

## USO DEL INDICE DE DIVERSIDAD RELATIVA DE KOSMAN PARA AGRUPAR FENOTIPOS EN POBLACIONES POLIMORFICAS DE FITOPATOGENOS: EL CASO DE *Phytophthora infestans*. \*

### Using Kosman's index of relative diversity to cluster phenotypes in polymorphic populations of plant pathogens: the case of *Phytophthora infestans*.

Elsa Janeth Gualtero<sup>1</sup>, Celsa García<sup>2</sup> y Enrique Torres<sup>3</sup>

#### RESUMEN

El polimorfismo en poblaciones de fitopatógenos aumenta los riesgos asociados con las enfermedades. La tecnología molecular ha permitido examinar esta variabilidad usando marcadores neutrales como proteínas totales e isoenzimas. El polimorfismo se puede medir y describir con indicadores de diversidad y de riqueza de información, y los varios fenotipos se pueden ordenar mediante análisis de agrupamiento. Este trabajo utilizó un estudio previo de proteínas totales en 55 aislamientos de *Phytophthora infestans*, el patógeno que causa la gota de la papa, y contrastó los agrupamientos obtenidos con el coeficiente de similitud de Jaccard, con las diferencias en número de proteínas entre aislamientos y con el índice relativo de diversidad propuesto recientemente por Kosman, basado en las diferencias máximas entre aislamientos. Los resultados indican que

las distancias y el índice relativo de diversidad complementan adecuadamente el análisis de agrupamiento.

**Palabras claves:** Genética de poblaciones, vigilancia de fitopatógenos.

#### SUMMARY

Polymorphism in populations of plant pathogens enhances the risks associated with diseases. Molecular techniques allow the examination of this polymorphism using neutral markers such as total proteins and isozymes. Polymorphism may be measured and described using several diversity and information richness indices, and the different phenotypes may be ordinated with standard cluster analysis. This work used data from a study on total proteins within 55 isolates of *Phytophthora infestans*, the pathogen of potato late blight, and compared the clustering of isolates resulting from standard cluster analysis, from differences among isolates in number of proteins and from *Ko*, Kosman's index of relative diversity, which is based on the maximum differences among isolates. Results indicate that both distances and *Ko* are suitable complements to cluster analysis.

**Key words:** Population genetics, pathogen surveillance.

\* Recibido en Enero de 1998

1 Lic. en Biología y Química, Universidad del Tolima, Maestra en Ciencias Agrarias, Universidad Nacional de Colombia.

2 Ing. Agr., Ph. D. Profesor Asociado, Universidad Nacional de Colombia.

3 Ing. Agr., Ph. D. Profesor Asociado, Universidad Nacional de Colombia.

## INTRODUCCION

Una población es un grupo de individuos de la misma especie que interactúan en un determinado tiempo y espacio. Cuando, en una población, solamente existe un fenotipo para un carácter determinado, la población es monomórfica para ese carácter. De lo contrario, la población es polimórfica. Las poblaciones de patógenos de plantas son frecuentemente polimórficas, debido a las variantes que resultan de los procesos de mutación, migración, deriva genética y reproducción sexual y al papel selectivo de factores ambientales dinámicos (Burdon y Silk, 1997). Este polimorfismo genético tiene implicaciones sobre la estabilidad de las poblaciones de fitopatógenos y sobre el manejo de las enfermedades que causan. Por lo tanto, la vigilancia continua de la estructura genética de poblaciones de patógenos es un pilar fundamental en la protección de cultivos.

Cuando el polimorfismo se expresa en características fisiológicas que no se observan directamente, como por ejemplo, tipo de apareamiento, virulencia o resistencia a fungicidas, hay que examinar detalladamente un número representativo de aislamientos muestreados de la población, y puede ser necesario recurrir a marcadores moleculares y al empleo de métodos numéricos para la descripción del eventual polimorfismo. Fry y Goodwin (1997) recomiendan el uso de marcadores moleculares y bioquímicos neutrales para estudios de genética de poblaciones en patógenos cuya variación patotípica (es decir, en caracteres asociados con virulencia y patogenicidad) es muy rápida, como es el caso de *Phytophthora infestans*, el agente causal de la gota de la papa.

Refiriéndose a la diversidad fenotípica de *Puccinia graminis*, el hongo de la roya del tallo de los cereales, Groth y Roelfs (1987) consideran que para describir la diversidad en una población polimórfica de fitopatógenos se deben integrar tres criterios (1) que la población sea rica, es decir, que se componga de un alto número de

fenotipos; (2) que las frecuencias de los fenotipos tengan una distribución equitativa y (3) que el grado de diferencias entre fenotipos sea amplio.

La ecología matemática ha adaptado y desarrollado indicadores y métodos analíticos para describir y medir la diversidad en poblaciones polimórficas. Dos indicadores sencillos son el índice de entropía de Shannon ( $Sh = -\sum p_i \ln p_i$ ) y el coeficiente de diversidad de Simpson ( $S = 1 / \sum p_i^2$ ), los cuales se computan con base en  $p_i$ , la frecuencia del  $i$ -ésimo fenotipo de  $n$  fenotipos diferentes (Whittaker, 1975). Ambos índices son estimadores válidos de la riqueza y la equidad de la diversidad, pero ninguno tiene la capacidad de estimar el grado de disimilitud entre los distintos fenotipos. Esta deficiencia es tanto más grave si se considera que es muy probable que las diferencias entre pares de aislamientos no idénticos, en poblaciones polimórficas de fitopatógenos, sean de diferente magnitud. Por esta razón, los índices  $Sh$  y  $S$  no son adecuados para describir la diversidad en poblaciones de fitopatógenos, según los criterios planteados por Groth y Roelfs (1987).

El análisis de gradientes, desarrollado por Whittaker (1975) en el contexto de ecología de comunidades, es un enfoque investigativo para el estudio de asociaciones polimórficas que examina la similitud o la diferencia de ubicación de sus elementos constitutivos con respecto a dos o más gradientes o dimensiones.

En el análisis de agrupamiento (cluster analysis), los elementos constitutivos de la asociación polimórfica se confrontan iterativamente en pares por la presencia o ausencia de ciertas características, con el fin de encontrar el par de registros más similar, o sea, el que tiene un mayor número de características en común. La presencia o ausencia de una característica se registra como una variable binaria que asume valores de 1 si la característica está presente y 0 si está ausente (Johnson y Wichern, 1992).

Recientemente, Kosman (1996) ha propuesto un índice para describir la diversidad de una población polimórfica de fitopatógenos, el

cual tiene en cuenta la disimilitud entre los fenotipos que constituyen esa población. Kosman considera una muestra de  $n$  aislamientos que representa una población de patógenos, de manera que, al considerar  $k$  caracteres descriptivos, cada aislamiento tenga un patrón de  $k$  respuestas binarias. La 'distancia' entre dos aislamientos está definida por el número de caracteres a los cuales responden diferencialmente los dos aislamientos y varía entre 0 (cuando los dos aislamientos presentan respuestas iguales a los  $k$  caracteres) y  $k$  (cuando los dos aislamientos presentan respuestas diferentes a cada carácter). Al comparar cada aislamiento contra todos los demás, se escoge la mayor de las  $n-1$  'distancias' obtenidas para cada aislamiento, y se suman las  $n$  'distancias' para obtener la suma máxima de las distancias,  $Ass_{max}(A,A)$ . El índice relativo de diversidad de la población,  $Ko$ , se obtiene dividiendo  $Ass_{max}(A,A)$  entre  $nk$ , y tiene valores que fluctúan entre 0 y 1. Un índice de diversidad 0 corresponde a una población que está limitada a un solo fenotipo. El máximo índice de diversidad relativa, 1, corresponde a una población con fenotipos que ocurren en frecuencias iguales y que presentan patrones de respuesta diferentes para los  $k$  factores diferenciales, especialmente si la población posee todos los  $2^k$  fenotipos posibles (Kosman, 1996).

La gota de la papa, probablemente, es la enfermedad más destructiva de este cultivo a nivel mundial. A partir de los años 1980, se ha observado un resurgimiento global de la enfermedad, asociado con la aparición de nuevos linajes de *P. infestans*, cuyo supuesto origen fueron migraciones iniciadas en México (Dagget *et al.*, 1993; Fry y Goodwin, 1997). Hasta antes de esta migración, las poblaciones de *P. infestans* por fuera de México estaban limitadas a dos o tres genotipos de una misma línea clonal pero ahora suman cientos.

Las características iniciales de la población migrante deducidas por comparación de varios marcadores entre aislamientos contemporáneos y aislamientos colectados antes de la detección del tipo A2 son las siguientes: tipo de apareamiento A1 GPI 86/

100 y 100/100; PEP 92/100 y 100/100. Previo a la detección del tipo A2, el alelo 90 para GPI y el alelo 83 para PEP no habían sido detectados en Europa. Análisis posteriores de DNA han corroborado los resultados de aloenzimas (Fry *et al.*, 1993).

Análisis realizados en Alemania, Corea, Gran Bretaña, Holanda, Japón y Polonia indican que los nuevos linajes desplazaron en pocos años a las poblaciones locales. La sustitución de los viejos linajes clonales por nuevas poblaciones indica que la población migrante está mejor adaptada que los viejos linajes. Las migraciones y la sustitución de poblaciones han agravado el problema de la pérdida de sensibilidad al fungicida metalaxyl (Fry y Goodwin, 1997). Bajo estas circunstancias, es imperativo disponer de técnicas sencillas y eficientes para cuantificar el polimorfismo genético de poblaciones y monitorear la aparición de nuevos linajes, utilizando marcadores neutrales cuya frecuencia no se vea afectada por los genes de resistencia del hospedante (Fry y Goodwin, 1997).

Este trabajo comparó la ordenación de aislamientos de *P. infestans* por sus valores de proteínas totales (Gualtero, 1997), según el análisis de agrupamiento, las distancias absolutas entre aislamientos y  $Ko$ , como una contribución hacia el establecimiento de métodos numéricos para describir la diversidad y ordenar o clasificar aislamientos en poblaciones de fitopatógenos,

## MATERIALES Y METODOS

El trabajo utilizó los datos de un estudio de proteínas totales en 55 aislamientos de *P. infestans* colectados entre 1995 y 1996 en lotes comerciales de papa de Cundinamarca y Boyacá, en el cual se visualizaron 71 bandas, que representan sendas proteínas con pesos moleculares relativos entre 14000 y 270000 (Gualtero, 1997). Los resultados generaron una matriz binaria de  $55 \times 71$ , en la cual la presencia de una banda determinada para cada aislamiento se registró como 1 y su ausencia, como 0.

Con estos resultados se calculó el coeficiente de similitud de Jaccard de acuerdo con la fórmula  $C_s = (a)/(a+b+c)$ , donde (a) representa el número de comparaciones de pares de aislamientos en que la proteína ocurre en ambos aislamientos y (a+b+c) el número de comparaciones en que la proteína ocurre, al menos, en un aislamiento. Los valores posibles de  $C_s$  van desde 0, cuando la proteína nunca ocurre en ambos aislamientos, hasta 1, cuando la proteína ocurre siempre en ambos aislamientos. El coeficiente de similitud de Jaccard se utilizó para realizar un análisis de agrupamiento, el cual generó un dendrograma que reconoció tres subgrupos de aislamientos, así: el subgrupo uno con 11 aislamientos cuyos coeficientes van de 0,43 a 0,84; el subgrupo dos con 22 aislamientos, cuyos coeficientes van de 0,5 a 0,85 y el subgrupo tres (3) con 22 aislamientos y cuyos coeficientes van de 0,5 a 0,84. (Figura 1). Para estos análisis, se utilizó el software Syn-Tax-pc versión 5.0 (Podani, 1993).

Para determinar el índice de Kosman, fue necesario establecer las 'distancias' entre cada aislamiento y los otros 54 aislamientos en términos de presencia o ausencia de cada una de las 71 diferentes bandas proteínicas y, luego, sumar las distancias máximas. Debido a la falta del software apropiado, las 1485 comparaciones pareadas entre los 55 aislamientos se efectuaron manualmente. Para facilitar el proceso y hacerlo menos dispendioso, se procedió de la siguiente manera: se agruparon las 71 bandas proteínicas en 18 con-

juntos de cuatro bandas cada uno, con excepción del último que solamente contenía tres bandas; las comparaciones entre dos aislamientos se realizaban al interior de cada conjunto de bandas, y el resultado se codificaba con una letra que indicaba en cuales de las cuatro bandas había diferencias (Cuadro 1).

Este procedimiento permitió reducir a la cuarta parte el tamaño de la matriz original de comparaciones. Finalmente, una Tabla Auxiliar, en la cual aparecían, en doble entrada, los Códigos de Diferencias, daba la 'distancia' parcial entre los dos aislamientos para cada uno de los 18 conjuntos de bandas. La distancia total entre los dos aislamientos se obtenía sumando las 18 distancias parciales. Los valores mínimos y máximos esperados de las distancias eran de 0 para aislamientos idénticos y de 71 para dos aislamientos que difirieran en todas las bandas proteínicas.

Después de haber calculado las 54 distancias para cada aislamiento, se seleccionó el valor de la mayor distancia y se sumaron los n valores para el grupo de aislamientos que se estuviera evaluando. Esta suma constituye el parámetro  $Ass_{max}(A,A)$ . El índice de diversidad relativa,  $K_o$ , definido como la relación entre  $Ass_{max}(A,A)$  y el producto del número de aislamientos n y el número de bandas {k = 71}, se calculó para la muestra de 55 aislamientos {n = 55}, para cada uno de los tres subgrupos reconocidos por el dendrograma generado a partir del análisis de agrupamiento {n = 11, n = 22 y n

**Cuadro 1.** Código de las 16 posibles diferencias entre aislamientos con base en cuatro bandas proteínicas, (x indica diferencias).

BANDA	CODIGO DE DIFERENCIAS															
	B	C	D	E	F	G	H	J	K	L	M	N	O	P	Q	R
Banda a									X	X	X	X	X	X	X	X
Banda b					X	X	X	X					X	X	X	X
Banda c			X	X			X	X			X	X			X	X
Banda d		X		X		X		X		X		X		X		X

= 22) y para pares de aislamientos ubicados a distancias cortas o largas en dicho dendrograma ( $n = 2$ ).

## RESULTADOS

La mayoría de las distancias entre aislamientos estuvo entre 15 y 25 bandas. Los aislamientos 26 y 32 mostraron la diferencia mínima de seis bandas y los aislamientos 37 y 46 tuvieron la máxima distancia de 31 bandas proteínicas. Los aisla-

mientos más contrastantes para cada uno de los 55 aislamientos y las correspondientes distancias máximas se presentan en el Cuadro 1, en el orden en que aparecen en el dendrograma (Figura 1). Para la muestra de 55 aislamientos,  $K_0$  tuvo un valor de 0,37 y para cada uno de los tres subgrupos del dendrograma fue de 0,34 para el subgrupo uno, y de 0,32 para los subgrupos dos y tres. El  $K_0$  para el par de aislamientos extremos de cada subgrupo (a saber, 1 y 11 para el subgrupo uno, 4 y 41

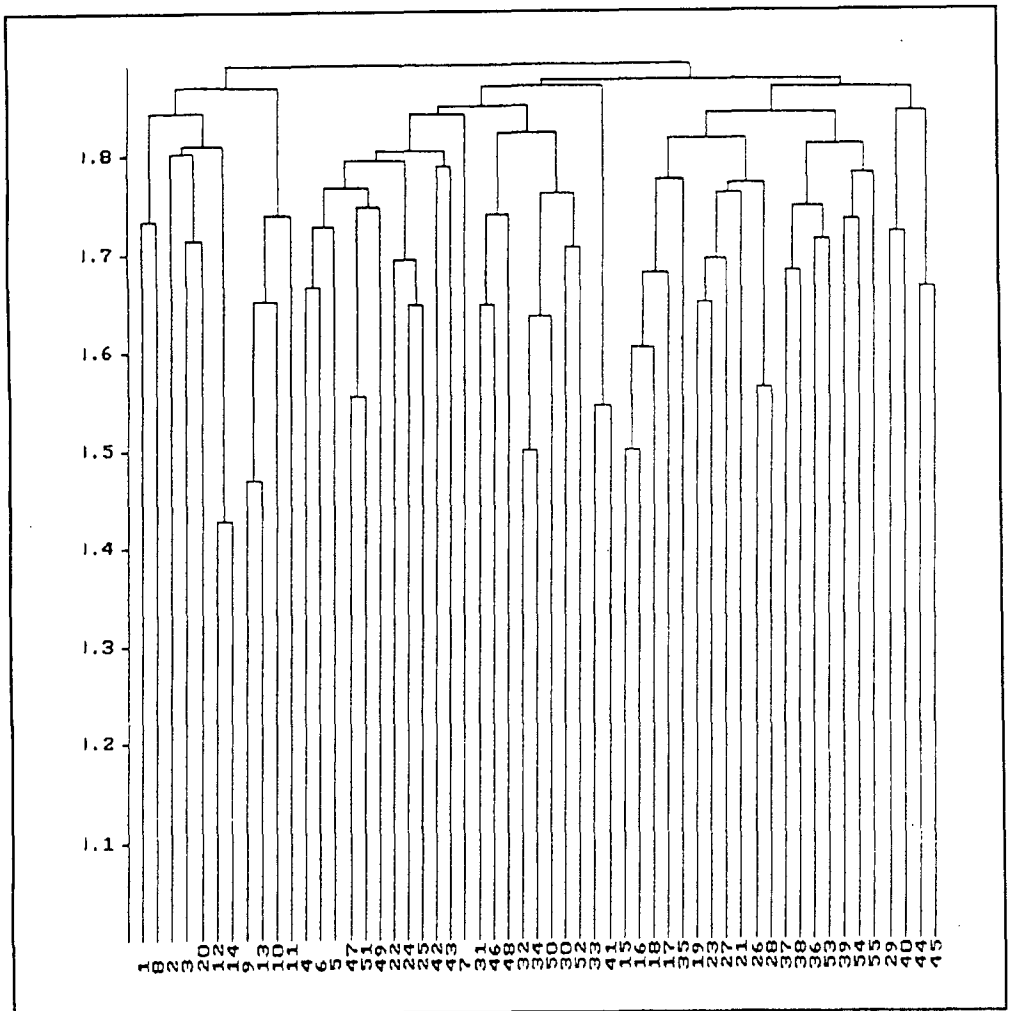


Figura 1. Agrupamiento de aislamientos de *Phytophthora infestans* del altiplano Cundiboyacense por similitud de proteínas solubles de acuerdo al coeficiente de similitud de Jaccard.

para el subgrupo dos y 15 y 45 para el subgrupo tres, véase la Figura 1) fue de 0,26 para los dos primeros pares y de 0,29 para el último.

## DISCUSION

El resurgimiento global de la goma de la papa, como resultado de recientes migraciones intercontinentales de *P. infestans*, pone

**Cuadro 2.** El aislamiento más distante para cada uno de los 55 aislamientos de *Phytophthora infestans* ordenados según el dendrograma de la figura 1. y la distancia que los separa, en número de bandas proteínicas.

Aislamiento	1	8	2	3	20	12	14	9	13	10	11	4	6	5
Aislamiento más distante	49	49	36	21	9	21	37	37	37	49	6	37	40	9
										12				19
														21
Distancia (Bandas)	24	24	27	25	27	30	29	27	25	27	26	26	30	23

Aislamiento	47	51	49	22	24	25	42	43	7	31	46	48	32	34
Aislamiento más distante	25	6	10	37	12	12	12	12	18	12	37	37	6	37
		12						19					12	
Distancia (Bandas)	21	27	27	26	25	25	27	25	26	25	31	26	22	23

Aislamiento	50	30	52	33	41	15	16	18	17	35	19	23	27	21
Aislamiento más distante	37	6	6	12	6	6	2	2	12	6	12	6	12	12
		12	12		9			7						
			37		12			22						
Distancia (Bandas)	23	27	23	24	24	29	24	26	27	28	28	28	28	30

Aislamiento	26	28	37	38	36	53	39	54	55	29	40	44	45
Aislamiento más distante	37	2	46	6	12	12	12	6	37	6	6	6	6
													12
Distancia (Bandas)	27	25	31	26	28	25	26	24	23	24	30	29	22

de manifiesto la vulnerabilidad de los sistemas de producción agrícola frente a movimientos de material vegetal favorecidos por la apertura de mercados. Por lo tanto, los fitopatólogos deben mantener una vigilancia cuidadosa por nuevos patógenos o por indicios de cambios en poblaciones de patógenos ya establecidos (Fry y Goodwin, 1997). Además, los marcadores moleculares neutrales, como el análisis de proteínas totales y de isoenzimas (Gualtero, 1997), permiten estudiar la estructura genética de las poblaciones sin ser afectados por cambios en características patotípicas (Fry y Goodwin, 1997). Teniendo en cuenta la trascendencia de estos trabajos de vigilancia, es importante que la presentación de sus resultados permita la descripción más precisa y completa posible del cuadro de diversidad en poblaciones de fitopatógenos.

Al comparar la localización de los aislamientos en el dendrograma de la Figura 1, ordenados según el coeficiente de similitud de Jaccard, con las distancias en número de bandas proteínicas, se observan discrepancias considerables. Los aislamientos 1 y 45, que fueron los más contrastantes en el dendrograma, apenas tuvieron una diferencia de 14 bandas, que es inferior a la diferencia mediana observada en el estudio y que corresponde a un índice de diversidad de 0,19, bastante inferior al índice de diversidad para la muestra, que fue de 0,37. De otra parte, los aislamientos 26 y 32, que presentaron la mínima distancia de seis bandas proteínicas, fueron ubicados en diferentes subgrupos por el análisis de agrupamiento (Figura 1). Esta discrepancia puede entenderse si se tiene en cuenta que el análisis de agrupamiento es un proceso iterativo que busca ordenar o agrupar los elementos más similares a lo largo de coordenadas, cada una de las cuales es el eje o gradiente de alguna variable. Whittaker (1975) advierte que, cuando se trabaja con más de tres variables, el resultado del análisis es un hiperespacio de tantas dimensiones cuantas variables se hayan considerado. Por lo tanto, el análisis de agrupación realizado con 71 bandas proteínicas genera un hiperespacio de 71 dimensio-

nes y el dendrograma de la Figura 1 es la proyección bidimensional de ese hiperespacio. Consecuentemente, se observa en el Cuadro 2 que los aislamientos 6, 12 y 37 son los más lejanos para 39 de los 55 aislamientos, con distancias que varían entre 22 y 31 bandas, a pesar de que ocupan posiciones diferentes en el dendrograma. Para comprender esta situación hiperespacial, la mejor analogía es recordar que la posición que ocupan las estrellas en el firmamento no guarda relación con las distancias cósmicas que las separan.

Por otra parte, se observó una reducción gradual de los valores del índice de Kosman, a medida que se pasa de la muestra global de 55 aislamientos a subgrupos de 11 y 22 aislamientos y, finalmente, a pares de aislamientos, lo cual se puede atribuir parcialmente a una reducción en la heterogeneidad de los subgrupos. No obstante, esta tendencia sugiere, también, que el índice de Kosman puede ser sensible al tamaño de la muestra.

En conclusión, la interpretación del perfil de proteínas totales en los 55 aislamientos de *P. infestans* mediante el índice de Kosman fue un complemento sencillo y expresivo al análisis de agrupamiento, y el procesamiento manual de los datos dio confianza en la significancia lógica de los resultados.

## BIBLIOGRAFIA

- BURDON, J. J. y SILK, J.** 1996. Sources and patterns of diversity in plant-pathogenic fungi. *Phytopathology* 87:664-669.
- DAGGETT, S. S., GOTZ, E., y THERRIEN, C. D.** 1993. Phenotypic changes in populations of *Phytophthora infestans* from eastern Germany. *Phytopathology* 83: 319-323.
- FRY, W. E. y GOODWIN, S. B.,** 1997. Re-emergence of potato and tomato late blight in the United States. *Plant Disease* 81:1349-1357

**FRY, W. E. y GOODWIN, S. B., DRENTH, A., TOOLEY, P. W., SUJKOWSKI, L. S., KOH, Y. J., COHEN, B. A., SPIELMAN, L. J., DEALT, K. L., INGLIS, D. A. Y SANDLAN, K. P.** 1993. Historical and recent migrations of *Phytophthora infestans*: Chronology, pathways and implications. Plant disease 77: 653-661

**GROTH, J. V. Y ROELFS, A. P.** 1987. The concept and measurement of phenotypic diversity in *Puccinia graminis* on wheat. Phytopathology 77:1395-1399.

**GUALTERO, E, J.** 1997. Caracterización aloenzimática de poblaciones de *Phytophthora infestans* en el altiplano Cundiboyacense. Tesis de Maestría en Ciencias Agrarias, Facultad de Agronomía, Universidad Nacional de Colombia sede Santafé de Bogotá.

**JOHNSON, R. Y WICHERN, D.** 1992. Applied multivariate statistical analysis. Págs. 573-627. Prentice Hall, Englewood Cliffs, N. J.

**KOSMAN, E.** 1996. Difference and diversity of plant pathogen populations: A new approach for measuring. Phytopathology 86:1152-1155.

**PODANI, J.** 1993. Software Syn-Tax-pc Versión 5.0 computer programs for multivariate data. Analysis in ecology and systematics. User's guide. Scientia Publishing, Budapest.

**WHITTAKER, R. H.** 1975. Communities and ecosystems. 2a. Edición. Macmillan, Nueva York.