

ENDOGAMIA DE SINTÉTICOS FORMADOS CON LÍNEAS Y CRUZAS SIMPLES

Jaime Sahagún-Castellanos¹; Juan Enrique Rodríguez-Pérez

Instituto de Horticultura, Departamento de Fitotecnia, Universidad Autónoma Chapingo, km 38.5 Carretera México-Texcoco,

C. P. 56230 Chapingo, Estado de México. C. P. 56230. MÉXICO. Correo-e: jsahagunc@yahoo.com.mx (¹ Autor para correspondencia)

RESUMEN

La explotación de la heterosis en cultivos como el jitomate (*Lycopersicon esculentum* Mill.), la cebolla (*Allium cepa* L.), el maíz (*Zea mays* L.), etc., ha sido limitada por el alto costo de la semilla F_1 . Para evitar éste, los productores han recurrido al uso de semilla F_2 , F_3 ,... Sin embargo, la depresión endogámica que esta práctica produce ha hecho que los fitomejoradores consideren las variedades sintéticas, particularmente para los cultivos alógamos, las cuales, de todos modos, no están exentas de problemas. En este estudio se propone la formación de una variedad sintética (VS) con L líneas y s cruza simples ($Sin_{L,CS}$) como una forma de reducir el número de progenitores (originalmente $L+2s$ líneas) y con esto el costo y trabajo para predecir los VSs que se pueden formar. Sin embargo, los indicadores de calidad más importantes del $Sin_{L,CS}$ deben conocerse para hacerles una valoración objetiva. Para el $Sin_{L,CS}$ los objetivos fueron determinar: 1) su coeficiente de endogamia ($FSin_{L,CS}$) y 2) la mejor combinación de valores de L y s . La derivación de fórmulas se basó en el arreglo genotípico del $Sin_{L,CS}$ ($AGESin_{L,CS}$) y en la sustitución del genotipo general de este arreglo por la coancestría de los individuos que aportan los genes que lo forman (para $FSin_{L,CS}$). Y la mejor combinación de valores de L y s fue la que produjo el menor coeficiente de endogamia. Para un número par fijo de $L + 2s$ líneas homocigóticas no emparentadas, se hicieron dos descomposiciones del $AGESin_{L,CS}$ y se encontró que $FSin_{L,SC} = \{1 + Ls/[2(L+s)^2]\}/(L+2s)$, que alcanza su mínimo cuando los progenitores son sólo líneas ($s = 0$) o sólo cruza simples ($L = 0$). Sin embargo, con sólo cruza simples la predicción demanda menos recursos que la basada en todas las líneas.

PALABRAS CLAVE ADICIONALES: *Allium cepa* L., *Zea mays* L., coancestría, predicción, arreglo genotípico.

INBREEDING OF SYNTHETIC VARIETIES DERIVED FROM LINES AND SINGLE CROSSES

ABSTRACT

The exploitation of heterosis in crops such as tomato (*Lycopersicon esculentum* Mill.), onion (*Allium cepa* L.) and corn (*Zea mays* L.) has been limited by the high cost of F_1 seed. To avoid this, producers have resorted to using F_2 , F_3 ,... seed. However, the inbreeding depression that this practice produces has led plant breeders to consider the synthetic varieties (SVs), particularly for cross-pollinated crops, which are not, however, without problems. In this work, the derivation of a SV whose parents are L lines and s single crosses ($Syn_{L,SC}$) is proposed as a way of decreasing the parent number (originally $L+2s$ lines) and thereby also the costs and work involved in prediction. However, the most important $Syn_{L,SC}$ quality indicators must be known in order to make an objective assessment. Thus, the study objectives related to $Syn_{L,SC}$ were to determine: 1) the inbreeding coefficient ($FSyn_{L,SC}$) and 2) the best combination of L and s values. The derivation of $FSyn_{L,SC}$ was based on the replacement of the general genotype of the genotypic array of $Syn_{L,SC}$ ($GASyn_{L,SC}$) by the coancestry between the individuals that contribute the genes that make up this genotype, and the best combination of L and s values was the one that produced the lowest inbreeding coefficient. From a fixed even number of $L + 2s$ fully inbred and unrelated lines and from two decompositions of $GASyn_{L,SC}$, it was found that $FSyn_{L,SC} = \{1 + Ls/[2(L+s)^2]\}/(L+2s)$. By inspection of the formula, $FSyn_{L,SC}$ reaches its minimum in two cases: when $L = 0$ and when $s = 0$. However, the use of $(L+2s)/2$ single crosses requires fewer resources for prediction.

ADDITIONAL KEY WORDS: *Allium cepa* L., *Zea mays* L., coancestry, prediction, genotypic array.

INTRODUCCIÓN

El origen de las variedades sintéticas tiene que ver con la depresión endogámica que se observó en las generaciones avanzadas de los primeros híbridos de cruza simple de maíz (*Zea mays* L.) cuando éstos irrumpieron en el escenario agrícola y, posteriormente, en los híbridos de cruza doble y en los trilineales (Márquez, 1992). La observación de este fenómeno fue posible en México debido a que algunos agricultores, para evitar el alto costo de la semilla de las variedades híbridas en cada ciclo agrícola, sembraban la semilla que cosechaban de un híbrido, primero, y después de sus generaciones avanzadas (Villanueva *et al.*, 1994). La depresión endogámica también ha sido observada en cultivos hortícolas como el jitomate (*Lycopersicon esculentum* Mill.), según reportan Mendoza *et al.* (2010) y Martínez *et al.* (2005), y en cebolla (*Allium cepa* L.). Si bien las cruza simples, híbridos trilineales y cruza dobles se construyen con dos, tres y cuatro líneas, respectivamente, la pregunta natural a ser contestada fue la relativa al comportamiento general de las variedades formadas por más líneas. Wright (1922) se refirió a la predicción de la media genotípica de las poblaciones que pueden construirse por el apareamiento aleatorio de las líneas de cada uno de los subconjuntos de dos o más elementos de un conjunto de n líneas homocigóticas. A cualquier población así formada y a cualquiera de sus sucesivas generaciones producidas por apareamiento aleatorio se le denomina variedad sintética. Posteriormente, Busbice (1970) desarrolló una fórmula de predicción más general que la de Wright (1922) para este tipo de variedades. Según Busbice (1970), la media genotípica del sintético en la generación t es una función lineal del coeficiente de endogamia de la variedad en esa generación.

Respecto a la forma de desarrollar una variedad sintética, Sahagún y Villanueva (1997) propusieron el uso de híbridos de cruza simple como progenitores, en lugar de las líneas involucradas. Con esto, se pensó, se reduce en un 50 % el número de progenitores, lo que a su vez reduce el número de cruza entre ellos que, una vez evaluadas experimentalmente junto con los progenitores, permiten predecir las medias genotípicas de las variedades que se pueden formar con las cruza simples. En efecto, si con un número par p de líneas se pueden construir $2^p - (p+1)$ sintéticos de dos o más progenitores, con $p/2$ cruza simples formadas con las p líneas el número de sintéticos que se puede formar con uno o más de este tipo de progenitores es $2^{p/2} - p/2$. En el primer caso el número de cruza (directas y recíprocas) que hay que hacer y después evaluar experimentalmente es $p(p-1)$ y en el segundo es $(p/2)[p/2 - 1]$; con $p=12$, por ejemplo, estos números son 132 y 30, respectivamente. Sin embargo, con el uso de híbridos de cruza simple o doble como progenitores de sintéticos no es posible la predicción de sintéticos

INTRODUCTION

The origin of synthetic varieties is linked to the inbreeding depression observed in advanced generations of the first single-cross corn (*Zea mays* L.) hybrids when they broke onto the agricultural scene and later in double-cross and trilinear hybrids (Márquez, 1992). This phenomenon could be observed in Mexico because some farmers, in order to avoid the high seed cost of hybrid varieties in each crop season, first planted the seed harvested from a hybrid and then that of its advanced generations (Villanueva *et al.*, 1994). Inbreeding depression has also been observed in vegetable crops such as tomato (*Lycopersicon esculentum* Mill.), as reported by Mendoza *et al.* (2010) and Martínez *et al.* (2005), and onion (*Allium cepa* L.). While single crosses, trilinear hybrids and double crosses are constructed with two, three and four lines, respectively, the obvious question to be answered was in regards to the general performance of varieties derived from more lines. Wright (1922) referred to the prediction of the genotypic mean of populations that can be constructed by random mating of the lines of each of the subsets of two or more elements of a set of homozygous n lines. Any population thus formed and any of its successive generations produced by random mating is called a synthetic variety. Subsequently, Busbice (1970) developed a more general prediction formula than Wright's (1922) for such varieties. According to Busbice, the genotypic mean of the synthetic in generation t is a linear function of the inbreeding coefficient of the variety in that generation.

Regarding how to develop a synthetic variety, Sahagún and Villanueva (1977) proposed the use of single-cross hybrids as parents, instead of the lines involved. This was thought to reduce the number of parents by 50 %, which in turn reduces the number of crosses among them that, once evaluated experimentally along with the parents, allow predicting the genotypic means of varieties that can be formed with single crosses. In effect, if with an even number p of lines $2^p - (p+1)$ synthetics can be constructed from two or more parents, with $p/2$ single crosses formed with p lines the number of synthetics that can be formed with one or more of these kinds of parents is $2^{p/2} - p/2$. In the first case, the number of crosses (direct and reciprocal) that must be done and then evaluated experimentally is $p(p-1)$ and in the second it is $(p/2)[p/2 - 1]$; with $p=12$, for example, these numbers are 132 and 30, respectively. However, with the use of single- or double-cross hybrids as parents of synthetics, it is not possible to predict synthetics formed with an odd number of lines or synthetics that disperse with any of the parental lines of these hybrids.

In corn, for example, a synthetic variety attains its optimum yield with no more than 12 lines (Kutka and Smith, 2007). Accordingly, in the search for the best synthetic variety of corn that can be formed if one starts with 12 lines, the performance of synthetic varieties that can be constructed with two or more of these lines should be investigated.

formados con un número no de líneas ni de sintéticos que prescindan de ninguna de las líneas progenitoras de esos híbridos.

En maíz, por ejemplo, una variedad sintética alcanza su rendimiento óptimo con no más de 12 líneas (Kutka y Smith, 2007). De acuerdo con esto, en la búsqueda de la mejor variedad sintética de maíz que se puede formar si se parte de 12 líneas habría que investigar el comportamiento de las variedades sintéticas que se pueden construir con dos o más de estas líneas. Una solución intermedia al problema del número elevado de sintéticos que se puede construir (4,083) se ejemplifica para este caso particular de 12 líneas. Con $s = 4$ y $L = 4$ las VSs posibles de ser construidas con dos o más líneas son 252, y con $s = 5$ y $L = 2$ este número se reduce a 126. Sin embargo, habrá que considerar que en una variedad sintética formada con L líneas y s cruza simples ($Sin_{L'CS}$) se genera un desbalance en las frecuencias de los genes de las líneas progenitoras con relación a las frecuencias de los genes de las líneas que forman las cruza simples. Este desbalance puede tener efecto en el coeficiente de endogamia y en la media genotípica de la variedad sintética; además, habrá que afrontar el problema de determinar la mejor combinación de valores de L y s . Con el fin de generar información que aporte al desvanecimiento de estas dudas, se planeó este trabajo teórico en torno a un sintético formado con s cruza simples y L líneas ($Sin_{L'CS}$). Los objetivos fueron: 1) derivar fórmulas generales para el coeficiente de endogamia de $Sin_{L'CS}$, y 2) determinar el mejor diseño de progenitores para $Sin_{L'CS}$ (la mejor combinación de valores para s y L).

MÉTODOS Y MARCO TEÓRICO

Sea $A_{pi1}A_{pi2}$ el genotipo del p – ésimo ($p = 1, 2, \dots, m$) individuo que representa al progenitor i . Si $i = 1, 2, \dots, L$ el progenitor es una línea, y si $i = L+1, L+2, \dots, L+s$, el progenitor es una cruza simple. En lo sucesivo se usará esta notación y los conceptos y enfoque metodológico utilizados por Sahagún (1994). El arreglo genotípico de los progenitores (AGEP) es

$$AGEP = \sum_{p=1}^m \sum_{i=1}^{L+s} [1/m[(L+s)]] A_{pi1}A_{pi2}$$

De acuerdo con Busbice (1970), esta población es el “sintético 0”.

Similarmente, el arreglo gamético de estos $L + s$ progenitores (AGAP) es

$$AGAP = \sum_{p=1}^m \sum_{i=1}^{L+s} \sum_{k=1}^2 [1/[2m(L+s)]] A_{pik}$$

Con este AGAP, el arreglo genotípico de las cruza (directas y recíprocas) entre los $L + s$ progenitores (AGCP) es

An intermediate solution to the problem of the large number of synthetics that can be constructed (4,083) is illustrated by the case of 12 lines. With $s = 4$ and $L = 4$, the potential number of SVs that can be constructed with two or more lines is 252, and with $s = 5$ and $L = 2$ this number drops to 126. However, one must consider that in a synthetic variety formed with L lines and s single crosses ($Syn_{L'SC}$), an imbalance is created in the gene frequencies of the parental lines with respect to the gene frequencies of the lines that form the single crosses. This imbalance may have an effect on the inbreeding coefficient and the genotypic mean of the synthetic variety; one will also have to face the problem of determining the best combination of L and s values. In order to generate information that helps resolve these doubts, this theoretical work was planned around a synthetic formed with s single crosses and L lines ($Syn_{L'SC}$). The objectives were: 1) derive general formulae for the inbreeding coefficient of $Syn_{L'SC}$ and 2) determine the best parental design for $Syn_{L'SC}$ (the best combination of s and L values).

METHODS AND THEORETICAL FRAMEWORK

Let $A_{pi1}A_{pi2}$ be the genotype of the p -th ($p = 1, 2, \dots, m$) individual that represents the parent i . If $i = 1, 2, \dots, L$, the parent is a line, and if $i = L+1, L+2, \dots, L+s$, the parent is a single cross. This notation and the concepts and methodological approach used by Sahagún (1994) will hereinafter be used. The genotypic array of the parents (GEAP) is

$$GEAP = \sum_{p=1}^m \sum_{i=1}^{L+s} [1/m[(L+s)]] A_{pi1}A_{pi2}$$

According to Busbice (1970), this population is the “synthetic 0.”

Similarly, the gametic array of these $L + s$ parents (GAAP) is

$$GAAP = \sum_{p=1}^m \sum_{i=1}^{L+s} \sum_{k=1}^2 [1/[2m(L+s)]] A_{pik}$$

With this GAAP, the genotypic array of the crosses (direct and reciprocal) among the $L + s$ parents (GAPC) is

$$AGCP = \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^{L+s} \sum_{j=1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{(2m)^2 Ls(Ls-1)} A_{pik}A_{qjl}$$

The population represented by GAPC is often called “synthetic 1” (Busbice, 1970).

Since the genotypic array of the synthetic ($GASyn_{L'SC}$) is that of the population resulting from the random mating of the parents or their crosses, it follows that

$$GASyn_{L'SC} = \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^{L+s} \sum_{j=1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 [2m(L+s)]^{-2} A_{pik}A_{qjl} \quad (2)$$

$$AGCP = \sum_{p=1}^m \sum_{q=1}^m \sum_{i \neq j}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{(2m)^2 Ls(Ls-1)} A_{pik} A_{qjl} \quad (1)$$

La población representada por $AGCP$ se suele llamar “sintético 1” (Busbice, 1970).

Como el arreglo genotípico del sintético ($AGESin_{L',CS}$) es el de la población que resulta del apareamiento aleatorio de los progenitores o de sus cruza, resulta que

$$AGESin_{L',CS} = \sum_{p=1}^m \sum_{q=1}^m \sum_{i \neq j}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 [2m(L+s)]^{-2} A_{pik} A_{qjl} \quad (2)$$

De la Ecuación 2 se pueden generar diferentes expresiones para el coeficiente de endogamia de $Sin_{L',CS}$ ($FSin_{L',CS}$) mediante descomposiciones de su miembro derecho en partes que faciliten el cálculo, la interpretación y/o el manejo de este coeficiente de endogamia. Por ejemplo, la descomposición se puede hacer en partes que incluyan: 1) subpoblaciones derivadas de un mismo tipo de progenitores, y 2) subpoblaciones generadas por un mismo tipo de evento reproductivo (por ejemplo, por autofecundación, cruza intrapaternales y cruza interpaternales). Cada uno de estos procedimientos para determinar el coeficiente de endogamia del sintético debe producir el mismo resultado, pero las diferencias permiten observar y estudiar las contribuciones que hacen a este coeficiente los diferentes componentes del sintético.

La determinación de $FSin_{L',CS}$, de acuerdo con Sahagún *et al.* (2005), se basó en la expresión que resulta de la sustitución del genotipo $A_{pik} A_{qjl}$ del arreglo genotípico del sintético (Ecuación 2) por la coancestría entre los dos individuos que aportan sendos genes, A_{pik} y A_{qjl} . Esta coancestría es la probabilidad (P) de que A_{pik} y A_{qjl} sean idénticos por descendencia ($\equiv$) [$P(A_{pik} \equiv A_{qjl})$].

Se consideró que la mejor combinación de valores de L y s es la que produce el $Sin_{L',CS}$ cuyo $FSin_{L',CS}$ es el menor porque, de acuerdo con Busbice (1970), hay una relación inversa entre $FSin_{L',CS}$ y la media genotípica de la variedad sintética.

DERIVACIÓN DE RESULTADOS Y DISCUSIÓN

El coeficiente de endogamia de una variedad sintética derivada de L líneas y s híbridos de cruce simple ($FSin_{L',CS}$) puede ser expresado como una combinación lineal de los coeficientes de endogamia de dos variedades sintéticas, una construida con las L líneas ($FSin_L$) y la otra con las s cruza simples ($FSin_{sc}$), y el de la coancestría promedio entre líneas e híbridos ($\Gamma_{L \times CS}$). En efecto, de la descomposición del arreglo genotípico (Ecuación 2) acorde con esta combinación lineal y de su adecuación al coeficiente de endogamia, resulta que

From Equation 2 different expressions can be generated for the inbreeding coefficient of $Syn_{L',SC}$ ($FSyn_{L',SC}$) through decompositions of its right member in parts to facilitate the calculation, interpretation and/or management of this inbreeding coefficient. For example, the decomposition can be done in parts that include: 1) subpopulations derived from the same type of parents, and 2) subpopulations generated by the same type of reproductive event (for example, by selfing, intraparental crosses and interparental crosses). Each of these procedures to determine the inbreeding coefficient of the synthetic must produce the same result, but the differences allow observing and studying the contributions that the different components of the synthetic make to this coefficient.

The determination of $FSyn_{L',SC}$ according to Sahagún *et al.* (2005), was based on the expression resulting from the replacement of the genotype $A_{pik} A_{qjl}$ of the genotypic array of the synthetic (Equation 2) by coancestry between the two individuals that provide separate genes, A_{pik} and A_{qjl} . This coancestry is the probability (P) that A_{pik} and A_{qjl} are identical by descent ($\equiv$) [$P(A_{pik} \equiv A_{qjl})$].

It was concluded that the best combination of L and s values is the one that produces the $Syn_{L',SC}$ whose $FSyn_{L',SC}$ is the lowest because, according to Busbice (1970), there is an inverse relationship between $FSyn_{L',SC}$ and the genotypic mean of the synthetic variety.

DERIVATION OF RESULTS AND DISCUSSION

The inbreeding coefficient of a synthetic variety derived from L lines and s single-cross hybrids ($FSyn_{L',SC}$) can be expressed as a linear combination of the inbreeding coefficients of two synthetic varieties, one constructed with L lines ($FSyn_L$) and the other with s single crosses ($FSyn_{sc}$), and the average coancestry between lines and hybrids ($\Gamma_{L \times SC}$). In effect, from the decomposition of the genotypic array (Equation 2) in accordance with this linear combination and its adaptation to the inbreeding coefficient, it follows that

$$\begin{aligned} FSyn_{L,SC} = & \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) + \\ & + \sum_{p=1}^m \sum_{q=1}^m \sum_{i=L+1}^{L+s} \sum_{j=L+1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) \\ & + 2 \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^L \sum_{j=L+1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) \end{aligned} \quad (3)$$

What has been done to arrive at Equation 3 can also be interpreted as a decomposition of the average coancestry of the individuals that by random mating produce the SV. In addition, the identification of $FSyn_L$ and $FSyn_{sc}$ in Equation 3 results in the formula

$$FSin_{L,CS} = \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) +$$

$$+ \sum_{p=1}^m \sum_{q=1}^m \sum_{i=L+1}^{L+s} \sum_{j=L+1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl})$$

$$+ 2 \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^L \sum_{j=L+1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) \quad (3)$$

Lo que se ha hecho para llegar a la Ecuación 3 también se puede interpretar como una descomposición de la coancestría promedio de los individuos que por apareamiento aleatorio producen la VS. Además, de la identificación de $FSin_L$ y $FSin_{SC}$ en la Ecuación 3, resulta la fórmula

$$FSin_{L,CS} = [L/(L+s)]^2 FSin_L + [s/(L+s)]^2 FSin_{SC}$$

$$+ [2Ls/(L+s)^2] \Gamma_{L \times SC} \quad (4)$$

Similarmente, $FSin_{L',SC}$ puede expresarse en términos del coeficiente de endogamia promedio de las poblaciones que resultan del apareamiento aleatorio de los m individuos de cada progenitor (F_{PAA}) y de la coancestría de las $(L+s)$ ($L+s-1$) cruza directas y recíprocas entre los $L+s$ progenitores Γ_{CP} . Así, de la Ecuación 2 y de su adecuación al $FSin_{L',SC}$, expresado como una combinación lineal de F_{PAA} y de Γ_{CP} resulta que

$$FSin_{L,CS} = \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) +$$

$$\sum_{p=1}^m \sum_{q=1}^m \sum_{i \neq j}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl})$$

$$= (L+s)^{-1} F_{PAA} + [(L+s-1)/(L+s)] \Gamma_{CP} \quad (5)$$

Además, el coeficiente de endogamia de cada variedad sintética en la Ecuación 4 se puede descomponer en las partes que involucran las progenies producidas por: 1) autofecundación, 2) cruza entre individuos de cada progenitor, y 3) cruza entre individuos de diferentes progenitores (Busbice, 1970; Márquez-Sánchez, 1992; Sahagún-Castellanos, 1994). Con base en esta descomposición adicional, $FSin_{L',SC}$, además de (coancestría promedio entre líneas y cruza), incluye tres Γ_{CP} términos relacionados con las líneas:

Γ_{LA} : Coeficiente de endogamia promedio de la población producida por autofecundación de los Lm individuos que representan a las líneas.

Γ_{WL} : Coancestría promedio entre dos individuos que pertenecen a una misma línea.

Γ_{BL} : Coancestría promedio entre dos individuos que pertenecen a líneas diferentes.

$$FSyn_{L,SC} = [L/(L+s)]^2 FSyn_L + [s/(L+s)]^2 FSyn_{SC}$$

$$+ [2Ls/(L+s)^2] \Gamma_{L \times SC} \quad (4)$$

Similarly, $FSyn_L$ can be expressed in terms of the average inbreeding coefficient of populations resulting from randomly mating the m individuals of each parent (F_{RMP}) and the coancestry of the $(L+s)(L+s-1)$ direct and reciprocal crosses among the $L+s$ parents (Γ_{PC}). Thus, from Equation 2 and its adaptation to $FSyn_{L',SC}$, expressed as a linear combination of F_{RMP} and Γ_{PC} , it follows that

$$FSyn_{L,SC} = \sum_{p=1}^m \sum_{q=1}^m \sum_{i=1}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl}) +$$

$$\sum_{p=1}^m \sum_{q=1}^m \sum_{i \neq j}^{L+s} \sum_{k=1}^2 \sum_{l=1}^2 \frac{1}{[2m(L+s)]^2} P(A_{pik} \equiv A_{qjl})$$

$$= (L+s)^{-1} F_{RMP} + [(L+s-1)/(L+s)] \Gamma_{PC} \quad (5)$$

Furthermore, the inbreeding coefficient of each synthetic variety in Equation 4 can be decomposed into parts involving the progeny produced by: 1) selfing, 2) crosses among individuals from each parent, 3) crosses among individuals from different parents (Busbice, 1970; Márquez-Sánchez, 1992; Sahagún-Castellanos, 1994). Based on this further decomposition, $FSyn_{L',SC}$, plus $\Gamma_{L \times SC}$ (average coancestry between lines and crosses), includes three terms associated with the lines:

F_{LS} : Average inbreeding coefficient of the population produced by selfing of the Lm individuals representing the lines.

Γ_{WL} : Average coancestry between two individuals belonging to the same line.

Γ_{BL} : Average coancestry between two individuals belonging to different lines.

Three other terms that include $FSyn_{L',SC}$ are those involving individuals from single crosses (SCs) and are analogous to F_{LS} , Γ_{WL} and Γ_{BL} ; i.e., they are: F_{SCS} and Γ_{WSC} and Γ_{BSC} .

According to the two decompositions of the first two terms of Equation 4, $FSyn_{L',SC}$ can also be expressed as

$$FSyn_{L,SC} = \frac{L}{m(L+s)^2} F_{LS} + \frac{(m-1)L}{m(L+s)^2} \Gamma_{WL} + \frac{(L-1)L}{(L+s)^2} \Gamma_{BL}$$

$$+ \frac{s}{m(L+s)^2} F_{SCS} + \frac{(m-1)s}{m(L+s)^2} \Gamma_{WSC} + \frac{s(s-1)}{(L+s)^2} \Gamma_{BSC}$$

$$+ \frac{2Ls}{(L+s)^2} \Gamma_{L \times SC} \quad (6)$$

In the particular case of practical importance in which the $L+2s$ lines that form the $L+s$ parents are unrelated homozygous lines, the terms of Equation 6 take the following numerical values:

Otros tres términos que incluye $FSin_{L'CS}$ son los que involucran a los individuos de las cruza simples (CSs) y son análogos a F_{LA} , Γ_{WL} y Γ_{BL} ; es decir, son: F_{CSA} , Γ_{WCS} y Γ_{BCS} .

De acuerdo con las dos descomposiciones de los primeros dos términos de la Ecuación 4, $FSin_{L'CS}$ también puede expresarse como

$$FSin_{L,CS} = \frac{L}{m(L+s)^2} F_{LA} + \frac{(m-1)L}{m(L+s)^2} \Gamma_{WL} + \frac{(L-1)L}{(L+s)^2} \Gamma_{BL} + \frac{s}{m(L+s)^2} F_{CSA} + \frac{(m-1)s}{m(L+s)^2} \Gamma_{WCS} + \frac{s(s-1)}{(L+s)^2} \Gamma_{BCS} + \frac{2Ls}{(L+s)^2} \Gamma_{LXCS} \quad (6)$$

En el caso particular de importancia práctica en que los $L + 2s$ líneas que forman los $L + s$ progenitores son líneas homocigóticas no emparentadas, los términos de la Ecuación 6 toman los valores numéricos siguientes:

$$\begin{aligned} F_{LA} &= 1 \\ \Gamma_{WL} &= 1 \\ \Gamma_{BL} &= 0 \\ F_{CSA} &= 1/2 \end{aligned} \quad (7)$$

$$\begin{aligned} \Gamma_{WCS} &= 1/2 \\ \Gamma_{BCS} &= 0 \\ \Gamma_{LXCS} &= 0 \end{aligned}$$

De acuerdo con estos resultados (Ecuaciones 6 y 7), $FSin_{L'CS}$ se puede expresar como

$$FSin_{L,CS} = \frac{L}{m(L+s)^2} + \frac{(m-1)L}{m(L+s)^2} + \frac{s/2}{m(L+s)^2} + \frac{(m-1)s}{m(L+s)^2} \quad (8)$$

Si en la Ecuación 8 se suman el primer y segundo términos y, por separado, los dos restantes se genera la expresión siguiente:

$$FSin_{L,CS} = [L/(L+s)]^2 (1/L) + [s/(L+s)]^2 [1/(2s)] \quad (9)$$

En la Ecuación 9, $[L/(L+s)]^2$ y $[s/(L+s)]^2$ son las proporciones con que Sin_L y Sin_{CS} participan en $Sin_{L'CS}$ (Ecuación 4), y $1/L$ y $1/(2s)$ son sus coeficientes de endogamia (Busbice, 1970; Sahagún y Villanueva, 1997).

Respecto al coeficiente de endogamia de las $L + s$ poblaciones producidas por el apareamiento aleatorio de los m individuos de cada progenitor (Ecuación 5), la coancestría entre dos individuos es 1 si el progenitor es una línea y $1/2$ si es una cruza simple. Por estos resultados,

$$\begin{aligned} F_{LS} &= 1 \\ \Gamma_{WL} &= 1 \\ \Gamma_{BL} &= 0 \\ F_{SCS} &= 1/2 \\ \Gamma_{WSC} &= 1/2 \\ \Gamma_{BSC} &= 0 \\ \Gamma_{LXSC} &= 0 \end{aligned} \quad (7)$$

According to these results (Equations 6 and 7), can be expressed as

$$FSyn_{L,SC} = \frac{L}{m(L+s)^2} + \frac{(m-1)L}{m(L+s)^2} + \frac{s/2}{m(L+s)^2} + \frac{(m-1)s}{m(L+s)^2} \quad (8)$$

If in Equation 8 the first and second terms are added and, separately, the remaining two, the following expression is generated:

$$FSyn_{L,SC} = [L/(L+s)]^2 (1/L) + [s/(L+s)]^2 [1/(2s)] \quad (9)$$

In Equation 9, $[L/(L+s)]^2$ and $[s/(L+s)]^2$ are the proportions with which Syn_L and Syn_{SC} participate in $Syn_{L'SC}$ (Equation 4), and $1/L$ and $1/(2s)$ are their inbreeding coefficients (Busbice, 1970; Sahagún and Villanueva, 1997).

Regarding the inbreeding coefficient of the $L + s$ populations produced by randomly mating the m individuals from each parent (Equation 5), the coancestry between two individuals is 1 if the parent is a line and $1/2$ if it is a single cross. For these results,

$$\begin{aligned} F_{RMP} &= \frac{4m^2L(1) + 4m^2s(1/2)}{4m^2(L+s)} \\ &= \frac{2L+s}{2(L+s)} \end{aligned} \quad (10)$$

Due to Equation 10 and because the coancestry among parents, the other component of $FSyn_{L'SC}$ according to Equation 5, is zero ($\Gamma_{PC} = 0$), Equation 5, for the conditions covering Equation 7, reduces to

$$FSyn_{L,SC} = \frac{2L+s}{2(L+s)^2} \quad (11)$$

Equations 9 and 11 describe the inbreeding coefficient of the same synthetic variety (formed by L lines and s single crosses) and, of course, produce the same result. Interes-

$$F_{PAA} = \frac{4m^2L(1) + 4m^2s(1/2)}{4m^2(L + s)} = \frac{2L + s}{2(L + s)} \quad (10)$$

Por la Ecuación 10 y porque la coancestría entre progenitores, el otro componente de $FSin_{L',CS}$ según la Ecuación 5, es cero ($\Gamma_{CP} = 0$), la Ecuación 5, para las condiciones que abarca la Ecuación 7, se reduce a la forma

$$FSin_{L,CS} = \frac{2L + s}{2(L + s)^2} \quad (11)$$

Las Ecuaciones 9 y 11 describen el coeficiente de endogamia de una misma variedad sintética (formada por L líneas y s cruza simples) y, evidentemente, producen el mismo resultado. Interesantemente, con líneas homocigóticas no emparentadas $FSin_{L',CS}$ no depende de m (Ecuaciones 9 y 11). Esto desde luego representa una ventaja, pues se pueden hacer las adecuaciones en términos del tamaño de m que la disponibilidad de recursos del programa de mejoramiento y objetivos requieran.

De la expresión $FSin_{L',CS} = (2L + s)/[2(L + s)^2]$ (Ecuación 11) se desprende que el número de líneas (L) es más importante que el número de híbridos de cruce simple (s) en la determinación del valor del coeficiente de endogamia del sintético. Una línea contribuye a $FSin_{L',CS}$ el doble de lo que contribuye una cruce simple de la misma variedad sintética. Una explicación de la contribución relativa de las líneas y las cruza, que involucra argumentos genéticos, se basa en que cada progenitor (línea o cruce simple) está representado por un mismo número de individuos (m) y que el apareamiento es aleatorio. De esto se debe esperar que los números de autofecundaciones y cruza entre individuos de un mismo progenitor sean iguales, independientemente de que éste sea línea o cruce simple. Además, como las $L + 2s$ líneas originales son homocigóticas, la coancestría entre dos individuos de una misma línea (o el coeficiente de endogamia de su progenie) es 1 y la de dos individuos de una misma cruce simple es 1/2. Similarmente, la probabilidad de identidad por descendencia en todos los genotipos producidos por autofecundación (que es la coancestría de un individuo consigo mismo) es 1 si los individuos representan a una línea y es 1/2 si los individuos representan a una cruce simple. Resumiendo, como las autofecundaciones y las cruza entre individuos de una misma línea son iguales en número a las que ocurren en una cruce simple, pero el coeficiente de endogamia de las progenies producidas en las líneas (1) es el doble que el de las progenies producidas en las cruza (1/2), la contribución de las líneas al coeficiente de endogamia de una misma variedad sintética es el doble de la de las cruza simples. Esto es importante porque los sintéticos más rendidores, según la ecuación de predicción de Busbice (1970), y el conocimiento empírico, suelen ser los que tienen los coeficientes de endogamia más pequeños.

tingly, with unrelated homozygous lines, $FSyn_{L',SC}$ does not depend on m (Equations 9 and 11). This is certainly an advantage because adaptations can be made in terms of the size of m that the availability of breeding program resources and objectives require.

From the expression $FSyn_{L',CS} = (2L + s)/[2(L + s)^2]$ (Equation 11), it follows that the number of lines (L) is more important than the number of single-cross hybrids (s) in determining the inbreeding coefficient of the synthetic. A line contributes twice as much to $FSyn_{L',SC}$ as a single cross of the same synthetic variety. An explanation of the relative contribution of lines and crosses, involving genetic arguments, is based on each parent (line or single cross) being represented by the same number of individuals (m) and mating being random. It should therefore be expected that the numbers of selfings and crosses among individuals of the same parent are equal, regardless of whether it is a line or single cross. Furthermore, as the original $L + 2s$ lines are homozygous, the coancestry between two individuals of the same line (or the inbreeding coefficient of their progeny) is 1 and that of two individuals of the same single-cross is 1/2. Similarly, the probability of identity by descent in all genotypes produced by selfing (the coancestry of an individual with itself) is 1 if the individuals represent a line and is 1/2 if the individuals represent a single cross. Summing up, as the selfings and crosses among individuals of the same line are equal in number to those that occur in a single cross, but the inbreeding coefficient of the progeny produced in the lines (1) is twice that of the progeny produced in the crosses (1/2), the contribution of the lines to the inbreeding coefficient of the same synthetic variety is twice that of the single crosses. This is important because the highest-yield synthetics, according to Busbice's prediction equation (1970) and empirical knowledge, are usually those with the smallest inbreeding coefficients.

Regarding the magnitudes of the inbreeding coefficient, Equation 11 implies that the two synthetic extremes [the two formed by parents of the same type; only lines ($s = 0$) or only single crosses ($L = 0$)] have inbreeding coefficients that are equal [$FSyn_L = 1/L$ and $FSyn_{sc} = 1/(2s)$]. This is because in these two cases the gene frequencies are equal, and random mating, therefore, produces the same genotypic array in the two synthetics. In light of these results, for an even number of $L + 2s$ lines, the two SV extremes must both have an inbreeding coefficient equal to $1/(L + 2s)$. On the other hand, for a SV formed by L lines and s single crosses, the inbreeding coefficient (Equation 11) is $(2L + s)/[2(L + s)^2]$. It can also be found that:

$$FSyn_{L,SC} = \frac{1}{L + 2s} + \left[1 + \frac{Ls}{2(L + s)^2} \right] \quad (12)$$

According to Equation 12 and the results related to the two SV extremes, it is clear that for a fixed number of lines $FSyn_{L',SC}$ is higher when $L > 0$ and $s > 0$ than in the cases of the SV extremes. This means that SV ex-

Respecto a las magnitudes del coeficiente de endogamia, la Ecuación 11 implica que los dos sintéticos extremos [los dos formados por progenitores de un mismo tipo; sólo líneas ($s = 0$) o sólo cruza simples ($L = 0$)] tienen coeficientes de endogamia que son iguales [$FSin_L = 1/L$ and $FSin_{sc} = 1/(2s)$]. Esto se debe a que en estos dos casos las frecuencias de genes son iguales, y el apareamiento aleatorio, en consecuencia, produce el mismo arreglo genotípico en los dos sintéticos. Por estos resultados, para un número par de $L + 2s$ líneas las dos VSs extremas deben, ambas, tener un coeficiente de endogamia igual a $1/(L + 2s)$. En cambio, para una VS formada por L líneas y s cruza simples, el coeficiente de endogamia (Ecuación 11) es $(2L + s)/[2(L + s)^2]$. También se puede encontrar que:

$$FSin_{L,CS} = \frac{1}{L + 2s} + \left[1 + \frac{Ls}{2(L + s)^2} \right] \quad (12)$$

De acuerdo con la Ecuación 12 y los resultados relativos a las dos VSs extremas, es claro que para un número fijo de líneas $FSin_{L,CS}$ es mayor cuando $L > 0$ y $s > 0$ que en los casos de las VSs extremas. Esto significa que las VSs extremas, cuyos coeficientes de endogamia son iguales, son las de menor coeficiente de endogamia y presumiblemente las de mayor media genotípica. Por ejemplo, con ocho líneas y con base en los coeficientes de endogamia, los dos mejores diseños de progenitores son: 1) $L = 8$ y $s = 0$, y 2) $L = 0$ y $s = 4$. Con éstos, según la Ecuación 11, $FSin_{L,SC} = 0.125$ (o bien $FSin_L = 1/L = 0.125$). Para el resto de diseños de progenitores que incluyan las ocho líneas, los coeficientes de endogamia son mayores y diferentes. Por ejemplo: 1) $L = 6$ y $s = 1$ (0.1327), 2) $L = 4$ y $s = 2$ (0.1378), y 3) $L = 2$ y $s = 3$ (0.1400).

Respecto a la predicción, si bien la fórmula de Busbice (1970) permite predecir y comparar las medias genotípicas de sintéticos formados con diferentes diseños de progenitores basados en un mismo conjunto de líneas, su valor predictivo depende del grado de realismo de la suposición de que la relación entre la media genotípica y $FSin_{L,CS}$ es lineal.

CONCLUSIONES

Para variedades sintéticas formadas con el apareamiento aleatorio de L líneas y s cruza simples se derivaron dos fórmulas para su coeficiente de endogamia ($FSin_{L,CS}$) basada en la descomposición de la coancestría promedio de los individuos que por apareamiento aleatorio producen el sintético en partes de facilidad interpretativa y/o manejo. Se derivó la fórmula

$$FSin_{L,CS} = [L/(L + s)]^2 FSin_L + [s/(L + s)]^2 FSin_{CS} + [2Ls/(L + s)(2s)^2] \Gamma_{L,CS}$$

; en donde $FSin_L$ y $FSin_{CS}$ son el coeficiente de

tremes, whose inbreeding coefficients are equal, have the lowest inbreeding coefficient and presumably the highest genotypic mean. For example, with eight lines and based on the inbreeding coefficients, the two best parental designs are: 1) $L = 8$ and $s = 0$, and 2) $L = 0$ and $s = 4$. With these, according to Equation 11, $FSyn_{L,SC} = 0.125$ (or $FSyn_L = 1/L = 0.125$ and $FSyn_{sc} = 1/(2s) = 0.125$). For all other parental designs that include the eight lines, the inbreeding coefficients are greater and different. For example: 1) $L = 6$ and $s = 1$ (0.1327), 2) $L = 4$ and $s = 2$ (0.1378), and 3) $L = 2$ and $s = 3$ (0.1400).

Regarding the prediction, although Busbice's formula (1970) allows predicting and comparing the genotypic means of synthetics derived from different parental designs based on the same set of lines, its predictive value depends on the degree of realism of the assumption that the relationship between the genotypic mean and $FSyn_{L,SC}$ is linear.

CONCLUSIONS

For synthetic varieties derived from randomly mating L lines and s single crosses, two formulae were derived for their inbreeding coefficient ($FSyn_{L,SC}$) based on the decomposition of the average coancestry of the individuals that, by random mating, produced the synthetic in easy to interpret and/or manage parts. The formula derived was

$$FSyn_{L,SC} = [L/(L + s)]^2 FSyn_L + [s/(L + s)]^2 FSyn_{SC} + [2Ls/(L + s)(2s)^2] \Gamma_{L,SC}$$

; where $FSyn_L$ and $FSyn_{SC}$ are the inbreeding coefficient of the synthetic formed with L lines and that formed with s crosses, respectively, and $\Gamma_{L,SC}$ is the average coancestry coefficient between lines and single crosses. For unrelated homozygous lines, the expression was

$FSyn_{L,SC} = \{1 + Ls/[2(L + s)^2]\}/(L + 2s)$. This formula implies that $FSyn_{L,SC}$ reaches its maximum when the synthetic is formed with only lines ($s = 0$) or with only single crosses ($L = 0$), which is when the gene frequencies of the parents are equal. However, when the parents are only single crosses the prediction requires fewer resources and is the best parental design.

End of English Version

endogamia del sintético formado con las L líneas y el del formado con las s cruza, respectivamente, y $\Gamma_{L,CS}$ es el coeficiente de coancestría promedio entre líneas y cruza simples. Para líneas puras no emparentadas, se llegó a

la expresión $FSin_{L,CS} = \{1 + Ls/[2(L + s)^2]\}/(L + 2s)$. Esta fórmula implica que $FSin_{L,SC}$ alcanza su mínimo cuando el sintético es formado con sólo líneas ($s = 0$) o con sólo cruza simples ($L = 0$), que es cuando las frecuencias génicas de los progenitores son iguales. Sin embargo,

cuando los progenitores son sólo cruza simples la predicción demanda menos recursos y constituye el mejor diseño de progenitores.

LITERATURA CITADA

- BUSBICE, T. H. 1970. Predicting yields of synthetic varieties. *Crop Science* 10: 260-269.
- 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: 1905-1913.
- MÁRQUEZ-SÁNCHEZ, F. 1992. Inbreeding and yield prediction in synthetic maize cultivars made with parental lines: Basic methods. *Crop Science* 32: 271-274.
- MARTÍNEZ-SOLÍS, J.; PEÑA-LOMELÍ, A.; RODRÍGUEZ-PÉREZ, J. E.; VILLANUEVA-VERDUZCO, C.; SAHAGÚN-CASTELLANOS, J. 2005. Comportamiento productivo en híbridos y en sus respectivas poblaciones F_2 . *Revista Chapingo Serie Horticultura* 11(2): 299-307.
- MENDOZA DE JESÚS, V.; SAHAGÚN-CASTELLANOS, J.; RODRÍGUEZ-PÉREZ, J. E.; LEGARIA-SOLANO, J. P.; PEÑA-LOMELÍ, A.; PÉREZ-GRAJALES, M. 2010. Heterosis intervarietal de jitomate (*Lycopersicon esculentum* Mill.) tipo saladete indeterminado. *Revista Chapingo Serie Horticultura* 16(1): 57-66.
- SAHAGÚN-CASTELLANOS, 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: 60-80.
- SAHAGÚN CASTELLANOS, J.; RODRÍGUEZ PÉREZ, J. E.; PEÑA LOMELÍ, A. 2005. Predicting yield of synthetics derived from double crosses. *Maydica* 50(2): 129-136.
- VILLANUEVA V., C.; SÁNCHEZ DEL CASTILLO, F. G.; 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, S. 1922. The effects of inbreeding and cross-breeding of guinea pigs. *Tech. Bull. U.S. Dep. Agric.* 1112. Washington D.C.