

# ¿VARIEDADES SINTÉTICAS DERIVADAS DE CRUZAS SIMPLES O DE CRUZAS DOBLES?

Jaime Sahagún-Castellanos\* y Clemente Villanueva-Verduzco

Departamento de Fitotecnia, Universidad Autónoma Chapingo. km 38.5 Carretera México-Texcoco, C. P. 56230, Chapingo Estado de México

(\*Autor para correspondencia) Correo-e: jsahagunc@yahoo.com.mx

## RESUMEN

Las variedades sintéticas (VSs) formadas con cruza simple ( $Sin_{cs}$ ) o dobles ( $Sin_{cd}$ ) permiten evitar el alto costo de semilla de variedades híbridas como las de maíz (*Zea mays* L.) y cebolla (*Allium cepa* L.). Además, con estas VSs se podrían predecir más económicamente las medias genotípicas (MGs) de las VSs que se generarían con las líneas progenitoras de tales híbridos ( $Sin_L$ ). Sin embargo, se sabe que el  $Sin_L$ ,  $Sin_{cs}$  y el  $Sin_{cd}$  generados con las mismas líneas pueden tener coeficientes de endogamia (CEs) diferentes, pero se desconoce por qué. Este desconocimiento limita al fitomejorador para hacer que el  $Sin_{cs}$  y  $Sin_{cd}$  representen mejor al  $Sin_L$  y para mejorar la estimación de su MG. El objetivo de este trabajo fue determinar qué hace que el  $Sin_L$ ,  $Sin_{cs}$  y  $Sin_{cd}$  puedan diferir. Con  $L$  líneas no emparentadas cuyo CE es  $F_L$  se formaron  $L/2$  cruza simple (CSs) y  $L/4$  cruza dobles (CDs); cada línea y cada híbrido se representó por  $m$  plantas, y por apareamiento aleatorio de los progenitores correspondientes se formó el  $Sin_L$ ,  $Sin_{cs}$  y  $Sin_{cd}$ . Se encontró que para valores fijos de  $m$  y de  $F_L$ , en general, los arreglos genotípicos de las tres VSs difieren porque con las mismas cuatro líneas progenitoras de una CD se aporta a la VS un número más grande de genes no idénticos por descendencia en la medida en que menos líneas integren un progenitor [línea (1), cruza simple (2), cruza doble (4)]. Así, el  $Sin_L$  y el  $Sin_{cd}$  tienen el menor y mayor CE, respectivamente, y se espera que las MGs tengan el orden de magnitud opuesto. Además, los arreglos genotípicos de las VSs difieren más a medida que  $F_L$  y  $m$  sean más pequeños. Por el contrario, cuando las líneas son puras, los del  $Sin_L$  y el  $Sin_{cs}$  son iguales y los coeficientes de endogamia deben ser iguales también. Para que el  $Sin_{cd}$  también sea igual al  $Sin_L$  se requiere, además de líneas puras, un valor grande de  $m$ . Sin embargo, aun con  $F_L < 1$ , en la medida en que  $m$  es más grande el  $Sin_{cs}$  y el  $Sin_{cd}$  también tienden a  $Sin_L$ , aunque más lentamente, en particular en el  $Sin_{cd}$ .

**PALABRAS CLAVE ADICIONALES:** *Allium cepa* L., *Zea mays* L., endogamia, variedades sintéticas, híbridos, arreglo genotípico.

## SYNTHETIC VARIETIES DERIVED FROM SINGLE OR DOUBLE CROSSES?

### ABSTRACT

The use of synthetic varieties (SVs) derived from single ( $Syn_{sc}$ ) or double ( $Syn_{dc}$ ) crosses is a way of avoiding the high seed cost of hybrid maize (*Zea mays* L.) and onion (*Allium cepa* L.) varieties, among others. In addition, these SVs could be used to more economically predict the genotypic means (GMs) of the SVs derived from the hybrid parental lines ( $Syn_L$ ). It is known, however, that the inbreeding coefficients (ICs) of the  $Syn_L$ ,  $Syn_{sc}$ , and  $Syn_{dc}$  derived from the same lines can differ, but the reason why is unknown. This lack of knowledge makes it difficult for the breeder to improve the representation of the  $Syn_L$  through the  $Syn_{dc}$  and  $Syn_{sc}$  and thereby to improve the estimation of the  $Syn_L$  GM. The objective of this study was to determine the reason why the  $Syn_L$ ,  $Syn_{sc}$ , and  $Syn_{dc}$  can differ. From unrelated  $L$  lines whose IC is  $F_L$ ,  $L/2$  single (SCs) and  $L/4$  double crosses (DCs) were derived; and each line, SC, and DC were represented by  $m$  individuals. By randomly mating each set of parents, the  $Syn_L$ ,  $Syn_{sc}$ , and  $Syn_{dc}$  were formed. It was found that the genotypic arrays of the three SVs differ because with the same 4 initial lines the number of non-identical by descent genes contributed to the genotypic array is larger as fewer lines are involved in a parent [lines (1), single crosses (2), and double crosses (4)]. Thus, the ICs of  $Syn_L$  and  $Syn_{dc}$  are the smallest and largest, respectively; and it is expected that the GMs show the opposite ordering. In addition, as  $m$  and  $F_L$  are smaller the differences among the genotypic arrays of the SVs are larger. On the contrary, when the  $L$  lines are pure, the genotypic arrays of  $Syn_L$  and  $Syn_{sc}$  are equal, and their IC must be equal as well. To be equal to  $Sin_L$ ,  $Syn_{dc}$  requires, in addition, a large  $m$ . However, even with  $F_L < 1$ , as  $m$  is larger the genotypic arrays of  $Syn_{sc}$  and  $Syn_{dc}$  tend to the genotypic array of  $Syn_L$ , although slowly, particularly for  $Syn_{dc}$ .

**ADDITIONAL KEYWORDS:** *Allium cepa* L., *Zea mays* L., inbreeding, synthetic varieties, hybrids, genotypic array.

## INTRODUCCIÓN

Recientemente, las variedades sintéticas (VSs) de cultivos como la cebolla (*Allium cepa* L.) y el maíz (*Zea mays* L.), cuyos progenitores son cruza simples (Sahagún y Villanueva, 1997) o cruza dobles (Márquez-Sánchez, 2008), han sido tema de estudio. Entre los argumentos para justificar estos estudios se encuentran los siguientes: i) que las VSs resultantes deben ser poblaciones muy similares a las que se formarían con las líneas que generan los híbridos progenitores del sintético resultante (Sahagún, 2011); ii) que algunos agricultores en México forman variedades sintéticas para evitar el alto costo de la semilla de las variedades híbridas, ya que la semilla cosechada de una VS regenera la misma variedad sin cambio en la media genotípica, lo que sí sucede en los híbridos (Villanueva *et al.*, 1994), y iii) que la reducción del número de progenitores de la VS que se genera con dos, tres o cuatro líneas para formar un solo híbrido reduce los costos y trabajo necesarios para predecir el rendimiento de las VSs posibles de ser formadas con las líneas progenitoras de tales híbridos. Cuando el número de progenitores potenciales se reduce en una unidad, el número de poblaciones que se tiene que formar y evaluar experimentalmente para efecto de predicción según el método de Wright (1981), disminuye en un 50 %, aproximadamente.

Con el apareamiento aleatorio de  $L/2$  cruza simples se forma una variedad sintética que tiene el mismo coeficiente de endogamia que el de la VS que se formaría con el apareamiento aleatorio de las  $L$  líneas progenitoras de esos  $L/2$  híbridos (Márquez-Sánchez, 2007). Sin embargo, la VS formada con el apareamiento aleatorio de  $L/4$  cruza dobles tiene un coeficiente de endogamia que difiere del de la VS formada con  $L/2$  cruza simples construidas con todas las mismas  $L$  líneas (Sahagún, 2011). Para generar información que contribuya a determinar el origen y la magnitud de tal diferencia, debe ser útil determinar qué es la VS derivada con  $L$  líneas en relación a las VSs cuyos progenitores son  $L/2$  cruza simples o  $L/4$  dobles formadas, en ambos casos, con la participación de todas las  $L$  líneas. La información que se obtenga debe ser útil para evaluar el potencial de las VSs construidas con híbridos y para hacer predicciones más precisas y económicas, y debe contribuir al desarrollo de mejores variedades sintéticas. El objetivo de este estudio es determinar por qué las VSs formadas con las líneas con que se formaría un conjunto de híbridos de cruce simple o doble puede diferir del de las VSs que se formarían con tales híbridos.

## MÉTODOS Y MARCO TEÓRICO

En este estudio se emplearon la notación y el enfoque utilizados por Sahagún (2011) para el modelo de un *locus* de una especie diploide que se reproduce por apareamiento aleatorio. Brevemente, en este enfoque se utilizaron dos conceptos: si en una población el gen  $A_i$  y el genotipo  $AA_{ij}$

## INTRODUCTION

Synthetic varieties (SVs) of crops such as onion (*Allium cepa* L.) and maize (*Zea mays* L.), whose parents are single crosses (Sahagún and Villanueva, 1997) or double crosses (Márquez-Sánchez, 2008), have been the subject of recent study. Arguments that have been advanced to justify these studies include the following: i) the resulting SV populations must be very similar to those that would be formed with the lines that are the parents of the parental hybrids (Sahagún, 2011), ii) some Mexican farmers form synthetic varieties to avoid the high seed cost of hybrid varieties because the harvested seed of a SV regenerates the same variety without change in the genotypic mean, which does happen in hybrids (Villanueva *et al.*, 1994), and iii) reducing the number of parents of the SV derived from two, three or even four lines to form a single hybrid reduces the costs and effort required to predict the performance of the potential SVs that can be formed with the parent lines of these hybrids. When the number of potential parents is reduced by one, the number of populations that must be generated and evaluated experimentally for the prediction effect, according to the Wright (1981) method, decreases by approximately 50 %.

Random mating of  $L/2$  single crosses forms a synthetic variety with the same inbreeding coefficient as the SV that would be formed by random mating the  $L$  parent lines of these  $L/2$  hybrids (Márquez-Sánchez, 2007). However, the SV formed by random mating the  $L/4$  double crosses has an inbreeding coefficient that differs from the SV formed with  $L/2$  single crosses built with all the same  $L$  lines (Sahagún, 2011). To generate information that helps determine the source and extent of this difference, it should be useful to determine the SV derived from  $L$  lines relative to the SVs whose parents are  $L/2$  single or  $L/4$  double crosses formed, in both cases, with the participation of all the  $L$  lines. The information obtained should be useful in assessing the potential of the SVs constructed with hybrids and in making more accurate and economical predictions, and should contribute to the development of improved synthetic varieties. The objective of this study is to determine why the SVs derived from the lines that would form a set of single or double cross hybrids may differ from the SVs that would be formed with these hybrids.

## METHODS AND THEORETICAL FRAMEWORK

This study employs the notation and approach used by Sahagún (2011) for modeling a *locus* of a diploid species that reproduces by random mating. Briefly, this approach uses two concepts: if in a population the gene  $A_i$  and the genotype  $AA_{ij}$  have frequencies  $P_i$  and  $P_{ij}$ , gametic (GAA) and genotypic (GEA) arrays in such a population are defined as:

$$GAA = \sum_i P_i A_i \text{ and } GEA = \sum_i \sum_j P_{ij} A_i A_j \quad (1)$$

tienen frecuencias  $P_i$  y  $P_j$ , los arreglos gamético (AGA) y genotípico (AGE) de tal población se definen como:

$$AGA = \sum_i P_i A_i \text{ y } AGE = \sum_i \sum_j P_i P_j A_i A_j \quad (1)$$

Si en particular la reproducción de una población es por apareamiento aleatorio, como en el caso de las VSs, el arreglo genotípico de su progenie (AGEP) es igual al cuadrado del AGA de tal población. Por ejemplo, para una población cuyo AGA es el de la Ecuación 1, el apareamiento aleatorio debe producir la progenie cuyo AGEP es:

$$AGEP = (AGA)(AGA) = \sum_i \sum_j P_i P_j A_i A_j \quad (2)$$

Además, si los arreglos gaméticos de las poblaciones  $P_1$  y  $P_2$  ( $AGAP_1$  y  $AGAP_2$ ) son  $\sum_i P_i A_i$  y  $\sum_j q_j A_j$ , su cruce produce una progenie cuyo arreglo genotípico [ $AGE(P_1 \times P_2)$ ] es:

$$AGE(P_1 \times P_2) = (AGAP_1) \times (AGAP_2) = \sum_i \sum_j p_i q_j A_i A_j \quad (3)$$

Para la derivación de las variedades sintéticas (VSs) se partió de  $L$  líneas, formadas comúnmente por autofecundación, no emparentadas y con un coeficiente de endogamia  $F_L$ . Con la participación de todas las  $L$  líneas en cada tipo de híbridos, se consideró la formación de  $L/2$  cruces simples (CSs) y de  $L/4$  cruces dobles (CDs), y cada línea y cada híbrido se representó por  $m$  plantas. Luego, por el apareamiento aleatorio de las líneas, de las CSs y de las CDs por separado, se formaron los tres tipos de sintéticos. Las diferencias entre las VSs se expresaron en términos de sus arreglos genotípicos y sus consecuencias en términos de sus coeficientes de endogamia.

## DERIVACIÓN DE RESULTADOS

En la VS formada con las  $L$  líneas, si  $A_{ip1} A_{ip2}$  es el genotipo del  $p$ -ésimo individuo ( $p = 1, 2, \dots, m$ ) de la línea  $i$  ( $i = 1, 2, \dots, L$ ), la frecuencia del gameto  $A_{ipk}$  ( $k=1, 2$ ) debe ser  $(2mL)^{-1}$ , y el arreglo gamético de los  $mL$  individuos que representan a los  $L$  progenitores ( $AGAP_L$ ) de la VS es:

$$AGAP_L = \sum_{i=1}^L \sum_{p=1}^m \sum_{k=1}^2 (2mL)^{-1} A_{ipk} \quad (4)$$

El arreglo genotípico de la VS que resulta del apareamiento aleatorio de los  $mL$  individuos que representan a los  $L$  progenitores ( $AGES_L$ ) es, por lo tanto,

$$AGES_L = (2mL)^{-2} \sum_{i=1}^L \sum_{j=1}^L \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (5)$$

A continuación se investigará con detalle en qué medida la variedad sintética (VS) cuyos progenitores son las  $L$  líneas, puede ser visualizada como una población que resulta del apareamiento aleatorio de un grupo de híbridos ( $L/2$  cruces simples o  $L/4$  cruces dobles).

If a population is reproduced by random mating, as in the case of the SVs, the genotypic array of its progeny (GEAP) is equal to the square of the GAA of such a population. For example, for a population whose GAA is that of Equation 1, random mating should produce progeny whose GEAP is:

$$GEAP = (GAA)(GAA) = \sum_i \sum_j P_i P_j A_i A_j \quad (2)$$

Furthermore, if the gametic arrays of the populations  $P_1$  and  $P_2$  ( $GAAP_1$  and  $GAAP_2$ ) are and, their cross produces a progeny whose genotypic array [ $GEA(P_1 \times P_2)$ ] is:

$$GEA(P_1 \times P_2) = (GAAP_1) \times (GAAP_2) = \sum_i \sum_j p_i q_j A_i A_j \quad (3)$$

The derivation of the synthetic varieties (SVs) was based on  $L$  unrelated lines, commonly formed by self-pollination, and with an inbreeding coefficient  $F_L$ . With the participation of all  $L$  lines in each type of hybrid, the formation of  $L/2$  single crosses (SCs) and  $L/4$  double crosses (DCs) was considered, and every line and hybrid was represented by  $m$  plants. Then, by random mating of the lines, SCs and DCs on a separate basis, three types of synthetics were formed. The differences between the SVs were expressed in terms of their genotypic arrays and their consequences in terms of their inbreeding coefficients.

## DERIVATION OF RESULTS

In the SV derived from the  $L$  lines, if  $A_{ip1} A_{ip2}$  is the genotype of the  $p$ -th individual ( $p = 1, 2, \dots, m$ ) of the line  $i$  ( $i = 1, 2, \dots, L$ ), the frequency of the gamete  $A_{ipk}$  ( $k=1, 2$ ) must be  $(2mL)^{-1}$ , and the gametic array of the  $mL$  individuals that represent the  $L$  parents ( $GAAP_L$ ) of the SV is:

$$GAAP_L = \sum_{i=1}^L \sum_{p=1}^m \sum_{k=1}^2 (2mL)^{-1} A_{ipk} \quad (4)$$

The genotypic array of the SV resulting from the random mating of the  $mL$  individuals representing the  $L$  parents ( $GEAS_L$ ) is, therefore,

$$GEAS_L = (2mL)^{-2} \sum_{i=1}^L \sum_{j=1}^L \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (5)$$

Next, the extent to which the synthetic variety (SV) whose parents are the  $L$  lines can be viewed as a population that results from the random mating of a group of hybrids ( $L/2$  single crosses or  $L/4$  double crosses) will be investigated in detail.

### Single crosses

Suppose that  $L$ , the number of initial lines, is a multiple of 4 (although the results for the synthetic derived from single crosses are equally valid for even numbers of

## Cruzas simples

Supóngase que  $L$ , el número de líneas iniciales, es un múltiplo de 4 (aunque los resultados para el sintético de cruza simple son igualmente válidos para números pares de líneas). Supóngase ahora que en forma aleatoria con muestreo iterativo sin reposición de tamaño 2 se forman  $L/2$  subconjuntos de 2 líneas cada uno. Supóngase, también, sin pérdida de generalidad, que el  $s$ -ésimo par de líneas es el conjunto  $\{L_{s-1}, L_s\}$ ,  $s = 2, 4, \dots, L$ ; y que el genotipo del  $p$ -ésimo individuo ( $p = 1, 2, \dots, m$ ) de la línea  $s-1$  es  $A_{ip1}A_{ip2}$  ( $i = s-1$ ); similarmente, el genotipo del  $q$ -ésimo individuo ( $q = 1, 2, \dots, m$ ) de la línea  $s$  es  $A_{jq1}A_{jq2}$  ( $j = s$ ). Así, de acuerdo con la Ecuación 5, el arreglo genotípico de la población producida por el apareamiento aleatorio de los  $2m$  individuos que representan a las líneas  $L_{s-1}$  y  $L_s$  [(AGES) $_s$ ] debe ser:

$$(AGES)_s = (4m)^{-2} \sum_{i=s-1}^s \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (6)$$

$s = 2, 4, 6, \dots, \text{ó } L$

(AGES) $_s$  (Ecuación 6) produce a su vez el arreglo gamético [(AGAS) $_s$ ]:

$$(AGAS)_s = [2(2m)]^{-1} \sum_{i=s-1}^s \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad (7)$$

$s = 2, 4, 6, \dots, \text{ó } L$

Por otra parte, los arreglos gaméticos de las líneas  $L_{s-1}$  y  $L_s$  [(AGAL) $_{s-1}$  y (AGAL) $_s$ ] deben ser:

$$(AGAL)_{s-1} = (2m)^{-1} \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad (8)$$

$i = s-1; s = 2, 4, \dots, \text{ó } L$

y

$$(AGAL)_s = (2m)^{-1} \sum_{q=1}^m \sum_{k=1}^2 A_{jqk} \quad (9)$$

$j = s; s = 2, 4, \dots, \text{ó } L$

Según las Ecuaciones 8 y 9, el arreglo genotípico de la cruce simple  $L_{s-1} \times L_s$ ,  $s = 2, 4, \dots, \text{ó } L$  [(AGECS) $_s$ ] es:

$$(AGECS)_s = (2m)^{-2} \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (10)$$

$i = s-1; j = s; s = 2, 4, 6, \dots, \text{ó } L$

Evidentemente, (AGECS) $_s$  (Ecuación 10) produce el arreglo gamético [(AGACS) $_s$ ]:

$$(AGACS)_s = [2(2m)]^{-1} \sum_{i=s-1}^s \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad (11)$$

$s = 2, 4, 6, \dots, \text{ó } L$

Como se puede observar, los arreglos gaméticos de

líneas). Now suppose that with random iterative sampling without replacement of size 2,  $L/2$  subsets of 2 lines each are formed. Suppose, also, without loss of generality, that the  $s$ -th pair of lines is the set  $\{L_{s-1}, L_s\}$ ,  $s = 2, 4, \dots, L$ , and that the genotype of the  $p$ -th individual ( $p = 1, 2, \dots, m$ ) of the line  $s-1$  is  $A_{ip1}A_{ip2}$  ( $i = s-1$ ); similarly, the genotype of the  $q$ -th individual ( $q = 1, 2, \dots, m$ ) of the line  $s$  is  $A_{jq1}A_{jq2}$  ( $j = s$ ). Thus, according to Equation 5, the genotypic array of the population produced by random mating the  $2m$  individuals representing lines  $L_{s-1}$  and  $L_s$  [(GEAS) $_s$ ] must be:

$$(GEAS)_s = (4m)^{-2} \sum_{i=s-1}^s \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (6)$$

$s = 2, 4, 6, \dots, \text{ó } L$

(GEAS) $_s$  (Equation 6) produces in turn the gametic array [(GAAS) $_s$ ]:

$$(GAAS)_s = [2(2m)]^{-1} \sum_{i=s-1}^s \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad (7)$$

$s = 2, 4, 6, \dots, \text{ó } L$

On the other hand, the gametic arrays of the lines  $L_{s-1}$  and  $L_s$  [(GAAL) $_{s-1}$  and (GAAL) $_s$ ] must be:

$$(GAAL)_{s-1} = (2m)^{-1} \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad (8)$$

$i = s-1; s = 2, 4, \dots, \text{ó } L$

and

$$(GAAL)_s = (2m)^{-1} \sum_{q=1}^m \sum_{k=1}^2 A_{jqk} \quad (9)$$

$j = s; s = 2, 4, \dots, \text{ó } L$

According to Equations 8 and 9, the genotypic array of the single cross  $L_{s-1} \times L_s$ ,  $s = 2, 4, \dots, \text{ó } L$  [(GEASC) $_s$ ] is:

$$(GEASC)_s = (2m)^{-2} \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (10)$$

$i = s-1; j = s; s = 2, 4, 6, \dots, \text{ó } L$

Evidently, (GEASC) $_s$  (Equation 10) produces the gametic array [(GAASC) $_s$ ]:

$$(GAASC)_s = [2(2m)]^{-1} \sum_{i=s-1}^s \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad (11)$$

$s = 2, 4, 6, \dots, \text{ó } L$

As can be seen, the gametic arrays of a SV formed by random mating of the  $2m$  individuals representing two lines (Equation 7) and that of the progeny produced by the cross between them (Equation 11) are equal, although they come from different populations. They are different in their genotypic array, but their gene frequencies are equal. In the first instance, this means that the random mating between members of each of the two populations separately described in Equations 6 and 10 produces a synthetic whose



una VS generada por el apareamiento aleatorio de los  $2m$  individuos que representan dos líneas (Ecuación 7) y el de la progenie que produce la cruce entre ellas (Ecuación 11) son iguales, no obstante que provienen de poblaciones diferentes. Son diferentes en su arreglo genotípico, sí, pero sus frecuencias génicas son iguales. En primera instancia, esto significa que el apareamiento aleatorio entre los integrantes de cada una de las dos poblaciones por separado descritas en las Ecuaciones 6 y 10 produce un sintético cuyo arreglo genotípico es el mismo  $[(AGESCS)_s]$ :

$$(AGESCS)_s = [2(2m)]^{-2} \sum_{i=s-1}^s \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (12)$$

$s = 2, 4, \dots, \text{ó } L$

En segunda instancia, la mencionada igualdad de arreglos gaméticos (Ecuaciones 7 y 11) implica que la cruce entre dos poblaciones cuyos arreglos genotípicos son  $(AGES)_s$  y  $(AGES)_r$  ( $r \neq s$ , Ecuación 6) debe producir la misma progenie que generaría la cruce entre las subpoblaciones cuyos arreglos genotípicos son  $(AGECS)_s$  y  $(AGECS)_r$  [Ecuación 10]. En particular, de la Ecuación 11, para  $r \neq s$ :

$$\begin{aligned} (AGES)_r X (AGES)_s &= (AGESCS)_r X (AGESCS)_s \\ &= [2(2m)]^{-1} \left[ \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{k=1}^2 A_{ipk} \right] [2(2m)]^{-1} \left[ \sum_{j=s-1}^s \sum_{k=1}^2 A_{jqk} \right] \\ &= [2(2m)]^{-2} \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \end{aligned} \quad (13)$$

$r, s = 2, 4, 6, \dots, \text{ó } L; r \neq s$

Con base en las Ecuaciones 12 y 13, se mostrará cómo el arreglo genotípico de la VS cuyos progenitores son  $L$  líneas (Ecuación 5) puede ser visualizado como el de la VS generada por el apareamiento aleatorio de  $L/2$  cruces simples formadas cada una con una pareja de líneas definida por muestreo aleatorio de tamaño 2 sin reemplazo a partir del conjunto inicial de  $L$  líneas. En efecto, con sólo argumentos algebraicos, de la Ecuación 5 se obtiene, en principio, que:

$$\begin{aligned} AGES_L &= (2mL)^{-2} \sum_{i=1}^L \sum_{j=1}^L \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \\ &= [2(2m)L/2]^{-2} \left\{ \sum_{r=2}^L \sum_{s=2}^L \left[ \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right\} \end{aligned} \quad (14)$$

$r, s = 2, 4, \dots, L$

Además, cuando  $r = s$  en la Ecuación 14, se puede visualizar que con cada valor de  $s$  ( $s = 2, 4, \dots, \text{ó } L$ ) se definen los elementos del arreglo genotípico que produciría el apareamiento aleatorio de la progenie de la cruce simple  $L_{s-1} \times L_s$ , representable como  $[(AGACS_s)]^2$  (Ecuación 11). Por otra parte, para cada caso en que  $r \neq s$ , los genotipos de la Ecuación 14 que satisfacen esta condición pueden ser visualizados como los de la progenie que produciría la cruce entre las cruces simples cuyos arreglos genotípicos son:  $(AGECS)_r$  y  $(AGECS)_s$ ,  $r \neq s$ ; ó  $(AGESCS)_r$  y  $(AGESCS)_s$ ,

genotypic array is the same  $[(GEASSC)_s]$ :

$$(GEASSC)_s = [2(2m)]^{-2} \sum_{i=s-1}^s \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad (12)$$

$s = 2, 4, \dots, \text{ó } L$

In the second instance, the aforementioned equality of gametic arrays (Equations 7 and 11) implies that the cross between two populations whose genotypic arrays are  $(GEAS)_s$  and  $(GEAS)_r$  ( $r \neq s$ , Equation 6) must produce the same progeny generated by the cross between the subpopulations whose genotypic arrays are  $(GEASC)_s$  and  $(GEASC)_r$  [Equation 10]. In particular, from Equation 11, for  $r \neq s$ :

$$\begin{aligned} (GEAS)_r X (GEAS)_s &= (GEASSC)_r X (GEASSC)_s \\ &= [2(2m)]^{-1} \left[ \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{k=1}^2 A_{ipk} \right] [2(2m)]^{-1} \left[ \sum_{j=s-1}^s \sum_{k=1}^2 A_{jqk} \right] \\ &= [2(2m)]^{-2} \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \end{aligned} \quad (13)$$

$r, s = 2, 4, 6, \dots, \text{ó } L; r \neq s$

Based on Equations 12 and 13, it will be shown how the genotypic array of the SV whose parents are  $L$  lines (Equation 5) can be viewed as that of the SV generated by random mating of  $L/2$  single crosses, each formed with a pair of lines defined by random sampling of size 2 without replacement from the initial set of  $L$  lines. In effect, with only algebraic arguments, from Equation 5 it follows, in principle, that:

$$\begin{aligned} GEAS_L &= (2mL)^{-2} \sum_{i=1}^L \sum_{j=1}^L \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \\ &= [2(2m)L/2]^{-2} \left\{ \sum_{r=2}^L \sum_{s=2}^L \left[ \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right\} \end{aligned} \quad (14)$$

$r, s = 2, 4, \dots, L$

Moreover, when  $r = s$  in Equation 14, one can see that with each value of  $s$  ( $s = 2, 4, \dots, \text{ó } L$ ) the genotypic array elements that would be produced by random mating of the progeny of the single cross  $L_{s-1} \times L_s$ , representable as  $[(GAACS_s)]^2$  (Equation 11), are defined. On the other hand, for each case where  $s \neq r$ , the genotypes of Equation 14 that satisfy this condition can be viewed as those of the progeny that would be produced by the cross between single crosses whose genotypic arrays are:  $(GEASC)_r$  and  $(GEASC)_s$ ,  $r \neq s$ , or  $(GEASSC)_r$  and  $(GEASSC)_s$ ,  $r \neq s$  (Equation 13). In summary, from Equation 14 and the latter two considerations, it follows that:

$$\begin{aligned} GEAS_L &= [2(2m)L/2]^{-2} \left\{ \sum_{s=2}^L \left[ \sum_{i=s-1}^s \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right. \\ &\quad \left. + \sum_{r=2}^L \sum_{s=2}^L \left[ \sum_{i=r-1}^r \sum_{j=s-1}^s \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right\} \end{aligned}$$

$s = 2, 4, 6, \dots, L$

$r \neq s$  (Ecuación 13). En resumen, de la Ecuación 14 y de estas últimas dos consideraciones, resulta que:

$$AGES_L = \left[ 2(2m)L / 2 \right]^2 \left\{ \sum_{s=2}^L \left[ \sum_{i=r-1}^s \sum_{j=s-1}^m \sum_{p=1}^m \sum_{q=1}^2 \sum_{k=1}^2 A_{ipk} A_{jqk} \right] \right. \\ \left. + \sum_{r=2}^L \sum_{s=2}^L \left[ \sum_{i=r-1}^r \sum_{j=s-1}^m \sum_{p=1}^m \sum_{q=1}^2 \sum_{k=1}^2 A_{ipk} A_{jqk} \right] \right\} \\ r, s = 2, 4, 6, \dots, L; r \neq s$$

De las Ecuaciones 12 y 13, la ecuación anterior para  $AGES_L$  se puede escribir como

$$AGES_L = (2/L)^2 \left\{ \sum_{s=2}^L (AGESCS)_s + \sum_{r=2}^L \sum_{s=2}^L [(AGESCS)_r X (AGESCS)_s] \right\} \\ r \neq s; r, s = 2, 4, \dots, L \quad (15)$$

La Ecuación 15 y su origen implican que el  $AGES_L$  es igual al arreglo genotípico de la VS cuyos progenitores son las  $L/2$  cruza simples hasta ahora consideradas ( $AGES_{CS}^x$ ); es decir:

$$AGES_L = AGES_{CS}^x \quad (16)$$

Por su origen, la Ecuación 16 debe ser válida para cualquier conjunto de  $L/2$  pares de líneas que pueda producir el muestreo aleatorio sin reemplazo considerado para formar las  $L/2$  cruza simples. Sin embargo, debe tenerse en cuenta que las cruza simples del  $AGES_{CS}^x$  son virtuales; es decir, aparentes, no se hicieron realmente. Además, en la Ecuación 14, el término  $2mL$  de la expresión inicial se escribió posteriormente como  $2(2m)L/2$ , en congruencia con la idea de representar el arreglo genotípico de un sintético formado con  $L/2$  cruza simples, cada una representada por  $2m$  genotipos, que a su vez están formados por sendos pares de genes.

### Cruzas dobles

En el segundo término de la Ecuación 15 se puede visualizar la presencia de los genotipos que produciría la cruce entre los híbridos simples  $r$  y  $s$  ( $r \neq s$ ). El arreglo genotípico de tal cruce doble  $[(AGECD^x)_{rs}]$ , de acuerdo con la Ecuación 13, es:

$$(AGECD^x)_{rs} = \left[ 2(2m) \right]^2 \sum_{i=r-1}^r \sum_{j=s-1}^m \sum_{p=1}^m \sum_{q=1}^2 \sum_{k=1}^2 A_{ipk} A_{jqk} \\ r \neq s; r, s = 2, 4, 6, \text{ ó } L \quad (17)$$

Para simplificar, sin pérdida de generalidad, supóngase que las líneas involucradas en la formación de la cruce doble  $d$  ( $d = 4, 8, \dots, L$ ) se representan por  $L_{d-3}$  y  $L_{d-2}$  si son las que forman la  $r$ -ésima cruce simple y por  $L_{d-1}$  y  $L_d$  a las que forman la  $s$ -ésima cruce simple ( $r \neq s; r, s = 2, 4, 6, \dots, \text{ ó } L$ ).

$$r, s = 2, 4, 6, \dots, L; r \neq s$$

From Equations 12 and 13, the previous equation for  $GEAS_L$  can be written as:

$$GEAS_L = (2/L)^2 \left\{ \sum_{s=2}^L (GEASSC)_s + \sum_{r=2}^L \sum_{s=2}^L [(GEASSC)_r X (GEASSC)_s] \right\} \\ r \neq s; r, s = 2, 4, \dots, L \quad (15)$$

Equation 15 and its origin imply that the  $GEAS_L$  is equal to the genotypic array of the SV whose parents are the  $L/2$  single crosses hitherto considered ( $GEAS_{SC}^x$ ); that is:

$$GEAS_L = GEAS_{SC}^x \quad (16)$$

Due to its origin, Equation 16 should be valid for any set of  $L/2$  pairs of lines that can be produced by random sampling without replacement to form the  $L/2$  single crosses. However, it should be borne in mind that the single crosses related to  $GEAS_{SC}^x$  are virtual, i.e., they were not actually made. Furthermore, in Equation 14, the term  $2mL$  of the initial expression was subsequently written as  $2(2m)L/2$ , which is consistent with the idea of representing the genotypic array of a synthetic formed with  $L/2$  single crosses, each represented by  $2m$  genotypes, which in turn are each formed by a pair of genes.

### Double crosses

In the second term of Equation 15, one can visualize the presence of the genotypes that would be produced by the cross between the single-cross hybrids  $r$  and  $s$  ( $r \neq s$ ). The genotypic array of such a double cross  $[(GEADC^x)_{rs}]$ , according to Equation 13, is:

$$(GEADC^x)_{rs} = \left[ 2(2m) \right]^2 \sum_{i=r-1}^r \sum_{j=s-1}^m \sum_{p=1}^m \sum_{q=1}^2 \sum_{k=1}^2 A_{ipk} A_{jqk} \\ r \neq s; r, s = 2, 4, 6, \text{ or } L \quad (17)$$

For simplicity, without loss of generality, assume that the lines involved in the formation of the double cross  $d$  ( $d = 4, 8, \dots, L$ ) are represented by  $L_{d-3}$  and  $L_{d-2}$  if they form the  $r$ -th single cross, and by  $L_{d-1}$  and  $L_d$  if they form the  $s$ -th single cross ( $r \neq s, r, s = 2, 4, 6, \dots, \text{ or } L$ ). With this notation, the genotypic array of Equation 17  $[(GEADC^x)_{rs}]$  can be expressed as:

$$(GEADC^x)_d = \left[ 2m(2) \right]^2 \sum_{i=d-3}^{d-2} \sum_{j=d-1}^d \sum_{p=1}^m \sum_{q=1}^2 \sum_{k=1}^2 A_{ipk} A_{jqk} \\ d = 4, 8, \dots, \text{ or } L \quad (18)$$

The genotypic array  $(GEADC^x)_d$  (Equation 18) produces the following gametic array  $[(GAADC)_d]$ :

$$(GAADC)_d = \left[ 8m \right]^{-1} \sum_{i=d-3}^{d-2} \sum_{p=1}^m A_{ipk}$$

L). Con esta notación, el arreglo genotípico de la Ecuación 17  $[(AGECD^x)_d]$  se puede expresar como:

$$(AGECD^x)_d = [2m(2)]^{-2} \sum_{i=d-3}^{d-2} \sum_{j=d-1}^d \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad d = 4, 8, \dots, \text{ ó } L \quad (18)$$

El arreglo genotípico  $(AGECD^x)_d$  (Ecuación 18) produce el arreglo gamético  $[(AGACD)_d]$  siguiente:

$$(AGACD)_d = [8m]^{-1} \sum_{i=d-3}^d \sum_{p=1}^m \sum_{k=1}^2 A_{ipk} \quad d = 4, 8, \dots, \text{ ó } L \quad (19)$$

Por analogía (Ecuación 19), el arreglo gamético de la craza doble  $e$  ( $e \neq d$ ) es de la forma:

$$(AGACD)_e = [8m]^{-1} \sum_{j=e-3}^e \sum_{q=1}^m \sum_{l=1}^2 A_{jqk} \quad e = 4, 8, \dots, \text{ ó } L \quad e \neq d \quad (20)$$

De acuerdo con las Ecuaciones 19 y 20, el arreglo genotípico de la craza entre las cruza dobles  $d$  y  $e$  ( $d \neq e$ ),  $[(AGECD^x)_{de}]$  producido a su vez por el producto de sus arreglos gaméticos, es:

$$(AGECD^x)_{de} = [8m]^{-2} \sum_{i=d-3}^d \sum_{j=e-3}^e \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad d \neq e; d, e = 4, 8, \dots, L \quad (21)$$

Evidentemente, si  $e = d$ ,

$$(AGECD^x)_{dd} = [8m]^{-2} \sum_{i=d-3}^d \sum_{j=d-3}^d \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad d = 4, 8, \dots, L \quad (22)$$

Del arreglo genotípico de la VS formado por las  $L$  líneas iniciales (Ecuación 5), de la consideración de las Ecuaciones 21 y 22, y en congruencia con el álgebra, se genera un resultado de importancia:

$$\begin{aligned} AGES_L &= (2mL)^{-2} \sum_{i=1}^L \sum_{j=1}^L \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \\ &= [2(4m)(L/4)]^{-2} \left\{ \sum_{d=4}^L \left[ \sum_{i=d-3}^d \sum_{j=d-3}^d \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right. \\ &\quad \left. + \sum_{d=4}^L \sum_{e=4}^L \left[ \sum_{i=d-3}^d \sum_{j=e-3}^e \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right\} \\ &\quad d \neq e; d, e = 4, 8, \dots, L \end{aligned} \quad (23)$$

Evidentemente, la Ecuación 23, de acuerdo con las Ecuaciones 21 y 22, está formada por dos grupos de genotipos representados en sendos paréntesis rectangulares por: 1) el que contiene los genotipos que resultan del apa-

$$d = 4, 8, \dots, \text{ ó } L \quad (19)$$

By analogy (Equation 19), the gametic array of the double cross  $e$  ( $e \neq d$ ) is of the form:

$$(GAADC)_e = [8m]^{-1} \sum_{j=e-3}^e \sum_{q=1}^m \sum_{l=1}^2 A_{jqk} \quad e = 4, 8, \dots, \text{ ó } L \quad e \neq d \quad (20)$$

According to Equations 19 and 20, the genotypic array of the cross between the double crosses  $d$  and  $e$  ( $d \neq e$ )  $[(GEADC^x)_{de}]$ , produced in turn by the product of their gametic arrays, is:

$$(GEADC^x)_{de} = [8m]^{-2} \sum_{i=d-3}^d \sum_{j=e-3}^e \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad d \neq e; d, e = 4, 8, \dots, L \quad (21)$$

Clearly, if  $e = d$ ,

$$(GEADC^x)_{dd} = [8m]^{-2} \sum_{i=d-3}^d \sum_{j=d-3}^d \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \quad d = 4, 8, \dots, L \quad (22)$$

Based on the genotypic array of the SV formed by the initial  $L$  lines (Equation 5), Equations 21 and 22, and algebra, an important result is generated:

$$\begin{aligned} GEAS_L &= (2mL)^{-2} \sum_{i=1}^L \sum_{j=1}^L \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \\ &= [2(4m)(L/4)]^{-2} \left\{ \sum_{d=4}^L \left[ \sum_{i=d-3}^d \sum_{j=d-3}^d \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right. \\ &\quad \left. + \sum_{d=4}^L \sum_{e=4}^L \left[ \sum_{i=d-3}^d \sum_{j=e-3}^e \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 A_{ipk} A_{jqk} \right] \right\} \\ &\quad d \neq e; d, e = 4, 8, \dots, L \end{aligned} \quad (23)$$

Clearly, Equation 23, in accordance with Equations 21 and 22, is formed by two genotype groups represented in two square brackets by: 1) one containing the genotypes resulting from random mating among individuals shown here as the representatives of the double cross  $d$  (Equation 22), and 2) one containing the genotypes resulting from the cross between the double crosses  $d$  and  $e$  ( $d \neq e$ , Equation 21). This leads to the conclusion that the SV formed by  $L$  lines can also be viewed as a SV formed by  $L/4$  double crosses, built with the participation of all the  $L$  lines, each represented in turn by  $m$  plants. In summary, if the genotypic array of this SV is denoted by  $GEAS_{DC}^x$ , it has been found (Equation 23) that

$$GEAS_L = GEAS_{DC}^x \quad (24)$$

In addition, in accordance with Equations 16 and 24,

$$GEAS_L = GEAS_{SC}^x = GEAS_{DC}^x \quad (25)$$

reamiento aleatorio entre los individuos aquí visualizados como los representantes de la cruce doble  $d$  (Ecuación 22), y 2) el que contiene los genotipos que resultan del cruzamiento entre las cruces dobles  $d$  y  $e$  ( $d \neq e$ , Ecuación 21). Esto permite concluir que la VS formada por  $L$  líneas también puede ser visualizada como una VS formada por  $L/4$  cruces dobles, construidas con la participación de todas las  $L$  líneas, cada una representada a su vez por  $m$  plantas. Resumiendo, si el arreglo genotípico de esta VS se denota con  $AGES_{CD}^x$ , se ha encontrado (Ecuación 23) que

$$AGES_L = AGES_{CD}^x \quad (24)$$

Además, de acuerdo con las Ecuaciones 16 y 24,

$$AGES_L = AGES_{CS}^x = AGES_{CD}^x \quad (25)$$

Como en el caso del  $AGES_{CD}^x$ , el resultado de la Ecuación 24 sería el mismo para cualquier conjunto de  $L/4$  cruces de este tipo que se pueda generar con la participación de todas las  $L$  líneas. Sin embargo, similar al caso del  $AGES_{CS}^x$ , las cruces dobles del  $AGES_{CD}^x$  son virtuales; cada cruce doble es en realidad un conjunto de cuatro líneas.

## DISCUSIÓN

Considérese la variedad sintética (VS) formada por  $L$  líneas no emparentadas de una o más autofecundaciones ( $Sin_L$ ). Si cada línea se representa por  $m$  individuos y su coeficiente de endogamia es  $F_L$ , el coeficiente de endogamia de tal VS ( $FSin_L$ ) es igual a  $(1 + F_L)/(2L)$ . Además, de acuerdo con Sahagún (2011) la VS cuyos progenitores son las  $m$  plantas de cada una de las  $L/2$  cruces simples convencionales ( $Sin_{CS}$ ) formadas con la participación de todas las líneas con que se formó el  $Sin_L$  tiene un coeficiente de endogamia ( $FSin_{CS}$ ) igual a  $(1 + F_L)/(2L) + (1 - F_L)/(2mL)$ . Finalmente, según Sahagún (2011), la VS cuyos  $L/4$  progenitores son cruces dobles convencionales, cada una representada por  $m$  plantas, tiene el coeficiente de endogamia ( $FSin_{CD}$ ):

$$FSin_{CD} = \frac{3+m(1+F_L)-F_L}{2mL}$$

Los resultados previos sobre el coeficiente de endogamia (CE) parecen contradecir a los generados en este estudio para los sintéticos cuyos arreglos genotípicos son  $AGES_{CS}^x$  y  $AGES_{CD}^x$  y que son iguales al del  $Sin_L$  (Ecuaciones 16 y 24). Por esto y por la consideración de que  $FSin_L = (1 + F_L)/(2L)$ , debe ocurrir que respecto a sus CE ( $FS_{in_{CS}}^x$  y  $FS_{in_{CD}}^x$ ):

$$FS_{in_{CS}}^x = FS_{in_{CD}}^x = (1 + F_L)/(2L) \quad (26)$$

Sin embargo,  $FSin_{CS}$  y  $FSin_{CD}$  difieren entre sí y de  $(1 + F_L)/(2L)$ . Diferencias importantes entre coeficientes de endogamia (CEs) en este estudio son:

$$FSin_{CS} - FS_{in_{CS}}^x = (1 - F_L)/(2mL) \quad (27)$$

y

$$FSin_{CD} - FS_{in_{CD}}^x = (3 - F_L)/(2mL) \quad (28)$$

As in the case of  $GEAS_{SC}^x$ , the result of Equation 24 will be the same for any set of  $L/4$  crosses of this type that can be generated with the participation of all the  $L$  lines. However, similar to the case of  $GEAS_{SC}^x$ , the double crosses of  $GEAS_{DC}^x$  are virtual; each double cross is actually a set of four lines.

## DISCUSSION

Consider the synthetic variety (SV) formed by  $L$  unrelated lines of one or more selfings ( $Syn_L$ ). If each line is represented by  $m$  individuals and their inbreeding coefficient is  $F_L$ , the inbreeding coefficient of such a SV ( $FSyn_L$ ) equals  $(1 + F_L)/(2L)$ . Furthermore, according to Sahagún (2011), the SV whose parents are the  $m$  plants of each of the  $L/2$  conventional single crosses ( $Syn_{SC}$ ), formed with the participation of all the lines that formed the  $Syn_L$ , has an inbreeding coefficient ( $FSyn_{SC}$ ) equal to  $(1 + F_L)/(2L) + (1 - F_L)/(2mL)$ . Finally, according to Sahagún (2011), the SV whose  $L/4$  progenitors are conventional double crosses, each represented by  $m$  plants, has the inbreeding coefficient ( $FSyn_{DC}$ ):

$$FSyn_{DC} = \frac{3+m(1+F_L)-F_L}{2mL}$$

Previous results on the inbreeding coefficient (IC) appear to contradict those generated in this study for synthetics whose genotypic arrays are  $GEAS_{SC}^x$  and  $GEAS_{DC}^x$  which are equal to that of  $Syn_L$  (Equations 16 and 24). Based on this and the fact that  $FSyn_L = (1 + F_L)/(2L)$ , it must occur regarding their ICs ( $FSyn_{CS}^x$  and  $FSyn_{DC}^x$ ):

$$FSyn_{DC} = FSyn_{DC}^x = (1 + F_L)/(2L) \quad (26)$$

However,  $FSyn_{SC}$  and  $FSyn_{DC}$  differ from one another and from  $(1 + F_L)/(2L)$ . Important differences between inbreeding coefficients (ICs) in this study are:

$$FSyn_{SC} - FSyn_{DC}^x = (1 - F_L)/(2mL) \quad (27)$$

and

$$FSyn_{SC} - FSyn_{DC}^x = (3 - F_L)/(2mL) \quad (28)$$

Equations 27 and 28 imply that when  $F_L$  is less than 1, the inbreeding coefficients of the synthetics whose parents are conventional single crosses ( $FSyn_{SC}$ ) or conventional double crosses ( $FSyn_{DC}$ ) are greater than those of the SVs denoted here as  $S^{*}yn_{SC}$  and  $S^{*}yn_{DC}$  ( $FS^{*}yn_{SC}$  and  $FS^{*}yn_{DC}$ ), respectively. Clearly, the differences in Equations 27 and 28 are reduced as  $F_L$  approaches 1 and  $m$  is greater. With  $F_L = 1$ , they are reduced to 0 and  $1/mL$ , respectively; and when  $m$  is "large," the differences (Equations 27 and 28) are negligible (Table 1), and  $Syn_L$ ,  $Syn_{SC}$  and  $Syn_{DC}$  are virtually equal. Under these conditions, the breeder may validly opt to use conventional single or double crosses to form the SVs that would occur with the parental lines of such hybrids



Las Ecuaciones 27 y 28 implican que cuando  $F_L$  es menor que 1, los coeficientes de endogamia de los sintéticos cuyos progenitores son cruza simples convencionales ( $FSin_{cs}$ ) o cruza dobles convencionales ( $FSin_{cd}$ ) son mayores que los de las VSs aquí denominadas  $S_{in_{cs}}^x$  y  $S_{in_{cd}}^x$  ( $FS_{in_{cs}}^x$  y  $FS_{in_{cd}}^x$ , respectivamente). Evidentemente, las diferencias en las Ecuaciones 27 y 28 se reducen a medida que  $F_L$  se acerca a 1 y  $m$  es mayor. Con  $F_L = 1$  se reducen a 0 y  $1/(mL)$ , respectivamente; y cuando  $m$  es “grande”, las diferencias (Ecuaciones 27 y 28) son despreciables (Cuadro 1), y  $Sin_L$ ,  $Sin_{cs}$  y  $Sin_{cd}$  son prácticamente iguales. En estas condiciones, el fitomejorador puede válidamente optar por el uso de cruza simples o dobles convencionales para formar las VSs que se producirían con las líneas progenitoras de tales híbridos (o viceversa). Con esta estrategia se lograría una reducción de trabajo y recursos para efectos de predicción, en caso de disponer de tales híbridos. Los problemas de esta estrategia con relación al uso de VSs convencionales son la reducción de la precisión en la estimación de la media de la VS (Sahagún y Villanueva, 1997; Sahagún, 2011) y de la flexibilidad en términos del número y de la elección de progenitores.

A continuación se hará un análisis del origen de las diferencias que muestran las Ecuaciones 27 y 28 en términos de arreglos genotípicos. De las reflexiones que en este contexto se hagan, comúnmente se encontrarán evidencias concretas en términos de coeficientes de endogamia en el Cuadro 1. Según estas diferencias, cuando  $F_L < 1$ , y los valores de  $m$  son pequeños, estrictamente ni el arreglo genotípico de un sintético generado con  $L/2$  cruza simples ( $Sin_{cs}$ ) ni el de uno generado con  $L/4$  cruza dobles ( $Sin_{cd}$ ) pueden ser iguales al del sintético cuyos progenitores son las  $L$  líneas iniciales, que son las mismas con que se forman el  $Sin_{cs}$  y el  $Sin_{cd}$ . En el caso del  $Sin_{cs}$ , para que el apareamiento aleatorio de los  $m(L/2)$  individuos progenitores permita que todos los genes que no son idénticos por descendencia (NIPD) de las  $L$  líneas iniciales formen parte del arreglo genotípico del  $Sin_{cs}$  es necesario que  $F_L = 1$ . Cuando  $F_L < 1$  y  $m$  es pequeño, esto puede no ser posible; por ejemplo, con  $m = 1$ , mientras que la planta que representa una cruce simple aporta dos genes NIPD, las dos plantas que representan a las líneas progenitoras de tal cruce simple pueden aportar conjuntamente dos, tres o hasta cuatro genes NIPD al  $Sin_L$ , dependiendo del valor de  $F_L$ . Y esto hace que los arreglos genotípicos del  $Sin_L$  y del  $Sin_{cs}$  sean diferentes. Es claro, sólo si  $F_L = 1$ , en ambos casos, el número de genes NIPD aportados al arreglo genotípico debe ser dos.

En el caso del  $Sin_{cd}$ , puede no ser posible que todos los genes NIPD de las cuatro líneas progenitoras de una cruce doble se muestren en su arreglo genotípico, cualquiera que sea el valor de  $F_L$  y un valor pequeño de  $m$ . Con  $m = 1$ , por ejemplo, sólo dos de los cuatro o más genes NIPD que portan las cuatro líneas progenitoras pueden formar parte del arreglo genotípico del  $Sin_{cd}$ , puesto que

(or vice versa). With this strategy, less work and fewer resources would be needed for prediction purposes, in case of such hybrids being available. The problems of this approach compared to the use of conventional SVs are reduced accuracy in the estimation of the SV mean (Sahagún and Villanueva, 1997; Sahagún, 2011) and reduced flexibility in terms of the number and choice of parents.

Next, the origin of the differences shown in Equations 27 and 28 in terms of genotypic arrays will be analyzed. Of the reflections that are made in this context, concrete evidence would be commonly found in terms of inbreeding coefficients in Table 1. According to these differences, when  $F_L < 1$ , and the values of  $m$  are small, neither the genotypic array of a synthetic generated with  $L/2$  single crosses ( $Syn_{sc}$ ) or that of one generated with  $L/4$  double crosses ( $Syn_{dc}$ ) can be equal to that of the synthetic whose parents are the initial  $L$  lines, which are the same with which the  $Syn_{sc}$  and  $Syn_{dc}$  are formed. In the case of the  $Syn_{sc}$ , for the random mating of the  $m(L/2)$  individual parents to allow all genes that are not identical by descent (NIBD) from the initial  $L$  lines to form part of the genotypic array of the  $Syn_{sc}$ , it is necessary that  $F_L = 1$ . When  $F_L < 1$  and  $m$  is small, this may not be possible; for example, with  $m = 1$ , while the plant representing a single cross provides two NIBD genes, the two plants representing the parent lines of this single cross can bring together two, three or even four NIBD genes to the  $Syn_L$ , depending on the value of  $F_L$ . And this makes the genotypic arrays of the  $Syn_L$  and  $Syn_{sc}$  different. Clearly, only if  $F_L = 1$ , in both cases, the number of NIBD genes contributed to the genotypic array must be two.

In the case of the  $Syn_{dc}$ , it may not be possible for all of the NIBD genes of the four parental lines of a double cross to be shown in its genotypic array, regardless of the value of  $F_L$  and a small value of  $m$ . With  $m = 1$ , for example, only two of the four or more NIBD genes carried by the four parental lines can be part of the genotypic array of the  $Syn_{dc}$ , as these lines are unrelated. Thus, with smaller values of  $m$ , the  $Syn_{sc}$  and  $Syn_{dc}$  cannot have the same genotypic arrays.

Clearly, the differences between the genotypic arrays of synthetics formed with conventional single and double crosses ( $Syn_{sc}$  and  $Syn_{dc}$ , respectively) compared to that of  $Syn_L$  (and consequently to those of  $FSyn_{sc}^x$  and  $FSyn_{dc}^x$ ) when  $F_L < 1$  and  $m$  is small, should be consistent with the differences between the corresponding inbreeding coefficients. The results in Table 1 illustrate these general trends with realistic numbers of lines for maize (Kutka and Smith, 2007).

Finally, the results of this study in terms of inbreeding coefficients and genotypic arrays establish that synthetic varieties whose parents are  $L$  pure lines ( $F_L = 1$ ) are equal to synthetic varieties constructed with  $L/2$  single crosses formed with the participation of all  $L$  lines, regardless of the size of  $m$  and the form of integration of the pairs of lines that form the  $L/2$  single crosses. When  $F_L < 1$ , the difference between the inbreeding coefficients of the synthetics made with

**CUADRO 1.** Coeficientes de endogamia (X1000) de sintéticos de tres tipos, formados con la participación de las mismas  $L$  líneas: 1) Con las  $L$  líneas ( $L$ ), 2)  $L/2$  cruza simples (CS), y 3)  $L/4$  cruza dobles (CD) ( $L = 4, 8, 12, 16$ ). Cada progenitor está representado por  $m$  plantas y  $F_L$  es el coeficiente de endogamia de las líneas iniciales.

**TABLE 1.** Inbreeding coefficients (X1000) of three types of synthetics formed with the participation of the same  $L$  lines: 1) With the  $L$  lines: ( $L$ ), 2)  $L/2$  single crosses (SC), and 3)  $L/4$  double crosses (DC) ( $L = 4, 8, 12, 16$ ). Each parent is represented by  $m$  plants and  $F_L$  is the inbreeding coefficient of the initial lines.

$m$	$L$	CS / SC	CD / DC	$L$	CS / SC	CD / DC	$L$	CS / SC	CD / DC	$L$	CS / SC	CD / DC
	4	2	1	8	4	2	12	6	3	16	8	4
$F_L = 0.50$												
1	188	250	500	94	125	250	63	83	167	47	63	125
5	188	200	250	94	100	125	63	67	83	47	50	63
10	188	194	219	94	97	109	63	65	73	47	48	55
20	188	191	203	94	96	102	63	64	68	47	48	51
50	188	189	194	94	95	97	63	63	65	47	47	48
100	188	188	191	94	94	95	63	63	64	47	47	48
200	188	188	189	94	94	95	63	63	63	47	47	47
$F_L = 1.00$												
1	250	250	500	125	125	250	83	83	167	63	63	125
5	250	250	300	125	125	150	83	83	100	63	63	75
10	250	250	275	125	125	138	83	83	92	63	63	69
20	250	250	263	125	125	131	83	83	88	63	63	66
50	250	250	255	125	125	128	83	83	85	63	63	64
100	250	250	253	125	125	126	83	83	84	63	63	63
200	250	250	251	125	125	126	83	83	83	63	63	63

tales líneas no tienen parentesco. Así, con valores pequeños de  $m$  el  $Sin_{CS}$  y el  $Sin_{CD}$  no pueden tener arreglos genotípicos iguales.

Evidentemente, las diferencias de los arreglos genotípicos de los sintéticos formados con cruza simples y con cruza dobles convencionales ( $Sin_{CS}$  y  $Sin_{CD}$ , respectivamente) respecto al del  $Sin_L$  (y consecuentemente a los del  $S^{x}_{in_{CS}}$  y  $S^{x}_{in_{CD}}$ ) cuando  $F_L < 1$  y  $m$  es pequeño, son consistentes con las diferencias entre los coeficientes de endogamia correspondientes. Los resultados del Cuadro 1 ejemplifican estas tendencias generales con números realistas de líneas para maíz (Kutka y Smith, 2007).

Finalmente, los resultados que se han generado en este estudio en términos de coeficientes de endogamia y de arreglos genotípicos permiten establecer que las variedades sintéticas cuyos progenitores son  $L$  líneas puras ( $F_L = 1$ ) son iguales a las variedades sintéticas construidas con  $L/2$  cruza simples formadas con la participación de todas las  $L$  líneas, independientemente del tamaño de  $m$  y de la forma de integración de las parejas de líneas que forman las  $L/2$  cruza simples. Cuando  $F_L < 1$ , la diferencia entre los coeficientes de endogamia de los sintéticos hechos uno con cruza dobles convencionales y con cruza dobles virtuales,  $FSin_{CD} - FS^{x}_{in_{CD}}$  (Ecuación 27), depende de la magnitud de  $F_L$  y de  $m$ ; a medida que éstos sean más grandes, se reduce según la expresión  $(3 - F_L)/$

conventional double crosses and with virtual double crosses,  $FSyn_{DC} - FS^{x}_{yn_{DC}}$  (Equation 27) depends on the magnitude of  $F_L$  and  $m$ ; as they become larger, it decreases according to the expression  $(3 - F_L)/(2mL)$ . Moreover, the random mating of  $L/4$  double crosses that can be defined by sampling without replacement of size 2 taken from the set of  $L/2$  single crosses until depleting the  $L/2$  single crosses generates a genotypic array that, even with pure lines ( $F_L = 1$ ), differs from the  $Syn_L$ . In terms of IC, the difference is  $1/(mL)$ , which tends to zero when  $m$  is larger. If  $F_L < 1$ , the difference  $FSyn_{DC} - FSyn_L$  is greater and the speed with which it tends to zero is slower than in the case of single crosses.

## CONCLUSIONS

It was found that the synthetic variety (SV) produced by random mating of  $L$  unrelated lines whose inbreeding coefficient (IC) is  $F_L$ , and are each represented by two sets of  $m$  plants, has an IC ( $FSyn_L$ ) that is never greater than that of the SV derived from random mating of  $L/2$  single crosses ( $FSyn_{SC}$ ) or  $L/4$  double crosses ( $FSyn_{DC}$ ) generated with the same  $L$  lines. Only with pure lines ( $F_L = 1$ ),  $FSyn_L = FSyn_{SC}$ . When  $F_L < 1$ , the differences  $FSyn_L - FSyn_{SC}$  and  $FSyn_L - FSyn_{DC}$ , which are  $(1F_L)/(2mL)$  and  $(3 - F_L)/(2mL)$ , respectively, become marginal and the genotypic arrays of the three types of SVs are virtually identical when  $F_L$  and  $mL$  are large. The differences between the SVs are be-

( $2mL$ ). Por otra parte, el apareamiento aleatorio de las  $L/4$  cruza dobles que se puedan definir mediante muestreo sin reemplazo de tamaño 2 tomadas del conjunto de  $L/2$  cruza simples hasta agotar las  $L/2$  cruza simples genera un arreglo genotípico que, aun con líneas puras ( $F_L = 1$ ), difiere del  $FSin_L$ . En términos del CE, la diferencia es  $1/(mL)$ , que tiende a cero cuando  $m$  es más grande. Si  $F_L < 1$ , la diferencia  $FSin_{CD} - FSin_L$  es mayor y la velocidad con que tiende a cero es más lenta que en el caso de las cruza simples.

### CONCLUSIONES

Se encontró que la variedad sintética (VS) producida por el apareamiento aleatorio de  $L$  líneas no emparentadas cuyo coeficiente de endogamia (CE) es  $F_L$ , y son representadas por sendos conjuntos de  $m$  plantas, tiene un CE ( $FSin_L$ ) que nunca es mayor que el de la VS derivada del apareamiento aleatorio de  $L/2$  cruza simples ( $FSin_{CS}$ ) o  $L/4$  cruza dobles ( $FSin_{CD}$ ) generadas con las mismas  $L$  líneas. Sólo con líneas puras ( $F_L = 1$ ),  $FSin_L = FSin_{CS}$ . Cuando  $F_L < 1$ , las diferencias  $FSin_L - FSin_{CS}$  y  $FSin_L - FSin_{CD}$ , que son  $(1 - F_L)/(2mL)$  y  $(3 - F_L)/(2mL)$ , respectivamente, se vuelven marginales y los arreglos genotípicos de los tres tipos de VSs son prácticamente iguales cuando  $F_L$  y  $mL$  son grandes. Las diferencias entre las VSs se deben a que cada conjunto de cuatro líneas iniciales aporta al arreglo genotípico de la VS un número más grande de genes no idénticos por descendencia en la medida en que menos líneas integren cada tipo de progenitor [líneas (1), cruza simples (2) o cruza dobles (4)]. Esto es lo que hace que las tres VSs puedan diferir.

cause each set of four initial lines gives the genotypic array of the SV a larger number of genes not identical by descent as fewer lines make up each type of parent [lines (1), single crosses (2) or double crosses (4)]. This is why the three SVs may differ.

**End of English Version**

### LITERATURA CITADA

- KUTKA F., J.; SMITH M.E. 2007. How many parents give the highest yield in predicted synthetic and composite populations of maize? *Crop Science* 47: 1907-1913.
- MÁRQUEZ-SÁNCHEZ, F. 2007. Predicción de sintéticos de maíz hechos con cruza simples. *Agrociencia* 41: 271-274.
- MÁRQUEZ S., F. 2008. Endogamia y predicción de sintéticos de maíz de cruza dobles. *Revista Fitotecnia Mexicana* 31: 1-4.
- SAHAGÚN C., J. 1994. Sobre el cálculo de coeficientes de endogamia de variedades sintéticas. *Agrociencia Serie Fitociencia* 5(2): 67-78.
- SAHAGÚN C., J.; VILLANUEVA V., C. 1997. Teoría de las variedades sintéticas formadas con híbridos de cruza simple. *Revista Fitotecnia Mexicana* 20: 69-79.
- SAHAGÚN C., J. 2011. Inbreeding and yield of synthetic varieties derived from single and double cross hybrids. *Maydica* 56(3): 265-272.
- VILLANUEVA V., C.; CASTILLO G., F.; MOLINA G., J. D. 1994. Aprovechamiento de cruzamientos dialélicos entre híbridos comerciales de maíz: Análisis de progenitores y cruza. *Revista Fitotecnia Mexicana* 17: 175-185.
- WRIGHT A., J. 1981. The quantitative genetics of diploid synthetic varieties. *Proc. Fourth meeting of the section biometrics in plant breeding*. Poitiers, France, Sept. 2-4. pp. 137-157.