Equation 2, $\left[{}^6_k(NF)^{12} \right]}$.

k	Juegos de seis frecuencias de seis genotipos						$\left[{}^6_k(NP)^{12} \right]$	$\left[{}^6_k(NF)^{12} \right]$	$\left[{}^6_k(NP)^{12} \right]$	$\left[{}^6_k(NF)^{12} \right]$
	G_1	G_2	G_3	G_4	G_5	G_6				
1	1	1	1	1	1	7	6	95 040	570 240	
2	1	1	1	1	2	6	30	332 640	9 979 200	
3	1	1	1	1	3	5	30	665 280	19 958 400	
4	1	1	1	2	2	5	60	997 920	59 875 200	
5	1	1	1	1	4	4	15	831 600	12 474 000	
6	1	1	1	2	3	4	120	1 663 200	199 584 000	
7	1	1	2	2	2	4	60	2 494 800	149 688 000	
8	1	1	1	3	3	3	20	2 217 600	44 352 000	
9	1	1	2	2	3	3	90	3 326 400	299 376 000	
10	1	2	2	2	2	3	30	4 989 600	149 688 000	
11	2	2	2	2	2	2	1	7 484 400	7 484 400	
Σ									953 029 440	

$$\begin{aligned} \left[{}^6_1 CG \right]^{12} &= \frac{6!}{5!} \frac{12!}{7!} (5-2F)/8^{12} \\ &= 570240(5-2F)/8^{12} \end{aligned}$$

$$\begin{aligned} \left[{}^6_1 CG \right]^{12} &= \frac{6!}{5!} \frac{12!}{7!} (5-2F)/8^{12} \\ &= 570240(5-2F)/8^{12} \end{aligned}$$

Para determinar la contribución total cuando $m=12$ y $g=6$ se debe considerar la aportación de los restantes 27 conjuntos de seis genotipos que puede producir el muestreo y la cantidad de genes *NIPD* que cada uno de éstos contiene (Cuadro 2), así como todos los datos que permiten la determinación de $E(Y_m)$ para cada uno de 11 valores de m (Cuadro 3).

De los 28 conjuntos diferentes de seis genotipos, cuatro contienen $5-2F$ genes *NIPD* cada uno y cada uno de los 24 restantes aporta $6-3F$ (Cuadro 2). Con las aportaciones de los 11 conjuntos de frecuencias genotípicas posibles ($k = 1, 2, 3, \dots, 11$; Cuadro 1) se puede calcular la contribución total cuando $m = 12$ y $g=6$ a $E(Y_m)$. Ésta es (Cuadros 1 a 3):

To determine the total contribution when $m = 12$ and $g = 6$, we should consider the contribution of the remaining 27 sets of six genotypes which may be produced by the sampling and the amount of *NIPD* genes that each one contains (Table 2), as well as all of the data that permit the determination of $E(Y_m)$ for each one of the 11 values of m (Table 3).

Of the 28 different sets of six genotypes, four contain $5 - 2F$ *NIPD* genes each, and each one of the 24 remaining sets supplies $6 - 3F$ (Table 2). With the contributions of the 11 sets of possible genotypic frequencies ($k = 1, 2, 3, \dots, 11$; Table 1), the total contribution can be calculated when $m = 12$ and $g = 6$ to $E(Y_m)$. That is (Tables 1 to 3):

Cuadro 2. Ejemplos de conjuntos de g genotipos ($g = 1, 2, 3, \dots, 8$) que contienen y_m genes no idénticos por descendencia y número de casos posibles. Los genotipos se obtienen por muestreo aleatorio de tamaño m con reemplazo del conjunto de los ocho genotipos que produce la cruza $(A_1A_2 \times B_1B_2) \times C_1C_2$ (Ecuación 1).

Table 2. Examples of sets of g genotypes ($g = 1, 2, 3, \dots, 8$) that contain y_m genes not identical by descent and number of possible cases. The genotypes are obtained by random sampling of m size with replacement of the set of the eight genotypes produced by the cross $(A_1A_2 \times B_1B_2) \times C_1C_2$ (Equation 1).

g	Ejemplos	y_m	Casos posibles	g	Ejemplos	y_m	Casos posibles
1	A_1C_1	2	8	5	$A_1C_1, A_1C_2, A_2C_1,$ A_2C_2, B_1C_1	$5 - 2F$	24
2	A_1C_1, A_1C_2 A_1C_1, B_1C_1 A_1C_1, A_2C_2 A_1C_1, B_1C_2	$3 - F^\dagger$ 3 $4 - 2F$ $4 - F$	8 8 4 8		$A_1C_1, A_1C_2, A_2C_1,$ B_1C_1, B_2C_1	$6 - 3F$	32
3	A_1C_1, A_1C_2, A_2C_1 A_1C_1, A_1C_2, B_1C_1 A_1C_1, A_2C_1, B_1C_2	$4 - 2F$ $4 - F$ $5 - 2F$	8 24 24	6	$A_1C_1, A_1C_2, A_2C_1,$ A_2C_2, B_1C_1, B_1C_2 $A_1C_1, A_1C_2, A_2C_1,$ A_2C_2, B_1C_1, B_2C_1	$5 - 2F$ $6 - 3F$	4 24
4	$A_1C_1, A_1C_2, A_2C_2, B_1C_1$ $A_1C_1, A_1C_2, B_1C_1, B_1C_2$ $A_1C_1, A_2C_1, B_1C_1, B_2C_2$ $A_1C_1, A_1C_2, A_2C_1, A_2C_2$	$5 - 2F$ $4 - F$ $6 - 3F$ $4 - 2F$	50 4 14 2	7	$A_1C_1, A_1C_2, A_2C_1,$ $A_2C_2, B_1C_1, B_1C_2,$ B_2C_1	$6 - 3F$	8
				8	$A_1C_1, A_1C_2, A_2C_1,$ $A_2C_2, B_1C_1, B_1C_2,$ B_2C_1, B_2C_2	$6 - 3F$	1

$^\dagger F$: coeficiente de endogamia de las líneas progenitoras de los híbridos trilineales ♦ F : inbreeding coefficient of the parent lines of the three-way hybrids.

Cuadro 3. Número de formas de ocurrencia de g genotipos en muestras aleatorias de tamaño m tomadas con reemplazo de los ocho genotipos que produce la cruza trilineal $(A_1A_2 \times B_1B_2) \times C_1C_2$. Si $m = g$, además del número de formas, entre paréntesis está la suma de genes no idénticos por descendencia que aportan los $8!/[(8 - g)!g!]$ conjuntos posibles de g genotipos. No se incluye el caso $m = 11$.

Table 3. Number of forms of occurrence of g genotypes in random samples of m size taken with replacement of the eight genotypes produced by the three-way cross $(A_1A_2 \times B_1B_2) \times C_1C_2$. If $m = g$, in addition to the number of forms, between parenthesis is the sum of genes not identical by descent supplied by the $8!/[(8 - g)!g!]$ possible sets of g genotypes. The case $m = 11$ is not included.

m	g							
	1	2	3	4	5	6	7	8
1	8 (2)							
2	8	$2!(96-24F)^\dagger$						
3	8	6	$3!(248-88F)$					
4	8	14	36	$4!(358-150F)$				
5	8	30	150	240	$5!(312-144F)$			
6	8	62	540	1560	1800	$6!(164-80F)$		
7	8	126	1800	8400	16 800	15 120	$7!(48-24F)$	
8	8	254	5796	40 824	126 000	191 520	141 120	$8!(6-3F)$
9	8	510	15 882	186 480	834 120	1 905 120	2 328 480	1 451 520
10	8	1022	55 980	818 520	5 103 000	16 435 440	20 635 200	30 240 000
12	8	4094	519 156	14 676 024	165 528 000	953 029 440	3 162 075 894	6 411 968 640

$^\dagger F$: coeficiente de endogamia de las líneas progenitoras de los híbridos trilineales ♦ F : inbreeding coefficient of the parent lines of the three-way hybrids.

$$\begin{aligned} \sum_{k=1}^{11} \left[CG \right]^{12} &= 953029440 \left[4(5-2F) + 24(6-3F) \right] / 8^{12} \\ &= \frac{156296828200 - 76242355200F}{68719476740} \\ &= 2.2744 - 1.1095F \end{aligned}$$

La obtención de este resultado involucró la consideración ya evidenciada de que las 28 combinaciones de seis de los ocho genotipos de la progenie de un HT (Ecuación 1) tienen la misma probabilidad de ocurrencia en la muestra. Además, con $m \geq 8$, los resultados en cuanto a genotipos obtenidos en la muestra aleatoria con reemplazo son todas las combinaciones posibles que pueden ser formadas con g de los ocho genotipos para cada uno de los valores posibles de g (1, 2, ..., 8).

La información de los Cuadros 2 y 3 hace posible el cálculo del numerador del cociente que permite calcular el promedio de genes *NIPD* que llegan a los m representantes de la crucea trilineal. El denominador (8^m) es el número de resultados igualmente posibles y mutuamente excluyentes que puede producir el muestreo; para definir el numerador se requiere calcular el número de genes *NIPD* que aporta cada uno de los 8^m casos posibles (con $8 \leq m$). Por ejemplo, de acuerdo con la información del Cuadro 3, para $m=12$, que permite la obtención de muestras que pueden contener 1, 2, 3, 4, ..., 7 u 8 genotipos:

$$\begin{aligned} E(Y_{12}) &= \frac{8(2) + 4094(96 - 24F) + 519156(248 - 88F) + 14676024(358 - 150F)}{8^{12}} \\ &+ \frac{165528000(312 - 144F) + 953029440(164 - 80F) + 3162075844(48 - 24F)}{8^{12}} \\ &+ \frac{6411968640(6 - 3F)}{8^{12}} \\ &= 5.8728 - 2.8733F \end{aligned}$$

Este resultado, $E(Y_{12}) = 5.8728 - 2.8733F$, es importante para los fitomejoradores y especialistas en recursos genéticos y es un indicador de estar próximo a lo que se debe esperar cuando m es "grande": ningún gen *NIPD* debe perderse; es decir, su número en la muestra debe ser $6-3F$. Del Cuadro 4 se desprenden varios hechos relevantes: para un valor

$$\begin{aligned} \sum_{k=1}^{11} \left[CG \right]^{12} &= 953029440 \left[4(5-2F) + 24(6-3F) \right] / 8^{12} \\ &= \frac{156296828200 - 76242355200F}{68719476740} \\ &= 2.2744 - 1.1095F \end{aligned}$$

The obtainment of this result involved the consideration already evidenced that the 28 combinations of six of the eight genotypes of the offspring of a TH (Equation 1) have the same probability of occurrence in the sample. Furthermore, with $m \geq 8$, the results with respect to genotypes obtained in the random sample with replacement are all of the possible combinations that can be formed with g of the eight genotypes for each one of the possible values of g (1, 2, 3, ..., 8).

The information of Tables 2 and 3 make it possible to calculate the numerator of the quotient that permits the calculation of the average of *NIBD* genes that reach the m representatives of the three-way cross. The denominator (8^m) is the number of results equally possible and mutually exclusive that can be produced by the sample; to define the numerator, it is necessary to calculate the number of *NIBD* genes supplied by each one of the 8^m possible cases (with $8 \leq m$). For example, according to the information of Table 3, for $m = 12$, which allows the obtainment of samples that may contain 1, 2, 3, 4, ..., 7 or 8 genotypes:

This result, $E(Y_{12}) = 5.8728 - 2.8733F$, is important for breeders and specialists in genetic resources and is an indicator that one is close to what should be expected when m is "big": no *NIBD* gene must be lost; that is, their number in the sample should be $6-3F$. From Table 4, various relevant facts are revealed: for a fixed value of F when m increases,

Cuadro 4. Números de genes no idénticos por descendencia esperados $E(Y_m)$ y perdidos $[(NIPD)_{pt}]$ en muestras de tamaño m con reemplazo tomadas del arreglo genotípico de una cruce trilineal hecha con líneas no emparentadas (Ecuación 1) para cinco coeficientes de endogamia (F). El caso $m=11$ no se incluye.

Table 4. Numbers of genes not identical by descent expected $[(E(Y_m))]$ and lost $[(NIPD)_{pt}]$ in samples of m size with replacement taken from the genotypic array of a three-way cross made with non-related lines (Equation 1) for five inbreeding coefficients (F). The case $m=11$ is not included.

m	$F = 0.00$		$F = 0.50$		$F = 0.75$		$F = 0.875$		$F = 1.00$	
	$E(Y_m)$	$(NIPD)_{pt}$	$E(Y_m)$	$(NIPD)_{pt}$	$E(Y_m)$	$(NIPD)_{pt}$	$E(Y_m)$	$(NIPD)_{pt}$	$E(Y_m)$	$(NIPD)_{pt}$
1	2.00	4.00	2.00	2.50	2.00	1.75	2.00	1.38	2.00	1.00
2	3.25	2.75	2.88	2.38	2.69	1.06	2.59	0.78	2.50	0.50
3	4.06	1.94	3.41	1.09	3.08	0.67	2.91	0.46	2.75	0.25
4	4.61	1.39	3.74	0.76	3.31	0.44	3.09	0.28	2.88	0.13
5	4.99	1.01	3.96	0.54	3.45	0.30	3.19	0.18	2.94	0.06
6	5.26	0.74	4.11	0.39	3.54	0.21	3.25	0.12	2.97	0.03
7	5.45	0.55	4.22	0.28	3.60	0.15	3.29	0.08	2.98	0.02
8	5.59	0.41	4.29	0.21	3.64	0.11	3.32	0.06	2.99	0.01
9	5.69	0.31	4.34	0.16	3.67	0.08	3.33	0.04	2.99	0.01
10	5.77	0.23	4.39	0.11	3.69	0.06	3.34	0.03	3.00	0.00
12	5.87	0.13	4.44	0.06	3.72	0.03	3.36	0.02	3.00	0.00

fijo de F , cuando m aumenta, $E(Y_m)$ crece, más rápidamente cuando F es menor. Además, con líneas puras ($F=1$) basta una muestra de 10 individuos de una cruce trilineal para que se alcance el equilibrio en tres genes $NIPD$ y, por lo tanto, cero genes perdidos. Con $F=0.5$, en cambio, esta condición empieza a ser ostensible con $m=12$. También se infiere que cuando $F=0.75$, a partir de $m=15$ no hay pérdida de genes $NIPD$.

Con una muestra cada vez más grande su número de genes $NIPD$ tiende a aumentar porque cada individuo adicional que se incluya, es una nueva oportunidad de que tenga un genotipo que aporte genes que aún no eran parte de esa muestra. Esto puede repercutir en el coeficiente de endogamia; Márquez-Sánchez (2010) encontró que una VS formada con t HTs (Sin_t) contruidos con líneas puras no emparentadas tiene coeficiente de endogamia ($FSin_t$) igual a $(3m + 1)/(8tm)$. Según este resultado m se relaciona inversamente con $FSin_t$ y se reduce el número de genes $NIPD$ perdidos, como se prediciría, cuando $FSin_t$ es más grande.

Para Busbice (1970), el coeficiente de endogamia de una VS se relaciona lineal e inversamente con su media genotípica. En este contexto, los coeficientes de endogamia de una variedad sintética formada con un número par de L líneas puras (Sin_L) y otra formada con $L/2$ cruza simples (Sin_{CS}) hechas con las mis-

$E(Y_m)$ increases, and more rapidly when F is smaller. Furthermore, with pure lines ($F=1$), a sample of 10 individuals of a three-way cross is enough to reach the equilibrium in three $NIPD$ genes, and therefore, zero genes are lost. On the other hand, with $F=0.5$, this condition begins to be ostensible with $m=12$. It is also inferred that when $F=0.75$, starting from $m=15$ there is no loss of $NIPD$ genes.

With an increasingly larger sample its number of $NIPD$ genes tends to increase because each additional individual included is a new opportunity for it to have a genotype that supplies genes that were not yet part of that sample. This can have repercussions on the inbreeding coefficient; Márquez-Sánchez (2010) found that an SV formed with t THs (Sin_t) constructed with pure non-related lines has an inbreeding coefficient ($FSin_t$) equal to $(3m + 1)/(8tm)$. According to this result m is related inversely to $FSin_t$ and the number of $NIPD$ genes lost is reduced, which would predictable when $FSin_t$ is larger.

For Busbice (1970), the coefficient of endogamy of an SV is linearly and inversely related to its genotypic mean. In this context, the inbreeding coefficients of a synthetic variety formed with an even number of pure L lines (Sin_L) and another formed with $L/2$ simple crosses (Sin_{CS}) made with the same pure L lines are not dependent on m , are equal to $1/L$, and the genotypic means of the two synthetic varieties

mas L líneas puras, no dependen de m , son iguales a $1/L$, y las medias genotípicas de las dos variedades sintéticas deben ser iguales. Sin embargo, si L es un número par y múltiplo de tres, el coeficiente de endogamia de un Sin_t hecho con $L/3$ cruza trilineales sería mayor que $1/L$ según el resultado reportado por Márquez-Sánchez (2010). Además, en consistencia con la igualdad de sus coeficientes de endogamia, si $F=1$ el Sin_{CS} y el Sin_L nunca pierden genes *NIPD* cuando se forma la muestra de los m representantes de cada progenitor; también, los m representantes de cada progenitor tienen el mismo genotipo. El Sin_p , en cambio, sí pierde genes *NIPD* (Cuadro 4). Sahagún-Castellanos y Villanueva-Verduzco (2012) reportan pérdidas de genes *NIPD* en sintéticos formados únicamente con cruza dobles o simples ($F<1$).

CONCLUSIONES

Desde la formación de los m individuos de cada una de las cruza trilineales progenitoras de una variedad sintética, puede ocurrir pérdida de genes no idénticos por descendencia (*NIPD*); ésta es mayor cuando m es más pequeña y sus líneas progenitoras tienen un coeficiente de endogamia (F) menor. En cambio, si $F=1$ basta una muestra de tamaño ocho ($m=8$) para que esta pérdida se reduzca a cero. Si $F=0.5$ y $m=12$ el número esperado de genes *NIPD* perdidos es 0.06. Además, el desbalance de las frecuencias génicas de los progenitores de cada cruza trilineal produce un coeficiente de endogamia mayor en el Sin_t que en el Sin_L y esto afecta negativamente la media genotípica del primero.

LITERATURA CITADA

- Busbice, T. H. 1970. Predicting yields of synthetic varieties. *Crop Science* 10: 260-269.
- Escalante-González, J. L., J. Sahagún-Castellanos, J. E. Rodríguez-Pérez, y A. Peña-Lomelí. 2013. Erosión genética en las cruza simples progenitoras de una variedad sintética. *Rev. Chapingo Serie Hort.* 19: 151-161.
- Espinosa, C. A., M. Sierra M., y N. Gómez M. 2002. Producción y tecnología de semillas mejoradas de maíz por el INIFAP en el escenario sin la PRONASE. *Agron. Mesoam.* 14: 117-121.
- Luna, M., A. Hinojosa R., O. Ayala G., F. Castillo G., y F. J. Mejía C. 2012. Perspectivas de desarrollo de la industria semillera de maíz en México. *Rev. Fitotec. Mex.* 35: 1-7.

should be equal. However, if L is an even number and multiple of three, the inbreeding coefficient of a Sin_t made with $L/3$ three-way crosses would be greater than $1/L$ according to the result reported by Márquez-Sánchez (2010). Furthermore, in consistence with the equality of its inbreeding coefficients, if $F=1$ the Sin_{CS} and the Sin_L never lose *NIPD* genes when the sample is formed from the m representatives of each one parent; also, the m representatives of each parent have the same genotype. On the other hand, the Sin_t loses *NIPD* genes (Table 4). Sahagún-Castellanos and Villanueva-Verduzco (2012) report losses of *NIPD* genes in synthetics formed only with double or single crosses ($F<1$).

CONCLUSIONS

Starting at the formation of the m individuals of each one of the parent three-way crosses, loss of *NIPD* genes may occur; this is higher when m is smaller and its parent lines have a lower inbreeding coefficient (F). In contrast, if $F=1$ a sample size of eight ($m=8$) is enough for this loss to be reduced to zero. If $F=0.5$ and $m=12$, the expected number of *NIPD* genes lost is 0.06. In addition, the imbalance of the genetic frequencies of the parents of each three-way cross produces a higher inbreeding coefficient in the Sin_t than in the Sin_L , which has a negative effect on the genotypic mean of the former.

—End of the English version—

-----*

- Márquez-Sánchez, F. 2010. Inbreeding coefficient and mean of maize synthetics of three-way lines hybrid. *Maydica* 55: 227-229.
- Pérez-López, F. J., R. Lobato-Ortiz, J. de J. García-Zavala, J. D. Molina-Galán, J. de J. López-Reynoso, y T. Cervantes-Santana. 2014. Líneas homocigóticas de maíz de alto rendimiento como progenitores de híbridos de cruza simple. *Agrociencia* 48: 425-437.
- Sahagún-Castellanos, J., y C. Villanueva-Verduzco. 2012. ¿Variedades sintéticas derivadas de cruza simples o de cruza dobles? *Rev. Chapingo Serie Hort.* 17(3):107-115.