

DIVERGENCIA GENETICA Y HETEROSIS

Tulio Cesar Lagos B. ¹

Hernando Criollo E. ¹

Oscar Checa C. ¹

RESUMEN

El análisis de la divergencia genética de la población base de un programa de mejoramiento, permite establecer progenitores o genotipos contrastantes, los cuales pueden dar lugar a combinaciones híbridas, descendientes con alto vigor híbrido o individuos transgresivos que pueden originar híbridos o variedades comerciales. En este artículo, se presentan los principios teóricos del análisis de divergencia genética y se analiza su importancia en la identificación de combinaciones híbridas, en las cuales se puede explotar la heterosis. Además, se discuten los fundamentos de la heterosis.

Palabras claves: Divergencia genética, vigor híbrido, distancias, similaridad, análisis de clasificación, análisis discriminante.

SUMMARY

The analysis of the population's genetic divergence bases of a program of improvement, it allows to establish progenitors or contrasting genotypes, which can generate in hybrid combinations, descending with high hybrid vigor or transgressive individuals that can originate hybrid or commercial varieties. In this paper, the theoretical principles of the analysis of genetic divergence are presented and their importance is analyzed in the identification of hybrid combinations, in which you can exploit the heterosis. Also, the foundations of the heterosis are discussed.

Key words: Genetic divergence, hybrid vigor, distances, similitude, classification analysis, discriminant analysis.

¹ Profesores Facultad de Ciencias Agrícolas, Universidad de Nariño, Pasto, Nariño, Colombia.

E- Mail: tclagos3@hotmail.com; hcriollo@hotmail.com

INTRODUCCION

La variabilidad genética es la materia prima con la cual trabaja el fitomejorador. De acuerdo con los objetivos del programa de mejoramiento, se debe contar con suficiente número de genotipos, en donde la selección pueda llevarse a cabo con la mayor eficiencia.

Es necesario tener suficiente variabilidad genética para mejorar características como rendimiento, resistencia a plagas, la composición nutricional de la semilla, la calidad del forraje, la tolerancia a estrés por deficiencia de minerales o por déficit ambiental, la adaptación a la mecanización, la resistencia al volcamiento y la respuesta al fotoperíodo, entre otras.

Pardos (1988) indica que si no existiera suficiente variabilidad genética en los caracteres de interés económico, los intentos de usar la genética para mejorar las especies vegetales serían vanos. Por ello resulta esencial cuando se inicia un programa de mejora genética determinar la cantidad, las causas y la naturaleza de la variación presente en el género y, sobre todo, en la especie objeto de estudio, a fin de poder manejarla en beneficio de los objetivos pretendidos.

Es un hecho, que la variación existe entre especies, razas e individuos; pero la determinación de las causas de su existencia es una tarea larga y costosa. El descubrir qué proporción de esta variación está controlada genéticamente, a fin de su máxima explotación, constituye parte esencial de dicha tarea.

Por lo tanto, el éxito de un programa de mejoramiento depende fundamentalmente de la escogencia de los parentales que se van a incluir en los planes de hibridación. Los objetivos del mejoramiento involucran el uso de progenitores que hacen posible obtener poblaciones que posean un alto promedio asociado a una base genética amplia, para las características que están bajo selección (Fehr, 1987).

Se puede afirmar que el análisis de diversidad genética se destina a la identificación de progenitores adecuados para la obtención de híbridos con mayor efecto heterótico o heterosis y que proporcionen mayor segregación de recombinantes, con la posibilidad de encontrar individuos transgresivos (Cruz y Souza, 2003).

La heterosis puede ser positiva o negativa. Ambos tipos de heterosis pueden ser útiles dependiendo de la característica. Por ejemplo: la heterosis positiva para rendimiento y negativa para la precocidad.

Existen dos formas de establecer la divergencia genética, siendo la primera de naturaleza cuantitativa y la segunda de naturaleza predictiva. Entre los métodos de evaluación de diversidad o de manifestación de heterosis de naturaleza cuantitativa, están los cruzamientos dialélicos.

El sistema de cruzamientos dialélicos se ha utilizado ampliamente ya sea para evaluar la aptitud combinatoria y establecer el potencial heterótico de líneas y variedades en cruzamientos como para realizar los estudios básicos de la estructura genética de poblaciones (Gerald y Miranda Filho, 1988).

Existen muchas formas de realizar el análisis de un dialelo. Se destacan las metodologías de Griffing (1956), Gardner y Eberhart (1966, 1967), Hayman (1954), Hallauer y Miranda (1995), Comstock y Robinson (1948). De acuerdo con Ceballos (1997), el principal problema no radica en el análisis per se de los datos, sino en inferencias injustificadas y/o la interpretación errónea que se le ha dado frecuentemente a los resultados.

Entre los métodos predictivos de heterosis están aquellos que se basan en las diferencias morfológicas, fisiológicas o moleculares. En estos, se cuantifica de alguna manera la disimilaridad que se expresa o el grado de diversidad genética entre los progenitores.

Otro ejemplo predictivo de heterosis, es la inferencia de diversidad genética con base en la diversidad geográfica. Sin embargo, el uso de diversidad geográfica como indicador de diversidad genética ha sido cuestionado debido a que no se cuantifica la diversidad existente entre las poblaciones y de que, en muchos casos, no se verifica la relación entre diversidad genética y diversidad geográfica.

Dos padres no distantes genéticamente tienden a compartir muchos genes o alelos. Cuando estos se cruzan, hay poca complementariedad y bajo vigor, en razón de la poca heterocigosidad alélica en el cruzamiento.

En cambio, cuando dos padres son más distantes genéticamente, ellos difieren en gran magnitud en el número de locis donde los efectos de dominancia son

evidentes, contribuyendo a una mayor manifestación de la heterosis (Cruz y Sousa, 2003).

Entre las herramientas utilizadas para establecer la diversidad genética de parentales, están las medidas de disimilaridad calculadas con base en distancias (Euclidiana, Mahalanobis), las técnicas de agrupamiento o de clasificación (optimizado, jerárquico), el Análisis de Componentes Principales y el Análisis Discriminante, entre otros.

HETEROSIS

La heterosis es un fenómeno por el cual los híbridos F_1 , derivados del cruzamiento entre dos padres divergentes muestran superioridad sobre sus progenitores en vigor, rendimiento, tamaño de panícula, entre otros. Este término fue propuesto por Shull en 1908, para denotar el estímulo del tamaño y el vigor en un híbrido como expresión del vigor híbrido. Ambos términos, vigor híbrido y heterosis, son sinónimos y pueden usarse indistintamente (Virmani *et al.*, 2003; Poehlman y Sleper, 2003).

De acuerdo con Hallauer y Miranda (1995), los efectos de la hibridación fueron descritos por Kolreuter en 1766, Knight en 1799 y Von Gärtner en 1849. La primera interpretación correcta de hibridación en maíz fue descrita en una carta por Cotton Mather en 1716. Naudín y Mendel en forma independiente, en 1865, describieron no solamente las leyes fundamentales de la herencia fundando la ciencia de la genética, sino que también detallaron los resultados de sus cruzamientos.

Ellos notaron y comentaron sobre el crecimiento lujurioso observado en los híbridos y, debido a la naturaleza influyente eventual de su trabajo, sus observaciones adicionales contribuyeron al desarrollo de un concepto de heterosis (Goldman, 1999).

En 1908, Shull desarrolló una perspectiva sobre la heterosis que expuso en una publicación, titulada: "Composición de un campo de maíz". Este artículo es considerado el punto de partida para el desarrollo de la naciente teoría heterótica que indica claramente que una variedad de maíz es una mezcla compleja de genotipos y que cada planta, en cierto sentido, se aísla, produciendo cada una, un genotipo esencialmente diferente.

Shull acuñó la palabra heterosis y consideró al vigor híbrido como un sinónimo descriptivo de esta, sin pensar en la implicación de mecanismo alguno (Crow, 1999; Goldman, 1999 y Virmani *et al.*, 2003).

Tipos de heterosis. El comportamiento de un híbrido respecto a sus parentales, puede ser expresado de tres formas, dependiendo del genotipo utilizado como referencia. Existen muchas formas de heterosis, sin embargo, acorde con Fehr (1987) y Virmani *et al.*, (2003), la heterosis puede ser media-parental (HMP), heterobeltiosis (H) y relativa (HR).

La heterosis media-parental es el incremento o decremento en el comportamiento de un híbrido comparado con el valor promedio de sus parentales. La heterobeltiosis es el incremento o decremento en el comportamiento de un híbrido en comparación con el mejor progenitor de la combinación híbrida. La heterosis estándar es el incremento o decremento en el comportamiento de un híbrido comparado con un testigo comercial o variedad estándar de la región.

Desde el punto de vista práctico, la heterosis estándar es de mayor importancia, debido a que el mejoramiento apunta a desarrollar híbridos que sean mejores que las variedades de alto rendimiento cultivadas comercialmente por los agricultores.

Estimación de la heterosis. La medida de la heterosis es muy simple. Generalmente se expresa como el porcentaje del incremento o decremento en el comportamiento de un híbrido en comparación con un genotipo referencia o un parámetro (Fehr, 1987; Vallejo y Estrada, 2002; Virmani *et al.*, 2003).

A continuación se presentan las fórmulas para calcular los tres tipos de heterosis.

$$\text{Heterosis media-parental (HMP): } HMP = \frac{F_1 - \bar{P}}{\bar{P}} \times 100$$

$$\text{Heterobeltiosis (H): } H = \frac{F_1 - MP}{MP} \times 100$$

$$\text{Heterosis relativa (HR): } HR = \frac{F_1 - \bar{T}}{\bar{T}} \times 100$$

Donde: F_1 = promedio del híbrido F_1 ,

\bar{P} = promedio de los progenitores que participan en el cruzamiento

\bar{T} = comportamiento del genotipo comercial

En la Tabla 1 se observan los resultados de HMP y H para caracteres de importancia económica para cultivos de especies alógamas, en las cuales, el uso del vigor híbrido es más generalizado debido a que la polinización cruzada es el sistema natural de reproducción.

Tabla 1. Heterosis media parental (HMP), heterobeliosis (HB) y heterosis estandar para algunos caracteres de importancia económica en algunos cultivos.

Especie	Carácter	Heterosis (%)	Fuente
<i>Physalis ixocarpa</i>	Rendimiento de fruto	138,7 ^{HB}	Sahagún <i>et al.</i> , 1999
<i>Solanum melongena</i>	Altura de planta	26,4 ^H	Saha <i>et al.</i> , 1991
	Ramas por planta	48,5 ^H	
<i>Solanum melongena</i>	Rendimiento de fruto	41,2 a 13,3 ^{HMP}	Sousa y Maluf, 1998.
	Peso de fruto	-11,5 a +26,17 ^{HMP}	
	Precocidad	+29 a +141,8 ^{HMP}	
	Rendimiento de fruto, producción total de fruto y peso de fruto	13,9 a 92,5 ^H	
<i>Solanum tuberosum</i>	Número de tubérculos	H y HMP positiva y significativa	González <i>et al.</i> , 1996
<i>Capsicum annum</i>	Rendimiento de fruto	-15,5 a 37,1 ^H	Lagos, 1998
	Peso de fruto	-19,8 a 28,5 ^{HB} -39,9 a 7 ^H -57,6 a 4,55 ^{HB}	
<i>Nicotiana tabacum</i>	Días a floración	-2,60 ^{HMP}	Legg, 1970
<i>Zea mays</i>	Rendimiento de grano	6,10 a 68,6 ^{HMP}	Melchinger <i>et al.</i> , 1998
	Rendimiento de grano	-6,7 a 16,7 ^{HMP}	Romero <i>et al.</i> , 2002
	Rendimiento de grano	8,0 a 92 ^{HMP}	Parentoni <i>et al.</i> , 2001

Bases genéticas de la heterosis. Se han propuesto dos hipótesis principales para explicar las bases genéticas de la heterosis: la sobredominancia formulada por Shull en 1908 y la dominancia formulada por Davenport en 1908. Al mismo tiempo en que la herencia Mendeliana llegó a ser aceptada, existieron estas dos alternativas.

Los estudios sobre las bases genéticas de la heterosis para características cuantitativas en varios cultivos, han mostrado que esta es el resultado de dominancia completa, parcial, sobredominancia o puede ser una combinación de todas ellas (Virmani *et al.*, 2003).

En 1908, Shull y East propusieron que diferentes tipos de germoplasma estimulan el desarrollo y que el estímulo aumenta con la diversidad de éstos. En términos Mendelianos, esto significa heterocigosidad. Esta es la hipótesis de la sobredominancia, la cual fue llamada por Fisher como Súper-dominancia, estimulación por heterocigosis, heterosis del gen simple y acción acumulativa de alelos divergentes (Márquez, 1988; Crow, 1999; Goldman, 1999 y Virmani *et al.*, 2004).

En realidad, la sobredominancia es un caso particular de interacción intra locus; cuando el grado de dominancia es superior a la unidad. Esta hipótesis tiene la desventaja de que no es posible probarla a nivel de un locus simple. Siendo en los caracteres cuantitativos donde es más clara la manifestación de la heterosis, por su misma naturaleza y definición no es posible separar los efectos de cada locus individual.

En estas condiciones, la sobre dominancia existirá cuando $Aa > AA$ ó aa , siendo Aa el heterocigoto y AA y aa los homocigotos de los genes involucrados en el carácter (Márquez, 1988 y Virmani *et al.*, 2004)

Alternativamente, la heterosis puede ser el resultado de la no manifestación de alelos deletéreos recesivos (enmascarados) por dominancia o dominancia parcial. Cada raza aporta al híbrido una colección algo diferente de alelos dominantes favorables.

Generalmente, la naturaleza deletérea de los alelos recesivos fue enfatizada primero por Davenport en 1908. La hipótesis de la dominancia, fue planteada primero en forma explícita por Bruce en 1910, aunque curiosamente el argumento algebraico de Bruce pudo haber sido usado para explicar la sobre dominancia (Crow, 1999).

La heterosis por dominancia se debe a la acumulación de genes dominantes en un híbrido proveniente de dos parentales divergentes, tal como se plantea a continuación:

Parentales	P_1	x	P_2
Genotipo	AAbbCCdd	x	aaBBccDD
	(2 genes dominantes)		(2 genes dominantes)
F_1	AaBbCcDd		
	(4 genes dominantes)		

MEDIDAS DE DISIMILARIDAD

Como ya se anotó, el éxito de un programa de fitomejoramiento reside en la existencia de variabilidad genética en la población base o de trabajo. Esa divergencia puede ser evaluada en características agronómicas, morfológicas, anatómicas, fisiológicas, bioquímicas, isoenzimáticas y moleculares, entre otras.

Las múltiples informaciones de cada genotipo son expresadas en medidas de disimilaridad, que representan la diversidad existente en el conjunto de accesiones estudiadas (Cruz y Souza, 2003).

Las medidas de disimilaridad son de gran importancia en estudios de diversidad genética en que se procura identificar parentales que serán utilizados en programas de hibridación.

Existe la expectativa que progenitores de buen comportamiento, con alto grado de diversidad, pueden presentar genes complementarios que proporcionarán una F_1 con mayor heterosis y generaciones con individuos transgresivos.

En otras situaciones, los estudios sobre diversidad genética se realizan con el objeto de identificar genotipos con mayor similitud que permitan la formación de multilíneas.

En el manejo de bancos de germoplasma, los coeficientes de similitud evidencian la existencia de duplicados que pueden ser eliminados, reduciendo los costos de conservación de las accesiones (Cruz y Souza, 2003).

De manera general, los estudios de diversidad genética han sido realizados de la siguiente forma:

- Medidas de disimilaridad obtenidas de variables cuantitativas
- Medidas de disimilaridad obtenidas de variables binarias
- Medidas de disimilaridad obtenidas de variables multicategóricas

Variabes cuantitativas. Cruz y Souza (2003), Crisci y López (1983) afirman que los valores obtenidos a partir de la aplicación de los coeficientes de disimilaridad o distancia varían de cero a infinito; cero es la máxima similitud; además manifiestan que los coeficientes de disimilaridad comúnmente utilizados, son los siguientes:

- *Mean Character Diffeerence (MCD)*. Propuesto como medida taxonómica por Cain y Harrison en 1958; se expresa como la diferencia del valor absoluto de la diferencia entre cada estado de los caracteres de los genotipos, dividido por el número de caracteres:

Este coeficiente recibe el nombre de distancia de Manhattan, cuando no se divide por el número de caracteres.

- *Distancia Taxonómica (DT)*. Es muy usado y fue propuesto por Sokal en 1961, y expresado en la fórmula:

$$DT = \left\{ \sum_{j=1}^n (X_{ij} - X_{ik})^2 \right\}^{1/2}$$

Debido a que el número de caracteres influye en la estimación de la distancia mediante la distancia taxonómica, es común que en los cálculos se utilice una distancia promedio.

- *Distancia Euclidiana (d_{ij})*. Considerando una observación Y_{ij} en el i -ésimo genotipo (clón, línea, etc.) para la j -ésima variable, la distancia euclidiana entre un par de genotipos i e i' , se expresa así:

$$d_{ii'} = \sqrt{\sum_j (X_{ij} - X_{i'j})^2}$$

Donde: $d_{ii'}$ = distancia entre el par de genotipos i e i'
 X_{ij} = observación en el i -ésimo genotipo para la j -ésima característica
 $X_{i'j}$ = observación en el i' -ésimo genotipo para la j -ésima característica

En todas las distancias antes citadas, la escala afecta el valor obtenido. Adicionalmente, ellas son cuantificadas en diferentes unidades (peso, longitud, porcentajes, etc.), siendo, en estos casos, recomendable el cálculo de las distancias utilizando los valores estandarizados.

- *Distancia de Mahalanobis (D_{ij}^2)*. Una crítica que se le ha hecho a la distancia euclidiana es que no considera las varianzas y covarianzas residuales que existen entre las características medidas, posibles de ser cuantificadas cuando las evaluaciones son hechas en los genotipos evaluados en el experimento.

Con X_1, X_2, \dots, X_p , como variables medidas en cada genotipo y d_1, d_2, \dots, d_p como:

$$\bar{X}_1 - \bar{X}_1^2, \bar{X}_2 - \bar{X}_2^2, \dots, \bar{X}_p - \bar{X}_p^2$$

las diferencias en las medias de dos poblaciones respectivamente, las D_{ii}^2 de Mahalanobis se define como:

$$D_{ii}^2 = b_1 d_1 + b_2 d_2 + \dots + b_p d_p$$

Aquí, el valor b_i será estimado, de tal forma, que la proporción de la varianza entre las poblaciones con respecto a la varianza dentro de las poblaciones se maximice (Singh y Chaudhary, 1985).

$$D_{ii}^2 = W^{ii} (\bar{X}_i^1 - \bar{X}_i^2)(\bar{X}_i^1 - \bar{X}_i^2)$$

En términos de varianzas y covarianzas el valor D_{ii}^2 se obtiene de la siguiente forma:

Donde, W^{ii} es la matriz inversa de la matriz de varianzas y covarianzas. Cuando se tienen resultados de un Análisis de Componentes Principales la D_{ii}^2 de Mahalanobis, se obtiene así:

$$D_{ii}^2 = \frac{\sum_{K=1}^p (X_{ik} - X_{rk})^2}{\lambda_K}$$

Donde: D_{ii}^2 = distancia entre los genotipos i e i'
 P = componentes principales retenidos
 X_{ik} = aporte de i -ésimo genotipo al K -ésimo componente principal
 λ_K = valor propio del K -ésimo componente principal

Variables binarias. Una forma de evaluar la diversidad genética de una colección de recursos filogenéticos es a través de marcadores moleculares, en donde se evalúa la presencia o ausencia del marcador (bandas), codificados con uno y cero, en su orden.

En este caso, para calcular el índice de disimilaridad entre los genotipos i e i' (G_i y $G_{i'}$), se construye una tabla de contingencia que cruza la i -ésima columna con la i' -ésima fila (Crivisqui, 1999).

En la Tabla 3, se observan algunas formas para calcular coeficientes de similaridad utilizados en estudios de diversidad que contemplan variables binarias.

Tabla 3. Coeficientes de similaridad utilizados en estudios de diversidad genética cuando se consideran variables binarias.

Coeficiente	Dado por
Coincidencias simples	$S_{ii} = (a + d) / (a + b + c + d)$
Roger y Tanimoto	$S_{ii} = (a + d) / (a + 2(b + c) + d)$
Sokal y Sneath	$S_{ii} = 2(a + d) / (2(a + d) + b + c)$
Russel y Rao	$S_{ii} = a / (a + b + c + d)$
Jaccard	$S_{ii} = a / (a + b + c)$
Sorenso y Nei Li	$S_{ii} = 2a / (2^a + b + c)$
Anderberg	$S_{ii} = a / (a + 2(b + c))$
Sokal y Michener	$S_{ii} = (a + d) / (a + b + c + d)$
Baroni, Urbani y Buser	$S_{ii} = (a + ad) / (a + b + c + ad)$
Haman	$S_{ii} = [(a + d) \cdot (b + c)] / (a + b + c + d)$
Yule	$S_{ii} = (ad - bc) / (ad + bc)$
Ochiai	$S_{ii} = \frac{a}{\sqrt{(a+b)(a+c)}}$
Ochiai II	$S_{ii} = \frac{ad}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$

Como se trata de medidas de similaridad, es recomendable, que al hacer el análisis de clasificación, se utilice las medidas de disimilaridad obtenidas por el complemento aritmético o por el inverso del coeficiente obtenido, o sea:

- El complemento aritmético de un coeficiente de similaridad es iguala $d_{ii} = 1 - S_{ii}$.
- El inverso del coeficiente está dado por $d_{ii} = 1 / (S_{ii} + 1)$.

Variables multicategóricas. Las variables multicategóricas están asociadas con particularidades morfológicas y estructuras de la planta que tienen que ver con la calidad del producto comercial, como la forma, la coloración, el sabor y la consistencia de la pulpa de los frutos, entre otros.

Cuando los valores de caracteres multicategóricos pueden ser ordenados, estableciéndose una escala, estas pueden ser analizadas como variables cuantitativas discretas (Cruz y Souza, 2003).

Una manera más simple de estimar la similaridad considerando un conjunto de variables multicategóricas, es por medio del siguiente índice de similaridad:

$$S_{ii} = \frac{C}{C+D}$$

Donde: C = concordancia de la categoría y D = discordancia de la categoría.

La disimilaridad está dada por: $d_{ii} = \frac{D}{C+D}$

Considerando varias categorías, se tendría:

$$d_{ii} = \frac{D_1}{C_1 + D_1} + \frac{D_2}{C_2 + D_2} + \dots + \frac{D_v}{C_v + D_v} = \sum_{j=1}^v \frac{D_j}{C_j + D_j}$$

En donde: C_j = número de concordancias entre categorías para la j-ésima variable multicatórica.

D_j = número de discordancias entre categorías para la j-ésima variable multicatórica.

En este índice el valor máximo de disimilaridad es igual al número de características evaluadas y el valor mínimo es cero.

ANÁLISIS DE CLASIFICACION O AGRUPAMIENTO

El análisis de clasificación o agrupamientos comprende técnicas que, siguiendo reglas más o menos arbitrarias, forman grupos de accesiones que se asocian por su grado de similitud. Esta definición es poco precisa y esto se debe a dos factores: primero, el escaso acuerdo entre los investigadores acerca de cómo reconocer los límites entre grupos, y segundo, la enorme variedad de técnicas propuestas (Crisci y López, 1983).

Según Crisci y López (1983), estas técnicas se clasifican con base en los siguientes criterios:

- Técnicas que forman grupos exclusivos versus técnicas que forman grupos no exclusivos
- Técnicas que forman grupos jerárquicos versus técnicas que forman grupos no jerárquicos
- Técnicas divisivas versus técnicas aglomerativas y
- Técnicas secuenciales versus técnicas simultáneas.

Método de clasificación de optimización. Este método, al igual que las técnicas jerárquicas, es uno de los métodos más utilizados; en este caso, los grupos son formados de acuerdo a algún criterio de agrupamiento, o sea, un objetivo es alcanzar una partición de los individuos que optimice (maximice o minimice) alguna medida predefinida.

Uno de los métodos más comúnmente utilizados en el área de fitomejoramiento es el propuesto por Tocher (Cruz y Souza, 2003). Este método requiere de una matriz de disimilaridad, sobre la cual es identificado un par de individuos más similares. Estos individuos formarán un grupo inicial. A partir de este grupo, hay la posibilidad de incluir nuevos individuos, teniendo en cuenta el criterio de que la distancia media intragrupo debe ser menor que la distancia media intergrupo.

La inclusión o no de un individuo k a un grupo, considera el hecho que:

- Sí $d_{(\text{grupo})k}/n \leq \epsilon$, el individuo k se incluye al grupo;
- Sí $d_{(\text{grupo})k}/n > \epsilon$, el individuo k no se incluye al grupo.

Siendo n el número de individuos que constituyen el grupo original. En este caso, la distancia entre un individuo k y un grupo formado por los individuos ij, está dada por:

$$d_{(ij)k} = d_{ik} + d_{jk}$$

Métodos de clasificación Jerárquicas. Estos métodos incluyen los métodos jerárquicos ascendentes, en los cuales los genotipos son agrupados por un proceso que se repite en varios niveles, hasta que queda establecido un dendrograma, el cual es una representación gráfica de la información contenida en la matriz de datos n x p (n = genotipos, p = variables). Dentro de estos métodos se destacan: el vecino más próximo y el de Ward (Crivisqui, 1999).

- *Vecino más próximo.* En este método, el dendrograma es establecido por los genotipos con mayor similitud, siendo la distancia entre un individuo k y un grupo formado por los genotipos i y j, dada por: $d_{(ij)k} = \min\{d_{ik}; d_{jk}\}$, o sea $d_{(ij)k}$ está dada por el menor valor del conjunto de distancias de pares de individuos (i y k) y (i y j) (Cruz y Souza, 2003).

Método de Ward. En este método se considera, para la formación inicial del grupo, aquellos genotipos que proporcionan una menor inercia o varianza.

El agrupamiento se hace a partir de sumas de cuadrado de desvíos entre los genotipos o, alternativamente, a partir del cuadrado de la distancia euclidiana.

En este análisis de agrupamiento se identifica una matriz D (cuyos elementos son los cuadrados de las distancias euclidianas - d_{ii}^2) o una matriz S (cuyos elementos son las sumas de cuadrados de los desvíos - SCD_{ii}), el par de genotipos que proporcionan una menor suma de cuadrados de los desvíos. Con estos genotipos agrupados, se construye una nueva matriz de disimilaridad de menor dimensión, y así, sucesivamente.

En este procedimiento, el análisis se realiza, proporcionando g-1 (g es el número de genotipos a ser agrupados) pasos de agrupamiento para que se forme el dendrograma.

Otra forma de calcular la matriz de distancias para el método de Ward, es aplicando la fórmula de varianza. De esta forma, se agrupan los individuos que conllevan a la mínima pérdida de inercia o variabilidad. La distancia entre el individuo i e i' (d_{ii}), se calcula de la siguiente forma:

$$d_{ii} = \sum X_j^2 - n\bar{X}^2$$

Donde:

X_j^2 = cuadrados de la variable j en los individuos del agrupamiento (i e i')

n = número de individuos involucrados en la agrupación

\bar{X} = promedio de la variable j en los individuos del agrupamiento (i e i')

ANÁLISIS DISCRIMINANTE

El análisis discriminante busca obtener funciones que permitan clasificar un individuo, con base en la información de un grupo de características medidas en una población entre varias conocidas, buscando minimizar las probabilidades de una clasificación errónea. Así, se deben obtener funciones que permitan asignar un individuo a la población que él realmente pertenece.

En estos análisis, los estudios son realizados preliminarmente, a partir de la información de poblaciones previamente conocidas o de una clasificación previa. Una vez se conoce la eficacia de la discriminación, las funciones pueden

ser utilizadas para ubicar nuevos individuos, de los cuales se desconoce su origen (Cruz y Souza, 2003). Por ejemplo, en fitomejoramiento, el análisis discriminante puede utilizarse para diferenciar entre genotipos tolerantes y susceptibles a un estrés abiótico.

Otra aplicación, es aquella donde los genotipos pueden discriminarse con base a su origen geográfico, como el caso de los frijoles mesoamericanos y andinos.

La función discriminante para establecer si un individuo pertenece al grupo 1 (G_1) o al grupo 2 (G_2), debe ser una función de d_i y d_j (D^2), donde:

$$d_i = \bar{X}_i - \bar{X}_i'$$

$$d_j = \bar{Y}_j - \bar{Y}_j'$$

Donde: d_i = diferencia en la variable X entre los grupos G_1 y G_2 ,
 d_j = diferencia en la variable Y entre los grupos G_1 y G_2 .

La función discriminante es igual a: $D^2 = a_i d_i + a_j d_j$. En forma vectorial sería:

$$D^2 = (a_i a_j) \begin{bmatrix} d_i \\ d_j \end{bmatrix}$$

De aquí, se tiene que:

$$(a_i a_j) = (d_i d_j) S^{-1}$$

Donde: S^{-1} = matriz inversa de varianzas y covarianzas S para las variables X y Y.

Por otro lado: $D^2 = [a_i \bar{X}_i - a_j \bar{Y}_j] - [a_i' \bar{X}_i' - a_j' \bar{Y}_j']$

Entonces: $D^2 = L_1 - L_2$, donde: $L = a_i X + a_j Y$ es la función discriminante de genotipos, para determinar si pertenecen al grupo G_1 (L_1) o al G_2 (L_2).

BIBLIOGRAFIA

COMSTOCK, R.E. AND ROBINSON, H.F. 1948. The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. *Biometrics*. 4: 254-266.

CEBALLOS LASCANO, H. 1997. *Genética Cuantitativa y Fitomejoramiento*. Palmira, Universidad Nacional de Colombia Sede Palmira. 330p.

CRISCI, J; Y LOPEZ, MARIA. 1983. *Introducción a la teoría práctica de la taxonomía numérica*. Washington, D.C., Secretaría General de la Organización de los Estados Americanos. 128 p.

CRIVISQUI EDUARDO. 1999. *Presentación de los métodos de clasificación*. Pasto, Universidad de Nariño, Université Libre de Bruxelles, Seminario de métodos estadísticos multivariados aplicados a la investigación. 57 p.

CROW, J.F. 1999. Dominance and overdominance. Chapter 5. In: Coors James, G. and Pandey Shivaji. *Genetics and exploitation of heterosis in crops*. Madison, Wisconsin, USA, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. 524 p.

CRUZ COSME DAMIÃO E SOUZA PEDRO CRESCÊNCIO. 2003. *Modelos biométricos aplicados ao melhoramento genético*. Volume 2. Viçosa, Editora UFV. 585p.

FEHR WALTHER. 1987. *Principles of cultivar development: theory and technique*. Volume I. New York, Iowa University, Macmillan. 536p.

GARDNER, C.O. AND EBERHART, S.A. 1967. Simplified methods for estimating constants and computing sums of squares for a diallel cross analysis. *Fitotecnia Latinoamericana*. 4: 1-12.

GOLDMAN, I.L. 1999. Inbreeding and outbreeding in the development of a modern heterosis concept. Chapter 2. In: Coors James, G. and Pandey Shivaji. *Genetics and exploitation of heterosis in crops*. Madison, Wisconsin, USA, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. 524 p.

GRIFFING BRUCE. 1956. Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Sciences*. 9(4): 463-493.

HALLAUER A. AND MIRANDA, F. 1995. *Quantitative genetic in maize breeding*. Second edition, second printing. Iowa, Iowa State University Press. 462 p.

HAYMAN, B.I. 1954. The theory and analysis of diallel crosses. *Genetics*. 39: 789-809.

LIEGG PAUL, D., COLLINS, G.B. AND LITTON, C.C. 1970. Heterosis and Combining Ability in Diallel Crosses of Burley Tobacco, *Nicotiana tabacum* L. *Crop Science*. 10: 705-707.

LAGOS, T. 1998. *Avance generacional y selección de líneas promisorias de pimentón *Capsicum annum* L., en los municipios de Palmira y Candelaria, departamento del Valle*. Tesis M.Sc. Palmira, Universidad Nacional de Colombia sede Palmira. 129p.

MARQUEZ, F. 1988. *Genotecnia Vegetal*. Tomo II. México, AGT. 665p.

MELCHINGER, A.E., GUMBER, R.K., LEIPERT, R.B., VUYLSTEKE, M. AND KUIPER, M. 1998. Prediction of testcross means and variances among F_3 progenies of F_1 crosses from testcross means and genetic distances of their parents in maize. *Theor Appl Genet*. 96: 503-512.

PARDOS, J.A. 1988. Variabilidad: sus clases y uso en la mejora forestal. En: Pardos J.A. (Ed). *Mejora genética de especies arbóreas forestales*. Madrid, FUCOVASA. Pp 30-46.

PARENTONI, S.N., MAGALHÃES, J.V., PACHECO, C.A.P., SANTOS, M.X., ABADIE, T., GAMA, E.E.G., GUIMARÃES, P.E.O., MEIRELLES, W.F., LOPES, M.A., VASCONCELOS, M.J.V. AND PAIVA, E. 2001. Heterotic groups based on yield-specific combining ability data and phylogenetic relationship determined by RAPD markers for 28 tropical maize open pollinated varieties. *Euphytica*. 121: 197-208.

POEHLMAN JOHN MILTON Y SLEPER DAVID ALLEN. 2003. *Mejoramiento*

genético de las cosechas. Segunda edición. Trad. del Inglés por Guzman, M. y Hernández, M.A. México, Limusa. 511 p.

ROMERO PEÑALOZA, J; CASTILLO GONZALEZ, F Y ORTEGA PACZKA , R. 2002. Cruzas de poblaciones nativas de maíz de la raza Chalqueño: II Grupos genéticos, divergencia genética y heterosis. Rev. Fitotec. Mex. 25(1):107-115.

SAHA, M.G., HOSSAIN, A.K.M.A., HOQUE, K.R. AND BHOWMIK, A. 1991. Genetic analysis of plant height and number of branches in brinjal (*Solanum melongena* L.). Annals of Bangladesh Agriculture. 1, 2: 91-97.

SAHAGUN, J., GOMEZ, F.Y PEÑA, A. 1999. Efectos de aptitud combinatoria en poblaciones de tomate de cáscara (*Physalis ixocarpa* Brot.). Revista Chapingo, Serie Horticultura. 5(1): 23-27.

SINGH, R.K. AND CHAUDHARY, B.D. 1985. Biometrical methods in quantitative genetic análisis. New Delhi, Kalyani. 318 p.

SOUSA DE, J.A. AND MALUF, W.R. 1998. Expression of heterosis for productive traits in F₁ eggplant (*Solanum melongena* L.) hybrids. Genetics and Molecular Biology. 21(1): 99-103.

VALLEJO FRANCO, A Y ESTRADA EDGAR, I. 2002. Mejoramiento genético de plantas. Palmira, Universidad nacional de Colombia-Sede Palmira. 402p.

VIRMANI, S.S., PANDEY, M.P., SINGH, I.S. AND WEI JUN XU. 2004. Classical and molecular concepts of heterosis. In: Jain, H.K. and Kharkwal, M.C. (Editors). Plant breeding: Mendelian to Molecular Approaches. New Delhi, Narosa Publishing House. 811 p.

VIRMANI, S.S., SUN, Z.X., MOU, T.M., JAUHAR ALI, A. AND MAO C.X. 2003. Two-Line hybrid rice breeding manual. Los Baños (Philippines), International Rice Research Institute. 88p.