

# COEFICIENTE DE ENDOGAMIA DEL MUESTREO COMBINADO PARA LA REGENERACIÓN DE RECURSOS FITOGENÉTICOS

**J. Sahagún-Castellanos.**

**Instituto de Horticultura, Departamento de Fitotecnia, Universidad Autónoma Chapingo;  
km 38.5, Carr. México-Texcoco, Chapingo, Estado de México. C. P. 56230, MÉXICO.  
Correo-e: jsahagun@correo.chapingo.mx.**

## RESUMEN

Para la regeneración de recursos fitogenéticos de especies monoicas, ciertos cruzamientos artificiales producen tamaños efectivos de población deseables, pero el costo es alto. Para evitar este costo, en lugar de tales cruzamientos se puede usar el apareamiento aleatorio de los individuos de la muestra obtenida completamente al azar, como en la población ideal (PI), o bien de los individuos de una muestra formada en dos etapas (PIDE) por los  $m$  elementos tomados al azar de cada una de  $n$  familias aleatorias de la población sujeta a regeneración. Adicionalmente, el concepto de tamaño efectivo de población como una medida de la eficiencia de las estrategias de muestreo y apareamiento de los individuos de la muestra para regenerar una población no es tan conocido como el coeficiente de endogamia. El objetivo del presente estudio fue determinar el coeficiente de endogamia del segundo ciclo de regeneración de PIDE ( $F_{2,E}$ ). La derivación se basó en el cálculo de la contribución a este coeficiente de cada una de seis fuentes, que ya se manifiestan en este ciclo, con base en un enfoque probabilístico aplicado al concepto de identidad por descendencia de dos genes. Se encontró que PIDE reduce  $F_{2,E}$  a medida que  $m$  es más pequeño, y alcanza su valor más bajo cuando  $m = 1$  y sólo cuando  $m = 1$  y  $n > 1$  es menor que el de PI. Esto implica que para un tamaño de muestra  $x$ , para efectos de regeneración de una población lo mejor es tomar una semilla de cada una de  $x$  familias.

**PALABRAS CLAVE ADICIONALES:** *Physalis ixocarpa* Brot.; familias de medios hermanos; identidad por descendencia; autofecundación; cruzas intrafamiliares.

## INBREEDING COEFFICIENT OF COMBINED SAMPLING FOR REGENERATION OF PLANT GENETIC RESOURCES

## ABSTRACT

For the regeneration of plant genetic resources of monoecious species certain artificial crosses produce desirable effective population sizes but the cost is high. To avoid the cost of these crosses, random mating of the sample can be used. The sample can be taken completely at random, as in the ideal population (IP), or randomly taken at each of two stages (IPTS): first  $n$  families and then  $m$  individuals of each selected family. Further, since the concept of effective population size as a measure of the effectiveness of the sampling and reproduction strategies to regenerate populations is not always well known as the concept of inbreeding coefficient is, it was derived for IPTS as a measure of the efficiency of the IPTS method of regeneration. This derivation was based on the calculation of the contribution of each of six sources (selfings and crosses) of the inbreeding coefficient already active in the second regeneration cycle ( $F_{2,E}$ ). The calculations were made based on a probabilistic approach applied to the concept of identity by descent. It was found that with IPTS,  $F_{2,E}$  decreases as  $m$  is closer to 1 and its minimum value occurs when  $m = 1$ , and only when  $m = 1$  and  $n > 1$  is smaller than IP's. This implies that, for regeneration purposes, for a sample of size  $x$ , a random individual from each of  $x$  random families should be included in the sample.

**ADDITIONAL KEY WORDS:** *Physalis ixocarpa* Brot.; half-sib families; identity by descent; selfing; intrafamiliar crosses.

## INTRODUCCIÓN

La eficiencia para conservar la variabilidad genética de una población, además de la tecnología necesaria para mantener viable el material genético, depende de la estrategia utilizada para muestrear la población y para aparear sus elementos a fin de regenerar la población. Idealmente, la muestra debe captar la variabilidad existente y maximizar la eficiencia de recursos utilizados. Para la regeneración de material genético también se debe usar una estrategia de apareamiento que minimice la probabilidad de pérdida de genes y el costo.

El tamaño finito de la muestra y el carácter incógnito de los alelos que ésta captura dificultan conocer el grado de eficiencia que se alcance en la colecta y regeneración de recursos fitogenéticos. Esto hace necesario que se tenga que recurrir a métodos que involucren medidas indirectas de la riqueza genética que la muestra contiene. Dos formas de evaluar la eficiencia de la regeneración de los recursos fitogenéticos de una población de individuos monoicos se basan en la determinación del tamaño (o número) efectivo de población, evaluados en términos de endogamia o en términos de la varianza de la frecuencia génica. Estos números son el tamaño de una población, denominada población ideal (Crow y Kimura, 1970), que tiene la misma tasa de endogamia, o la misma varianza de la frecuencia génica, respectivamente.

Las magnitudes de los tamaños efectivos dependen del tamaño de la muestra, del tipo y frecuencias de los apareamientos utilizados para regenerar la población, de la cantidad y viabilidad de los gametos que aporta cada individuo de la muestra, etc. Con un buen manejo de esos factores se puede incrementar la eficiencia de los métodos de conservación. Por ejemplo, en términos generales, en especies monoicas que se reproducen por apareamiento aleatorio, como el tomate de cáscara (*Physalis ixocarpa* Brot.) y la cebolla (*Allium cepa* L.), el aumento del tamaño de muestra hace crecer el tamaño efectivo de población tanto en términos de endogamia como en términos de varianza de la frecuencia génica. Similarmente, si el apareamiento de los miembros de la muestra se hace de manera tal que cada familia contribuya con la misma cantidad de progenie, los dos tamaños efectivos también aumentan (Hallauer y Miranda, 1981). Estos incrementos son positivos para la conservación y regeneración de recursos genéticos. Por el lado de los costos, sin embargo, estas dos estrategias pueden ser demasiado costosas; además de que los conceptos de tamaño efectivo en términos de endogamia y en términos de la varianza de la frecuencia génica suelen no ser muy familiares en el argot de los profesionales involucrados en la colecta, conservación y uso de recursos fitogenéticos.

Dos formas de tomar la muestra son las basadas en el muestreo completamente al azar de individuos y en el

muestreo de familias de medios hermanos seguido del de individuos dentro de cada familia seleccionada. Si la regeneración se hace por apareamiento aleatorio, se evitará el gasto de la polinización artificial. Adicionalmente, existe un concepto que permite evaluar debidamente la eficiencia de la regeneración de la variabilidad genética y que tiene una mayor penetración en el argot de los profesionales especializados que el concepto de tamaño efectivo de población: el coeficiente de endogamia. Éste puede ser calculado de manera que permita identificar la contribución de cada una de las fuentes de endogamia existentes, de una manera similar a la utilizada por Sahagún (2006) para la población ideal. La determinación por separado de las contribuciones a la endogamia de las autofecundaciones y las cruzas intrafamiliares (cuando hay familias) puede ser utilizada como una guía para el diseño de una eficiente estrategia de regeneración de la muestra.

El objetivo del presente trabajo fue determinar el coeficiente de endogamia de una población monoica como la población ideal (Falconer, 1989), excepto que se muestrea en una primera etapa a familias y en una segunda a individuos dentro de cada familia. Esta determinación se hizo para el segundo ciclo de regeneración en el que ya se manifiestan las diferentes formas de aportar a la endogamia (autofecundaciones, cruzas entre medios hermanos, etc.) y se calculó la aportación de cada una de las diversas fuentes.

## MÉTODOS Y MARCO TEÓRICO

La población a preservar se consideró como un conjunto de individuos monoicos, diploides que se reproducen por apareamiento aleatorio en ausencia de cualquier factor que afecte las frecuencias génicas. También se supuso que los individuos que la forman no son endogámicos ni tienen relación de parentesco. De esta población se consideró la extracción de una muestra al azar de  $mn$  individuos cuyo apareamiento al azar produce un conjunto de  $mn$  familias de medios hermanos. El conjunto de  $(2mn)^2$  genotipos así formado es el ciclo 1 ( $C_1$ ). Si, como lo ha considerado Sahagún (2006), en poblaciones formadas de manera similar el genotipo  $A_{pi1}A_{pi2}$  de  $C_1$  representa al individuo  $p$  ( $p = 1, 2, \dots, m$ ) de la familia  $i$  ( $i = 1, 2, \dots, n$ ), el arreglo genotípico del ciclo 1 ( $AGC_1$ ) se expresa como:

$$AGC_1 = \sum_{i=1}^n \sum_{j=1}^n \sum_{p=1}^m \sum_{q=1}^m \sum_{k=1}^2 \sum_{l=1}^2 \left[ 1/(2mn)^2 \right] A_{pi1} A_{qj1}$$

De acuerdo con Crow y Kimura (1970), Márquez-Sánchez (1992), Sahagún-Castellanos (1994) y Falconer (1989), y como se puede verificar directamente en el  $AGC_1$ , el coeficiente de endogamia del  $C_1$  es  $1/(2mn)$ . Y ésta es producida sólo por la autofecundación de los individuos que ocurre cuando el  $C_0$  se reproduce por apareamiento aleatorio.

Para formar el ciclo 2, de  $AGC_1$  se toma una muestra aleatoria de  $n$  familias de medios hermanos y de cada una

de éstas, también al azar, se toma una muestra de  $m$  individuos (medios hermanos). Los ciclos o regeneraciones del germoplasma subsecuentes se forman de la misma manera.

A diferencia del de  $C_1$ , el coeficiente de endogamia del  $C_2$  ( $F_{2,E}$ ) ya refleja las contribuciones que también se manifestarán en las generaciones subsecuentes, como son las autofecundaciones, las cruzas entre medios hermanos y las cruzas entre individuos que tienen otras relaciones de parentesco. Por esta razón, se hizo la derivación del coeficiente de endogamia para este ciclo, identificando la aportación a él de cada una de sus fuentes.

El apareamiento aleatorio de la muestra tomada del  $C_1$  para formar el  $C_2$ , formada en principio por  $m$  semillas seleccionadas al azar de cada uno de  $n$  individuos, puede producir genotipos formados por dos genes idénticos por descendencia. Estos genotipos son producidos por autofecundaciones, cruzas entre medios hermanos y cruzas entre individuos emparentados de diferentes familias. Para calcular el coeficiente de endogamia del  $C_2$  ( $F_{2,E}$ ), se determinó la frecuencia esperada con que pueden ocurrir las autofecundaciones, las cruzas entre medios hermanos y las cruzas entre individuos emparentados de familias diferentes. También se calcularon las probabilidades de que los individuos involucrados en las autofecundaciones y en las cruzas mencionadas formen parte de la muestra que generará el siguiente ciclo. Considerando que ya estaban en la muestra, se calculó la probabilidad de que ocurrieran tales autofecundaciones y cruzas y de que, dada la ocurrencia de estos eventos reproductivos, se produjeran genotipos formados por dos genes idénticos por descendencia para el  $C_2$ .

En primera instancia, los cálculos se enfocaron a la determinación de la probabilidad de la presencia en el  $C_2$  de un genotipo cualquiera formado por dos genes idénticos por descendencia ( $A_{pik}A_{pik}$ ). Esta probabilidad se multiplicó por el número de genes diferentes. El producto obtenido fue interpretado como el coeficiente de endogamia del  $C_2$  ( $F_{2,E}$ ).

## DERIVACIÓN DE RESULTADOS Y DISCUSIÓN

Primero se derivará la frecuencia esperada del genotipo  $A_{pik}A_{pik}$  en el ciclo dos ( $C_2$ ) producido por autofecundación, y después la producida por cada una de ciertas cruzas del  $C_1$ . Luego, estas frecuencias esperadas se multiplicarán por  $2mn$ ; el resultado debe ser el coeficiente de endogamia del  $C_2$  ( $F_{2,E}$ )

### Autofecundaciones

En el arreglo genotípico del  $C_1$ ,  $\sum_{i=1}^m \sum_{j=1}^n \sum_{p=1}^m \sum_{q=1}^n \sum_{k=1}^m \sum_{l=1}^n \left[ \frac{1}{(2mn)^2} \right] A_{pik}A_{qjl}$ , hay genotipos que por autofecundación pueden producir el genotipo  $A_{pik}A_{pik}$  del ciclo siguiente ( $C_2$ ) ( $p = 1, 2, \dots, m$ ;  $i = 1, 2, 3, \dots, n$  y  $k = 1, 2$ ). Entre ellos se incluye el

genotipo  $A_{pik}A_{pik}$ . Éste aparece en el arreglo genotípico del  $C_1$ , y la probabilidad de que sea incluido en la muestra debe ser el producto de la probabilidad de que la familia que lo contiene,

$$\frac{\binom{1}{1} \binom{mn-1}{n-1}}{\binom{mn}{n}},$$

sea parte de la muestra por la probabilidad de que entre los  $m$  individuos tomados al azar de esa familia de tamaño  $4mn$  se incluya el genotipo  $A_{pik}A_{pik}$ ,

$$\frac{\binom{1}{1} \binom{4mn-1}{m-1}}{\binom{4mn}{m}},$$

El producto de estas dos probabilidades es  $1/(4mn)$ . Ya en la muestra de  $mn$  individuos y dado que la reproducción ocurriría por apareamiento aleatorio, la frecuencia esperada del genotipo  $A_{pik}A_{pik}$  generado por la autofecundación de  $A_{pik}A_{pik}$  es  $(mn)^{-2}$ . Resumiendo, con lo que contribuye la autofecundación del genotipo  $A_{pik}A_{pik}$  del  $C_1$  a la frecuencia del genotipo  $A_{pik}A_{pik}$  en el  $C_2$   $[P(A_{pik}A_{pik})]$  es

$$P(A_{pik}A_{pik}) = (4mn)^{-1} (mn)^{-2} = 16(4mn)^{-3} \quad 1)$$

En el  $C_1$  también hay genotipos heterocigóticos de la forma  $A_{pik}A_{qjl}$  ( $qj \neq pik$ ) que por autofecundación pueden contribuir al  $C_2$  con genotipos  $A_{pik}A_{pik}$ . De acuerdo con el arreglo genotípico del  $C_1$ , estos genotipos son los contenidos en las expresiones

$$\sum_{q=1}^m \sum_{j=1}^n \sum_{l=1}^2 A_{pik}A_{qjl} \quad y \quad \sum_{q=1}^m \sum_{j=1}^n \sum_{l=1}^2 A_{qjl}A_{pik} \quad qj \neq pik$$

Esto implica que en el  $C_1$  hay  $2(2mn-1)$  genotipos heterocigóticos que contienen el gen  $A_{pik}$ . Como en el caso anterior, la probabilidad de que cada uno de éstos sea incluido en la muestra es  $(4mn)^{-1}$ ; la frecuencia esperada con que en el apareamiento aleatorio de la muestra ocurría la autofecundación de cada uno de estos genotipos es  $(mn)^{-2}$ , y de que en cada caso ésta produzca  $A_{pik}A_{pik}$  es  $1/4$ . Estos resultados implican que la aportación a la frecuencia esperada del genotipo  $A_{pik}A_{pik}$  en el  $C_2$  por esta vía  $[P(A_{pik}A_{pik})_2]$  es

$$P(A_{pik}A_{pik})_2 = 2(2mn-1)(4mn)^{-1} (mn)^{-2} (1/4) = (2mn-1)/(2mn)^3 \quad 2)$$

## Cruzas intrafamiliares

El apareamiento aleatorio del  $C_0$  produce familias de medios hermanos. En particular, en el  $C_1$  estará la familia cuyo progenitor común es  $A_{pik}A_{pik'}$  ( $k \neq k'$ ). Esta familia tiene el arreglo genotípico (AG):

$$AG = \left[ \frac{1}{2}A_{pik} + \frac{1}{2}A_{pik'} \right] \left[ \sum_{q=1}^m \sum_{j=1}^n \sum_{l=1}^2 \left( 1/(2mn) \right) A_{qjl} \right]$$

$$= \sum_{q=1}^m \sum_{j=1}^n \sum_{l=1}^2 \left[ 1/(4mn) \right] A_{pik}A_{qjl} + \sum_{q=1}^m \sum_{j=1}^n \sum_{l=1}^2 \left[ 1/(4mn) \right] A_{pik'}A_{qjl}$$

Para formar el  $C_2$  ocurrirá apareamiento aleatorio entre todos los individuos del  $C_1$ . Respecto a las cruzas entre los individuos de la familia del  $C_1$  cuyo progenitor común es  $A_{pik}A_{pik'}$  ( $k, k' = 1, 2; k \neq k'$ ) capaces de producir  $A_{pik}A_{pik}$  se incluyen las cruzas de la forma (directas y recíprocas)  $A_{pik}A_{pik} \times A_{pik}A_{qjl}$  ( $qj \neq pik$ ), además de la crusa  $A_{pik}A_{pik} \times A_{pik'}A_{pik}$  y la de su recíproca  $A_{pik'}A_{pik} \times A_{pik}A_{pik}$  ( $k' \neq k$ ); el total de éstas es  $4mn$ . Las probabilidades de que los dos individuos involucrados en cada crusa sean incluidos en la muestra, de que ya en la muestra se crucen, y de que, dado que se han cruzado, produzcan el genotipo  $A_{pik}A_{pik}$  son:

$$\frac{\left[ \binom{1}{1} \binom{mn-1}{n-1} \right] \left[ \binom{2}{2} \binom{4mn-2}{m-2} \right]}{\binom{mn}{n} \binom{4mn}{m}} = \frac{m-1}{4mn(4mn-1)},$$

$1/(mn)^2$  y  $1/2$ , respectivamente. Por lo tanto, la contribución de este tipo de cruzas a la frecuencia de  $A_{pik}A_{pik}$  en el  $C_2$   $\left[ P(A_{pik}A_{pik})_3 \right]$  es

$$P(A_{pik}A_{pik})_3 = 4mn \left[ \frac{m-1}{4mn(4mn-1)} \right] \frac{1}{(mn)^2} \frac{1}{2}$$

$$= \frac{m-1}{2(4mn-1)(mn)^2} \quad 3)$$

Otro tipo de cruzas que pueden producir genotipos  $A_{pik}A_{pik}$  para el  $C_2$  son todas las posibles entre cada dos de los  $2mn$  genotipos del  $C_1$  que tienen un solo gen  $A_{pik}$  y que pertenecen a la familia cuyo progenitor común es  $A_{pik}A_{pik'}$  ( $k \neq k'$ ). El número total de estas cruzas, incluyendo directas y recíprocas, es  $2mn(2mn-1)$ . Además de estas cruzas intrafamiliares hay que considerar que en cada una de las  $mn-1$  familias restantes hay dos genotipos de la forma  $A_{qj1}A_{pik}$  y  $A_{qj2}A_{pik}$  ( $qj \neq pi$ ); de manera que considerando directas y recíprocas habrá  $2(mn-1)$  cruzas intrafamiliares adicionales de la forma  $A_{qj1}A_{pik} \times A_{qj2}A_{pik}$  (y su recíproca). Cada una de todas estas

$2mn(2mn-1)+2mn-2 = (2mn)^2-2$  cruzas intrafamiliares involucra dos individuos cuya probabilidad de que sean incluidos en la muestra es  $(m-1)[4mn(4mn-1)]^{-1}$ , y las probabilidades de que se crucen y de que produzcan un genotipo  $A_{pik}A_{pik}$  son  $(mn)^{-2}$  y  $1/4$ , respectivamente. Por lo tanto, la contribución a la frecuencia de genotipos  $A_{pik}A_{pik}$  en el  $C_2$  de todas estas  $(2mn)^2-2$  cruzas  $\left[ P(A_{pik}A_{pik})_4 \right]$  es:

$$P(A_{pik}A_{pik})_4 = \left[ (2mn)^2 - 2 \right] (m-1) \left[ 4mn(4mn-1) \right]^{-1} (2mn)^2$$

$$= \frac{\left[ (2mn)^2 - 2 \right] (m-1)}{16(mn)^3 (4mn-1)} \quad 4)$$

## Cruzas intrafamiliares

En este apartado se hará referencia a cruzas entre individuos que aunque son de diferente familia tienen una relación de parentesco que hace posible la generación de individuos de genotipo  $A_{pik}A_{pik}$ . Un tipo de estas cruzas involucra al genotipo  $A_{pik}A_{pik}$  de la familia cuyo progenitor común es  $A_{pi1}A_{pi2}$  y al genotipo  $A_{qj1}A_{pik}$  ( $qj \neq pi$ ) de una cualquiera de las  $mn-1$  familias restantes. Como cada una de estas familias tiene dos genotipos  $A_{qj1}A_{pik}$  ( $A_{qj1}A_{pik}$  y  $A_{qj2}A_{pik}$ ), el número total de cruzas (incluyendo directas y recíprocas) es  $4(mn-1)$ . La probabilidad de que los dos individuos de cada una de estas cruzas se incluyan en la muestra de  $mn$  individuos es:

$$\frac{\left[ \binom{2}{2} \binom{mn-2}{n-2} \right] \left[ \binom{1}{1} \binom{4mn-1}{m-1} \right]^2}{\binom{mn}{n} \binom{4mn}{m}^2} = \frac{n-1}{16mn^2(mn-1)}$$

Además, las probabilidades de que se realice cada crusa y de que ya en la muestra cada una de estas cruzas produzca un genotipo  $A_{pik}A_{pik}$  son  $1/(mn)^2$  y  $1/2$ , respectivamente. Con base en la información generada para este primer tipo de crusa interfamiliar, la aportación de ésta a la frecuencia de  $A_{pik}A_{pik}$  en el  $C_2$   $\left[ P(A_{pik}A_{pik})_5 \right]$  es:

$$P(A_{pik}A_{pik})_5 = \frac{4(mn-1)(n-1)}{16mn^2(mn-1)} \frac{1}{2(mn)^2}$$

$$= \frac{n-1}{n} \frac{1}{8(mn)^3} \quad 5)$$

El segundo tipo de crusa interfamiliar que puede producir  $A_{pik}A_{pik}$  para el  $C_2$  es de la forma  $A_{pik}A_{qj1} \times A_{rst}A_{pik}$ ,  $qj \neq pik$  y  $rs \neq pi$ , en donde  $A_{pik}A_{qj1}$  es uno cualquiera de

los  $2mn-1$  genotipos ordenados que portan un gen  $A_{pik}$  y que pertenecen a la familia de medios hermanos cuyo progenitor común es  $A_{pi1}A_{pi2}$  (o  $A_{pik}A_{pik'}, k \neq k'$ ) y  $A_{rst}A_{pik}$  es uno cualquiera de los dos genotipos que portan un gen  $A_{pik}$  en cada una de las  $mn-1$  familias restantes. En este tipo de cruzas se deben incluir las cruzas interfamiliares de la forma  $A_{pik}A_{pik} \times A_{rst}A_{pik}$ . Esta inclusión es lo que hace que el número total de cruzas interfamiliares de este tipo, incluyendo directas y recíprocas, sea  $2[2mn(2mn-2)]$ .

Finalmente, las cruzas interfamiliares de la forma  $A_{qj1}A_{pik} \times A_{rst}A_{pik}$  en donde  $qj \neq pi$ ,  $rs$  y  $rs \neq pi$  (es decir, cruzas que involucran dos individuos que son de diferente familia, pero no involucran a la familia cuyo progenitor común es  $A_{pi1}A_{pi2}$ ) también producen genotipos  $A_{pik}A_{pik}$  para el  $C_2$ . El número de cruzas de este tipo, incluyendo directas y recíprocas, es  $4(mn-1)(mn-2)$ . En este caso, similarmente a lo ocurrido en el segundo tipo de cruzas interfamiliares, las probabilidades de que los dos participantes en cada cruzada sean incluidos en la muestra, de que ya en la muestra se crucen y de que produzcan genotipos  $A_{pik}A_{pik}$  son:

$$\frac{(n-1)}{n} \frac{1}{16mn(mn-1)}, \frac{1}{(mn)^2} \text{ y } \frac{1}{4},$$

respectivamente. Así, las  $8mn(mn-1)+4(mn-1)(mn-2) = 12(mn)^2-20mn+8$  cruzas interfamiliares del segundo y tercer tipos contribuyen con genotipos  $A_{pik}A_{pik}$  al  $C_2$  con la frecuencia relativa esperada  $[P(A_{pik}A_{pik})_6]$  siguiente:

$$\begin{aligned} P(A_{pik}A_{pik})_6 &= [(12mn-8)(mn-1)] \frac{(n-1)}{8n} \frac{1}{(2mn)^3(mn-1)} \\ &= \frac{n-1}{n} \frac{[12mn-8]}{8(2mn)^3} \\ &= \frac{n-1}{n} \frac{[3mn-2]}{2(2mn)^3} \end{aligned} \quad 6)$$

Todas las fuentes de producción de genotipos  $A_{pik}A_{pik}$  para el  $C_2$  ya han sido cuantificadas en términos de su aportación en las Ecuaciones 1 a 6. De acuerdo con éstas, la frecuencia esperada de  $A_{pik}A_{pik}$  en el  $C_2$   $[P(A_{pik}A_{pik})]$  es:

$$\begin{aligned} P(A_{pik}A_{pik}) &= \sum_{y=1}^6 [P(A_{pik}A_{pik})_y] \\ &= \frac{1}{4(mn)^3} + \frac{2mn-1}{(2mn)^3} + \frac{m-1}{2(4mn-1)(mn)^2} + \frac{[2(mn)^2-1](m-1)}{8(mn)^3(4mn-1)} \\ &\quad + \frac{n-1}{n} \frac{1}{8(mn)^3} + \frac{n-1}{n} \frac{[3mn-2]}{2(2mn)^3} \end{aligned}$$

$$\begin{aligned} &= \frac{2+2mn-1}{(2mn)^3} + \frac{4mn(m-1)+[2(mn)^2-1](m-1)}{8(mn)^3(4mn-1)} \\ &\quad + \frac{n-1}{n} \left[ \frac{2+3mn-2}{16(mn)^3} \right] \\ &= \frac{2mn+1}{(2mn)^3} + \frac{(m-1)[4mn+2(mn)^2-1]}{8(mn)^3(4mn-1)} + \frac{n-1}{n} \frac{3}{16(mn)^2} \end{aligned} \quad 7)$$

La Ecuación 7 describe la contribución al coeficiente de endogamia del  $C_2$  del genotipo  $A_{pik}A_{pik}$ . Debe considerarse que: a)  $p = 1, 2, \dots, m$ ; b)  $i = 1, 2, \dots, n$ ; y c)  $k = 1, 2$ . De acuerdo con esta consideración, el coeficiente de endogamia del  $C_2$  ( $F_{2,E}$ ) debe expresarse como el producto del lado derecho de la Ecuación 7 por  $2mn$ ; es decir:

$$F_{2,E} = \frac{2mn+1}{(2mn)^2} + \frac{m-1}{mn} \frac{[4mn+2(mn)^2-1]}{4(mn)(4mn-1)} + \frac{n-1}{n} \frac{3}{8(mn)} \quad 8)$$

Las contribuciones a  $F_{2,E}$  por autofecundación, cruzas intrafamiliares y cruzas interfamiliares corresponden al primero, segundo y tercer términos de la Ecuación 8. Los dos últimos términos muestran con claridad las condiciones obvias requeridas para que las cruzas intra e interfamiliares tengan contribuciones a  $F_{2,E}$  mayores que cero:  $m > 1$  y  $n > 1$ , respectivamente.

Falconer (1989) se refiere al coeficiente de endogamia del  $C_2$  de la población ideal ( $F_2$ ) en términos de genotipos formados por dos genes idénticos por descendencia por dos vías: la nueva duplicación de genes y la duplicación previa, que ocurren con frecuencias de  $1/(2N)$  y  $[1-1/(2N)]F_1$ , respectivamente. De acuerdo con esto, este autor expresa el coeficiente de endogamia del  $C_2$  ( $F_2$ ) de la población ideal (PI) como:

$$F_2 = \frac{1}{2N} + \left[ 1 - \frac{1}{(2N)} \right] F_1$$

En esta fórmula,  $N$  es el tamaño de muestra tomada de una manera completamente aleatoria del ciclo 1 para que por apareamiento aleatorio produzca el ciclo siguiente. Tanto en la PI como en la población ideal con muestreo en dos etapas (PIDE) que aquí se estudia, la contribución de la nueva duplicación de genes es la mitad del recíproco del tamaño de muestra; es decir,  $1/(2mn)$ . Sin embargo, con respecto a la contribución al coeficiente de endogamia del  $C_2$  por duplicación previa hay diferencias entre PIDE y PI. Para PIDE, el coeficiente de endogamia es el que se describe en la Ecuación 8. Después de un trabajo algebraico relativamente extenso en la Ecuación 8, se puede encontrar que ésta también se expresa como:

$$F_{2,E} = \frac{1}{2mn} + \left(1 - \frac{1}{2mn}\right) \left[ \frac{1}{2mn} + \frac{4 - m(n+3)}{4mn(2mn-1)} \right. \\ \left. + \frac{2(m-1)[4mn + 2(mn)^2 - 1]}{4mn(4mn-1)(2mn-1)} \right] \quad 9)$$

Si se omiten los dos últimos términos del paréntesis rectangular de la Ecuación 9, queda la fórmula del  $F_2$  de la PI que describe Falconer (1989) para un  $C_2$  cuyo tamaño de muestra es  $mn$ . La constitución de estos dos últimos términos del paréntesis rectangular implica que el muestreo estratificado produce un mayor coeficiente de endogamia que el muestreo completamente aleatorio, excepto cuando  $m = 1$  y  $n > 1$ ; en este caso,  $F_{2,E} < F_2$ . Esto se debe a que con  $m = 1$ , que asegura que no ocurrirán apareamientos entre medios hermanos, se cancela el último término del paréntesis rectangular de la Ecuación 9; y con  $m = 1$  y  $n > 1$  el segundo término de este paréntesis se hace negativo y ocasiona que  $F_{2,E} < F_2$ . Es decir, sólo con un individuo por familia el muestreo combinado en dos etapas (PIDE) es mejor que el muestreo completamente al azar. En el caso, quizás sólo de interés académico, en que  $m = n = 1$  los coeficientes de endogamia de PIDE y de PI son iguales.

## CONCLUSIONES

En la población ideal (PI) que describe Falconer (1989), cada ciclo se forma por el apareamiento al azar de los individuos de una muestra completamente aleatoria de tamaño  $mn$  del ciclo anterior. Si la muestra es en dos etapas (PIDE) de  $m$  individuos aleatorios tomados de cada una de  $n$  familias también seleccionadas al azar, en relación a PI el coeficiente de endogamia de PIDE es mayor que el de PI excepto cuando  $m = 1$  y  $n > 1$  (en este caso el de PI es mayor). En el caso extremo en que  $m = 1$  y  $n = 1$  los dos coeficientes de endogamia son iguales.

## LITERATURA CITADA

- CROW, J. F.; KIMURA, M. 1970. An Introduction to Population Genetics Theory. Harper and Row Publishers. New York. 591 p.
- FALCONER, D. S. 1989. Introduction to Quantitative Genetics. Third edition. Longman. 438 p.
- HALLAUER, A. R.; MIRANDA FO. J.B. 1981. Quantitative Genetics in Maize Breeding. Iowa State University. Ames IA. USA. 468 p.
- MÁRQUEZ-SÁNCHEZ, F. 1992. Inbreeding and yield prediction in synthetic maize cultivars made with parental lines: I. Basic Methods. *Crop Sci.* 32:271-274.
- SAHAGÚN-CASTELLANOS, J. 1994. Sobre el cálculo de coeficientes de endogamia de variedades sintéticas. *Revista Agrociencia serie Fitociencia* 5:67-78.
- SAHAGÚN C., J. 2006. Determinación de las fuentes endogámicas de la población ideal bajo muestreo continuo y apareamiento aleatorio. *Agrociencia* 40(4):471-482.