

COMPONENTES DE VARIANZA EN VARIEDADES SINTÉTICAS. III. COMPONENTES PARENTALES ENDOGÁMICOS EMPARENTADOS

Sahagún Castellanos, J.

Departamento de Fitotecnia, Universidad Autónoma Chapingo, Chapingo, México. C.P. 56230.

RESUMEN. La estimación de los componentes de varianza genética de una variedad sintética es una forma de obtener información útil para tomar la decisión de emprender, o no, un programa de selección recurrente y, en su caso, el tipo de estrategia a seguir. Cuando la variedad sintética se ha formado por n componentes parentales endogámicos y emparentados, la estimación de componentes de varianza genética siguiendo el procedimiento que ignora los procesos endogámicos puede producir resultados diferentes de los que se obtendrían considerando la situación real. En este estudio se deriva la genética de

una variedad sintética del tipo en discusión y se analiza el efecto que tendría el estimar la varianza aditiva (σ_A^2) de esta población como si se tratara de una población con un coeficiente de endogamia F_0 igual a cero, encontrando que con este último: (1)

se produce un sesgo positivo de una magnitud de $F_0 \sigma_A^2$ en la estimación, y (2) el estimador resulta tener una precisión menor que la del estimador apropiado (el que considera la presencia de endogamia). La varianza del estimador que ignora la presencia del proceso endogámico es $2F_0 + F_0^2$ veces mayor que la del estimador insesgado.

PALABRAS CLAVE: arreglo genotípico, varianza aditiva, varianza de dominancia, coeficiente de coancestría, coeficiente de endogamia.

VARIANCE COMPONENTS OF SYNTHETIC VARIETIES. III. INBRED AND RELATED PARENTAL COMPONENTS.

SUMMARY. Estimation of the genetic components of variance of a synthetic variety is a way of obtaining useful information for making the decision as to undertake, or not, a program of recurrent selection on given population and, if so, the strategy to be applied. When the synthetic variety has been formed by n inbred and related parental lines, the estimation of the genetic variance components that ignores the presence of the inbreeding processes might yield results different from those obtained considering the actual situation. In this study the genetic variance of the synthetic variety under discussion is derived and the effect of neglecting the inbreeding coefficient (F) in the estimation of the additive variance (σ_A^2) is analyzed. Results indicate that:

(1) this estimation produces a positive bias of $F \sigma_A^2$ and (2) the precision of the estimation is smaller relative to the appropriate estimator's; the variance of the former is $2F + F^2$ times larger than the one corresponding to the unbiased estimator.

KEY WORDS: genotypic array, additive variance, dominance variance, coancestry coefficient, inbreeding coefficient.

INTRODUCCION

La constitución de una variedad sintética en ocasiones se basa en componentes parentales endogámicos que tienen cierta relación de parentesco. En este caso, como en cualquier otro en que se desea investigar las posibilidades de mejoramiento por la vía

de selección de una población, resulta bastante informativo desde un punto de vista genotécnico el conocer su estructura de variabilidad genética. De particular importancia en este sentido es la determinación de la varianza genética de este tipo de variedad sintética que, por su origen, involucra la presencia de procesos endogámicos de diferente naturaleza.

La presencia de endogamia en estas poblaciones tiene, obviamente, efectos en el arreglo genotípico de la población que lo hacen diferente en su esencia del arreglo genotípico de una población en donde no se efectúan procesos endogámicos. Esta diferencia a su vez, debe tener efecto en la estructura de variabilidad genética y, consecuentemente, en las fórmulas de estimación de las componentes de varianza genética aditiva y de dominancia.

En un estudio previo sobre el tema (Sahagún, 1994) se ha hecho referencia a la variedad sintética que resulta de n componentes parentales no endogámicos ni emparentados y a la variedad sintética producto del apareamiento aleatorio entre n componentes endogámicos no emparentados. En estos tipos de variedad sintética, particularmente en la del primer caso, la presencia del fenómeno endogámico puede representar una proporción tan pequeña en la reproducción de la variedad que su no consideración puede tener un efecto de poca importancia. Sin embargo, cuando se produce endogamia por la parte del apareamiento aleatorio que se da por autofecundación y por el apareamiento entre parientes, el ignorar la presencia del coeficiente de endogamia tanto en la determinación de la varianza genética de la población como en la forma de estimar sus componentes, puede ser de importancia considerable.

En este estudio se planteó como un objetivo, la determinación de la varianza genética de una variedad sintética formada por n componentes con un coeficiente de endogamia F y con un coeficiente de coancestría Γ_{oB} . La determinación del impacto que generaría el ignorar la presencia de los procesos endogámicos en la calidad del estimador de la componente de la varianza aditiva así generado también fue un objetivo de la presente investigación.

CONSIDERACIONES TEORICAS Y DERIVACIONES

Si el sintético se forma por n componentes parentales y el componente i (i=1,2,...,n) está representado por una planta de genotipo A_i1A_i2 , el apareamiento aleatorio de estos individuos producirá una población cuyo arreglo genotípico (AG Sin-2), el de la variedad sintética, es

$$AGE \text{ Sin-2} = \frac{1}{2n} \sum_i \sum_j \sum_k \sum_l A_{ik} A_{jl} \quad (1)$$

Ahora, definiendo a Y_{ikjl} como el valor genotípico codificado asociado al genotipo $A_{ik}A_{jl}$ y explicando a este valor genotípico en la forma

$$Y_{ikjl} = \alpha_{ik} + \alpha_{jl} + d_{ikjl}$$

en donde α_{ik} y α_{jl} son, respectivamente, los efectos de los alelos A_{ik} y A_{jl} , mientras que d_{ikjl} es la desviación de dominancia, Sahagún (en revisión) partió para derivar la varianza de un sintético $[V(Y_{ikjl})]$ de una expresión básica de la forma

$$\begin{aligned} V(Y_{ikjl}) = & E(\alpha_{ik}^2) + E(\alpha_{jl}^2) + E(d_{ikjl}^2) - [E(d_{ikjl})]^2 \\ & + 2E(\alpha_{ik}\alpha_{jl}) + 2E(\alpha_{ik}d_{ikjl}) \\ & + 2E(\alpha_{jl}d_{ikjl}) - 2E(\alpha_{ik})E(d_{ikjl}) \\ & - 2E(\alpha_{jl})E(d_{ikjl}). \end{aligned} \quad (2)$$

Si el coeficiente de endogamia promedio de los individuos parentales es F_0 y si su coeficiente de coancestría promedio es Γ_{oB} , aún en condiciones de apareamiento aleatorio, habrá endogamia por tres vías: (1) Por la fracción de autofecundación (1/2n) que se genera por este tipo de apareamiento; (2) por la fracción de apareamientos entre parientes (Γ_{oB}) y por la endogamia ya existente en los componentes parentales (F_0). Otro elemento a considerar para la derivación de la varianza es la probabilidad de que dos alelos A_{ik} y A_{jl} sean idénticos por descendencia, o no lo sean,

$$P(A_{ik} = A_{jl}) \text{ y } P(A_{ik} \neq A_{jl}),$$

respectivamente. Así, como el sintético se forma por apareamiento aleatorio y la identidad por descendencia, atendiendo al tipo de apareamiento y de alelos que forman el genotipo, se puede generar por tres eventos mutuamente excluyentes: (1) autofecundación y mismo alelo; (2) autofecundación y los dos alelos del genotipo autofecundado y (3) cruzamiento entre individuos emparentados y alelos idénticos por descendencia,

$$\begin{aligned} P(A_{ik} = A_{jl}) = & (1/2n) + (1/2n)F_0 + 2\left(\frac{n}{2}\right)\Gamma_{oB}/n^2 \\ = & (1/2n)(1 + F_0) + \frac{2n(n-1)}{2n^2}\Gamma_{oB} \end{aligned} \quad (3)$$

$$= (1/2n)(1 + F_0) + [(n-1)/n]\Gamma_{oB}$$

y obviamente,

$$P(A_{ik} \neq A_{jl}) = 1 - P(A_{ik} = A_{jl}) \quad (4)$$

La endogamia de un sintético como el que aquí se analiza ya ha sido presentada en otros estudios (Márquez, 1992).

Por otra parte, en una población bajo apareamiento aleatorio el valor esperado del efecto promedio de un gene siempre es cero. También, para alelos no idénticos por descendencia, el promedio de las desviaciones de dominancia y la covarianza entre estas desviaciones y los efectos de los correspondientes alelos son cero.

De acuerdo con lo anterior, para los términos expresados en (2) se tiene que:

$$E(\alpha_{ik}^2) = \sum_i \sum_k (1/2n) \alpha_{ik}^2 \quad (5)$$

$$E(d_{ikjl}^2) = P(A_{ik} = A_{jl}) E(d_{ikjl}^2 / A_{ik} = A_{jl}) +$$

$$P(A_{ik} \neq A_{jl}) E(d_{ikjl}^2 / A_{ik} \neq A_{jl})$$

$$= P(A_{ik} = A_{jl}) \sum_i \sum_k (1/2n) d_{ikik}^2 +$$

$$P(A_{ik} \neq A_{jl}) \sum_i \sum_j \sum_k \sum_l (1/2n) d_{ikjl}^2 \quad (6)$$

$$E(d_{ikjl}) = P(A_{ik} = A_{jl}) E(d_{ikjl} / A_{ik} = A_{jl}) +$$

$$P(A_{ik} \neq A_{jl}) E(d_{ikjl} / A_{ik} \neq A_{jl})$$

$$= P(A_{ik} = A_{jl}) \sum_i \sum_k (1/2n) d_{ikik}, \quad (7)$$

$$E(\alpha_{ik} \alpha_{jl}) = P(A_{ik} = A_{jl}) E(\alpha_{ik} \alpha_{jl} / A_{ik} = A_{jl}) +$$

$$P(A_{ik} \neq A_{jl}) E(\alpha_{ik} \alpha_{jl} / A_{ik} \neq A_{jl})$$

$$= P(A_{ik} = A_{jl}) \sum_i \sum_k (1/2n) \alpha_{ik}^2, \quad (8)$$

y

$$E(\alpha_{ik} d_{ikjl}) = P(A_{ik} = A_{jl}) E(\alpha_{ik} d_{ikjl} / A_{ik} = A_{jl}) +$$

$$P(A_{ik} \neq A_{jl}) E(\alpha_{ik} d_{ikjl} / A_{ik} \neq A_{jl})$$

$$= P(A_{ik} = A_{jl}) \sum_i \sum_k (1/2n) \alpha_{ik} d_{ikik} \quad (9)$$

Ahora, substituyendo los resultados (3) y (4) en los términos de las expresiones (5)-(9) según corresponda, se generarán cinco nuevas expresiones. Substituyendo los términos de (2), según corresponda, por estos nuevos resultados y considerando que con alelos múltiples en una población bajo apareamiento aleatorio como la que forma el sintético, los componen-

tes de varianza aditiva (σ_A^2) y de dominancia (σ_D^2) resultan ser expresables como

$$\sigma_A^2 = 2 \sum_i \sum_k (1/2n) \alpha_{ik}^2$$

y

$$\sigma_D^2 = \sum_i \sum_j \sum_k \sum_l (1/2n) d_{ikjl}^2$$

se genera una expresión para la varianza del sintético de la forma

$$V(Y_{ikjl}) = [1 + (1/2n)(1 + F_0) + (n-1)\Gamma_{ow}/n] \sigma_A^2 +$$

$$[1 - (1/2n)(1 + F_0) + (n-1)\Gamma_{ow}/n] \sigma_D^2 +$$

$$4[(1/2n)(1 + F_0) + (n-1)\Gamma_{ow}/n] \sum_i \sum_k (1/2n) \alpha_{ik} d_{ikik}$$

$$[(1/2n)(1 + F_0) + (n-1)\Gamma_{ow}/n] \sum_i \sum_k (1/2n) d_{ikik}^2$$

$$- [(1/2 n) (1 + F_0) + (n - 1) \Gamma_{ow}/n] \sum_i \sum_k (1/2 n) d_{ik}^2 \quad (10)$$

Esta varianza, es evidente, se desvía más y más de la varianza genética correspondiente a la de una población en donde de origen no han ocurrido procesos endogámicos ($\sigma_A^2 + \sigma_D^2$). En particular, la varianza aditiva se incrementa en la misma medida en que la varianza de dominancia decrece. Esta medida es el coeficiente de endogamia.

DISCUSION

Una de las finalidades de estimar las componentes de varianza genética de una población puede ser la estimación de la heredabilidad en sentido estrecho de uno o varios caracteres de importancia y con ella la respuesta a la selección que se esperaría bajo la aplicación de una u otra estrategia de mejoramiento recurrente. Esta información, a su vez, al lado de consideraciones adicionales que incluyen desde aspectos socioeconómicos y culturales hasta el grado de desarrollo tecnológico presentes en la región a la que está dirigido el producto de la actividad de fitomejoramiento, constituye una base objetiva sobre la conveniencia, o no, de emprender un programa de mejoramiento genético y, en su caso, el tipo de cultivar que concilie la explotación de la estructura de variabilidad genética con los aspectos sociales, económicos, culturales y tecnológicos que prevalezcan. De acuerdo con esto, resulta evidente que realizar una buena estimación es de importancia singular; más aún si se considera que la estimación de componentes de varianza usualmente presenta problemas de precisión.

Con componentes parentales endogámicos emparentados y en número reducido, una variedad sintética posee un coeficiente de endogamia que debe tener efecto en las proporciones de las componentes de varianza genéticas. La estimación de los componentes de varianza aditiva y de dominancia, desde luego, puede realizarse de acuerdo con metodologías que para este efecto ya han sido desarrolladas. Márquez-Sánchez y Sahagún-Castellanos (en revisión) han propuesto un nuevo procedimiento para estimar componentes de varianza, para el caso de una población bajo apareamiento aleatorio con ausencia de endogamia. Esta metodología se basa en la formación natural de familias de medios hermanos que, en cultivos agrícolas como el maíz (*Zea mays* L.) se da por polinización abierta. Estas familias son las que se evalúan en experimentos de campo. Una característica que diferencia a este procedimiento del que se sigue de acuerdo con el diseño I de Carolina del Norte

(Comstock y Robinson, 1948) es que en éste cada familia de medios hermanos maternos se forma mediante el apareamiento de un macho con h hembras (guardándose la identidad de cada cruce), siendo las familias de hermanos completos las que se asignan a las parcelas en experimentos de campo. Márquez-Sánchez y Sahagún-Castellanos (en revisión) demostraron que con familias de medios hermanos maternos la estimación de componentes de varianza es más precisa que la que se logra con el diseño I.

Para el caso de un sintético formado a partir de n componentes con un coeficiente de endogamia y coancestría F_0 y Γ_{oB} , respectivamente, en la estimación de las componentes de varianza mediante la formación natural de familias de medios hermanos maternos, la varianza entre familias de medios herma-

nos (σ_m^2), que se puede expresar como la varianza entre los individuos de la misma familia (COV(MHM)), resulta ser expresable como $[(1+F)/4] \sigma_A^2$ (Falconer, 1989); es decir,

$$\sigma_m^2 = COV(MHM) = [(1 + F) / 4] \sigma_A^2 \quad (11)$$

en donde F es el coeficiente de endogamia del sintético, o sea,

$$F = (1/2 n) (1 + F_0) + [(n - 1) / n] \Gamma_{oB}$$

De la expresión en (11) resulta que

$$\sigma_A^2 = 4 \sigma_m^2 / (1 + F) \quad (12)$$

Así, para estimar la varianza aditiva de una variedad sintética como la aquí descrita se requiere

tener un estimador de σ_m^2 . En el análisis de varianza de

la información producida en un experimento en, por ejemplo, un diseño en bloques al azar con r repeticiones, para evaluar las familias de medios hermanos maternos se genera información para esti-

mar a σ_m^2 en la forma

$$\hat{\sigma}_m^2 = \frac{CM(M) - CM(E)}{r}$$

en donde CM(M) y CM(E) son el cuadrado medio de familias de medios hermanos y del error, respectiva-

mente. Así, el estimador $\hat{\sigma}_A^2$ sería de la forma

$$\hat{\sigma}_A^2 = 4 \frac{\hat{\sigma}_m^2}{(1+F)}. \quad (13)$$

Ahora, pudiera considerarse la estimación de la varianza aditiva como si se tratara de una variedad de polinización libre en donde el coeficiente de endogamia fuera cero. si se ignorara la existencia de la endogamia de la variedad sintética ocurrirían dos hechos indeseables: sesgo y menor precisión en la estimación.

Con respecto al sesgo éste resultaría ser $4F \frac{\hat{\sigma}_m^2}{(1+F)}$ o bien, en términos de varianza aditiva, $F \frac{\hat{\sigma}_m^2}{A}$; es decir,

entre mayor sea el coeficiente de endogamia la sobreestimación de la varianza aditiva será de mayor magnitud.

Con relación a la precisión de la estimación, resulta que cuando se considera la presencia de la endogamia, la varianza del estimador [$V(\hat{\sigma}_A^2)$] es

$$V(\hat{\sigma}_A^2) = V\left[\frac{4[CM(M) - CM(E)]}{r(1+F)}\right]$$

$$= \frac{32}{r^2(1+F)^2} \left[\frac{[E[CM(M)]]^2}{f-1} + \frac{[E[CM(E)]]^2}{(f-1)(r-1)} \right] \quad (14)$$

En cambio, con el estimador que no considera la presencia de endogamia, la varianza del estimador es

$$V(\hat{\sigma}_A^2) = \frac{32}{r^2} \left[\frac{[E[CM(M)]]^2}{f-1} + \frac{[E[CM(E)]]^2}{(f-1)(r-1)} \right] \quad (15)$$

Esto es, la varianza del estimador que ignora la presencia de endogamia tiene una varianza que excede a la del estimador insesgado que da cuenta de la endogamia que se produce en el sintético en una cantidad de

$$\frac{32(2F+F^2)}{r^2(1+F^2)} \left[\frac{[E[CM(M)]]^2}{f-1} + \frac{[E[CM(E)]]^2}{(f-1)(r-1)} \right] \quad (16)$$

De otra forma: El estimador en (15) tiene una varianza que es $2F+F^2$ veces mayor que la del estimador apropiado. De cualquier forma, ignorando la endogamia la falta de precisión en la estimación será mayor en la medida en que F sea mayor. Obviamente, cuando $F=0$ este exceso de varianza será 0 también. Así pues, la estimación debería ser con el estimador que se describe en (13). Para este caso la varianza de una estimación se puede estimar de acuerdo con la fórmula

$$V(\hat{\sigma}_A^2) = \frac{32(2F+F^2)}{r^2(1+F^2)} \left[\frac{[CM(M)]^2}{f+1} + \frac{[CM(E)]^2}{fr-f-r+3} \right] \quad (17)$$

La varianza de dominancia de acuerdo con el método de medios hermanos maternos se podría estimar si de cada unidad experimental se evaluaran n individuos. Siguiendo un procedimiento relativamente análogo al descrito para el caso de la estimación de la varianza aditiva, se puede describir un procedimiento para estimar la varianza de dominancia de una variedad sintética con componentes parentales endogámicos y con relación de parentesco.

LITERATURA CITADA

- COMSTOCK, R.E.; H.F. ROBINSON. 1948. The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. *Biometric*. 4:254-256.
- FALCONER, D.S. 1989. *Introduction to Quantitative Genetics*. Third edition. Longman, London. 340 pp.
- MARQUEZ S., F. 1992. La obtención y la predicción del rendimiento de variedades sintéticas y variedades compuestas de maíz. *Ciencia* 43:413-428.
- MARQUEZ-SANCHEZ, F; J. SAHAGUN-CASTELLANOS. 199. Estimation of genetic variances in maize with maternal half-sib families. *Maydica* (en revisión).
- SAHAGUN C., J. 1994. Componentes de varianza en variedades sintéticas. I. Componentes parentales no endogámicos ni emparentados. *Revista Chapingo, Serie Horticultura* 2. p.