

## EMPLEO DEL METODO DE ANALISIS DE COMPONENTES PRINCIPALES PARA LA CLASIFICACION DE GERMOPLASMA DE CAFE (*Coffea arabica* L.)

SILVIA MONTES<sup>1</sup>, MARIA T. CORNIDE<sup>2</sup> Y J.V. MARTIN<sup>1</sup>

### RESUMEN

Este trabajo se realizó con el objetivo de estudiar y clasificar 14 genotipos de *Coffea arabica* L. resistentes a la royay la variedad Caturra utilizada como testigo. El experimento fue plantado en 1979 en la zona de Caracusey, Sancti Spiritus. Los datos de las variables morfoagronómicas evaluadas en 1981 y 1982 fueron procesadas mediante el método de análisis de componentes principales. Los resultados obtenidos permitieron diferenciar los genotipos estudiados, debido al aporte de

las variables y sus componentes principales a la explicación de la variabilidad fenotípica total; desde el primer año se observó la agrupación de 6 de los individuos evaluados; para el segundo año las variables relacionadas con el período de floración fueron las más estables; las variables relacionadas con la cosecha y el rendimiento mostraron la mayor contribución en C<sub>2</sub> y C<sub>3</sub>; se obtuvieron fuertes asociaciones entre la floración y las restantes variables relacionadas directamente con esta.

### INTRODUCCION

*La clasificación y el reconocimiento de la variabilidad genética se convierten en una tarea compleja, en la medida en que aumenten los individuos a valorar así como los caracteres en cualquier programa de mejoramiento.*

*El análisis multivariado ha probado ser una herramienta potencial para la diferenciación de genotipos en diversos grupos, debido a que permite incluir un gran número de variables para efectuar la clasificación así como reconocer la magnitud de su contribución en relación con el resto de los caracteres; este método ha sido empleado por diversos autores: Seal, 1964; Whitehouse, 1969; Ottaviano et al,*

*1974, Rao et al, 1980; Berthaud, 1985 entre otros.*

*El análisis de componentes principales se ha utilizado de igual manera en el estudio de variedades, María T. Cornide et al (1977). Miriam Alvarez y Ana Estévez (1979), Miriam Alvarez (1981), Miriam Alvarez y María M. Hernández (1982), todas ellas de Cuba, corroboran la factibilidad en el empleo del mismo.*

*El objetivo del presente trabajo fue estudiar y clasificar genotipos de café plantados en una localidad del Occidente del país, de acuerdo a diversas variables analizadas.*

### MATERIALES Y METODOS

*Este estudio se efectuó con observaciones llevadas a cabo en un experimento de café (*Coffea arabica* L.), cuyas informacio*

*nes generales se presentan en las Tablas I y II.*

<sup>1</sup>Instituto Nacional de Ciencias Agrícolas, ISCAH, La Habana.

<sup>2</sup>Academia de Ciencias de Cuba.

Tabla I: Datos generales sobre el experimento.

Localidad	Fecha de plantación	Fecha de evaluación	Número de genotipos	Variables analizadas	Abreviaturas
Caracusey	1979	1981, 1982	15	$X_1$ Rendimiento de café cereza kg/pl.	(Rend.)
				$X_2$ Número de floraciones.	(No. flor.)
				$X_3$ Duración $\bar{x}$ de la floración (días).	(Dur. $\bar{x}$ flor)
				$X_4$ Inicio de floración (días).	(Inicio flor)
				$X_5$ Duración total de floración (días).	(Dur. tot. flor)
				$X_6$ Máximo intervalo de flora- ción (días).	(máx. int. flor)
				$X_7$ Número de pases.	(No. pases)
				$X_8$ Duración $\bar{x}$ de los pases (días).	(Dur. $\bar{x}$ pases)
				$X_9$ Inicio de la cosecha (días).	(Inicio cos.)
				$X_{10}$ Duración total de la cose- cha (días).	(Dur. tot. cos.)

Tabla II: Relación de genotipos.

Designación					
1. 3-1 R	4. 7-2 R	7. 10-17 R	10. 20-15 R	13. 24-1 R	
2. 4-21 R	5. 8-12 R	8. 11-3 R	11. 22-25 R	14. 29-25 R	
3. 5-5 R	6. 9-8 R	9. 16-13 R	12. 23-16 R	15. Caturra (test.)	
R: resist. Roya					

La localidad Caracusey pertenece a la Empresa de Cultivos Varios de Trinidad, ubicada a 50 m s.n.m.; es una zona cafetalera establecida con sombra.

El marco de plantación fue de 3 x 2 m, se empleó un diseño experimental de Bloques al azar con 4 réplicas y 5 plantas por parcela; las atenciones culturales realizadas fueron las orientadas por el Instructivo Técnico para el cultivo (MINAGRI, 1975).

Con los datos obtenidos se llevó a cabo un análisis de componentes principales (Kendall, 1957), con el cual se calculó la matriz de correlaciones y los valores y vectores propios.

Durante el año 1981 sólo se consideraron cinco variables ( $X_1, X_7 - X_{10}$ ), debido a que hubo una sola floración y por ello todo lo derivado de esta se eliminó, para un mejor ajuste y análisis.

## RESULTADOS Y DISCUSION

En la Tabla III aparece la matriz de correlaciones fenotípicas entre las variables analizadas.

Se obtuvieron valores positivos y significativos para el coeficiente de correlación lineal simple, entre el ren-

dimiento y el número de pases. El inicio de la cosecha se relacionó, pero de forma negativa, con la duración x pases.

La duración total de la cosecha no se correlacionó con el resto de las variables estudiadas.

Tabla III: Estudio de colecciones: Caracusey 1981. Matriz de correlaciones fenotípicas entre caracteres estudiados.

Variabes	(1)	(7)	(8)	(9)	(10)
(1) Rendimiento	-				
(7) No. pases	0,68 **	-			
(8) Dur. $\bar{x}$ pases	0,47	0,54 **	-		
(9) Inicio cos.	-0,19	-0,44	-0,51 **	-	
(10) Dur. tot. cos.	0,17	0,29	0,38	-0,37	

(\*) Significativo para  $P < 0,05$ .

(\*\*) Significativo para  $P < 0,01$ .

Al analizar las componentes (Tabla IV), se pudo apreciar que las  $C_1, C_2, C_3$  explicaron el 85,06 % de la variabilidad total; el valor de cada variable era bastante similar en la  $C_1$ , por lo que no es posible caracterizar este componente por una sola variable. El inicio de la cosecha presentó un valor negativo.

La  $C_2$  discrimina los genotipos en dependencia de la variable duración total de la cosecha y  $C_3$  además de esta, añade al inicio de la misma:

La distribución de los individuos (genotipos) analizados aparece en la Figura 1. Para los componentes seleccionados, se puede observar que los genotipos se distribuyen según la diagonal en tres grupos (I, II y III), sobre la base de las cuatro características señaladas, las cuales decrecen según el sentido de la diagonal.

Tabla IV: Análisis de componentes principales: Caracusey 1981. Valores y vectores propios de las tres primeras componentes.

	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>
$\lambda_i$	2,6539	0,9974	0,6492
%	53,08	19,95	12,98
% acum.	53,08	72,08	85,06
X <sub>1</sub> Rend.	0,4382	-0,5915	0,2295
X <sub>7</sub> No. pases	0,5146	-0,3182	-0,0045
X <sub>8</sub> Dur. $\bar{x}$ pases	0,5019	0,0702	-0,1763
X <sub>9</sub> Inicio cosecha	-0,4175	-0,4501	0,6203
X <sub>10</sub> Dur. tot. cos.	0,3419	0,5842	0,7290

C<sub>1</sub> ..... C<sub>3</sub>: Componentes principales.

$\lambda_i$ : Valor propio de la *i*-ésima componente.

% acumulado: Por ciento acumulado de la variabilidad total representada por las componentes consideradas.

Dentro de los grupos I y II aparecen dos pequeños subgrupos formados por los genotipos (6, 3, 2) y (1, 7) respectivamente, que se explicarán a continuación.

La duración total de la cosecha decrece desde el grupo I hasta el III. En el grupo I, esta variable determina que se encuentran incluidos los genotipos 6, 3 y 2 a pesar de que tienen valores menores para el rendimiento, en relación con los restantes miembros de ese grupo.

En el grupo II, los genotipos 1 y 7 se encuentran más cercanos por presentar iguales valores en cuanto al número de pases y rendimiento, por lo cual pueden constituir un sub-grupo; el genotipo 12 se caracteriza por tener un número de pases y rendimientos menores en comparación con el resto de los miembros del grupo, pero su permanencia dentro de esta se determina por la duración total de la cosecha.

En el grupo III se ubica un solo genotipo, el cual presenta la menor duración total de la cosecha.

Debe resaltarse que los genotipos ubicados en el extremo superior derecho son los más productivos; sin embargo, el 4 aparece en el extremo inferior izquierdo, debido a su valor tan bajo para la duración total de la cosecha.

La productividad del material fue evaluada por Silvia Montes (1985). Los resultados coinciden con los anteriores. En el trabajo la variedad Caturra se caracterizó por tener buena producción. Obsérvese que la variedad Caturra testigo (No. 15) cae en este primer grupo.

Los genotipos que se encuentran en el grupo II tienen un comportamiento medio para este carácter.

Año 1982

La Tabla V expresa los resultados obtenidos a través del estudio de las correlaciones lineales simples entre las variables empleadas.

Puede observarse que se obtuvo el máximo valor para las correlaciones fenotípicas entre:

- Número de floraciones y duración  $\bar{x}$  floración.
- Número de floraciones y duración total de la floración.
- Número de floraciones y máximo intervalo de floración.
- Duración  $\bar{x}$  floración y duración total de la floración.
- Duración  $\bar{x}$  floración y máximo intervalo de floración.

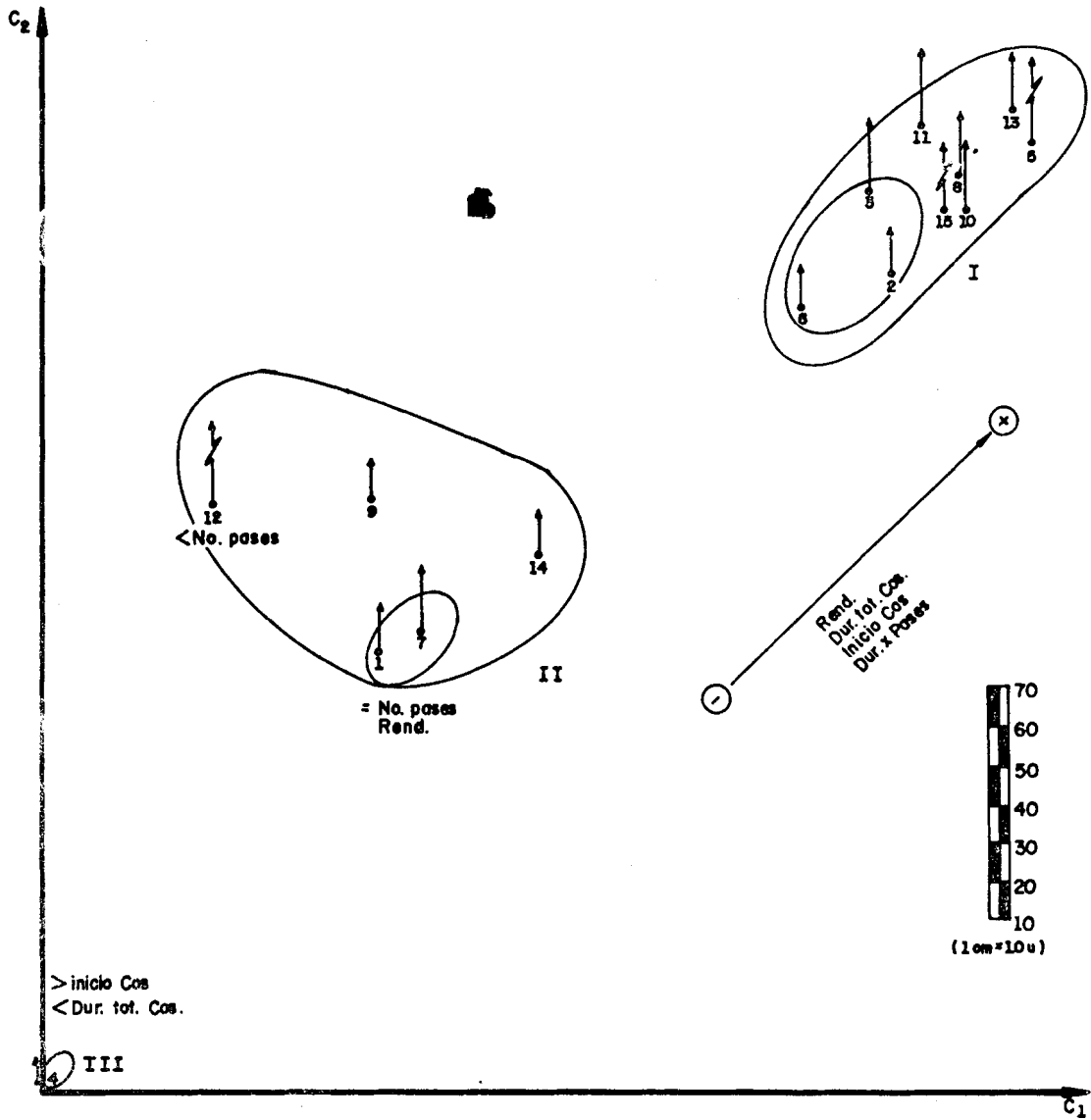


Figura 1: Componentes principales de las variables analizadas. (Caracusey-81).

Tabla V: Estudio de correlaciones: Caracusey 1982. Matriz de correlaciones fenotípicas entre caracteres estudiados.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Rend.	-									
(2) No. flor	-0,09	-								
(3) Dur. $\bar{x}$ flor.	-0,08	1,00 **	-							
(4) Inicio flor.	-0,09	0,29	-0,29	-						
(5) Dur. tot. flor.	-0,09	1,00 **	1,00 **	-0,29	-					
(6) Máx. int. flor.	-0,09	1,00 **	1,00 **	0,88	1,00 **	-				
(7) No. pases	-0,06	0,00	0,00	0,28	0,00	0,00	-			
(8) Dur. $\bar{x}$ pases	0,17	0,18	0,18	-0,03	0,18	0,18	0,28	-		
(9) Inicio cos.	-0,03	-0,47	-0,48	0,51 *	-0,48	-0,48	-0,08	-0,42	-	
(10) Dur. tot. cos.	-0,41	0,14	0,15	-0,11	0,15	0,15	-0,36	-0,39	-0,18	-

(\*) Significativo para  $P \leq 0,05$ .

(\*\*) Significativo para  $P \leq 0,01$ .

Los resultados anteriores indican que las variables relacionadas con el proceso de floración pueden considerarse como componentes de este carácter, de modo que a mayor número de floración corresponde mayor duración  $\bar{x}$  y duración total de la floración y un mayor intervalo de floración.

El inicio de la floración presentó correlación significativa con el inicio de la cosecha.

El rendimiento no mostró correlación significativa con ninguna de las variables estudiadas.

El análisis de componentes principales arrojó que los tres primeros componentes responden por el 76,43 % de la variabilidad total (Tabla VI).

La  $C_1$  se caracteriza principalmente por las variables de la floración que resultaron altamente correlacionados entre sí (No. 2, 3, 5 y 6).

La  $C_2$  puede caracterizarse por el rendimiento y las variables relacionadas con la cosecha (No. 1, 7, 8 y 10).

La  $C_3$  se explica por el paso positivo del rendimiento y negativo del inicio de la floración (variables 1 y 4).

En la Figura 2 aparecen los grupos obtenidos mediante este análisis.

Se pueden apreciar siete grupos (I - VII) de izquierda a derecha y con valores decrecientes para las variables explicadas por la  $C_1$ .

El grupo I se encuentra en el extremo izquierdo inferior y presenta los valores extremos para las variables de la cosecha representados por la  $C_2$ .

La  $C_3$  expresa mejor las diferencias para el inicio de la floración. De este modo, los genotipos 1 y 4 presentan el mayor retraso en el inicio de la floración.

De nuevo la variedad testigo Caturra (No. 15) se ubica en el grupo de mayor número de individuos.

Del presente estudio de colección se pueden considerar los siguientes planteamientos:

1. En las condiciones de Caracusey fue posible diferenciar los genotipos estudiados, debido al aporte de las variables y sus componentes principales a la explicación de la variabilidad fenotípica total.
2. Desde el primer año, se observó la agrupación de 6 de los individuos evaluados.
3. Para el segundo año de análisis, las variables relacionadas con el período de floración fueron las más estables

para su comportamiento y respondieron con un mayor peso en la primera componente.

Las variables relacionadas con la cosecha y el rendimiento mostraron

la mayor contribución en  $C_2$  y  $C_3$ .

4. Se obtuvieron fuertes asociaciones entre la floración y las restantes variables relacionadas directamente con esta.

Tabla VI: Análisis de componentes principales: Caracusey 1982. Valores y vectores propios de los tres primeros componentes.

	$C_1$	$C_2$	$C_3$
$\lambda_i$	4,5104	1,8342	1,2988
%	45,10	18,34	12,99
% acumulado	45,10	63,44	76,43
1 Rend.	0,0551	0,3396	0,5677
2 No. flor.	-0,4620	0,0363	-0,0879
3 Dur. $\bar{x}$ flor.	-0,4620	0,0363	-0,0879
4 Inicio flor.	0,1996	0,1335	-0,6175
5 Dur. tot. flor.	-0,4620	0,0363	-0,0879
6 M $\acute{a}$ x. int. flor.	-0,4620	0,0363	-0,0879
7 No. pases	0,0107	0,4559	-0,4269
8 Dur. $\bar{x}$ pases	-0,0939	0,5300	0,0455
9 Inicio cos.	0,2906	0,0202	-0,2559
10 Dur. tot. cos.	-0,0996	-0,6103	0,1233

$C_1$  .....  $C_3$  : Componentes principales.

$i$  : Valor propio de la  $i$ -ésima componente.

% : Por ciento de variabilidad total representada por la  $i$ -ésima componente.

% acum. : Por ciento acumulado de la variabilidad total representada por los componentes considerados.

## REFERENCIAS

- ALVAREZ, MIRIAM. 1981. Aplicación del análisis de componentes principales efectuado a un grupo de variedades de arroz (*Oriza sativa* L.) atendiendo a la variable de rendimiento agrícola e industrial. *Cultivos Tropicales* 3 (2): 157-167.
- ALVAREZ, MIRIAM Y ANA ESTEVEZ. 1979. Clasificación de un grupo de variedades de papa (*Solanum tuberosum* L.) con respecto al rendimiento y sus componentes. *Cultivos Tropicales* 1(3): 139-149.
- ALVAREZ, MIRIAM y MARIA M. HERNANDEZ. 1982. Estudio de los componentes principales en un grupo de variedades de plátano. *Cultivos Tropicales* 4 (2): 227-240.
- BERTHAUD, J. 1985. Propositions pour une nouvelle stratégie d'amélioration des caféiers de l'espèce *C. Canephora*, basée sur les résultats de l'analyse des populations sylvestres. XI<sup>ème</sup> Colloque ASIC, Lomé (Togo).
- CORNIDE, MARIA T.; MIRIAM ALVAREZ Y SILVIA MONTES. 1977. Classification de lignées de *Coffea arabica* L. moyen l'analyse des composantes principales. Résumé. VIII Colloque Scientifique International sur le Café (ASIC) Abidjan. pp. 447-453.

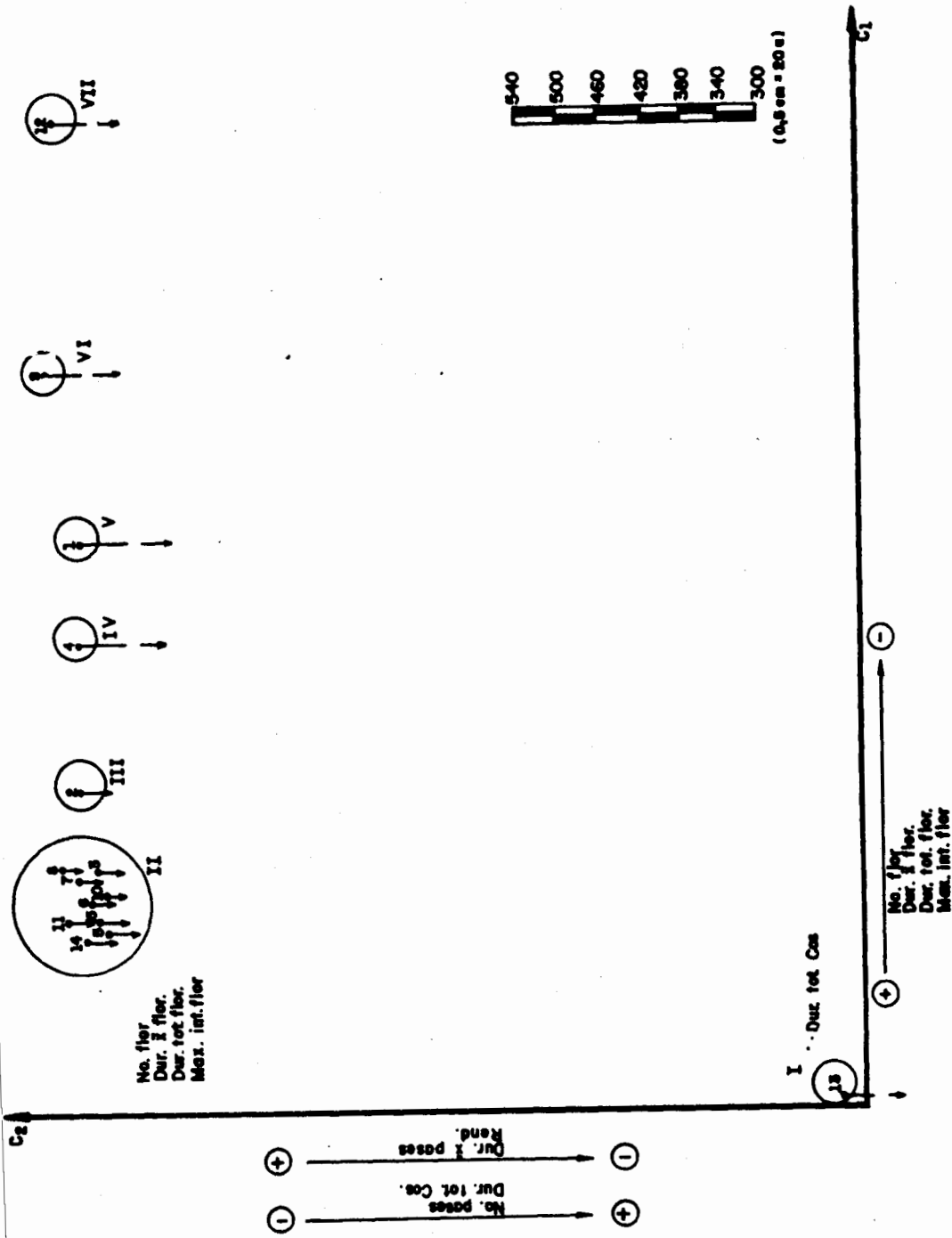


Figura 2: Componentes principales de las variables analizadas. (Caracusey-82).



- CUBA - INRA. 1975. *Normas Técnicas para el cultivo de café. La Habana.*
- KENDALL, M. 1967. *A Course in Multivariate Analysis.* London. Griffin pp: 185.
- MONTES, SILVIA. 1985. *Estudio de la interacción genotipo x ambiente en café. Sin publicar.*
- OTTAVIANO, E.T. AND S.M. SARI-GORLA. 1974. *Analisi multivariata per lo studio delle differenze genetiche in popolazione di Gerbera Jammesonn.* Genet. Agrar. 28: 292-306.
- RANGA RAO, V.; M. RAMACHANDRAM Y JAWAHAR RAIN SHARMA. 1980. *Multivariate Analysis of Genetic Divergence in Sunflower.* Indian J. Genet. and Plant Breed. 40 (1): 73-85.
- SEAL, H.L. 1964. *Multivariate Statistical Analysis for Biologists.* London, Methver. pp. 208.
- WHITEHOUSE, R. 1969. *Canonical Analysis as an Aid for Plant Banding.* 2nd. Barley Genetic Symposium.

## ABSTRACT

### THE USE OF AN ANALYSIS METHOD OF MAIN COMPONENTS FOR SCREENING COFFEE GERM PLASM.

This research study was aimed at studying and screening 14 rust-resistant genotypes from *Coffea arabica* L., using Caturra cv. as a check. This experiment was planted in Sancti Spiritus (Caracusey zone) in 1979. Data derived from all morphoagronomical variables evaluated in 1981 and 1982 were processed through an analysis

method of main components. Results enabled to differ every genotype studied, as variables and its main components accounted for the whole phenotypical variability. Six individuals were grouped since the first year of evaluations whereas for the second year, the most stable variables were those related to flowering; variables related to harvest and yield showed the greatest contribution to C<sub>2</sub> and C<sub>3</sub>. Strong associations were recorded between flowering and other variables directly related to it.

*Manuscrito recibido el 7/VII/87.*