

# ANÁLISIS DE ASOCIACIÓN DE CARACTERES EN EL CULTIVO DE ARROZ (*Oryza sativa* L.) EMPLEANDO TÉCNICAS MULTIVARIADAS

R. Morejón<sup>✉</sup> y Sandra H. Díaz

**ABSTRACT.** The work was developed in areas from “Los Palacios” Rice Research Station, belonging to the National Institute of Agricultural Sciences (INCA), with the objective of providing the commissioned persons for the management and preservation of plant genetic resource collections, the necessary tools for the analysis of character association in rice genotypes and obtaining in this way as much information and comprehension as possible by means of multivariate techniques. The associations among variables identified by the correlation matrix, by the discrimination degree identified in the Principal Components Analysis and confirmed by the Conglomerate Analysis for 65 rice genotypes used, constitutes a useful information for breeding programs in this crop.

*Key words:* rice, agronomic characters, statistical methods

**RESUMEN.** El trabajo se desarrolló en áreas de la Estación Experimental del Arroz “Los Palacios”, perteneciente al Instituto Nacional de Ciencias Agrícolas (INCA), con el objetivo de proporcionar a las personas encargadas del manejo y la conservación de las colecciones de recursos fitogenéticos las herramientas necesarias para el análisis de la asociación de caracteres en genotipos de arroz y, de esta forma, obtener la mayor información y comprensión posible de estos a partir de técnicas multivariadas. Las asociaciones entre variables que fueron identificadas por la matriz de correlaciones, por el grado de discriminación identificado en el Análisis de Componentes Principales y confirmadas por el Análisis de Conglomerados para los 65 genotipos de arroz utilizados, constituyen una información de utilidad para los programas de mejoramiento genético en este cultivo.

*Palabras clave:* arroz, características agronómicas, métodos estadísticos

## INTRODUCCIÓN

En la caracterización de una especie se estima la variabilidad existente en el genoma de la población de individuos que la conforman. Así, el genoma de las especies de animales o plantas contiene toda la información codificada en forma de genes que se necesitan, tanto para establecer su identidad morfológica como para desarrollar todos los procesos y funciones vitales para su supervivencia. Se estima que las plantas superiores poseen un poco más de 400 000 genes con funciones particulares dentro de la especie y un buen número de ellos ha creado variantes por efectos evolutivos y del medio ambiente. Esas variantes se van acumulando entre los diferentes miembros de la especie y la suma de todos los efectos de los genes y sus variantes es lo que se denomina variabilidad genética de una especie (1).

Existe una variabilidad que se detecta a simple vista y otra que requiere de técnicas especiales para ser detectadas; por ello, es primordial cuál de ellas se intenta

medir para poder elegir los métodos estadísticos adecuados para analizar los datos resultantes de un estudio de caracterización. Por lo general, las publicaciones al respecto contienen poca información sobre estos métodos y su empleo no resulta práctico para la mayoría de los usuarios (2).

Aunque existen trabajos de tesis y artículos científicos, su forma es puntual y no destacan la importancia del análisis estadístico en la caracterización de los recursos genéticos, ni sirven de guía para que los investigadores seleccionen las técnicas apropiadas según sus necesidades.

Por todo lo anterior, este trabajo tiene como objetivo proporcionar a las personas encargadas del manejo y la conservación de las colecciones de recursos filogenéticos, las herramientas necesarias para el análisis de la asociación de caracteres en genotipos de arroz y, de esta forma, obtener la mayor información y comprensión posible de estos a partir de técnicas multivariadas.

## MATERIALES Y MÉTODOS

El trabajo se desarrolló en áreas de la Estación Experimental del Arroz “Los Palacios”, perteneciente al Instituto Nacional de Ciencias Agrícolas (INCA), en condiciones de aniego sobre un suelo Hidromórfico Gley

Ms.C. R. Morejón y Ms.C. Sandra H. Díaz, Investigadores Agregados de la Estación Experimental del Arroz “Los Palacios”, Instituto Nacional de Ciencias Agrícolas, Gaveta Postal 1, San José de las Lajas, La Habana, Cuba, CP 32 700.

✉ shdiaz@inca.edu.cu

Nodular Ferruginoso (3), en el cual se sembraron 65 accesiones de arroz de diferente procedencia.

La siembra se realizó de forma directa a chorrillo, en parcelas de 4 m<sup>2</sup> y se utilizó un Diseño Aumentado Modificado (DAM), el cual se estructuró mediante un Cuadrado Latino (3x3), con tres líneas controles y 66 líneas de prueba (las 65 a comparar más una de ellas que se repite), distribuidas al azar en las subparcelas correspondientes (4); se realizaron las atenciones culturales de fertilización, riego y tratamientos fitosanitarios según lo estipulado en el Instructivo técnico del cultivo del arroz (5).

Se evaluaron 10 caracteres durante el ciclo de desarrollo del cultivo:

• 50 % de floración (días)	[50%F]
• Altura final de las plantas (cm)	[AP]
• Largo de la hoja (cm)	[LH]
• Ancho de la hoja (cm)	[AH]
• Longitud de la panícula (cm)	[LP]
• Panícula por m <sup>2</sup> (no.)	[Pm <sup>2</sup> ]
• Granos llenos por panícula (no.)	[Gll]
• Granos vanos por panícula (no.)	[Gv]
• Peso de 1000 granos (g)	[Pgr]
• Rendimiento agrícola (t.ha <sup>-1</sup> )	[R]

Para las evaluaciones realizadas se emplearon las siguientes metodologías:

- ♦ Sistema de Evaluación Estándar para Arroz (6).
- ♦ Formulario de Descripción Varietal para Arroz (7).

Las panículas por metro cuadrado se muestrearon una vez por parcela en un área de 0.1 m<sup>2</sup> y los granos llenos y vanos por panícula junto al peso de 1000 granos se determinaron en 20 panículas centrales tomadas al azar; asimismo, el rendimiento agrícola fue calculado en 1 m<sup>2</sup>.

La matriz de datos obtenida (genotipos x variables) fue procesada mediante las técnicas multivariadas de Componentes Principales, herramienta útil para conocer la relación existente entre las variables cuantitativas consideradas, con el fin de saber cuáles variables están o no asociadas, cuáles caracterizan en el mismo sentido o en el sentido contrario, también permite seleccionar las variables cuantitativas más discriminatorias para limitar el número de mediciones en caracterizaciones posteriores, y Conglomerados, método analítico que se aplica para clasificar las variables en grupos relativamente homogéneos con base en alguna similitud existente entre ellas (2).

## RESULTADOS Y DISCUSIÓN

La matriz de correlaciones entre cada par de característica se presenta en la Tabla I. Para este caso, se consideraron coeficientes significativos aquellos superiores a 0.4433 ( $p \leq 0.001$ ).

El rendimiento se correlacionó con todos sus componentes [panículas por m<sup>2</sup> (-0.485), granos llenos (0.826), peso de 1000 granos (0.775)], lo que indica que las plantas con menos panículas por m<sup>2</sup> y mayor número de granos llenos y peso de 1000 granos fueron las de mayor rendimiento; al evaluar el comportamiento de otras variedades de arroz se han obtenido resultados similares (8, 9, 10, 11, 12).

Además, el rendimiento se correlacionó de forma positiva con la longitud de la panícula (0.612). En trabajos realizados en Indonesia se sugiere que la selección para alto rendimiento puede ser dirigida hacia el tamaño de las panículas (13).

Una correlación positiva se encontró entre el peso de 1000 granos con la longitud de la panícula (0.544) y los granos llenos por panícula (0.606), esta última variable también mostró una correlación positiva con la longitud de la panícula (0.518) y negativa con las panículas por m<sup>2</sup> (-0.701).

La longitud de la hoja, la altura de la planta y el 50 % de floración están correlacionadas fuertemente y de forma directa entre sí con valores superiores a 0.900. Otros han referido que la altura de una planta de arroz en floración está positivamente correlacionada con la longitud del ciclo (14).

Del Análisis de Componentes Principales, la Tabla II muestra los valores propios, el porcentaje de contribución de los componentes y su acumulado, así como la proporción de la varianza explicada por cada variable original sobre los componentes.

La selección del número de componentes que se debe tomar para el análisis es aún tema de discusión entre especialistas, ya que no existen pruebas estadísticas inferenciales que permitan probar la significación de dichos valores. Existen diversos criterios de selección, entre ellos el criterio de Cliff (1987), que indicó que se deben considerar como aceptables los componentes cuyos valores propios expliquen un 70 % o más de la varianza total y el criterio de Kaiser (1960), que consiste en la selección de los componentes cuyo valor propio sea mayor o igual a uno (15). Siguiendo estos criterios, se decidió incluir en el análisis a las primeras tres componentes.

**Tabla I. Matriz de correlaciones fenotípicas**

	50% F	AP	LH	AH	LP	Pm <sup>2</sup>	Gll	Gv	Pgr
AP	0.920								
LH	0.913	0.947							
AH	0.016	-0.099	-0.048						
LP	-0.178	-0.273	-0.157	0.249					
Pm <sup>2</sup>	0.030	0.089	0.057	-0.087	-0.392				
Gll	-0.137	-0.121	-0.103	0.101	0.518	-0.701			
Gv	0.116	0.110	0.137	0.061	-0.435	0.308	-0.386		
Pgr	-0.242	-0.255	-0.250	0.029	0.544	-0.393	0.606	-0.360	
R	-0.182	-0.238	-0.199	0.131	0.612	-0.485	0.826	-0.400	0.775

**Tabla II. Valores propios, porcentaje de contribución y acumulado de las componentes, las correlaciones y proporción de la varianza explicada de cada variable original sobre las componentes seleccionadas**

	CP1	CP2	CP3	Proporción de la varianza
Valores propios	4.0636	2.9529	1.0750	
% contribución	40.6	31.5	9.1	
% acumulado	40.6	72.1	81.2	
50 % de floración	0.281	-0.897	-0.065	0.887795
Largo de la hoja	-0.079	-0.034	-0.915	0.844622
Granos llenos por panícula	-0.883	-0.146	0.104	0.811821
Altura de la planta	-0.783	-0.283	0.030	0.694078
Peso de 1000 granos	0.759	0.147	-0.223	0.647419
Rendimiento	-0.754	-0.159	-0.197	0.632606
Panículas por m <sup>2</sup>	0.285	-0.701	-0.015	0.572851
Longitud de la panícula	-0.713	-0.221	-0.002	0.557214
Ancho de la hoja	0.303	-0.486	0.060	0.331605
Granos vanos por panícula	0.293	0.261	-0.011	0.154091

La interpretación de la correlación entre las variables originales y las componentes originales se debe centrar en los coeficientes; mientras más altos sean estos, independientemente del signo, más eficientes serán en la discriminación de las accesiones. Sobre este aspecto, se sabe que las cargas que se distribuyen en los componentes indican el peso de cada variable asociada o el grado de contribución al componente: por tanto, se recomienda tomar en cuenta el comportamiento observado en las accesiones durante el trabajo de caracterización en relación con cada variable considerada en el estudio (15).

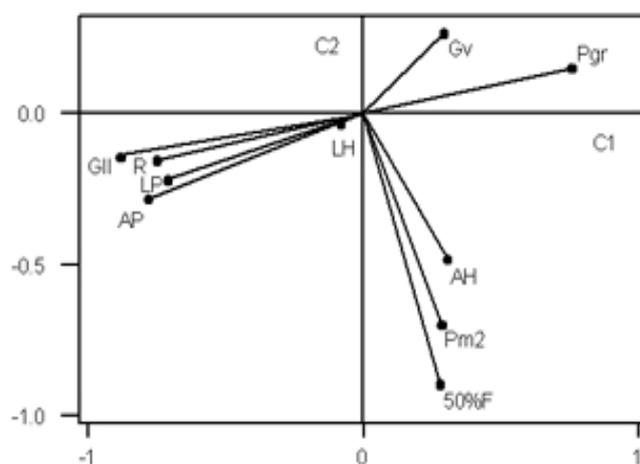
El primer componente contribuyó con más del 40 % de la varianza total explicada, mientras que las correlaciones con las variables originales indican que el rendimiento, el número de granos llenos por panícula, la altura de la planta y la longitud de la panícula fueron las variables que más aportaron, en forma negativa y positivamente en el peso de 1000 granos.

El segundo componente contribuyó con más del 31% de la varianza total y es explicado por los caracteres panículas por m<sup>2</sup> y 50 % de floración, en forma negativa. La contribución del tercer componente fue del 9 % y es explicado por el largo de las hojas.

Por otra parte, las coordenadas de las variables sobre cada componente principal son iguales a la correlación entre las variables originales y las componentes; por tanto, es posible representar la proyección de las variables originales sobre los dos primeros ejes principales (Figura 1).

La proyección opuesta del peso de 1000 granos sobre el primer eje, en relación con los granos llenos, el rendimiento, la altura de la planta y el largo de la panícula, significa que en la medida que sea mayor el peso de 1000 granos, menor será el rendimiento, el número de granos llenos, la altura de la planta y la longitud de la panícula.

La figura también proporciona el grado de asociación entre las variables, que está determinado por la separación angular que forman sus proyecciones; además, considera las distancias de cada una de estas a partir del origen, siendo su contribución mayor mientras más distantes se encuentren.

**Figura 1. Distribución de las variables originales sobre el primer y segundo componente principal en la caracterización de una colección de arroz**

De acuerdo con la separación angular, la mejor asociación está constituida por el rendimiento y los granos llenos y entre ellos dos y la altura de la planta, el largo de la hoja y la longitud de la panícula. Al estudiar el efecto de dos densidades de siembra (120 y 160 kg de semilla.ha<sup>-1</sup>) han comprobado que la mayoría de los caracteres de la panícula, entre ellos su longitud y el número de granos llenos, disminuyen al incrementar la densidad de siembra; sin embargo, el rendimiento (t.ha<sup>-1</sup>) fue superior con la densidad de siembra más alta, demostrando que el efecto detrimental sobre el rendimiento basado en la disminución de los caracteres de la panícula puede ser compensado o superado con el incremento del número de plantas por unidad de área (16).

Sobre la distancia al origen, las variables más importantes son 50 % de floración, granos llenos por panícula, altura de la planta, peso de 1000 granos, rendimiento, panícula por m<sup>2</sup> y longitud de la panícula.

Mediante el Análisis de Componentes Principales, se pudo determinar también el grado de discriminación de las variables, con el objetivo de conocer tanto las de mayor como las de menor variación dentro del

germoplasma, cuantificando la proporción de varianza explicada por cada variable original sobre las tres componentes seleccionadas; para ello, se efectuó la suma de cuadrados de la correlación que forma cada variable original con los componentes (Tabla IV). Esta operación es factible debido a que los componentes no están correlacionados entre sí (15).

En esta tabla se presentan los caracteres ordenados de forma descendente. Aquellos que explican una mayor proporción de varianza son los más discriminatorios del germoplasma y, por tanto, su importancia es mayor. En este caso, se encuentra la variable fenológica 50 % de floración, cuyo comportamiento depende principalmente del genotipo, igualmente las variables que describen la arquitectura de la planta, el largo de la hoja y la altura de la planta en segundo y cuarto lugar.

Otras variables importantes que discriminan suficientemente el germoplasma de arroz son los granos llenos, el peso de 1000 granos, el rendimiento, las panículas por metro cuadrado y la longitud de la panícula en ese orden.

Otra vía empleada para conocer la relación de caracteres para estos 65 genotipos fue a partir del Análisis de Conglomerados. El dendrograma obtenido a partir de esta técnica (Figura 2) para la formación de clases de las variables, muestra que las mejores asociaciones están formadas por el número de granos llenos con el rendimiento, la longitud de las panículas y el peso de 1000 granos, además de 50 % de floración con altura de la planta y longitud de la hoja, en comparación con el resto de las variables consideradas en el análisis.

En este sentido, se conoce que el rendimiento del arroz está en función de sus componentes y en ocasiones depende de las condiciones climáticas y composición varietal empleada, de ahí que existan diversas opiniones en cuanto a las correlaciones que se establecen entre el rendimiento y sus componentes, recomendándose el estudio de las causas de la variación para cada situación dada (17).

Según este resultado, se puede concluir que si se mejora o modifica uno de estos caracteres, podría

modificarse uno del resto. Varios autores plantean que aunque la disminución en algunos de los componentes se compensa en parte con pequeños incrementos de los valores de los restantes, se debe procurar que las variedades alcancen sus valores óptimos para cada uno de los componentes del rendimiento, a fin de lograr la máxima producción (18).

Las asociaciones confirman las variables que fueron identificadas por la matriz de correlaciones y por el grado de discriminación identificado en el Análisis de Componentes Principales y considerando la significación de cada uno de los grupos de variables, es posible obtener información de utilidad para los programas de mejoramiento genético en el cultivo del arroz.

Se debe destacar que no siempre es necesaria la aplicación de varios métodos para estudiar las variables consideradas; en algunos casos, cuando se conoce bien el germoplasma, es suficiente la aplicación de solo uno de los métodos descritos para obtener y comprobar resultados válidos.

## REFERENCIAS

1. Hidalgo, R. Variabilidad genética y caracterización de especies vegetales. *IPGRI. Boletín Técnico*, 2003, no. 8, p. 2.
2. Franco, T. e Hidalgo, R. Análisis estadístico de datos de caracterización morfoagronómica de recursos fitogenéticos. *IPGRI. Boletín Técnico*, 2003, no. 8.
3. Cuba. MINAGRI. Instituto de Suelos. Nueva versión de clasificación genética de los suelos de Cuba. La Habana: Agrinfor, 1999. 64 p.
4. Morejón, R. Estudio de tres métodos de ajuste en un diseño aumentado modificado. *Cultivos Tropicales*, 1998, vol. 19, no. 2, p. 59-63.
5. Cuba. MINAGRI. Instructivo técnico del arroz. La Habana: Instituto de Investigaciones del Arroz, 2000.
6. IRRRI. Standard evaluation system for rice. 4 ed. Filipinas, 1996. 51 p.
7. Cuba. MINAGRI. Formulario de descripción varietal para arroz (*Oryza sativa* L.). Dirección de certificación de semillas. Registro de variedades comerciales. 1998, 12 p.

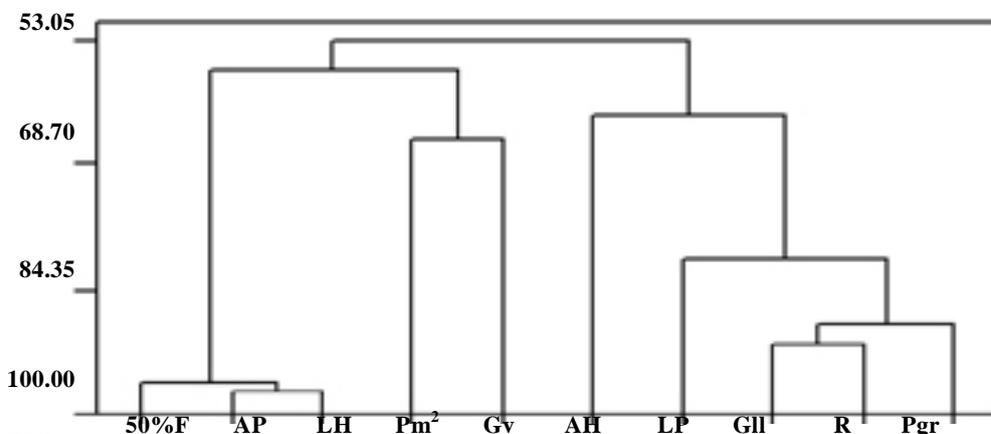


Figura 2. Dendrograma para las 10 variables en estudio según el Análisis de Conglomerados

8. Díaz, S. y Morejón, R. Comportamiento de variedades de arroz de diferente procedencia en la localidad de Los Palacios. *Cultivos Tropicales*, 2002, vol. 23, no. 2, p. 63-66.
9. Díaz, S. H. Caracterización morfoagronómica y genético-bioquímica de 19 accesiones de arroz (*Oryza sativa* L.). [Tesis de Maestría]; Universidad de La Habana, 2000. 98 p.
10. Díaz, S. H.; Morejón, R.; Castro, R.; Pérez, N. y González, M. C. Evaluación de variedades de arroz (*Oryza sativa* L.) para la época de primavera en Pinar del Río. *Cultivos Tropicales*, 2004, vol. 25, no. 4, p. 77-82.
11. Díaz, S. H.; Pérez, N. y Morejón, R. Evaluación del germoplasma de arroz (*Oryza sativa* L.). *Cultivos Tropicales*, 2000, vol. 21, no. 2, p. 55-58.
12. Alfonso, R.; Ramírez, E. y Rodríguez, S. Influencia del estrés hídrico sobre el rendimiento y sus componentes en el cultivo del arroz. En: Seminario Científico Internacional de Agrotecnia Tropical y Congreso de Arroz de Riego y Secano del Área del Caribe. (1: 1999: Camagüey), p. 93.
13. Zen, S. Contribution of agronomic characters and yield components of high elevation rice. En: Proceeding of the seminar on food crops research results in Sukarami Research Institute for Food Crop, west Sumatra. Risalah Seminar hasil penelitian. Balittan Sukarami (Indonesia). Bolittan (1992:Indonesia), 1992. p. 36-41.
14. Aguilar, M. Influencia de la dosis del abonado nitrogenado de fondo sobre los componentes del rendimiento y el comportamiento agronómico del arroz. En: Cultivo del arroz en clima mediterráneo. Junta de Andalucía. España, 1998.
15. Rojas, W. Análisis de la variabilidad genética en quinua. En: Análisis estadístico de datos de caracterización morfoagronómica de recursos fitogenéticos. *IPGRI. Boletín Técnico*, 2003, no. 8.
16. Ramírez, E. M.; Alfonso, R. V. y Medina, V. Efecto de dos densidades de siembra en las características morfológicas de la panícula de arroz (*Oryza sativa* L.). En: Encuentro Internacional de Arroz (1:1998:La Habana), 1998. p. 194-195.
17. López, L. Arroz. Cultivos herbáceos. Cereales. Madrid: Ed. Mundi-Prensa, 1991. p. 419.
18. Aguilar, M. Cultivo del arroz en el Sur de España. Sevilla : Centro de Investigación e Información Agraria. 2001. 189 p.

Recibido: 27 de noviembre de 2003

Aceptado: 30 de septiembre de 2004

# CURSOS DE POSGRADO

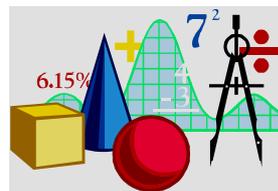
Precio: 350 CUC

## Estadística y diseño experimental

*Coordinador: Dr.C. Alberto Caballero Núñez*

*Fecha: a solicitud*

*Duración: 80 horas*



### SOLICITAR INFORMACIÓN

**Dr.C. Walfredo Torres de la Noval**  
**Dirección de Educación, Servicios Informativos**  
**y Relaciones Públicas**  
**Instituto Nacional de Ciencias Agrícolas (INCA)**  
**Gaveta Postal 1, San José de las Lajas,**  
**La Habana, Cuba. CP 32700**  
**Telef: (53) (64) 86-3773**  
**Fax: (53) (64) 86-3867**  
**E.mail: posgrado@inca.edu.cu**