

# ORDENACIÓN Y CLASIFICACIÓN NUMÉRICA EN NOPAL TUNERO MEDIANTE ATRIBUTOS DE FRUTO

**R. D. Valdez-Cepeda<sup>1</sup>; F. Blanco-Macías; C. Gallegos-Vázquez.**

Centro Regional Universitario Centro Norte, MCDRR-Sede Zacatecas, Universidad Autónoma Chapingo, Apdo. Postal 196, Zacatecas, Zacatecas. C. P. 98000. México. Correo-e: rivalce@terra.com.mx y vacrida@hotmail.com (<sup>1</sup>Autor responsable)

## RESUMEN

Las clasificaciones numéricas en biología se han basado en la genética y la filogenia. En este reporte de investigación se presenta un caso de desarrollo de una estrategia de ordenación y clasificación numérica objetiva en términos de masa y forma de fruto (tuna o nochtli) del género *Opuntia*. Se describen las ventajas del método de componentes principales (CP) para ordenar accesiones de *Opuntia* y se argumenta la conveniencia de usar el método de análisis por conglomerados para clasificar accesiones de *Opuntia*. También se describe una forma de definir el número de grupos con fundamentos estadísticos. Asimismo, se considera la importancia de involucrar los atributos de fruto con mayor valor de predicción, así como una estrategia correcta de escalas y ponderaciones de los atributos. Los dos primeros CP's explican el 62 % de la variación total de ocho variables (peso total, peso de cáscara, peso de la parte comestible, diámetro polar, diámetro ecuatorial, proporción entre DP y DE (DP/DE), volumen, y contenido de azúcares (°brix) de las 29 accesiones de nopal tunero. El espacio ortogonal definido por los CP1 y CP3 identifica a las accesiones 'Mango', 'Burrona 1' y 'Naranjón Legítimo' como los genotipos con más cáscara, aspecto atractivo para el manejo en postcosecha. El análisis por conglomerados con la distancia euclidiana como índice de similitud permitió diferenciar cinco grupos estadísticamente diferentes y la accesión 'Tapón Aguanoso' conforma por sí sola un grupo, lo cual es explicado porque es el único material silvestre de los 29 estudiados.

**PALABRAS CLAVE ADICIONALES:** *Opuntia* spp., tuna o nochtli, análisis de componentes principales y conglomerados.

## ORDERING AND NUMERICAL CLASSIFICATION IN PRICKLY PEAR CACTUS USING FRUIT ATTRIBUTES

### SUMMARY

Numerical classifications in biology have been based on genetics and phylogenetics. This research report presents the case of developing a strategy for an objective ordering and numerical classification in terms of fruit mass and shape (prickly pear, tuna, or nochtli) of the genus *Opuntia*. The advantages of principal components (PC's) in ordering *Opuntia* accessions are described, and the convenience of using cluster analysis for the same purpose is discussed. A way of defining the number of groups on statistical fundamentals is described. Likewise, the importance of involving fruit attributes with higher prediction value is considered, as well as a correct strategy attribute scale and weighing. The first two PC explain 62 % of the total variation from eight variables (total weight, skin weight, edible portion weight, polar diameter (PD), equatorial diameter (ED), volume ratio, and sugar content (°brix) of 29 accessions of prickly pear cactus. The orthogonal space defined by PC1 and PC3 identifies the accessions 'Mango', 'Burrona 1' and 'Naranjón Legítimo' as the genotypes with the most skin, an attractive feature for post-harvest handling. Cluster analysis, using Euclidean distance as a similarity index, helped to differentiate five statistically different groups, and the accession 'Tapon Aguanoso' made a group of its own, which was explained because it's the only wild material of the 29 studied.

**ADDITIONAL KEY WORDS:** *Opuntia* spp., cactus, pear, tuna or nochtli, principal component and cluster analysis.

### INTRODUCCIÓN

La clasificación es una actividad humana muy influenciada por los objetivos del usuario (Dale *et al.*, 1989). Sin embargo, el propósito esencial y fundamental de cualquier clasificación es organizar los miembros de una población en grupos o clases para que su naturaleza y las

interacciones entre ellos sean fácilmente entendidas (Arkley, 1976). En biología, la taxonomía numérica ha involucrado a las clasificaciones tomando como base a la genética y/o la filogenia de las especies (Villaseñor y Murguía-Romero, 1992), e inclusive, de esta forma los rangos categóricos de la taxonomía Lineana han sido refutados (e.g. Webb, 1954; Green, 1991). Los fundamen-

tos principales, hasta ahora argumentados en ese sentido, son: i). la mayoría de las especies presentan algún grado de separación por el patrón genético (Webb, 1954); y ii). la explicación y el entendimiento de la física del proceso evolutivo (Green, 1991).

En consecuencia, se han desarrollado clasificaciones numéricas que contemplan la variabilidad de los atributos. Desafortunadamente, en muchas de esas clasificaciones no se ha considerado el establecimiento de las clases como una operación distinta a la definición de las reglas para poder incluir nuevos atributos que permitan diferenciar clases o grupos. Otros problemas asociados a las clasificaciones numéricas son las definiciones apropiadas de escalas y ponderaciones de las restricciones contextuales y condicionantes (Dale *et al.*, 1989).

Lo anteriormente expuesto, el papel relevante del género en la agricultura y la ecología en las zonas áridas y semiáridas, su valiosa diversidad y la seriedad del problema de identificación de especies de *Opuntia* mediante el sistema de clasificación Lineana, dado el polimorfismo producto del proceso de hibridación natural presente en cactáceas (Rzedowski, 1964), dieron la pauta para que con fundamentos teóricos se pretendan lograr los objetivos: i). Ordenar, y ii). clasificar cuantitativamente 29 accesiones de *Opuntia* con base en la masa y forma de sus fruto (tuna o nochtli).

### Fundamentos teóricos de ordenación y clasificación

#### Selección de especies, accesiones, genotipos, cultivares, etc.

La cantidad de especies, accesiones, genotipos y/o cultivares de *Opuntia* a ser incluidos debe ser grande, de forma tal que sea una muestra representativa. Por ejemplo, se deben incluir aquellos(as) que sean caracterizados tanto por hábitos de crecimiento rastrero como de crecimiento erecto; pues si principalmente se incluyen los de hábito de crecimiento rastrero; entonces, con la inclusión de pocos de hábito de crecimiento arbustivo, se podría interferir con el análisis para la agrupación porque aquellos atributos asociados con el hábito de crecimiento erecto son no representativos del rango de variación de dichos atributos.

#### Definición de atributos

La selección de los atributos es invariablemente más importante que la selección de las especies, accesiones, genotipos y/o cultivares. Si el propósito de la clasificación es básica (para uso general), se deben incluir tantos atributos como sea posible para que sirva a una amplitud de objetivos. Sin embargo, debe quedar claro que con la finalidad de que la clasificación sea efectiva y comprensible, los atributos (que reflejan la genética, *e. g.* nivel de ploidía y número cromosómico; la filogenia; la morfología; y la productividad) usados para formar las clases deberán

contener el máximo de información posible (Arkley, 1976). Esto significa que deben ser aquellos que tengan: mayor valor de predicción de la naturaleza y del comportamiento de las plantas de *Opuntia* cuando son sometidas a influencias externas (diferentes ambientes). Es claro que el conocimiento de las preferencias ambientales de las especies, así como de su distribución geográfica, es en grado sumo importante si el uso de los genotipos colectados es el de mejoramiento genético (Pengelly y Eagles, 1995). Algunos investigadores han evidenciado que la morfología de las plantas de una misma especie no siempre está fuertemente relacionada con la distribución geográfica (*e. g.* Pengelly y Eagles, 1995). Además, la covarianza con otros atributos no considerados para definir los grupos o clases. Un gran número de atributos no es necesario para clasificar el género *Opuntia*.

### Escalas y ponderaciones de atributos

Algunos tipos de medición de los atributos de las especies, genotipos y/o cultivares son más fáciles de realizar que otros; unos atributos son de mayor interés para el usuario por lo que pueden ser sobre-representados e indebidamente ponderados; y también, el uso de un gran número de atributos involucra una gran inversión en tiempo y dinero, especialmente si se involucra una muestra grande de cada genotipo. Por ello, es conveniente que: en la primera etapa se contemple a un buen número de variables estandarizadas para ponderarlas igualmente; es imprescindible también que se pondere de forma tal que se iguale la contribución de atributos discretos y continuos. En este rubro se debe evitar el problema de sesgo mediante alguna transformación de los valores del o los atributo(s), y en la etapa final de la clasificación se reduzca el número de atributos a una cantidad manejable pero todavía de magnitud efectiva para realizar análisis de covarianza entre ellos.

### Fundamentos teóricos de la ordenación

El método de componentes principales (CP's) es de gran utilidad como una técnica de ordenación en taxonomía numérica en plantas (Sneath y Sokal, 1973), porque permite reducir la dimensionalidad (Pla, 1986; James y McCulloch, 1990; Parent *et al.*, 1994) del problema al eliminar los atributos que aportan poca información (Pla, 1986) y remover la redundancia y el efecto aleatorio entre variables correlacionadas (Parent *et al.*, 1994). Este método se basa en la maximización de la varianza de las transformaciones lineales de los valores de los atributos, es decir, en la generación de vectores y valores propios. Es recomendable la generación de esas nuevas variables (CP's) tomando como base la matriz de correlación producto de los valores estandarizados de los atributos para evitar la sensibilidad a las diferencias de escala en que se registran y a los valores extraños (Valdez-Cepeda, 1997). Así se producen CP's independientes entre sí, y cada uno sintetiza la máxima variabilidad residual contenida en los datos (Pla, 1986) y

explica una proporción de la variación total, misma que será igual al número de variables estandarizadas, es decir, todos los CP's explican el 100 % de la variación del conjunto total de atributos y observaciones. Por consiguiente, se considera que los CP's importantes son aquellos que explican más del 1 % de la variación, y los atributos también importantes son aquellos que se correlacionan más con los CP's importantes, estos atributos son útiles para definir la nomenclatura de los grupos de especies, genotipos y/o cultivares. La estructura de los CP's puede sugerir algún significado biológico (Iezzoni y Pritts, 1991) al considerar atributos lógicamente biológicos.

### Método y número de grupos o clases

Para la clasificación de *Opuntia* se propone hacer uso del análisis por conglomerados ("cluster analysis" o "agglomerative hierarchical cluster analysis", en inglés). Ha sido demostrado que dicho método es tan robusto para la reconstrucción de relaciones filogenéticas jerárquicas como los métodos cladísticos (James y McCulloch, 1990). Sin duda alguna, uno de los principales problemas de las técnicas de clasificación numérica es la determinación del número de grupos o clases. En el caso de los métodos numéricos jerárquicos dicho problema es el equivalente a definir el nivel o punto al cual se interrumpe la jerarquía.

Un ejemplo es el caso de las pruebas de mejoramiento genético en las cuales se desconoce el número de grupos diferentes que representan cierta composición genética (Bull *et al.*, 1992). Bull *et al.* (1992; 1993; 1994) han resuelto satisfactoriamente dicho inconveniente al considerar el supuesto de que la variabilidad de los patrones genéticos esperados a través del error experimental en cada ambiente es avalada mediante la repetición de un genotipo como testigo; así el número de grupos o clases es definido al truncar la jerarquía cuando las repeticiones del genotipo testigo se ubican en grupos diferentes. En consecuencia, se recomienda la estrategia de Bull *et al.* (1992, 1993, 1994) para el caso del género *Opuntia* cuando el objetivo fundamental sea el mejoramiento genético.

Los resultados de Pengelly y Eagles (1995) sugieren que debido a las diferencias (agronómicas y morfológicas) dentro de los grupos y a que la variación no siempre es discreta (pero puede ser continua), es factible que algunos genotipos sean ubicados en dos grupos o clases. Ello significa que puede haber casos en los que si un genotipo es parte de dos grupos o clases sea un aspecto de importancia relativa. Sin embargo, es posible que los análisis de varianza para comparar los grupos se puedan constituir como una herramienta valiosa para definir el número de grupos. Es decir, es recomendable identificar el número de grupos en el cual el mayor número de variables tengan realmente un efecto en la diferenciación de los grupos.

Definitivamente, la forma en como se han estado definiendo los números de grupos o el nivel de interrupción

de la jerarquía es subjetiva, con algunas excepciones como los trabajos de Bull *et al.* (1992, 1993, 1994), pero la conformación de los grupos o clases ha estado sustentada en los coeficientes de similitud o disimilitud entre genotipos, como se propone en el presente caso, y los métodos de agrupamiento o de definición de máxima similitud entre los grupos o clases empleados. Por lo tanto, enseguida se describen brevemente los más convenientes para apoyar el objetivo planteado.

El coeficiente de correlación ( $r$ ) de Pearson es muy usado como coeficiente de similitud porque es una medida de patrón más que de magnitud de diferencia (Arkley, 1976). Este coeficiente es entonces de gran utilidad si se desea agrupar en función del patrón de alguna parte de las plantas de *Opuntia* spp.

Otro coeficiente de similitud de uso común es la "distancia euclidiana" debido a que es muy sensible a la magnitud y es métrica (aspecto valioso si las variables son continuas). Sin embargo, debe tomarse con cuidado ya que por su sensibilidad a la magnitud puede contribuir desproporcionadamente sobre la disimilitud entre pares de genotipos si la diferencia entre ambos es grande.

El agrupamiento generalmente se hace considerando la similitud o disimilitud. Quizás, el método de agrupamiento más ampliamente usado es el que considera la "distancia de unión promedio" y se conoce como el método de "unión simple". Ello se debe probablemente a que los genotipos o grupos de genotipos se unen con base al promedio de similitud entre todos los pares de genotipos en un grupo y aquellos en algún otro grupo, antes de que cada par o pares iniciales sean unidos. Éste método entonces se basa en el promedio de las diferencias entre genotipos y grupos. Otras variantes se basan en aspectos modales de genotipos y grupos.

Otra ventaja, del método de análisis por conglomerados, es que la "distancia genética" puede ser usada como atributo de clasificación e índice de similitud o disimilitud en clasificaciones más objetivas (*i. e.* mejoramiento genético). Recientemente, Mienie *et al.* (1995) usaron este atributo considerando 37 variedades de soya de Sudáfrica para identificar cuatro grupos principales en un fenograma (dendrograma que muestra la relación de los genotipos de acuerdo a la similitud de polimorfismo).

Es muy posible que el método produzca grupos, mediante los dendrogramas, cuya codificación sea inestable y/o confusa (James y McCulloch, 1990); pero su compatibilidad con métodos de ordenación (*e. g.* análisis por componentes principales) facilita la interpretación de las relaciones dentro y entre los grupos.

Finalmente, el proceso de asignación del nombre genérico (nomenclatura) para los grupos o clases deberá

ser característico y flexible. En ese sentido parece conveniente resaltar el atributo que contribuya más en cada nivel de jerarquía, tal y como lo hicieron Pengelly y Eagles (1995). Este aspecto puede ser apoyado por los resultados del método de ordenación taxonómica.

## MATERIALES Y MÉTODOS

### Datos

En las instalaciones del Centro Regional Universitario Centro Norte de la Universidad Autónoma Chapingo se inició el establecimiento de un banco de germoplasma en 1987. Desde 1995 se cuenta con una "Colección de Nopal" conformada por 88 accesiones, cada una representada por seis u ocho especímenes. La mayoría de las accesiones produjeron tuna en 1999; sin embargo, solamente se presentan resultados de 29 de ellas que fueron las que produjeron más frutos, pues debido a la helada de diciembre de 1997 todavía no se estabiliza la producción de frutos. En cada una de esas colecciones se cosecharon cinco tunas para estimar peso total de fruto (PT), peso de cáscara (PC), peso de la parte comestible (PPC), diámetro polar (DP), diámetro ecuatorial (DE), volumen (VOL), contenido de azúcares en °brix (GB) y la relación entre diámetro polar y diámetro ecuatorial (DP/DE).

### Métodos

La matriz de correlaciones simples (Broschat, 1979) entre las variables o atributos estandarizadas(os) de fruto fue la base para transformarlas en combinaciones lineales o componentes principales (CP's).

Los promedios de esas variables se consideraron para ordenar las accesiones en grupos relativamente uniformes y definir sistemáticamente la similitud de las accesiones, mediante las técnicas de análisis de componentes principales y análisis por conglomerados.

Los factores, de las combinaciones lineales importantes, generados en el análisis de componentes principales se usaron en el análisis por conglomerados. El método de unión simple se utilizó para agrupar a las accesiones. La distancia euclidiana fue el índice de similitud y/o disimilitud usado en el análisis por conglomerados.

Todos los análisis estadísticos se hicieron con el programa Statistica 5.1 (StatSoft, Inc., 1998).

## RESULTADOS Y DISCUSIÓN

### Correlación

Se encontraron las correlaciones significativas ( $P \leq 0.05$ ) y positivas son PT con PC, PPC, DP, DE y VOL, PC con PPC, DP, DE y VOL, PPC con DP, DE y VOL, DP

con DE, DP/DE y VOL y DE con VOL (Cuadro 1); ello sugiere una gran interdependencia entre esos pares de variables.

Por otra parte, la correlación significativa ( $P \leq 0.05$ ) negativa entre DP/DE y GB (Cuadro 1) sugiere que las tunas más alargadas son menos dulces. Este resultado coincide con lo ya reportado por Valdez-Cepeda *et al.* (1995, 1996, 1997) y Fernández *et al.* (1999) mediante análisis de componentes principales por lo que se constituye en una relación con significado biológico, tal y como lo consignaron lezzoni y Pritts (1991).

El gran número de correlaciones significativas entre pares de esas variables indica un alto grado de intercorrelación entre ellas, de manera que para agrupar a las accesiones es necesario no incluirlas a todas o bien reducir la dimensionalidad del problema mediante el análisis de componentes principales (Pla, 1986; James y McCulloch, 1990; Parent *et al.*, 1994).

**CUADRO 1. Matriz de coeficientes de correlación (r) de Pearson entre atributos de tuna de 29 accesiones de nopal tunero (n=145).**

	PT	PC	PPC	DP	DE	DP/DE	VOL
PC	<b>0.84</b>						
PPC	<b>0.88</b>	<b>0.50</b>					
DP	<b>0.79</b>	<b>0.62</b>	<b>0.74</b>				
DE	<b>0.89</b>	<b>0.75</b>	<b>0.79</b>	<b>0.48</b>			
DP/DE	0.28	0.17	0.30	<b>0.80</b>	-0.13		
VOL	<b>0.99</b>	<b>0.84</b>	<b>0.87</b>	<b>0.76</b>	<b>0.88</b>	0.26	
GB	-0.09	-0.03	-0.12	-0.32	0.11	<b>-0.43</b>	-0.08

<sup>a</sup>Correlaciones significativas a una  $P \leq 0.05$  en negritas

PT: peso total de fruto; PC: peso de cáscara; PPC: peso de la parte comestible; DP: diámetro polar; DE: diámetro ecuatorial; DP/DE: proporción de diámetros polar y ecuatorial; VOL: volumen de fruto; GB: grados brix.

### Análisis de componentes principales

La matriz de correlaciones, con las observaciones estandarizadas, sirvió de fuente para el análisis de componentes principales. Casi el 38 % de la variación de las ocho variables y 29 accesiones es explicada por el CP1 (Cuadro 2).

La estructura del CP1 es definida por las variables PT, PPC, DE y VOL de manera que dicha combinación lineal puede ser etiquetada como de masa y volumen de tuna.

El CP 2 está definido por PC y explica el 24 % de la variación (componente de masa de cáscara).

El CP3 es definido por DP y DP/DE (componente de morfometría de fruto) y explica el 21.5 % de la variación total.



**CUADRO 2. Correlaciones entre variables de tuna y componentes principales (CP's) considerando 29 accesiones de nopal tunero.**

	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8
PT	<b>0.760537</b>	0.564713	0.216538	0.026964	0.017198	0.017120	0.029190	0.231572
PC	0.327001	<b>0.935340</b>	0.130613	-0.003829	0.022809	-0.008342	0.021146	-0.009894
PPC	<b>0.955643</b>	0.138552	0.234623	0.047711	-0.060504	-0.050934	0.024843	-0.057908
DP	0.492733	0.346246	<b>0.741543</b>	0.143307	0.041398	0.023410	0.253192	0.023187
DE	<b>0.777234</b>	0.501419	-0.148946	-0.080425	0.339638	-0.003511	0.020256	0.007569
DP/DE	0.043939	0.029649	<b>0.974784</b>	0.201968	-0.036997	0.007268	-0.068980	0.003868
VOL	<b>0.751088</b>	0.569997	0.195150	0.025888	-0.007210	0.267600	0.020015	0.012406
GB	-0.014521	0.003765	-0.212750	<b>-0.976940</b>	0.007732	-0.001913	-0.005945	-0.001376
Varianza explicada	3.011752	1.910057	1.724618	1.025896	0.123025	0.075184	0.071627	0.057842
Proporción del total	0.376469	0.238757	0.215577	0.128237	0.015378	0.009398	0.008953	0.007230
% del total	37.6469	23.8757	21.5577	12.8237	1.5378	0.9398	0.8953	0.7230

<sup>a</sup>En negritas correlaciones >0.7

PT: peso total de fruto; PC: peso de cáscara; PPC: peso de la parte comestible; DP: diámetro polar; DE: diámetro ecuatorial; DP/DE: proporción de diámetros polar y ecuatorial; VOL: volumen de fruto; GB: grados brix.

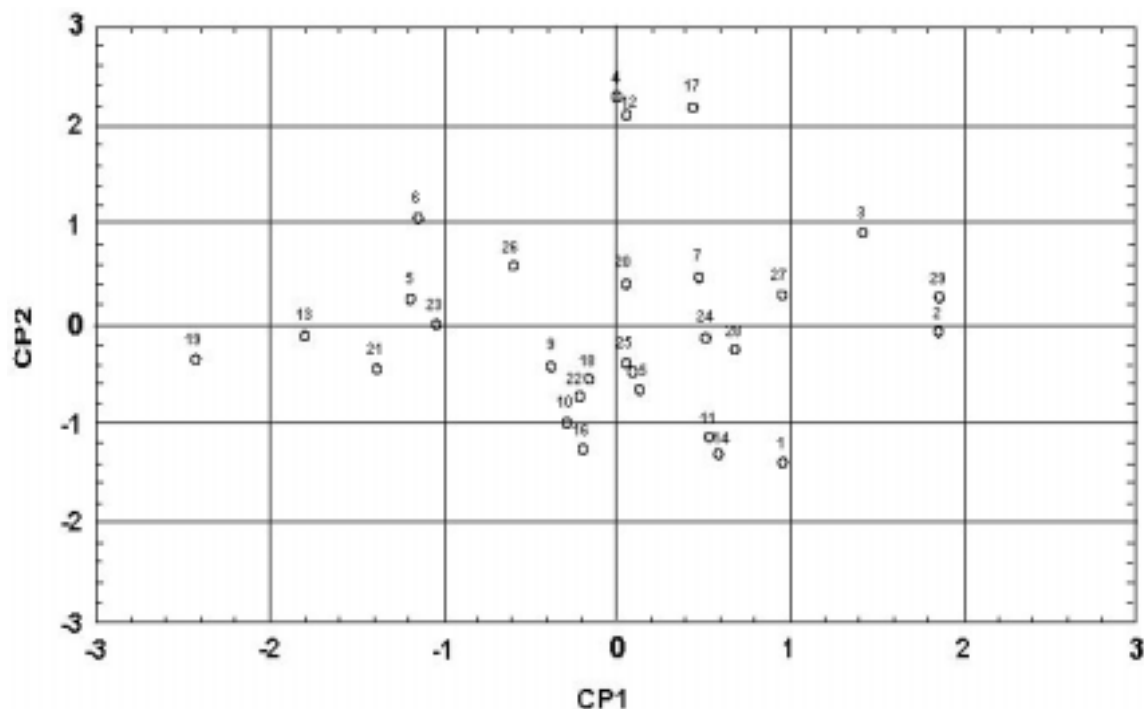
El contenido de azúcares (GB) domina en la estructura del CP4, el cual explica el 13 % de la variación total.

En resumen, los dos primeros CP's explican el 62 % de la variación total. Si se considera al componente de morfometría de fruto (CP3) también se explica casi el 84 % de la variación total de las 29 accesiones. Y si se contempla al CP4, el componente de contenido de azúcar, se explica el 97 % de la variación total. Estas nuevas cuatro variables son el producto del análisis de componentes principales, es decir con el uso de este método se redujo la dimensionalidad del problema de ocho a cuatro variables, y sus factores son de gran utilidad en la ordenación de los

objetos de estudio (accesiones de nopal en el presente caso) y en el análisis de conglomerados para agruparlos o clasificarlos numéricamente.

#### Ordenación de accesiones

Se aprecia que las accesiones 'Pachón', 'Camuezo' y 'Copena 15' producen frutos pequeños (Figura 1), pero 'Cristalina' y 'Tapón Aguanoso' son las que producen tunas de mayor volumen y peso; 'Cristalina' es cultivada en la región Centro-Norte del territorio nacional mientras que la 'Tapón Aguanoso' es silvestre, de manera que es



**Figura 1. Posiciones de 29 accesiones de nopal tunero en el plano ortogonal definido por los componentes principales de masa y volumen totales de tuna (CP1) y de su masa de cáscara (CP2).**

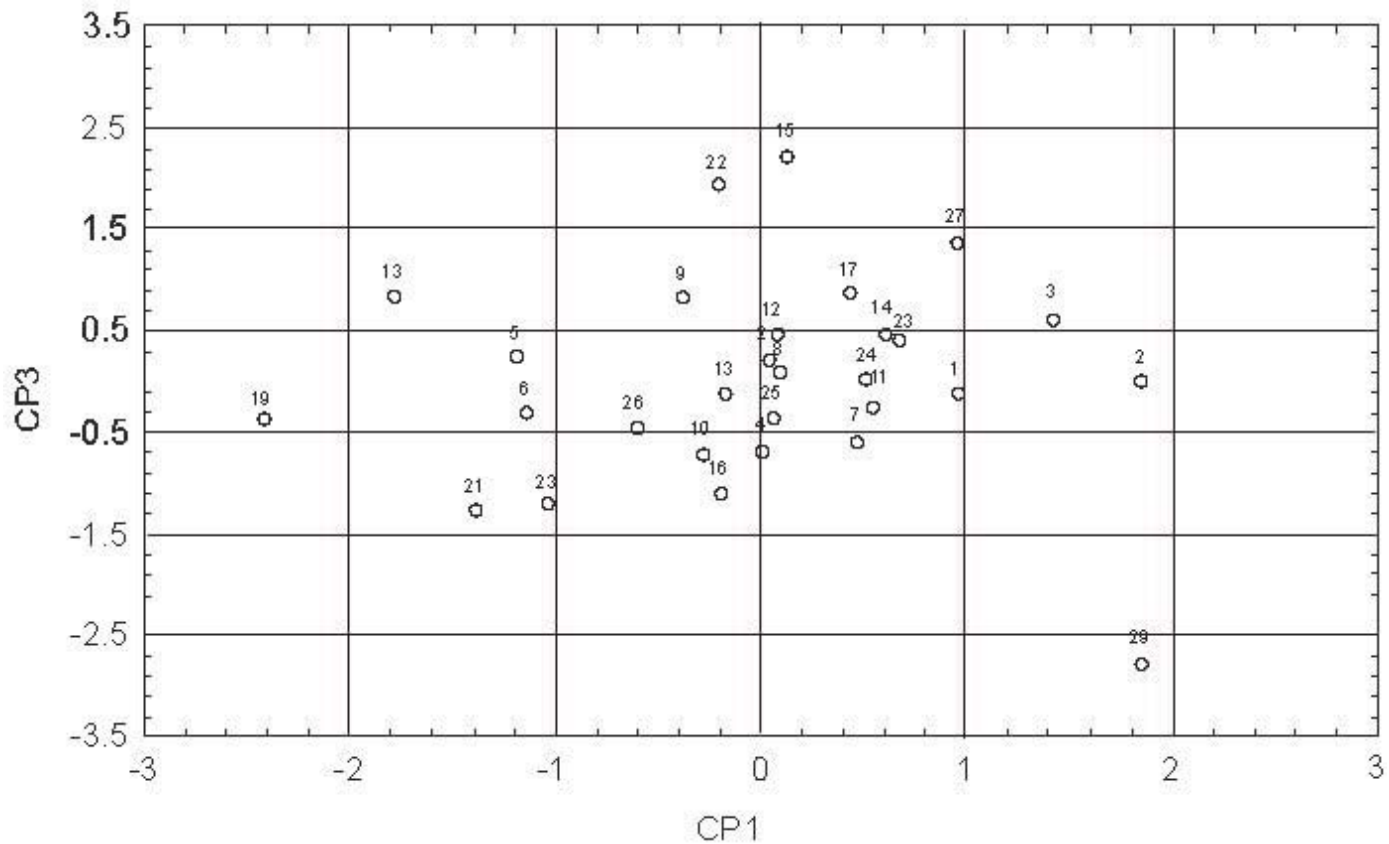


Figura 2. Posiciones de 29 accesiones de nopal tunero en el plano ortogonal definido por los componentes principales de masa y volumen totales de tuna (CP1) y de su morfometría (CP3).

conveniente no agruparlas juntas por lo que es necesario involucrar un tercer CP para su ordenación (Figura 2), en otras palabras, los dos primeros CP's no permiten ubicarlas gráficamente en posiciones distintas.

Las accesiones 'Mango', 'Burrón 1' y 'Naranjón Legítimo' producen las tunas con más masa de cáscara; mientras que 'Copena Z-2', 'Copena 5', 'Copena 14' y 'Copena Z-1' las que producen fruta con menos masa de cáscara (Figura 1).

Las accesiones 'Amarilla Montesa', 'Rubí Reyna' y 'Cristalina 2' producen las tunas más alargadas, aunque

cabe aclarar que el pedúnculo de dichas accesiones es prominente (Figura 2).

Al ordenar las accesiones en un gráfico (no se presenta) que involucre al CP4 y cualquier otro de los tres primeros, las Copenas se posicionan para formar el grupo que produce las tunas más dulces, lo cual ya ha sido previamente evidenciado por Valdez *et al.* (1995, 1996, 1997).

Es difícil identificar un patrón de agrupamiento de las accesiones, pues se ubican junto a vecinos diferentes en un gráfico simple respecto de los factores de los cuatro

CUADRO 3. Análisis de varianza para el agrupamiento de 29 accesiones de nopal tunero en cinco grupos considerando ocho atributos de sus tunas.

	Suma de Cuadrados	Grados de Libertad	Suma de Cuadrados (Dentro de Grupos)	Grados de Libertad	Estimated F	F $P \leq$
CP1	20.49925	4	7.57294	24	16.24145	0.000001
CP2	15.72946	4	11.48189	24	8.21962	0.000253
CP3	12.27392	4	15.70525	24	4.68910	0.006129
CP4	14.96875	4	12.76521	24	7.03572	0.000676

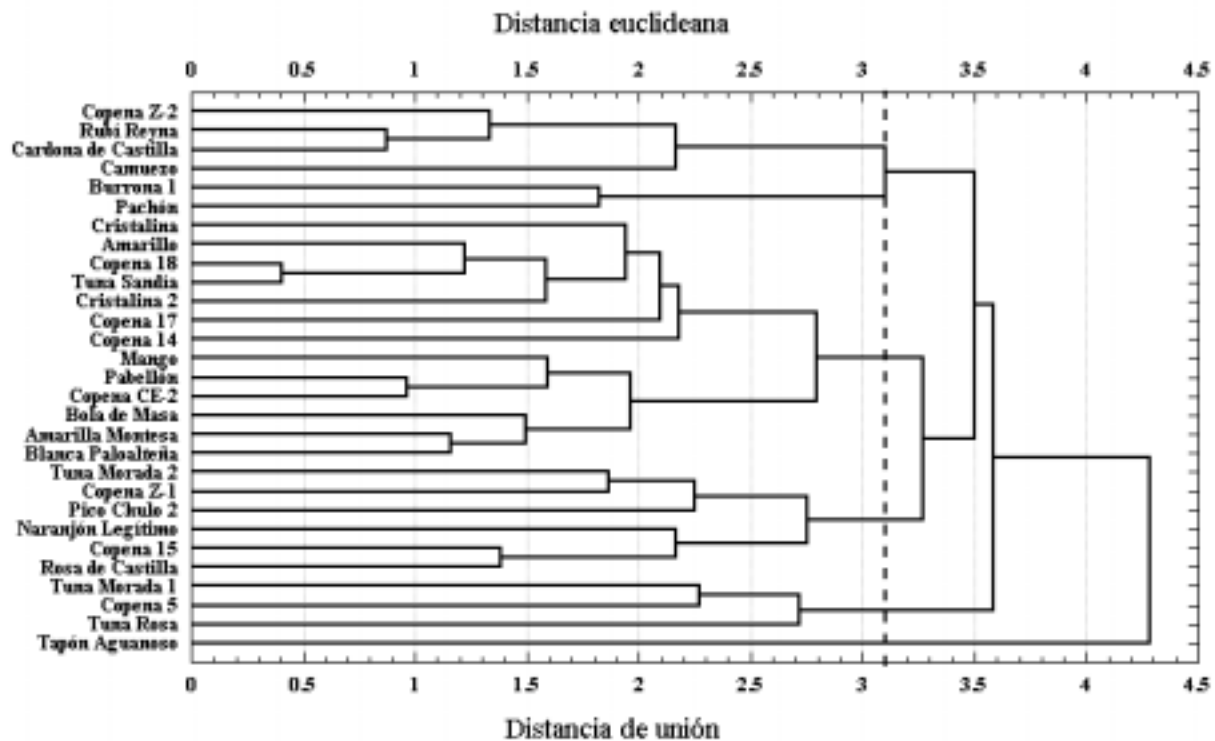


Figura 3. Dendrograma de similitud y clasificación de 29 accesiones de nopal tunero con base en atributos de sus frutos o tunas.

primeros CP's. Sin embargo, ese patrón de agrupamiento se puede identificar al usar los factores de los componentes principales importantes en el análisis por conglomerados (*i.e.* Boer *et al.*, 1993).

### Análisis por conglomerados

La clasificación de las accesiones con esta técnica puede hacerse mediante el uso de un gran número de métodos de agrupamiento e índices de similitud, pero debe considerarse bases estadísticas fundamentales. En el presente caso, el análisis de varianza en el que todas las variables (los cuatro primeros CP's) presentan diferencias significativas estadísticamente corresponde a la clasificación de las 29 accesiones en cinco grupos (Cuadro 3) o más. Sin embargo, es notorio que el componente de morfometría de tuna (CP3) es el que presenta el menor valor de significancia (Cuadro 3) y es explicado porque las diferencias morfométricas de tuna entre accesiones no se apreciaron marcadamente, ya que el año fue relativamente seco y los materiales que producen frutos grandes y/o alargados no manifestaron su potencial. Aún así, ello demuestra que el uso de los factores de esos CP's, como datos en el análisis por conglomerados, evita distorsiones en las coordenadas (Edmonds *et al.*, 1985) de las accesiones en el espacio definido por los atributos de tuna.

Considerando al índice de similitud denominado distancia euclidiana como distancia de unión y un valor de 3.1 se aprecian claramente cinco grupos de accesiones

de nopal tunero en el dendrograma (Figura 3): el primer grupo lo conforman 'Copena Z-2', 'Rubí Reyna', 'Cardona de Castilla', 'Camuezo', 'Burróna 1' y 'Pachón'; el segundo grupo lo integran 'Cristalina', 'Amarillo', 'Copena 18', 'Tuna Sandía', 'Cristalina 2', 'Copena 17' y 'Copena 14' en un subgrupo, mientras que en el otro subgrupo se asocian 'Mango', 'Pabellón', 'Copena CE-2', 'Bola de Masa', 'Amarilla Montesa' y 'Blanca Paloalteña'; el tercer grupo lo constituyen 'Tuna Morada 2', 'Copena Z-1', 'Pico Chulo 2', 'Naranjón Legítimo', 'Copena 15' y 'Rosa de Castilla'; el cuarto grupo lo forman 'Tuna Morada 1', 'Copena 5' y 'Tuna Rosa'; y finalmente el último grupo se identifica únicamente con 'Tapón Aguanoso', *Opuntia robusta*, material silvestre típico en la región Centro-Norte de México.

La heterogeneidad de cada uno de los tres grupos centrales en la Figura 3 es de magnitud parecida, valor de la distancia euclidiana entre 2.7 y 2.8 (Figura 3), pues el sustento teórico del método es claro en el sentido de que la heterogeneidad dentro de los grupos se incrementa conforme el coeficiente de similitud es mayor (Edmonds *et al.*, 1985). Así entonces, considerando este resultado y el análisis de varianza (Cuadro 3), se infiere que la conformación de los cinco grupos de accesiones de nopal tunero con base a los atributos de tuna es estadísticamente robusta.

### CONCLUSIONES

Los métodos estadísticos usados para ordenar y clasificar 29 accesiones de nopal tunero, con base en ocho

atributos de fruto, permitieron confirmar resultados previos respecto de que los frutos o tunas más grandes y alargadas contienen menor cantidad de azúcares, y esos atributos se asocian estrechamente con algunas accesiones.

Los componentes principales más importantes en la ordenación de esos genotipos son los relacionados a la masa y volumen de fruto, masa de cáscara y morfometría de la tuna.

La mejor estrategia, en el presente caso, para clasificar las 29 accesiones consistió en usar los factores de los componentes principales en el análisis de conglomerados con el método de unión simple y la distancia euclidiana como índice de similitud. Ello permitió identificar cinco grupos de accesiones de nopal considerando una distancia de unión de 3.1.

El genotipo silvestre 'Tapón Aguanoso' (*Opuntia robusta*) conformó por sí solo una clase. Es evidente que las accesiones nativas de la región Centro Norte de México tienden a conformar grupos o clases con una gran variabilidad fenotípica de fruto. Las accesiones Copena también tienden a agruparse en clases aunque con menor heterogeneidad fenotípica de fruto que las de la región Centro Norte.

### AGRADECIMIENTOS

Ricardo David Valdez Cepeda agradece al CONACyT el financiamiento parcial de la investigación mediante el contrato I-26976-N, y al Programa de Investigación en Fruticultura de la Universidad Autónoma Chapingo mediante el proyecto No. 230201.

### LITERATURA CITADA

ARKLEY, R. J. 1976. Statistical methods in soil classification research. *Adv. Agron.* 28: 37-70.

BOER, R.; FLETCHER, D. J.; CAMPBELL, C. 1993. Rainfall patterns in a major wheat-growing region of Australia. *Aust. J. Agric. Res.* 44: 609-624.

BROSCHAT, T. K. 1979. Principal component analysis in horticultural research. *HortScience* 14(2): 114-117.

BULL, J. K.; COOPER, M.; DELACY, I. H.; BASFORD, K. E.; WOODRUFF, D. R. 1992. Utility of repeated checks for hierarchical classification of data from plant breeding trials. *Field Crops Res.* 30: 79-95.

BULL, J. K.; BASFORD, K. E.; DELACY, I. H.; COOPER, M. 1993. Determining appropriate group number and composition for data sets containing repeated check cultivars. *Field Crops Res.* 31: 369-383.

BULL, J. K.; BASFORD, K. E.; COOPER, M.; DELACY, I. H. 1994. Enhanced interpretation of pattern analyses of environments: The use of blocks. *Field Crops Res.* 37: 25-32.

DALE, M. B.; MCBRATNEY, A. B.; RUSSELL, J. S. 1989. On the role of expert systems and numerical taxonomy in soil classification. *J. Soil Sci.* 40:223-234.

EDMONDS, W. J.; CAMPBELL, J. B.; LENTNER, M. 1985. Taxonomic variation within three soil mapping units in Virginia. *Soil Sci. Soc. Am. J.* 49: 394-401.

FERNÁNDEZ, M. R.; LUNA, J.; SÁENZ, L. A.; MÉNDEZ, S. J.; MONDRAGÓN, C.; ZEBBE, J. 1999. Estudio de cultivares de nopal tunero por componentes principales. *In: Aguirre, J.R.; Reyes, J.A. Memoria: VIII Congreso Nacional y VI Internacional Sobre Conocimiento y Aprovechamiento de el Nopal.* San Luis Potosí, S. L. P., México. pp. 4-5.

GREEN, D. M. 1991. Chaos, fractals and nonlinear dynamics in evolution and phylogeny. *Trends in Ecology and Evolution* 6(10): 333-337.

IEZZONI, A. F.; PRITTS, M. V. 1991. Applications of principal component analysis to horticultural research. *HortScience* 26(4): 334-338.

JAMES, F. C.; McCULLOCH, CH. E. 1990. Multivariate analysis in ecology and systematics: Panacea or Pandora's box? *Annual Rev. Ecol. Syst.* 21: 129-166.

MIENIE, C. M. S.; SMITH, M. A.; PRETORIUS, P. J. 1995. Use of random amplified polymorphic DNA for identification of South Africa soybean cultivars. *Field Crops Res.* 43: 43-49.

PARENT, L. E.; ISFAN, D.; TREMBLAY, N.; KARAM, A. 1994. Multivariate nutrient diagnosis of the carrot crop. *J. Amer. Soc. Hort. Sci.* 119(3): 420-426.

PLA, L. 1986. Análisis Multivariado: Método de Componentes Principales. Sría Gral. OEA. Washington, D. C., USA. Serie Matemática. Monografía No. 27. 94 p.

PENGELLY, B. C.; EAGLES, D. A. 1995. Geographical distribution and diversity in a collection of the tropical legume *Macroptilium gracile* (Poepig ex Benth) urban. *Aust. J. Agric. Res.* 46: 569-580.

SNEATH, P. H. A.; SOKAL, R. R. 1973. Numerical Taxonomy. W. H. Freeman & Co. San Francisco, USA.

STATSOFT, INC. 1998. Statistica for Windows 5.1. Computer Program Manual. Tulsa, OK, USA.

VALDEZ-CEPEDA, R. D. 1997. Análisis estocástico espacial de procesos edáficos y de plantas en *Zea mays* L. Universidad Autónoma de Nuevo León-Facultad de Agronomía. Tesis de Doctor en Ciencias. Marín N. L., México. 91 p.

VALDEZ-CEPEDA, R. D.; GALLEGOS-VÁZQUEZ, C.; BLANCO-MACÍAS, F. 1995. Agrupamiento jerárquico de veintinueve genotipos de *Opuntia* spp. mediante características del fruto (tuna), pp. 15-18. *In: Conocimiento y Aprovechamiento del Nopal: Memorias 6to. Congreso Nacional y 4to. Congreso Internacional.* Pimienta B.; Neri, E. C.; Muñoz, A.; Huerta, F. M. (Comps.). Las Agujas, Nextipac, Zapopan, Jal., México.

VALDEZ-CEPEDA, R. D.; GALLEGOS-VÁZQUEZ, C.; BLANCO-MACÍAS, F. 1996. Clasificación numérica de *Opuntia* spp. mediante características de su fruto (tuna). *Rev. Geografía Agrícola* 22-23: 287-293.

VALDEZ-CEPEDA, R. D.; GALLEGOS-VÁZQUEZ, C.; BLANCO-MACÍAS, F. 1997. Análisis multivariado en once variedades de nopal tunero: atributos de fruto, pp. 118-119. *In: Memorias del VII Congreso Nacional, V Internacional Sobre Conocimiento y Aprovechamiento del Nopal.* Vázquez, R.; Gallegos, C.; Treviño, N.; Díaz, Y. (Comps.). FA-UANL, Monterrey, N.L., México.

VILLASEÑOR, J. L.; MURGUÍA-ROMERO, M. 1992. La computadora en la identificación botánica. *Ciencia y Desarrollo* 18(104): 130-137.

WEBB, D. A. 1954. Is the classification of plant communities either possible or desirable? *Saertryk af Botanisk Tidsskrift* 51: 361-370.