

CAPÍTULO 4

Taxonomía numérica en el estudio de las especies

Luciana Lugones, María del Rosario Robles, Joaquín Cochero y Javier Ricardo Garcia de Souza

Introducción

La consideración de diferentes evidencias como morfología, datos ecológicos, marcadores moleculares, etc, en forma independiente o en simultáneo, permite la delimitación de unidades taxonómicas. Cada fuente de datos proporciona información relevante y pueden ser exploradas a partir de diferentes herramientas estadísticas descriptivas (Lanteri y Cigliano, 2006; Palacio *et al.*, 2020).

Cuando la fuente de evidencia es solo morfológica, los estudios morfométricos que comprenden la descripción, análisis e interpretación cuantitativa de la forma y del cambio morfológico, suelen ser muy útiles (Rohlf, 1990; Rohlf y Marcus, 1993). Tradicionalmente se obtiene información de las estructuras anatómicas desde el análisis de conjuntos de medidas lineales (*e.g.* distancias, proporciones) y desde coordenadas en el espacio anatómico (definidas como *landmarks* y *semi-landmarks* en el marco de estudios de morfometría geométrica) (Bookstein, 1991; Toro Ibacache *et al.*, 2010).

Otros datos como los obtenidos desde fuentes ecológicas y fisiológicas pueden ser considerados para sumar a la exploración de la delimitación de los taxones, traducidos en datos cualitativos o cuantitativos (Nei y Li, 1979; Palacio *et al.*, 2020).

Los datos sobre diversidad genética pueden recuperarse del ADN, ARN y de las proteínas. La interpretación de las variaciones genéticas para la delimitación de taxones puede valerse de diferentes datos cualitativos o cuantitativos. La mayoría de estos estudios presuponen modelos evolutivos de sustitución nucleotídica y son abordados en un contexto filogenético (*e.g.* Tamura *et al.*, 2007), pero también algunas propuestas sugieren considerar a estos datos en el marco de análisis estadísticos descriptivos que reflejan similitud (Núñez-Colín y Escobedo-López, 2011; Konishi *et al.*, 2019).

La [Microtaxonomía](#) se define como una rama de la taxonomía cuyo objetivo es identificar, describir y delimitar a las especies, considerando la presencia de variaciones infraespecíficas. Al

enfoque **feneticista** de la taxonomía se lo denomina **taxonomía numérica**, ya que emplea operaciones matemáticas que permiten establecer relaciones de **similitud global** (o fenéticas) entre unidades de estudio, sobre la base de la evidencia que brindan sus caracteres (e.g. morfológicos, ecológicos, moleculares) (Sneath y Sokal, 1973; Sokal y Rohlf, 1986), es decir, agrupa a los organismos sin tener en cuenta sus relaciones filogenéticas. Dado que cada organismo está caracterizado por múltiples atributos o caracteres (cuantitativos y cualitativos codificados), son los análisis multivariados (o también llamadas técnicas multivariadas) los que proveen las principales herramientas para estudios taxonómicos de tipo numérico (Arce *et al.*, 2009).

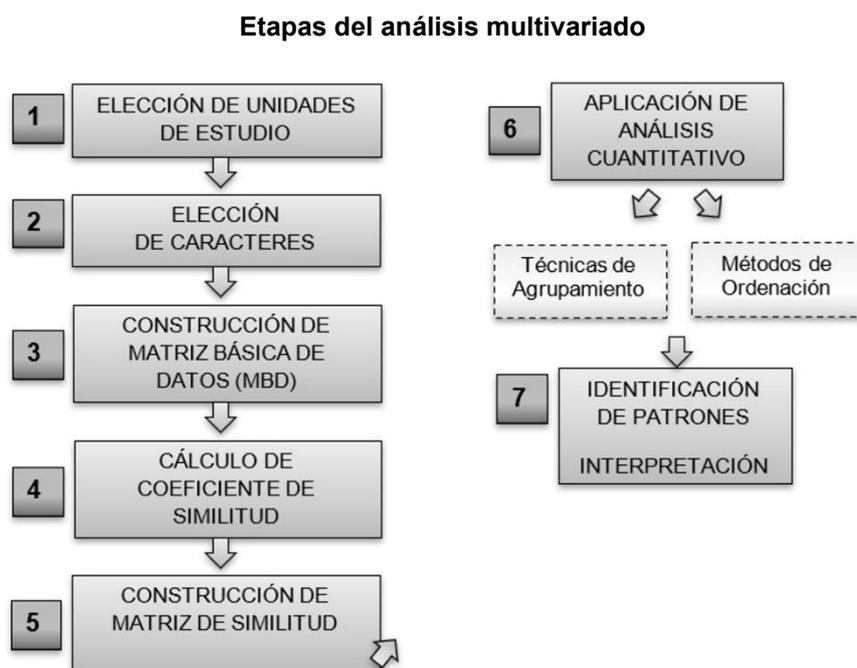
Técnicas multivariadas

Las técnicas y métodos, incluidos en estos análisis, se emplean tanto de forma exploratoria para describir tendencias y principales fuentes de variación en los datos; como de modo confirmatorio, para responder a hipótesis referidas a la variabilidad de las unidades de estudio y sus causas de variación (James y McChulloch, 1990; Palacio *et al.*, 2020).

A partir de los resultados obtenidos (e.g. un fenograma), se pueden adoptar decisiones taxonómicas con respecto a la delimitación de especies y las variaciones dentro de las mismas.

Los principales análisis son técnicas de agrupamiento y métodos de ordenación. La elección del tipo de análisis dependerá de la pregunta de el/la investigador/a y del tipo de datos. Ambos análisis cuantitativos tienen una serie de pasos en común (Figura 4.1).

Figura 4.1. Etapas comunes para la mayoría de los estudios que emplean técnicas multivariadas.



Descripción de cada etapa:

1. Elección de las unidades de estudio u *Operational Taxonomic Units (OTU)*

- Por lo general las *OTUs* son individuos, poblaciones, o linajes genéticos de una misma especie, subespecies o especies próximas dependiendo del problema a resolver.

2. Elección de los caracteres o variables

- Pueden ser de diferente tipo según su fuente (morfológicos, ecológicos, genéticos) y se registran como valores (datos) para cada *OTU*. Si bien se emplean caracteres cualitativos, la mayor parte de los caracteres utilizados en este tipo de estudio son variables cuantitativas continuas (datos morfométricos, frecuencias de alelos dominantes, frecuencias alélicas de loci polimórficos para sistemas enzimáticos, frecuencias cromosómicas, secuencias de ADN, etc.).

3. Construcción de una matriz básica de datos (MBD) (Tabla 4.1a y b).

- Se construye una MBD en la cual se colocan los valores de cada carácter correspondientes a cada *OTU*.

Tabla 4.1a. MBD de 3 OTUs y 7 caracteres discretos doble estado.

	1	2	3	4	5	6	7
A	1	1	0	1	0	1	1
B	1	0	1	1	0	0	1
C	0	1	1	0	1	1	0

Tabla 4.1b. MBD de 3 OTUs, 2 caracteres discretos multiestado y 5 caracteres continuos (matriz mixta).

	1	2	3	4	5	6	7
A	3	1	0,45	1,6	10,0	20,5	40,4
B	1	2	0,49	1,8	12,5	25,6	45,8
C	2	3	0,61	1,6	13,5	28,9	44,7

4. Cálculo de un coeficiente de similitud a partir de la MBD para cada par posible de OTU (Tabla 4.2 y 4.3).

- Para la elección de un coeficiente se debe tomar en cuenta el tipo de datos que conforman la matriz (e.g. fuente y tipo de caracteres doble estado, multiestado, cuantitativos), como así también, otros aspectos relacionados con los objetivos del trabajo a realizar.

Tabla 4.2. Ejemplos de coeficientes de similitud.

Tipo de coeficiente	Nombre	Fórmula	Datos sobre los que se aplica	Máxima similitud	Mínima similitud
Distancia	<i>Manhattan distance</i>	$\sum_{i=1}^n [X_{ij} - X_{ik}]$	Doble estado	0	∞
	<i>Mean character difference</i>	$1/n \sum_{i=1}^n [X_{ij} - X_{ik}]$	Multiestado Mixtos		
Correlación	<i>Momento producto de Pearson</i>	$\sum_{i=1}^n (X_{ij} - \bar{X}_j) \cdot (X_{ik} - \bar{X}_k)$	Multiestado cuantitativos	1	-1
		$\sum_{i=1}^n (X_{ij} - \bar{X}_j)^2 - \sum_{i=1}^n (X_{ik} - \bar{X}_k)^2$	Mixtos (con mayoría de multiestado cuantitativos)		
Asociación	<i>Simple Matching</i>	$\frac{a + d}{a + b + c + d}$	Doble estado	1	0
	<i>Jaccard</i>	$\frac{a}{a + b + c}$			

Nota. Adaptado de Coeficiente de similitud de uso más frecuente en análisis de agrupamiento, de Lanteri y Cigliano, 2006. En las fórmulas de los coeficientes de distancia y correlación, los términos X_{ij} y X_{ik} , indican el valor del carácter "i" en la OTU "j", y el valor del carácter "i" en la OTU "k", respectivamente. En las fórmulas de los coeficientes de asociación se emplean los términos: a= presencias compartidas; b= carácter presente en una de las OTU comparadas y ausente en la otra; c= inversa de b; d= ausencias compartidas,

- Los **coeficientes de similitud** más utilizados pueden dividirse en tres grupos: coeficientes de distancia, de correlación y de asociación. Un mayor detalle sobre los coeficientes puede consultarse en Sneath y Sokal (1973) y Palacio *et al.* (2020). Algunos de los coeficientes de uso frecuente para datos morfológicos son ejemplificados en la Tabla 4.2.

Ejemplo de cálculo de coeficientes de distancia y asociación a partir de la MBD (Tabla 1a)

Manhattan Distance: A, B= 3; A, C= 5; B, C= 6

Jaccard:
 A, B= $3 / (3+2+1) = 3/6 = 0,50$
 A, C= $2 / (2+3+2) = 2/ 7 = 0,28$
 B, C= $1 / (1+3+3) = 1/7 = 0,14$

Momento producto de Pearson (mediante el uso de software estadístico):

A, B= 0,94; A, C= 0,96; B, C= 0,98

- Cuando se analizan datos basados en morfometría geométrica o aquellos expresados en el marco de estudios moleculares (e.g. cromosómicos, alozímicos, de alelos dominantes, etc),

se aplican generalmente otros coeficientes (Tabla 4.3). En relación con la morfometría geométrica se emplea, por ejemplo, un coeficiente de distancia generalizada o de Mahalanobis D2 que considera las varianzas y covarianzas de los datos (*landmarks*) y su aplicación requiere análisis estadísticos previos (ver detalles sobre Morfometría geométrica al final de la introducción, Marcus *et al.*, 1996; Lanteri y Cigliano, 2006; Palacio *et al.*, 2020).

Tabla 4.3. Índices utilizados en morfometría geométrica y estudios sobre distancias genéticas.

Tipo de coeficiente	Nombre	Fórmula	Tipo de datos sobre los que se aplica
Distancia	Mahalanobis D2*	$\sqrt{(\bar{X}_{11} - \bar{X}_{12} / \sigma_1)^2 + \bar{X}_{21} - \bar{X}_{22} / \sigma_2)^2}$	Cuantitativos continuos y discretos (<i>landmarks</i>)
	Nei**	$-\ln [J_{xy} / (J_x J_y)^{1/2}]$	Cuantitativos continuos y discretos (bases nitrogenadas, isoenzimas, etc)

Nota. Donde * \bar{X} = media, σ = desvío estándar y ** \ln es el logaritmo natural, "x" es la frecuencia del alelo "i", en un locus particular de la unidad "x", e "y" es la frecuencia de ese mismo alelo "i", en la unidad "y". J_x , J_y , y J_{xy} son medias aritméticas de $\sum x^2$, $\sum y^2$, $\sum xy$

- Las variaciones en el ADN observadas en diferentes taxones o poblaciones son mutaciones resultantes de la sustitución, inserción, delección, duplicación o inversión de fragmentos de bases nitrogenadas (desde uno a varios miles de nucleótidos) (Ver capítulo 1). Los análisis de distancias genéticas que evalúan las mutaciones pueden ser divididos en dos grandes grupos (a) los que reflejan la similitud entre poblaciones sin considerar *a priori* modelos evolutivos (distancias geométricas) (Cavalli-Sforza y Edwards, 1967; Rogers, 1972; Nei *et al.*, 1983), y (b) los que presuponen modelos evolutivos *a priori* de sustitución nucleotídica (Tamura y Nei, 1993; Takezaki y Nei, 1996). En el primer caso (a), por ejemplo, si lo que se analiza son matrices de presencia/ausencia de bandas de ADN o de proteínas, obtenidas por electroforesis, el coeficiente más adecuado es el desarrollado por Dice en 1945, entre otros (Nei y Li, 1979; Núñez-Colín y Escobedo-López, 2011). En el segundo caso (b), uno de los coeficientes utilizados con mayor frecuencia es el de distancia de Nei (1972), que considera una tasa de mutación constante e igual para todos los *loci*, poblaciones con tamaños efectivos iguales y constantes a través de las generaciones y poblaciones en equilibrio mutación-deriva (Takezaki y Nei, 1996).

5. Construcción de una matriz de similitud (MS) (Tablas 4.4a, b, c).

Se construye una MS a partir de los valores de similitud calculados en el paso anterior. La diagonal principal incluye los valores de máxima similitud (0 en el caso de coeficientes de distan-

cia y 1 en los de correlación y asociación) y los restantes casilleros contienen los valores registrados para cada par de OTU. Solo se completa una mitad de la matriz, dado que la otra es su imagen especular.

Tabla 4.4a. *Matriz de similitud basada en coeficiente de distancia Manhattan Distance.*

	A	B	C
A	0		
B	3	0	
C	5	6	0

Nota. (Matriz de Tabla 4.1a)

Tabla 4.4b. *Matriz de similitud basada en coeficiente de asociación Jaccard.*

	A	B	C
A	1		
B	0,50	1	
C	0,28	0,14	1

Nota. (Matriz de Tabla 4.1a)

Tabla 4.4c. *Matriz de similitud basada en coeficiente de correlación.*

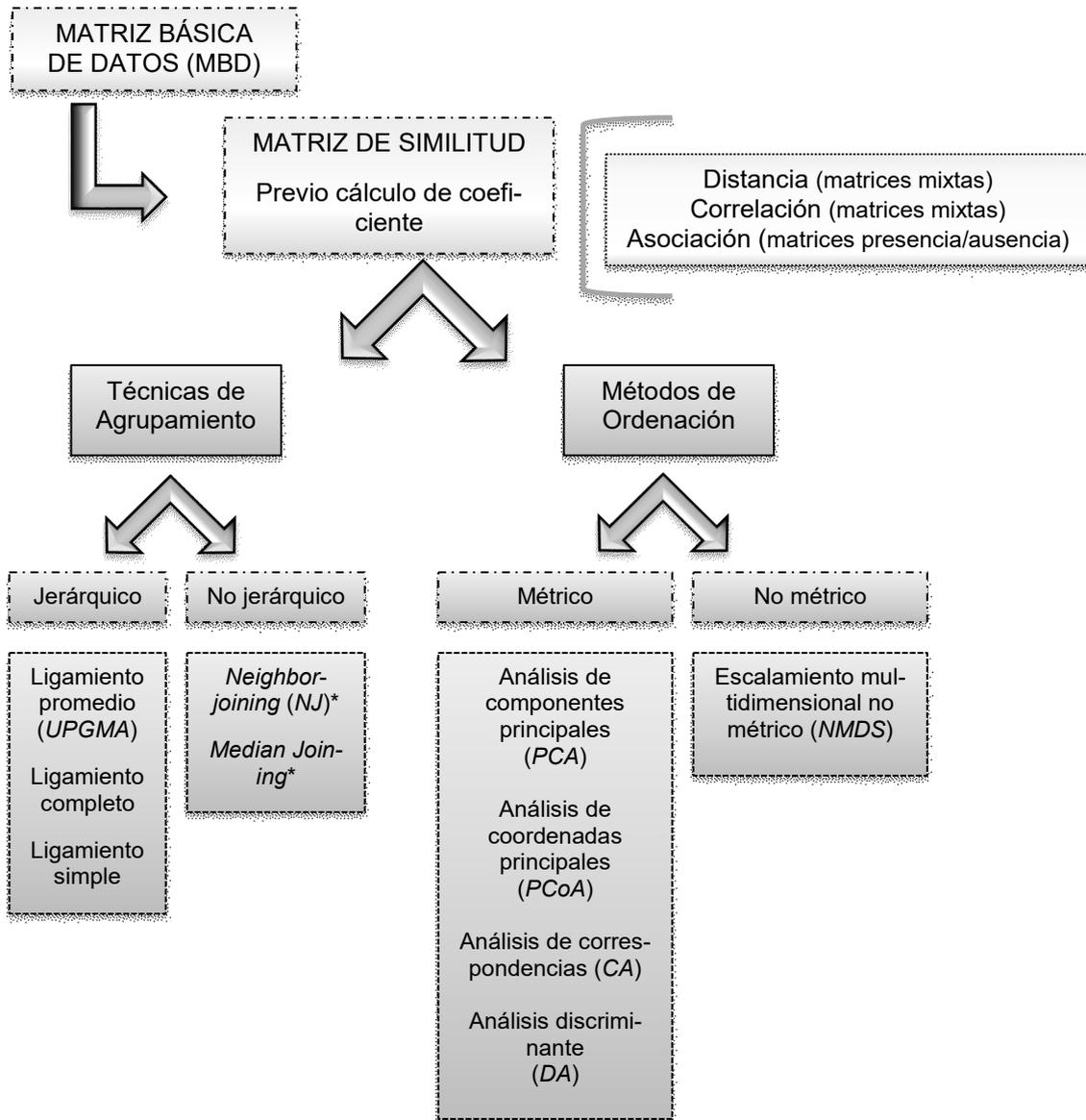
	A	B	C
A	1		
B	0,94	1	
C	0,96	0,98	1

Nota. (Matriz de Tabla 4.1b)

- Si bien en matrices pequeñas es posible hacer el cálculo manual de los índices, diferentes *softwares* permiten su cálculo generando las MS correspondientes (e.g. PAST, R.) (Palacio *et al.*, 2020).

6. Aplicación de un análisis multivariado mediante Técnicas de agrupamiento y Métodos de ordenación (Figura 4.2).

Figura 4.2. Técnicas de agrupamiento y métodos de ordenación.



**Análisis que presuponen modelos evolutivos y una interpretación filogenética.*

Una MBD puede ser estudiada desde dos puntos de vista (Cattell, 1952; Legendre y Legendre, 1998): el de asociación de las OTUs (**técnica o modo Q**) o el de asociación de las variables (**técnica o modo R**). A veces no resulta obvio si un análisis fue realizado mediante el modo Q o R dado que la representación de los resultados expresa relaciones de similitud entre OTUs. El resultado del análisis de agrupamientos suele construirse mediante el modo Q y el análisis de ordenación generalmente mediante el modo Q (en el caso de CoPA) o R (en el caso de PCA) (Palacio *et al.*, 2020).

Descripción de análisis multivariados:

Técnicas de agrupamiento

Este tipo de análisis permite representar a las unidades de estudio (individuos, muestras, especies, etc.) como grupos asociados por su grado de similitud (Palacio *et al.*, 2020).

1. Fenograma (por método de UPGMA)

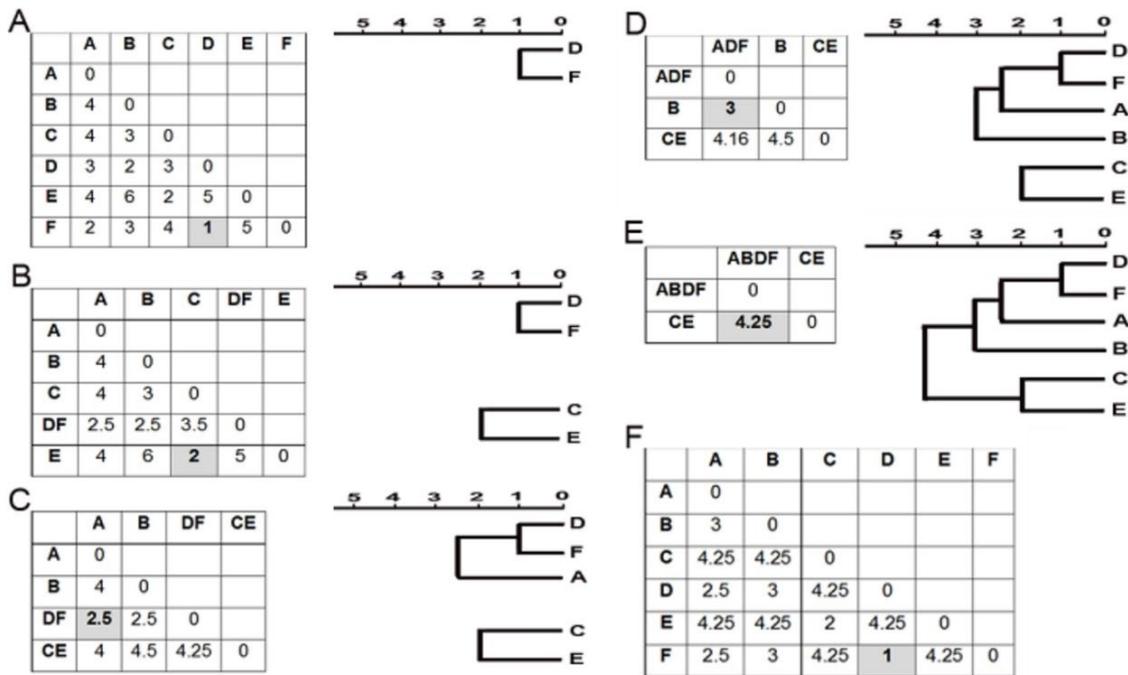
Las técnicas más utilizadas son jerárquicas, aglomerativas y secuenciales. Estas técnicas dan como resultado un fenograma (también llamado dendrograma ó cluster) (Figura 4.3), que es una representación gráfica en forma de árbol que muestra las relaciones jerárquicas entre las OTUs según sus valores de similitud. El método más sencillo es el del “grupo par” (*pair group*), en el cual solo puede ser admitida una OTU o un grupo de OTUs por nivel (Palacio *et al.*, 2020).

A continuación, se describe el método general (Figura 4.3):

- Se identifica el mayor valor de similitud en la MS, descartando la diagonal principal. Se identifican las dos OTUs que formarán el primer núcleo (grupo de dos OTUs) (Figura 4.3A). En algunos casos puede haber más de un par de OTUs con un valor máximo de similitud; en ese caso se elige uno, o se construyen los dos grupos por separado (Palacio *et al.*, 2020). Los valores de similitud se expresan en una escala ubicada generalmente sobre el fenograma que comenzará en 1 (si se está utilizando un coeficiente de correlación o de asociación) o en 0 (si se está utilizando un coeficiente de distancia) (Figura 4.3).
- Se construye una nueva MS derivada de la original, que al menos puede obtenerse por tres métodos diferentes: ligamiento simple, completo y promedio. En los ejemplos de este capítulo se utiliza el algoritmo UPGMA o técnica de ligamiento promedio no ponderado, que es la más utilizada, dado que produce una menor distorsión del fenograma con respecto a los valores de la matriz original. La MS derivada, presenta una columna y fila menos que la MS original, en la cual el núcleo ya ingresado al fenograma ocupará una sola fila y columna, y el valor correspondiente al casillero se calcula por UPGMA, siendo el promedio de los valores originales (Figura 4.3B).
- Luego se busca el próximo valor de mayor similitud o menor de distancia (Figura 4.3B) en la MS derivada. En las primeras etapas del proceso de agrupamiento, el hallazgo de este nuevo valor puede llevar a: (i) la formación de nuevos grupos (Figura 4.3B), (ii) la incorporación de una OTU a un grupo ya existente, o (iii) la fusión de los grupos preexistentes. Siempre se vuelve a la MS original para calcular los valores promedios que se incorporan a las siguientes MS derivadas (Figura 4.3C-E). Se repite el proceso hasta que todos los grupos estén unidos y se incluya la totalidad de las OTUs (Figura 4.3E).
- Al finalizar, se puede realizar una matriz cofenética que indica la distorsión del fenograma. Se construye una matriz derivada a partir de los valores de similitud de cada OTU observados en el fenograma (Figura 4.3F), y luego se aplica el coeficiente de correlación del Momento-Producto de Pearson, para comparar las matrices de similitud original y la cofenética. Si el valor

resultante se halla próximo a 1 indica una escasa distorsión, y si se acerca a 0 aumenta la distorsión (Sneath y Sokal, 1973; Crisci y López Armengol, 1983; Lanteri y Cigliano, 2006).

Figura 4.3. Pasos para la construcción de un fenograma de seis OTU, UPGMA.



Nota. Adaptado de Pasos para la construcción de un fenograma de seis OTU, de Lanteri y Cigliano 2006. (A-E) pasos según técnica de ligamiento promedio no ponderado (UPGMA); (F) Matriz cofenética.

2. Árboles basados en distancias genéticas

Estas técnicas dan como resultado un árbol sin raíz (aunque la misma puede ser incorporada a posteriori, ver capítulo 7). Y el método de construcción es similar al que se utiliza para construir fenogramas (Takezaki y Nei, 1993; Lemey *et al.*, 2009; Roch, 2010).

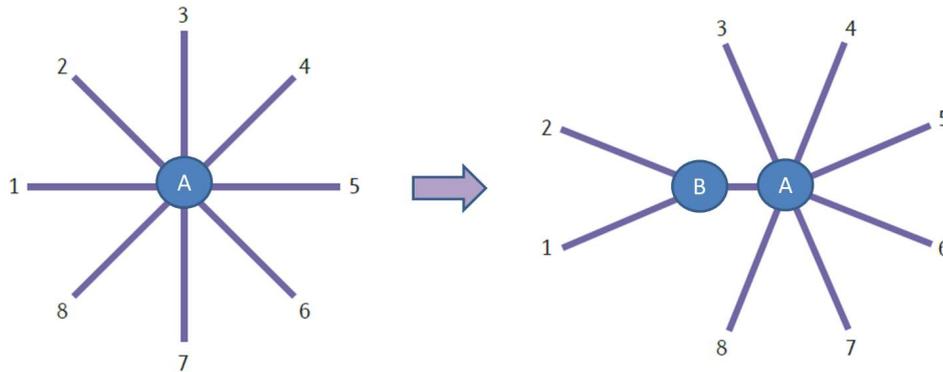
Uno de los métodos clásicos es el de **Neighbor-Joining (NJ)** que representa un árbol con longitudes de rama diferentes que reflejan las distancias evolutivas calculadas utilizando un modelo de sustitución nucleotídica *a priori* (e.g. Tamura *et al.*, 2004) (Figura 4.4).

Para construir el árbol hay que transformar la matriz de datos alineada en una matriz de distancias, aplicando un coeficiente de distancia genética, como el de Nei (1972). Luego *NJ* actúa como un método de agrupamiento o *clustering* en donde la distancia que existe entre el único nodo (punto) y cada *OTUs* comienza siendo igual (se representa a las *OTUs* unidas a un nodo en un árbol conocido como árbol estrella) (Figura 4.4A). Las *OTUs* que tengan menor distancia genética entre ellos se unirán en un nuevo nodo y surgirá una nueva rama entre este nodo y el resto del árbol estrella (Figura 4.4). Este proceso se repite secuencialmente, siempre de acuerdo con sus mínimas distancias, hasta incorporar a todas las *OTUs* y todas las distancias calculadas en la matriz de distancias. Las distancias genéticas se observan en el largo de las ramas y las

unidades correspondientes en una escala que representa el número acumulado de sustituciones de bases por sitio (Saitou y Nei, 1987).

Otra técnica utilizada para tratar datos de ADN es el método [Median-Joining](#) (un tipo de árbol de distancia), que agrupa de manera secuencial los taxones más similares (Bandelt *et al.*, 1999). El resultado de este análisis es un árbol no enraizado y aditivo.

Figura 4.4. Esquema de construcción del árbol de Neighbor Joining.



Nota. Adaptado de *The neighbour joining (NJ) algorithm*, de Yang y Rannala, 2012. El árbol ubicado a la izquierda es el árbol estrella con ocho taxones terminales (números). A partir de este árbol se construye un primer nodo con las OTUs que presentan menor distancia (1 y 2) surgiendo así una nueva rama (A-B) en el árbol ubicado a la derecha. Este proceso se itera hasta resolver todas las relaciones.

Estos árboles son empleados para estudios de naturaleza fenética (Cheema y Dicks, 2009), aunque también suelen ser utilizados para resolver aspectos filogenéticos (ver Capítulo 7).

Similitudes y diferencias entre las técnicas de agrupamiento descriptas

Los resultados obtenidos por el método UPGMA y por los métodos de árboles basados en distancia genética como *NJ* se representan como árboles. El método UPGMA no considera la tasa de evolución (la velocidad a la que una población adquiere y fija mutaciones) y el *NJ* la considera durante la construcción del árbol. El método UPGMA no presenta ramas con diferentes longitudes y el de *NJ* muestra diferentes longitudes de ramas indicando el número acumulado de sustituciones de pares de base (pb) por sitio.

Métodos de ordenación

Estos métodos permiten proyectar y representar las unidades de estudio (individuos, muestras, especies, etc.) como puntos en un espacio bi o tridimensional, donde cada dimensión representa un carácter (James y McCulloch, 1990). Asimismo, estos métodos reducen, sin gran pérdida de información, el número total de dimensiones (p), facilitando la representación de las OTUs y sus relaciones en función de los caracteres empleados (Palacio *et al.*, 2020).

1. Análisis de componentes principales (PCA, del inglés *Principal Component Analysis*)

Representa, según un modelo lineal, un conjunto de variables (caracteres) mediante un número reducido de variables hipotéticas denominadas componentes principales (*PC* del inglés *principal component*). Los *PC* no están correlacionados entre sí y, por lo tanto, se interpretan independientemente unos de otros. La contribución individual de cada carácter a cada *PC* esta expresada por el eigen-vector, aquellos caracteres que más contribuyan tendrán *eigen-vectores* con valores absolutos más elevados (Tabla 4.5). La sumatoria de las varianzas de todos los caracteres para un determinado *PC* se conoce como eigen-valor. La suma de todos los *eigen-valores* constituye la varianza total de la MBD original (Lanteri y Cigliano, 2006; Palacio *et al.*, 2020) (Tabla 4.6).

El primer *PC* (*PC1*) es el que contiene la mayor variabilidad, y de la variabilidad restante, el segundo (*PC2*) es el que incluye más información (Tabla 4.6). El tercer *PC* (*PC3*) posee la mayor variabilidad no contenida en los componentes anteriores. Así se continúa hasta que toda la variabilidad haya sido distribuida diferencialmente entre los componentes. Cada componente contiene información de todos los caracteres, pero en diferentes proporciones. Se reconocen tantos componentes como caracteres se consideren en el estudio. Los caracteres que están más correlacionados con los primeros componentes son los más importantes para explicar la variabilidad total de la MBD y presentan los *eigen-vectores* con valores absolutos más elevados (Kas-sambara, 2017). Las combinaciones más comunes para la elaboración de los gráficos bidimensionales son: *PC1* vs. *PC2*, *PC1* vs. *PC3* y *PC2* vs. *PC3*. Los gráficos tridimensionales se basan en los tres primeros componentes (Palacio *et al.*, 2020).

Tabla 4.5. *Eigen-vectores correspondientes a los tres primeros PC.*

	<i>Eigen-vector</i> : contribución individual de cada carácter a cada componente		
	PC1	PC2	PC3
Carácter 1	0,717	-0,482	0,262
Carácter 2	0,551	0,195	0,562
Carácter 3	0,007	0,462	0,759
Carácter 4	0,947	0,059	-0,132
Carácter 5	0,933	-0,128	-0,099
Carácter 6	0,840	-0,246	-0,193
Carácter 7	0,371	0,611	-0,261

Tabla 4.6. *Eigen-valores correspondientes a los PC I a III (= i), indicando porcentaje de variación explicada y porcentaje acumulado.*

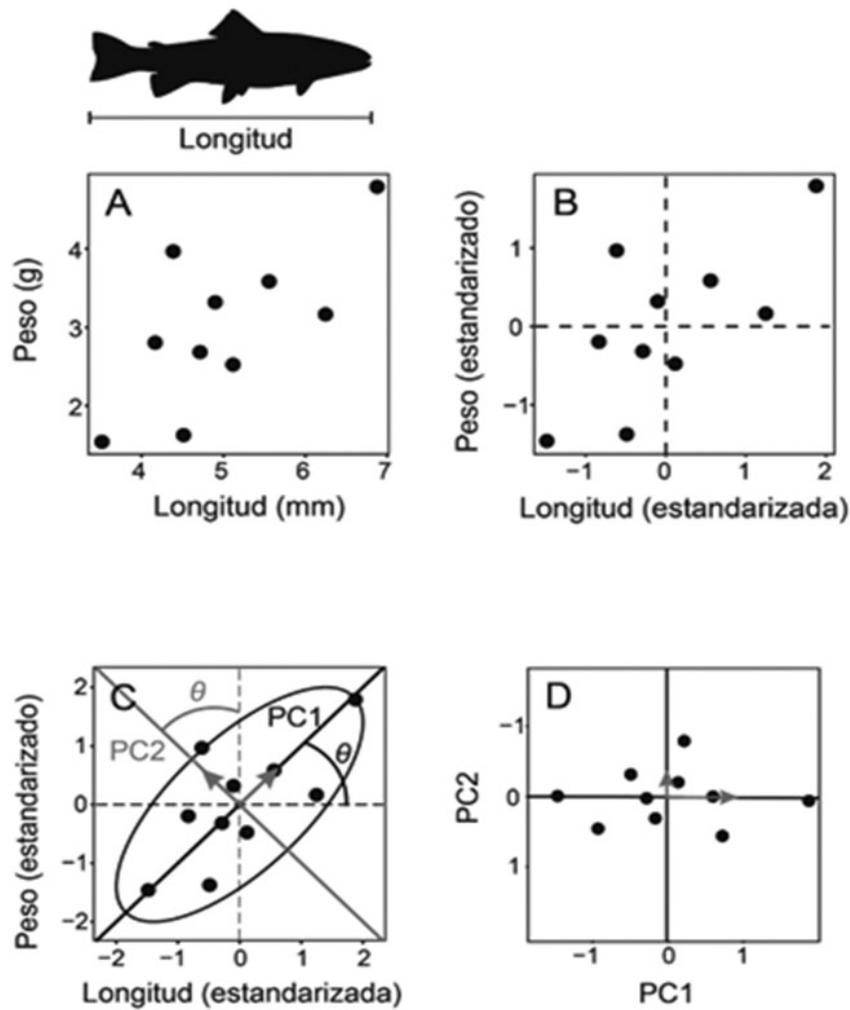
	<i>Eigen-valor</i>	% var. explicada	% acumulado
I	18,73	45,68	45,68
II	8,62	21,02	66,71
III	5,55	13,53	80,24

Nota. Adaptado de Tabla resumen del PCA para la MBD de especies de Bulnesia, de Palacio et al., 2020.

Se indican a continuación en términos generales los principios básicos de esta técnica. Para más detalles ver Jolliffe y Cadima (2016) y Husson *et al.*, (2017).

A modo de ejemplo, en la Figura 4.5 se observan a diez *OTUs* en el espacio bidimensional originado por dos caracteres (longitud y peso corporal), con una correlación positiva, a medida que aumenta la longitud también aumenta el peso (Figura 4.5A). Es necesario que ambos caracteres estén expresados en la misma unidad de medida, así es que es conveniente estandarizarlos. La consecuencia de la estandarización es que las *OTUs* están referidas ahora a un nuevo par de ejes ortogonales (perpendiculares entre sí) (Figura 4.5B). Cada *OTU* tendrá un nuevo par de coordenadas definidas en función de unidades de desvío estándar. Geométricamente, la disposición espacial de las *OTUs* para dos caracteres correlacionados es la de una nube elíptica. El eje de mayor variación coincide con el eje mayor de la elipse y corresponde al primer componente principal (Figura 4.5C). Para determinar la máxima variación entre las *OTUs* en una segunda dimensión perpendicular a la primera, el eje buscado coincidirá con el eje menor de la elipse. Este eje corresponde al segundo componente principal. Ambos componentes corresponden a una rotación de la ordenada y la abscisa un cierto ángulo θ . Los vectores que definen la ubicación y la dirección de los ejes mayor y menor son los *eigen-vectores* (Figura 4.5C). Por otra parte, los *eigen-valores*, reflejan la varianza de cada componente principal (mayor en el primero, menor en el segundo, y así sucesivamente). Si se analizaran tres caracteres, la elipse se transformaría en un elipsoide y el tercer componente principal estaría representado por el tercer eje del elipsoide, perpendicular a los dos primeros. Si el estudio incluye más de tres caracteres, se necesitarán dimensiones adicionales cuya representación geométrica no puede ser visualizada, pero de igual forma podrá aplicarse el tratamiento matemático. La contribución relativa de los primeros componentes principales representan el mayor porcentaje de variación total de la MBD. Se resuelve el análisis y se observan las *OTUs* en dos dimensiones (Figura 4.5D) (Palacio *et al.*, 2020).

Figura 4.5. Representación de OTUs en el espacio bidimensional.

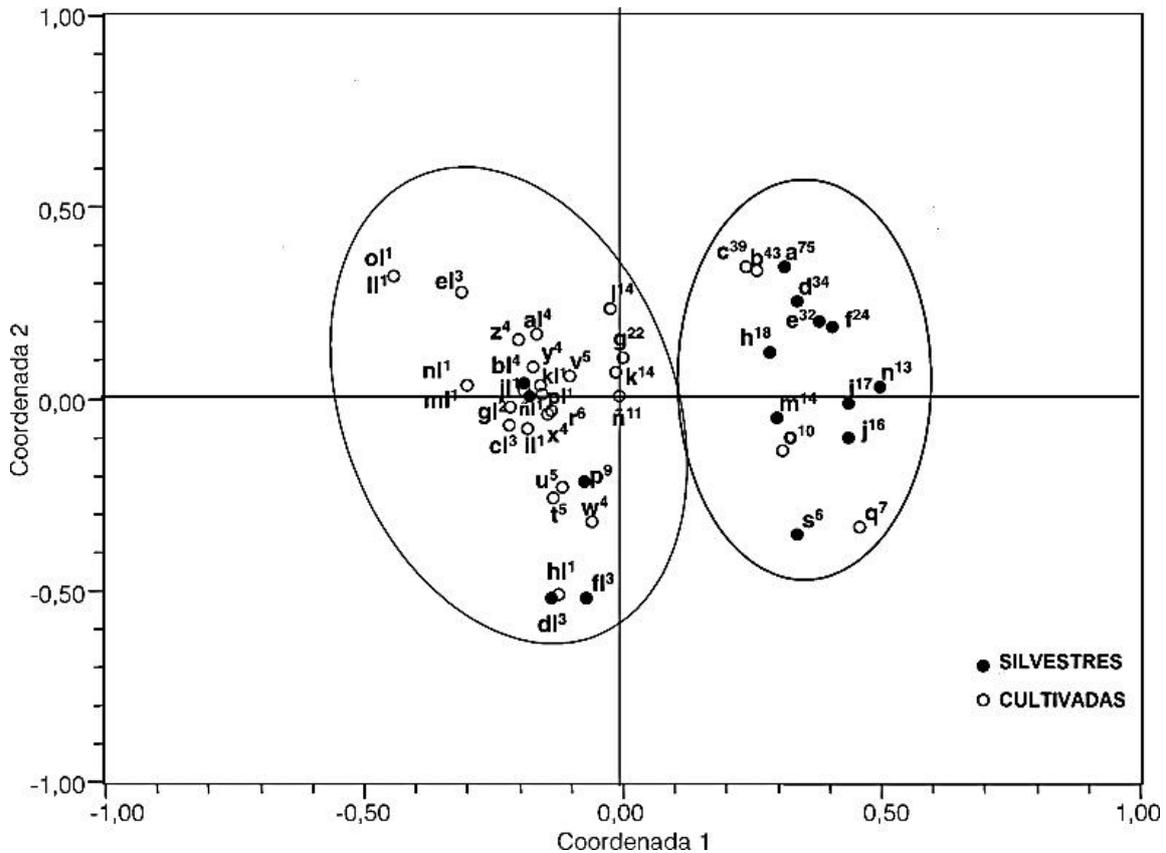


Nota. Adaptado de PCA, de Palacio et al., 2020. (A) Gráfico de dispersión de 10 individuos de una especie de pez y dos variables (longitud y peso); (B) estandarización de las variables a media 0 y varianza 1; (C) elipse que engloba las OTUs y componentes principales (PC) 1 y 2, que corresponden a la rotación de la ordenada y la abscisa un cierto ángulo θ , los vectores (flechas grises) representan los eigen-vectores del PCA; (D) resultado del PCA en dos dimensiones.

2. Análisis de Coordenadas Principales (PcoA del inglés *Principal Coordinate Analysis*)

También se trata de un análisis cuyo resultado se proyecta en un gráfico bidimensional (Figura 4.6). La dimensionalidad del espacio de coordenadas principales depende del número de caracteres y del coeficiente de distancia utilizado (Legendre y Legendre, 1998).

Figura 4.6. Ejemplo de grafico bidimensional de PCoA de especies silvestres y cultivadas de plantas de uso medicinal.



Nota. Adaptado de *Análisis de coordenadas principales de las especies utilizadas con respecto a la frecuencia de mención*, de Hernández et al., 2005.

3. Escalado Multidimensional No Métrico (NMDS del inglés *Nonmetric MultiDimensional Scaling*)

El objetivo del NMDS es explicar la información de la MBD original en unas pocas dimensiones, de manera de poder visualizarla e interpretarla. Esta técnica es recomendable cuando se trabaja con datos de distribución no normal o bien en escalas arbitrarias o discontinuas (McCune, 1991). Al ser no métrico, tiene la ventaja de poder aplicarse a una MS basada en diferentes tipos de coeficiente (Sneath y Sokal, 1973). El NMDS es un método iterativo de búsqueda para la ordenación y ubicación de OTUs con un mínimo *stress* en la configuración final. *Stress* es un concepto que se refiere a la magnitud de la desviación entre las distancias existentes entre los puntos en la matriz de distancia original y las distancias en el espacio reducido de ordenación (Palacio et al., 2020).

Similitudes y diferencias entre los métodos de ordenación descriptos

Todos los métodos de ordenación producen una representación de las *OTUs* en el espacio euclideo, preservando las distancias originales de la mejor forma posible. En todos los casos, cada dimensión representa un carácter, y se consideran tantas dimensiones como caracteres se encuentren en la MBD. Cada método emplea diferentes coeficientes y construyen diferentes matrices de similitud. En los tres métodos descriptos, las similitudes entre las *OTUs* están reflejadas en la posición en que se disponen en ese espacio, cuanto más próximas más similares

Dependiendo del tipo de datos, puede recomendarse uno u otro tipo de análisis. El *PCoA* no brinda información sobre la contribución de los caracteres a cada coordenada principal, para lo cual el *PCA* resulta de gran utilidad identificando los caracteres más discriminativos. A diferencia del *PCA* o *PCoA* que son métodos basados en *eigen-valores* y *eigen-vectores*, el *NMDS* no busca maximizar la variabilidad asociada a los ejes principales y posee la propiedad de preservar con mayor fidelidad las distancias pequeñas. En *NMDS* los ejes son arbitrarios, por lo que pueden ser rotados, centrados o invertidos.

7. Identificación de patrones de relaciones y grupos entre las *OTUs*, y elección de variables discriminatorias. Interpretación de los resultados.

En un [fenograma](#), las *OTUs* se relacionan entre sí formando grupos cada vez más inclusivos, cuyos valores de similitud se expresan en una escala ubicada generalmente en su parte superior. En el ejemplo de la figura 4.3 la escala comienza en 0, lo cual indica que se ha utilizado un coeficiente de distancia, representando el 0 la mayor similitud. En el caso de haberse empleado coeficientes de correlación o asociación, la escala comenzaría en 1. Así, cuanto más cerca al inicio de la escala se encuentre la unión entre las *OTUs*, mayor será su similitud o menor su distancia.

En la interpretación del fenograma se reconocen visualmente los grandes grupos, es decir los que se han originado a bajos niveles de similitud (ver figura 4.3, grupo DFAB y CE). Luego, se analizan dichos grupos separándolos en subgrupos hasta llegar a los grupos que presentan la máxima similitud entre las *OTUs*. En general se describen como núcleos cuando los grupos están formados por dos *OTUs*.

Los resultados de un [árbol basado en distancias genéticas](#) indican la distancia entre *OTUs* (largo de ramas) como reflejo de las mutaciones entre estas. En este sentido, mayor recorrido en las ramas de una *OTU* a la otra, indica mayor distancia genética. La escala indica la cantidad de cambios (mutaciones) entre las *OTUs*.

Los resultados del [PCA](#) se grafican sobre ejes cartesianos ortogonales que representan los *PC*, y delimitan un espacio bi o tridimensional según se utilicen dos o tres ejes, respectivamente. Las *OTUs* se sitúan dentro del espacio delimitado por los *PC* según los valores de sus coordenadas. Las relaciones entre las *OTUs* se establecen por su proximidad en el espacio delimitado por los *PC*, mayor proximidad, mayor similitud. Sin embargo, es válida la interpretación si la representación en el espacio de menor dimensión captura relativamente bien la información contenida en la MBD original (alto porcentaje de variación explicada).

Además, las *OTUs* y los caracteres pueden visualizarse simultáneamente en un gráfico denominado *biplot*. Las reglas de interpretación de un *biplot* son las siguientes: (1) las *OTUs* cercanas en el espacio tienen características similares; (2) una *OTU* que está cercana a un carácter tiene un alto valor para ese carácter; y (3) una *OTU* que se encuentra opuesta a un carácter tiene un bajo valor para ese carácter.

En el caso del *PCoA* también se observan las *OTUs* sobre ejes cartesianos ortogonales que representan las coordenadas principales, y delimitan un espacio bidimensional. Del mismo modo que en el *PCA*, las relaciones entre las *OTUs* se establecen por su proximidad en el espacio delimitado por las coordenadas, mayor proximidad, mayor similitud.

Para la interpretación de valores de *eigen-vectores* y *eigen-valores*, y porcentajes de variabilidad explicada en cada *PC* se da un ejemplo en las Tablas 4.5 y 4.6. En la tabla 4.5 se observa que el carácter que contribuye en mayor proporción al *PC1* es el 4, seguido por el 5 (mayores valores, correlacionado positivamente), y al *PC2* es el 7, seguido por el 1 (mayores valores, correlacionado positiva y negativamente). También se puede observar en la tabla 4.6 que el *PC1* es el que reúne el mayor porcentaje de variación (45,68%) y en los restantes componentes la variabilidad decrece en forma progresiva. Los *PC1* y *PC2* expresan un total del 66,71% de la variación total. Los tres primeros componentes un 80%, por lo cual son indicados para realizar el análisis, la gráfica y realizar las interpretaciones.

La interpretación de las *OTUs* en el espacio desde el análisis de *NMDS* es similar al *PCA* y *PCoA*, las relaciones entre las *OTUs* se establecen por su proximidad en el espacio delimitado por las coordenadas, mayor proximidad, mayor similitud.

Ventajas y desventajas de las técnicas multivariadas:

Las técnicas multivariadas se dividen en *análisis de agrupamiento* (permiten la búsqueda de grupos similares lo más homogéneos posible) y *métodos de ordenación* (buscan explicar la mayor parte de la variabilidad total existente en la muestra, en un número reducido de dimensiones).

Las técnicas de agrupamiento constituyen herramientas útiles para el análisis de datos continuos frecuentemente utilizados para el estudio de la variación infraespecífica y para resolver problemas a nivel de especie (Lanteri y Cigliano, 2006). Una desventaja de los fenogramas (*UPGMA*) es que en los resultados siempre aparecen grupos delimitados, aun cuando éstos no se ajusten a la estructura de los datos.

La ventaja del *PCA* y del *PCoA* es que representan gráficamente a las *OTUs* (*modo Q*) y/o a los caracteres (*modo R*) en simultáneo. A diferencia de los análisis de agrupamiento, la mayoría de los métodos de ordenación no trazan límites en el espacio que separen a los grupos, esta tarea corresponde al investigador.

La aplicación de diferentes técnicas puede conducir a decisiones taxonómicas distintas. Así es que estudios donde se aplican, por ejemplo, componentes principales, coordenadas principales o escalas multidimensionales no métricas a los mismos datos pueden producir distintos resultados (Moss, 1968; Arce *et al.*, 2009).

Morfometría Geométrica

La geometría de la estructura se puede capturar siguiendo distintos métodos de análisis: coordenadas de Landmarks (puntos de referencia) y coordenadas de contornos (Bookstein, 1991; Rohlf, 1996; Marcus *et al.*, 1996; Cardini, 2019) que básicamente, siguen tres pasos esenciales resumidos en la Figura 4.7: 1) obtención de los datos, 2) obtención de la información que describe la forma y 3) análisis exploratorios y confirmatorios de covariación de la forma con factores causales (Toro Ibacache *et al.*, 2010). Estos datos de coordenadas pueden ser obtenidos con diferentes procedimientos, como cámaras digitales, tabletas de digitalización, *scanners*, etc. y permiten estudiar la forma de los organismos independientemente del tamaño en dos o tres dimensiones (Bookstein, 1991). Para conocer más detalle de estos procedimientos remitirse a Bookstein (1991), Lanteri y Cigliano (2006) y Toro Ibacache *et al.* (2010).

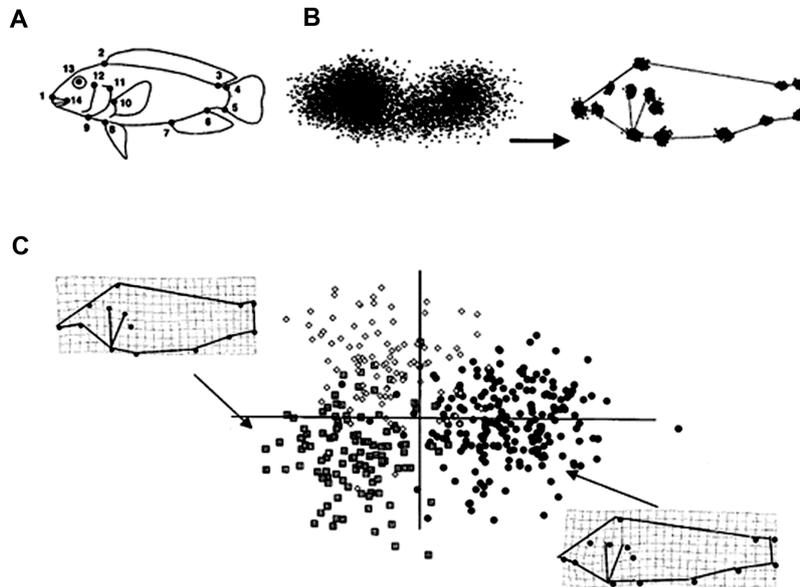
1) Obtención de datos (Figura 4.7A). Se reconocen tres tipos de landmarks: (a) yuxtaposición de tejidos que corresponden a puntos homólogos; (b) máxima curvatura o procesos morfológicos locales; y (c) puntos extremos o de inflexión que corresponden a *landmarks* geométricos o funcionales. La elección de los *landmarks* depende del tipo de objetivo planteado, y pueden estar definidos por dos coordenadas, en un espacio bidimensional (2D), o por tres coordenadas en un espacio tridimensional (3D) (Lanteri y Cigliano, 2006). Representar un organismo únicamente por landmarks deja de lado algunos aspectos de la curvatura, así es que el término *semi-landmarks* fue utilizado por Bookstein (1997) para referirse a una serie de puntos que se encuentran a lo largo de una curva (coordenadas de contornos).

2) Obtención de la información que describe la forma (Figura 4.7B). Se realizan una serie de procedimientos geométricos y estadísticos basados en la forma de una estructura, definida por su configuración de *landmarks*. Los mismos, analizan la forma, libre del efecto de tamaño, rotación y translación, considerando distintos “espacios morfométricos”, constructos geométricos y matriciales que permiten la obtención de los elementos que, finalmente, diferencian una forma de otra. Los procedimientos estadísticos incluyen medias y varianzas, coordenadas de la forma, métodos de superposición (*procrustes*), *thin plate spline* (TPS) y análisis de Fourier que permiten obtener resultados preliminares como tablas, gráficos y coeficientes (Lanteri y Cigliano, 2006; Toro Ibacache *et al.*, 2010).

3) Análisis exploratorios y confirmatorios de covariación, aplicación de método de ordenación (Figura 4.7C). La información que describe la forma en estudio puede ser estudiada de diversas maneras: *e.g.* PCA, CVA (del inglés *canonical variant analysis*), análisis de deformación relativa (*relative warps análisis*), MANDEVA (análisis multivariado de la varianza) y varios test no paramétricos. Así, las coordenadas permiten probar las diferencias entre formas promedio, el grado de correlación y covariación de la forma con el tamaño, la covariación de la forma con factores externos, la covariación entre distintas formas a través del análisis parcial de cuadrados mínimos

(*partial least squares analysis*), el estudio de la variación de formas simétricas y pruebas de la presencia de los distintos tipos de asimetría (Toro Ibacache *et al.*, 2010).

Figura 4.7. Representación gráfica resumida de las etapas de un estudio de morfometría geométrica.

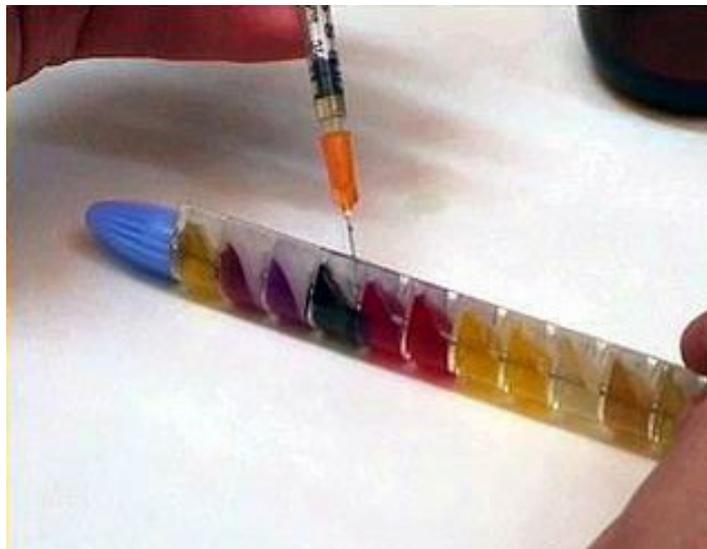


Nota. Adaptado de *Graphical representation of the four-step morphometric protocol*, de Adams *et al.*, 2004. A. Obtención de datos. B. Ejemplo de Métodos de superposición (antes y después del empleo de un Análisis Procrustes Generalizado). C. Aplicación del método de ordenación (CVA).

Ejercitaciones

Ejercicio 1

Figura 4.8. SEQ Ilustración * ARABIC Cultivo de bacterias.



Para la identificación de microorganismos se utilizan pruebas bioquímicas que determinan la actividad de una vía metabólica a partir de un sustrato que se incorpora en un medio de cultivo y que las bacterias al crecer incorporan o no (Figura 4.8). A partir de los resultados de las reacciones se pueden identificar los distintos organismos presentes. Estas variables cualitativas consideran la presencia o ausencia de un determinado carácter; por ejemplo, si una bacteria fermenta o no un determinado hidrato de carbono. Con el objetivo de evaluar la similitud entre distintos grupos de organismos, se establecen coeficientes que dan cuenta de los grados de similitud entre las unidades taxonómicas, en este caso cepas de microorganismos.

Se desea analizar dos cepas de *Escherichia coli* (I y II) aisladas de una muestra de carne de un frigorífico con el objetivo de evaluar si alguna de ellas corresponde a una cepa enterohemorrágica, cuyas toxinas pueden ocasionar síndrome urémico hemolítico. Para esto se realiza un análisis de agrupamiento utilizando dos cepas de referencia (*E. coli* 1- cepa no patógena y *E. coli* 2- cepa enterohemorrágica). Utilizando un sistema de identificación comercial, y se obtienen los siguientes resultados:

Tabla 4.7. Matriz básica de datos.

Pruebas realizadas	Fermentación de										VP	RM	Leci	DNasa	Catalasa	COX	Red. de NO ₃ ⁻
	Glu	Sac	Lac	FruC	Man	Ara	Gal	Ram	Cel								
Cepas (OTUs)	I	+	+	+	+	-	-	-	+	+	+	+	+	-	-	-	-
	II	+	+	-	+	+	+	-	+	-	+	-	+	+	+	+	-
	<i>E. coli</i> 1	+	+	+	+	+	+	+	+	+	-	+	-	-	-	-	-
	<i>E. coli</i> 2	+	-	-	-	-	+	-	-	-	-	-	+	+	+	+	+

- Construya la matriz de similitud utilizando el coeficiente *Simple Matching*.
- A partir de la matriz obtenida en el inciso anterior, construya un fenograma, utilizando el método del ligamiento promedio (*UPGMA*).
- ¿Considera que las cepas halladas en la muestra puedan llegar a ser perjudiciales para el consumidor? ¿Por qué?
- Construya una matriz presencia/ ausencia en Excel (reemplazando los símbolos – y + por 0 y 1 respectivamente), y mediante el uso del *software* Past (ver tutorial [Anexo 1](#)) realice los puntos a y b utilizando otro coeficiente de similitud (*Jaccard o Manhattan Distance*).
- Compare los resultados con aquellos obtenidos mediante el método manual. ¿Cuáles son las similitudes y diferencias observadas?
-

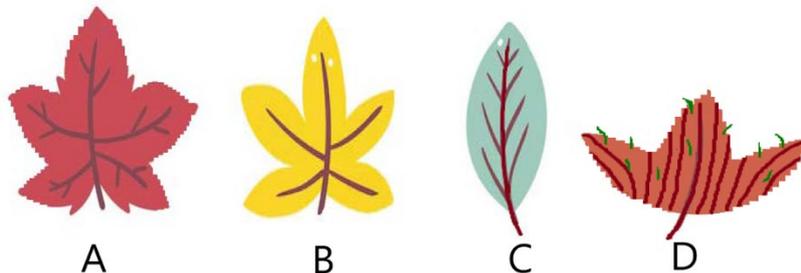
Ejercicio 2

Con el objetivo de realizar un análisis exploratorio en relación con la similitud morfológica entre cuatro especies hipotéticas de plantas vasculares (*OTUs*-Figura 4.9) se consideraron ocho caracteres morfológicos doble estado de la hoja: 1. lámina (simple/compuesta), 2. forma (palmada/ lanceolada), 3. tipo de margen (dentado/liso), 4. tamaño del peciolo ($\leq 1\text{cm}$ / $>1,1\text{cm}$), 5. patrón de coloración (variegado/uniforme), 6. relación largo/ancho de la lámina (dos veces más largo que ancho/ menos de dos veces más largo que ancho), 7. venación (reticulada/paralela), 8. pelos glandulares (ausente/presente) y se generó la siguiente matriz de datos (Tabla 4.8).

Tabla 4.8. Matriz de *OTUs* x caracteres de plantas vasculares hipotéticas.

	1	2	3	4	5	6	7	8
A	0	0	0	0	1	0	0	0
B	0	0	1	0	1	0	0	0
C	0	1	1	1	1	0	0	0
D	0	0	0	1	1	1	1	1

Figura 4.9. Hojas pertenecientes a cuatro especies de plantas vasculares hipotéticas.



A partir de la matriz de datos (Tabla 4.8).

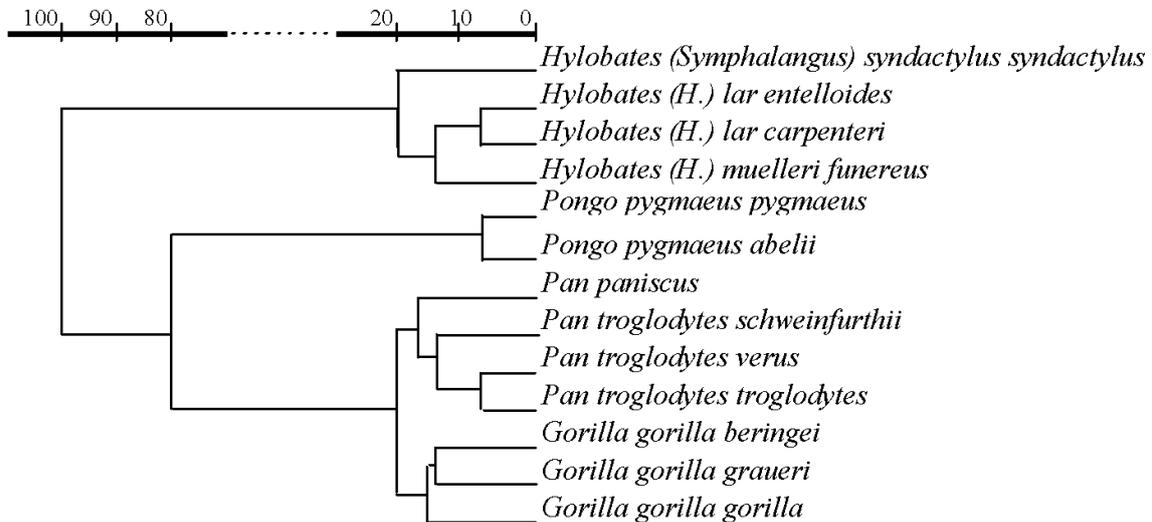
- Obtenga una matriz de similitud, aplicando un coeficiente de distancia y otro de asociación.
- A partir de alguna de las matrices obtenidas en el punto anterior, construya un fenograma entre *OTUs* mediante el método de ligamiento *UPGMA*.
- Describa el fenograma obtenido.

Ejercicio 3

El fenograma de la Figura 4.10 pertenece a 13 taxones específicos y subespecíficos de hominoideos no humanos vivientes. *Hylobates* (gibón) y *Pongo* (orangután) son asiáticos, *Gorilla* (gorila) y *Pan* (chimpancé) son africanos. La matriz de datos que dio origen al fenograma fue

obtenida a partir de 19 *landmarks* del complejo maxilo-facial, analizados por el método de *relative warps* (= deformaciones relativas).

Figura 4.10. Fenograma obtenido para 13 taxones de hominoideos no humanos vivientes



Nota. Adaptado de UPGMA phenogram, de Guy et al., 2003.

- Describe las relaciones de similitud expresadas en el fenograma.
- ¿Qué tipo de coeficiente de similitud se usó para construir el fenograma? Justifique.
- ¿Existe correlación entre los principales agrupamientos y la distribución de los hominoideos asiáticos y africanos? Explique

A partir de nuevas fuentes de evidencia se generó la siguiente matriz de datos, con 15 caracteres adicionales relevados para los mismos 13 taxones (Tabla 4.9).

- Copie en Excel la matriz (Tabla 4.9) y empleando el PAST (ver tutorial [Anexo 1](#)), realice un análisis para obtener un fenograma (Análisis de *Cluster*, utilizando UPGMA y empleando un coeficiente de asociación -*Jaccard*- y uno de distancia -*Manhattan Distance*). Describa los resultados de estos nuevos análisis, explicando si son congruentes con los observados en la Figura 4.10.

Realice un *PCA* (ver tutorial [Anexo 1](#)) considerando la misma matriz y responda:

- ¿Se observan grupos similares a los conformados por las OTUs en el fenograma?
- Mencione cuáles son los 3 caracteres que contribuyen en mayor medida a la variabilidad de los tres primeros componentes.

Tabla 4.9. Matriz de OTUS x caracteres de los hominoideos estudiados.

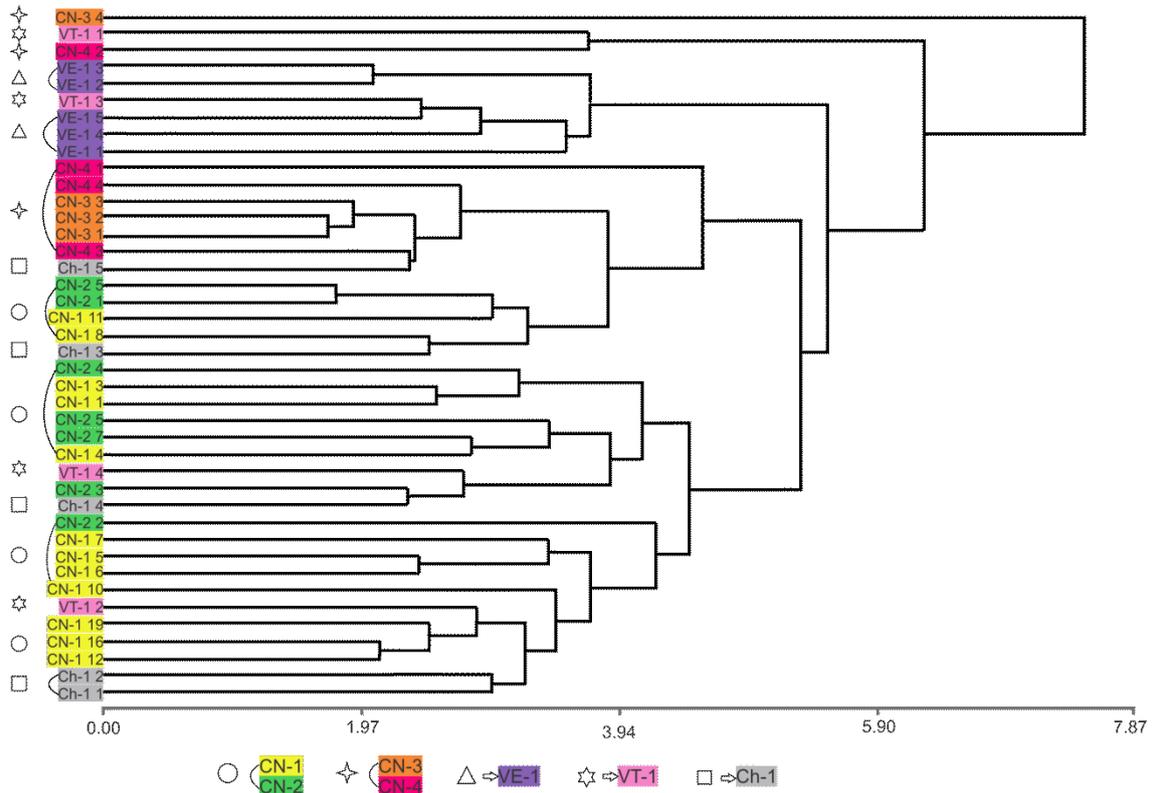
	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
<i>H. syndactylus syndactylus</i>	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
<i>H. entelloides</i>	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
<i>H. carpenteri</i>	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
<i>H. muelleri funereus</i>	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
<i>P. pygmaeus pygmaeus</i>	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
<i>P. abelti</i>	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
<i>P. paniscus</i>	0	1	0	0	0	0	1	1	0	1	1	0	0	0	0
<i>P. troglodytes schweinfurthii</i>	0	1	0	0	0	0	1	1	0	1	1	1	1	0	0
<i>P. troglodytes verus</i>	0	1	0	0	0	0	1	1	0	1	1	1	1	1	0
<i>P. troglodytes troglodytes</i>	0	1	0	0	0	0	1	1	0	1	1	1	1	1	0
<i>G. gorilla beringei</i>	0	1	0	0	0	0	1	1	1	0	0	0	0	0	1
<i>G. gorilla graueri</i>	0	1	0	0	0	0	1	1	1	0	0	0	0	0	1
<i>G. gorilla gorilla</i>	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0

Ejercicio 4

El uso de especies que son parientes silvestres de cultivos (PSC) ayuda a asegurar la existencia de una adecuada diversidad en el patrimonio genético de un cultivo en particular, dado que se relacionan genéticamente con especies de importancia socioeconómica, alimentos, cultivos forrajeros, plantas medicinales, condimentos y especies forestales. Así, los PSC son recursos fundamentales para abordar las necesidades de seguridad alimentaria, proporcionando diversidad genética para el mejoramiento de la productividad y la sostenibilidad de los sistemas agrícolas (Jansky *et al.*, 2012). La especie *Solanum brevicaule* Bitter se distribuye al oeste de Sudamérica, desde el sur de Perú y Bolivia hasta el noroeste de Argentina y es una PSC de *Solanum tuberosum* L. (papa cultivada).

Se analizaron los genotipos de siete poblaciones de *S. brevicaule* (CN1, CN2, CN3, CN4, VE, VT y Ch) halladas en cinco unidades de vegetación diferentes (Cerro Negro (CN), Valle Encantado (VE), Valle de Tronco (VT) y Churcal (CH)) emplazadas desde los 2819 a 3926 m.s.n.m, obteniendo el siguiente fenograma.

Figura 4.11. Fenograma de genotipos de *Solanum brevicaule*.



Unidades de vegetación: CN=Cerro Negro (1 y 2 Arbustal de transición Monte-Puna, ○), CN= Cerro Negro (3 y 4 Arbustal Puneño con vegas de altura, ✧), VE=Valle Encantado (Pastizal de neblina yungueño, △), VT=Valle de Tonco (Arbustal de transición Monte-Puna, ☆), Ch=Churcal (Bosque de *Prosopis ferox*, □)

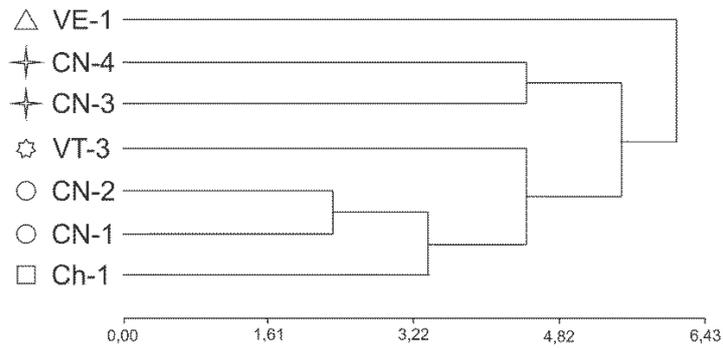
Nota. Adaptado de Fenograma de genotipos de *Solanum brevicaule* aplicando la técnica de agrupamiento UPGMA y la distancia Euclidiana Promedio, indicando con símbolos, acrónimos y colores las poblaciones a las que pertenecen y las unidades de vegetación en las que se hallan como sigue en cuadro, de Peña Cervera, 2020.

A partir de la Figura 4.11.

- a. ¿Observa agrupamientos de las OTUs (genotipos) en correspondencia con cada unidad de vegetación?
- b. A partir de los resultados observados ¿Es posible inferir la existencia de variabilidad intrapoblacional?

Por otro lado, se realizó un análisis de la variabilidad morfológica de las 7 poblaciones de *Solanum brevicaule* (Figura 4.12).

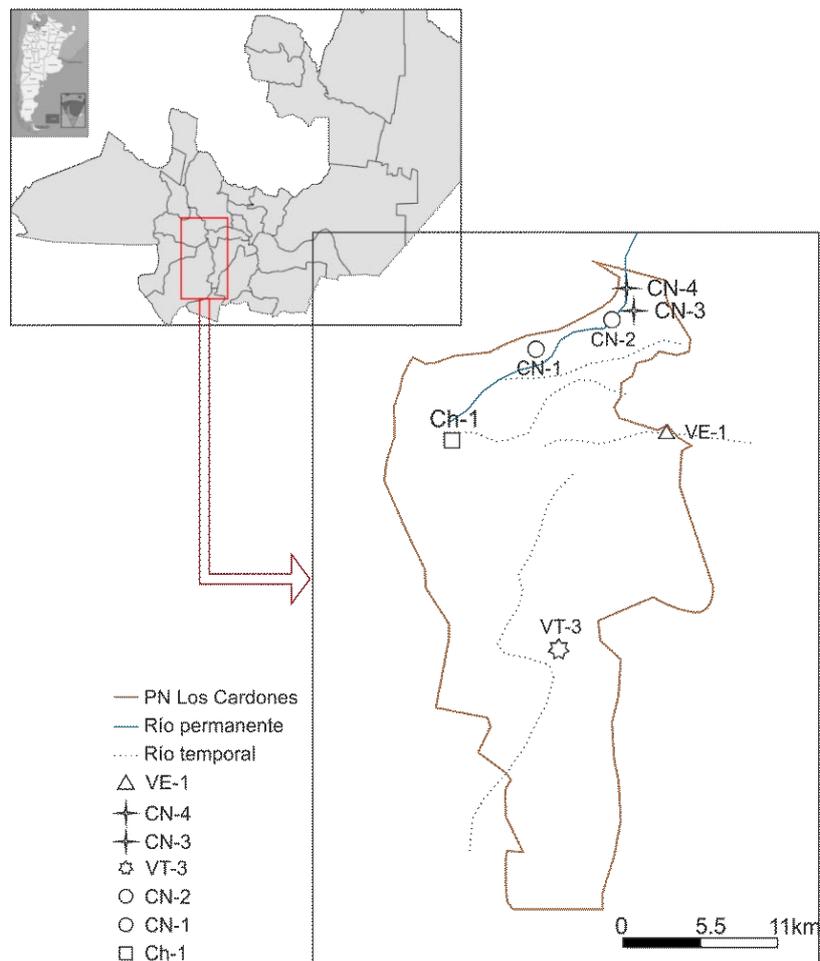
Figura 4.12. Fenograma de variabilidad morfológica de poblaciones de *Solanum brevicaulle* aplicando la técnica de agrupamiento UPGMA y la distancia Euclidiana Promedio.



Nota. Adaptado de Variabilidad morfológica en poblaciones de *Solanum brevicaulle* del Parque Nacional Los Cardones, de Peña Cervera, 2020.

c. Describa el fenograma y explique que indica la escala que aparece en la Figura 4.12. ¿Encuentra una relación entre los grupos/núcleos y las unidades de vegetación (ver en la zona inferior de la figura 4.11 la descripción de estas unidades)?

Figura 4.13. Distribución geográfica de *Solanum brevicaulle*.



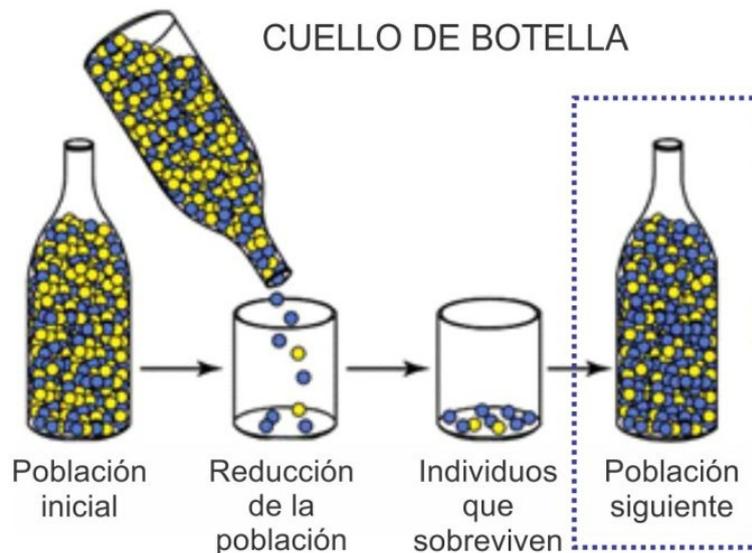
Nota. Adaptado de Distribución geográfica de las poblaciones de *Solanum brevicaulle* en el Parque Nacional Los Cardones, Salta, de Peña Cervera, 2020.

- d. ¿Encuentra una asociación entre los grupos/núcleos de la Figura 4.12 y la distribución geográfica de la Figura 4.13?
- e. Con el fin de establecer el ajuste de los datos, se calculó el coeficiente de correlación cofenético, el cual resultó ser de 0,826. ¿Qué indica el valor obtenido?

Ejercicio 5

Las planarias terrestres (Platyhelminthes, Tricladida) se consideran un buen modelo para estudios paleoclimáticos ya que, por su baja capacidad de dispersión y alta dependencia del bosque que habitan, los cambios producidos en el ambiente dejan huellas en su carga genética. Álvarez-Presas *et al.* (2011) examinaron poblaciones de dos especies de planarias, que habitan en el Bosque Atlántico en el sur de Brasil, a fin de contrastar si dichas poblaciones fueron afectadas por las glaciaciones mostrando huellas de “efecto cuello de botella” (Figura 4.14).

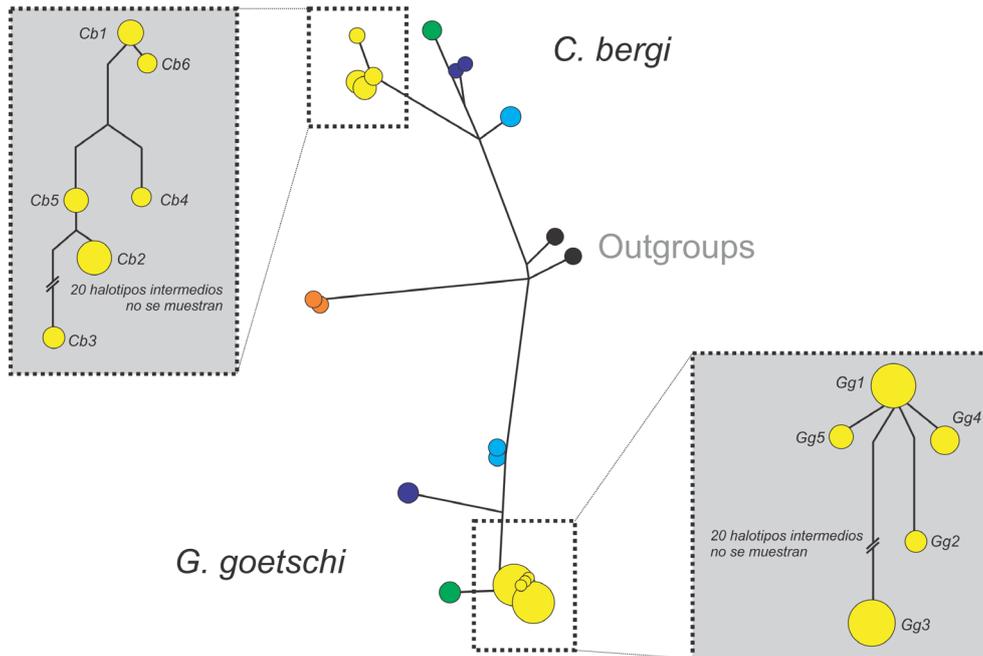
Figura 4.14. Esquema representativo de una población resultante del efecto cuello de botella.



Nota. Adaptado de Anjile, Public domain, via Wikimedia Commons.

El supuesto básico sobre el que trabajaron es que las poblaciones afectadas por las glaciaciones quedaron restringidas a refugios y muestran huellas de “cuello de botella” y por lo tanto baja diversidad nucleotídica, mientras que las que no se vieron afectadas muestran una alta diversidad genética. Para ello analizaron, entre otros genes, el mitocondrial COI de *Cephaloflexa bergi* Graff, 1899 y *Geoplana goetschi* Riester, 1938 principalmente del parque *Serra da Bocaina*. Entre otros resultados obtuvieron un árbol de *median joining* (Figura 4.15).

Figura 4.15



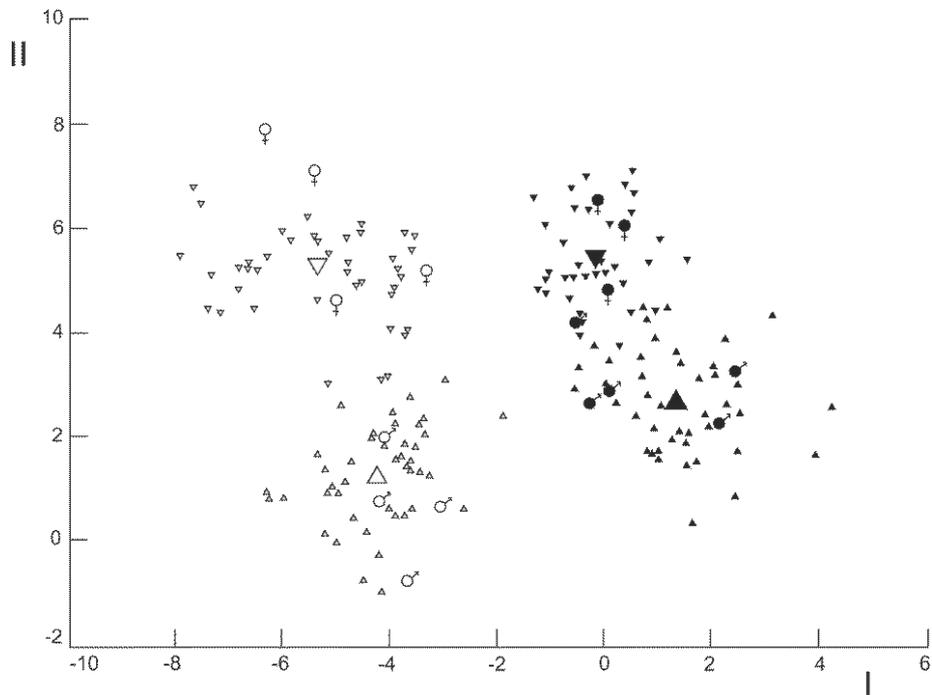
Nota. Adaptado de Median joining network based on COI sequences, de Álvarez-Presas et al., 2011. Cada círculo representa un haplotipo y el tamaño de cada círculo su frecuencia relativa. Los colores de los círculos representan las distintas localidades, de norte a sur: amarillo (dentro de los recuadros): Serra da Bocaina; verde: Estación Biológica de Boracéia; celeste: Parque Estadual de Intervales; azul: Parque Nacional de Saint-Hilaire/Lange.

- Realice la descripción del árbol sin raíz.
- ¿Considera que las poblaciones de *Serra da Bocaina* fueron afectadas por las glaciaciones del pleistoceno?, ¿por qué?

Ejercicio 6

Campbell y Mahon (1974) utilizaron un análisis multivariado para estudiar la morfología de dos especies de cangrejos (cangrejo azul y cangrejo naranja) del género *Leptograpsus*. Se colectaron individuos machos y hembras de ambas especies. De cada individuo se midieron los siguientes caracteres: ancho de la región frontal del caparazón, ancho de la región posterior del caparazón, largo del caparazón, ancho máximo del caparazón y alto total del cuerpo (Tabla 4.10). A partir de estos datos se realizó un PCA, obteniendo el siguiente resultado (Figura 4.16).

Figura 4.16. Análisis multivariado del género *Leptograpsus*.



Nota. Adaptado de *Discriminación de cangrejos*, de Campbell y Mahon, 1974. Triángulos claros= cangrejo azul; triángulos negros= cangrejo naranja; triángulos hacia arriba = machos; triángulos hacia abajo= hembras. Porcentaje de la varianza explicada: I = 69%, II = 29%.

Tabla 4.10. Eigen vectores calculados para cada carácter para los PC I y II.

CARACTERES	I	II
Ancho de la región frontal del caparazón	0,58	0,31
Ancho de la región posterior del caparazón	0,17	0,81
Largo del caparazón	0,06	0,65
Ancho máximo del caparazón	0,12	0,22
Alto total del cuerpo	0,76	-0,12

Observando los resultados obtenidos en la Figura 4.16 y la Tabla 4.10.

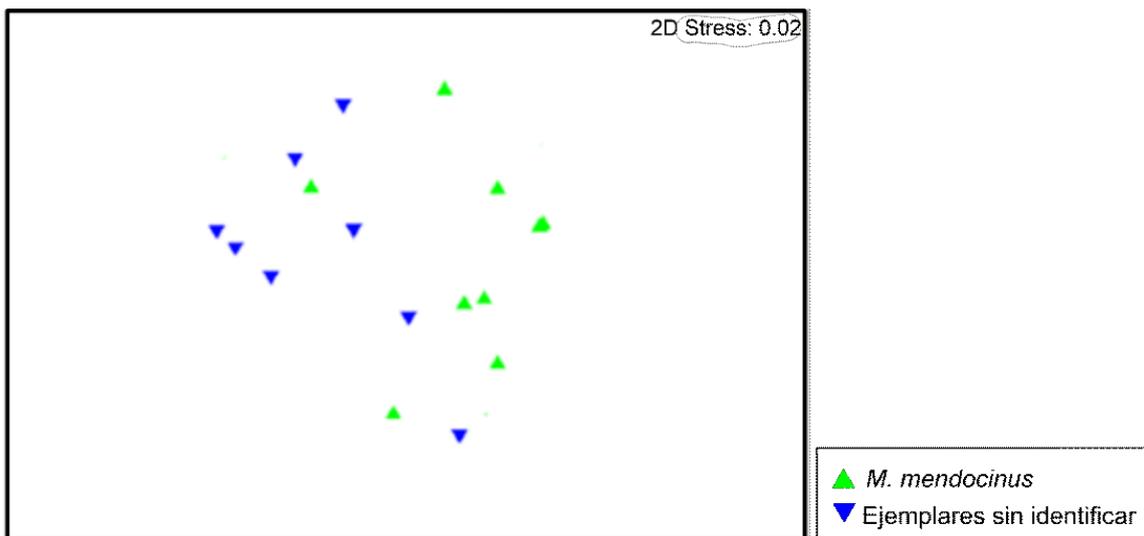
- ¿Se pueden distinguir ambas especies? Justifique.
- ¿Se pueden distinguir machos de hembras? Justifique.
- ¿Qué caracteres son los que más contribuyen para diferenciar a las dos especies? ¿Y a los sexos?
- ¿Cuál es la variabilidad total explicada por ambos componentes?

Ejercicio 7

En las lagunas asociadas a la cuenca del Río Salado (Provincia de Buenos Aires) existen al menos tres especies de copépodos del Orden Cyclopoida. Recientemente se han encontrado en una de estas lagunas (La Salada de Monasterio, partido de Lezama), numerosos ejemplares de estos microcrustáceos que inicialmente no han podido ser identificados a nivel específico, pero que coinciden en muchas características diagnósticas con la especie *Metacyclops mendocinus*.

En un análisis exploratorio de caracteres fisiológicos y de comportamiento empleando un *NMDS* (Escalado multidimensional no métrico), se observa la siguiente distribución de especímenes macho de *M. mendocinus* y especímenes macho de los copépodos no identificados a nivel específico, procedentes de la misma laguna (Figura 4.17).

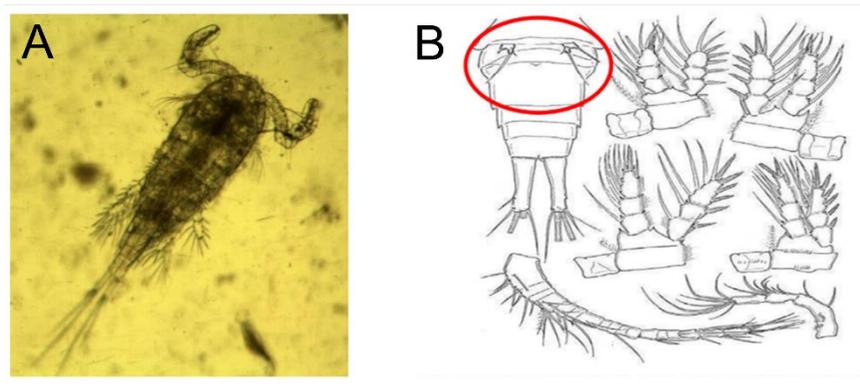
Figura 4.17. *NMDS* realizado utilizando caracteres fisiológicos y de comportamiento de especímenes macho de *M. mendocinus* y copépodos no identificados procedentes de La Salada de Monasterio.



- Describe la figura considerando los agrupamientos que se observan en la misma.
- ¿Considera que los ejemplares sin identificación pertenecen a *M. mendocinus*? Justifique.

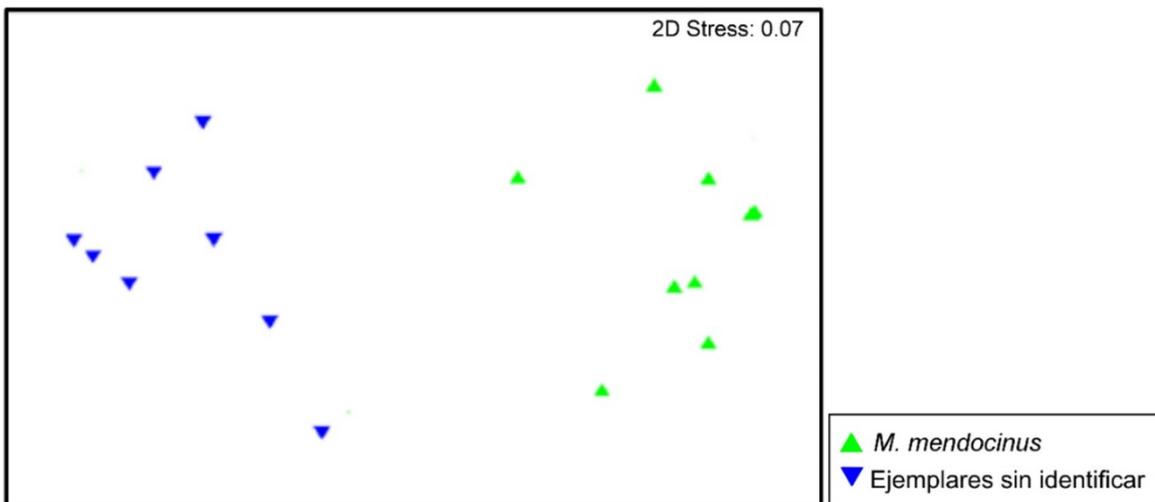
En otro análisis de *NMDS*, se consideraron sólo caracteres morfológicos externos, entre éstos, la distribución de las espinas en las patas nadadoras y en la quinta pata torácica (caracteres diagnósticos a nivel específico) (Figuras 4.18 y 4.19).

Figura 4.18. Copépedo *Cyclopoida*.



Nota. A) Fotografía de un copépedo *Cyclopoida* tomada por García de Souza. B) Adaptado de Ilustración que indica el detalle de la ubicación de las quintas patas torácicas, de Mercado-Salas y Suárez-Morales, 2012.

Figura 4.19. NMDS realizado utilizando caracteres morfológicos de especímenes macho de *M. mendocinus* y copépidos no identificados procedentes de La Salada de Monasterio.

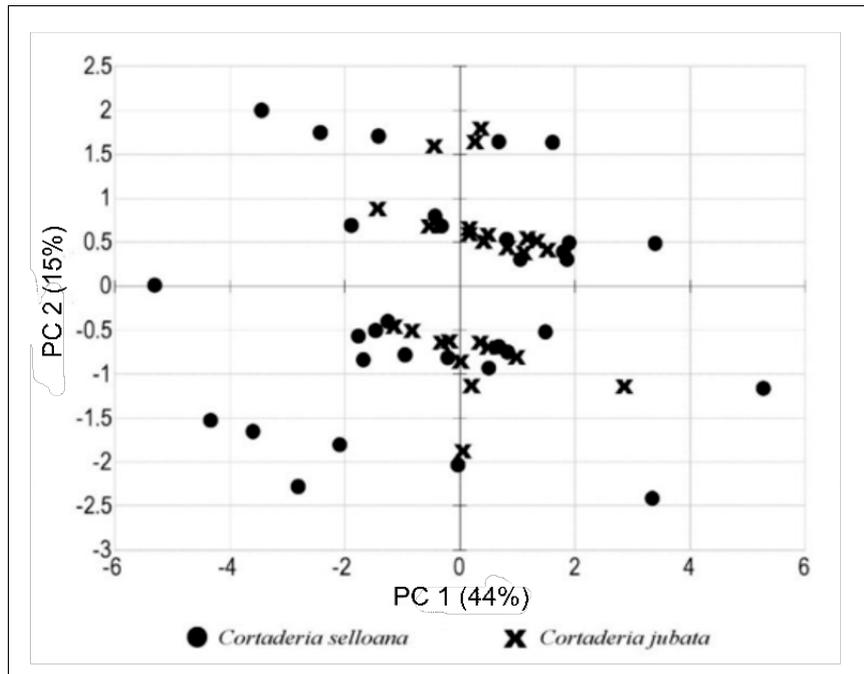


- c. Considerando esta nueva evidencia, ¿se modifica la decisión taxonómica tomada en el punto b. Justifique.

Ejercicio 8

El género *Cortaderia* (Poaceae) incluye 19 especies sudamericanas separadas en cuatro secciones. Se estudiaron 2 especies incluidas en la sección *Cortaderia* sect. *Cortaderia*: *Cortaderia selloana* (Schult. y Schult.f. 1827) Asch. y Graebn 1900 y *Cortaderia jubata* (Lemoine 1878) Stapf. 1898. A partir de caracteres tradicionalmente utilizados para delimitar las especies (morfológicos de estructuras reproductivas), se realizó un análisis de coordenadas principales (PCoA) (Figura 4.20).

Figura 4.20. Resultado del análisis de coordenadas principales (PCoA) que incluye 27 ejemplares de *Cortaderia jubata* y 32 de *C. selloana*.



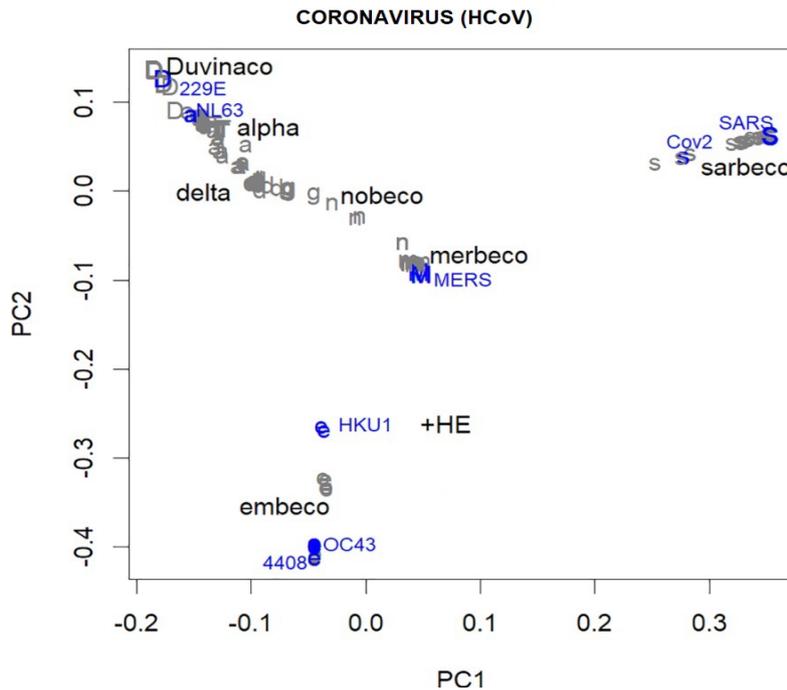
Nota. Adaptado de Análisis de componentes principales (ACP), representación en los ejes I y II del ordenamiento de ejemplares de *Cortaderia jubata* y de *C. selloana*, de Testoni y Villamil, 2014.

- ¿Cuál es la variabilidad total explicada por ambos componentes de la Figura 4.20?
- En relación con los resultados obtenidos ¿cuál es el estatus taxonómico de *C. selloana* y *C. jubata*?
- Considerando las conclusiones a las que llego en el inciso anterior ¿debería realizar alguna otra acción en su rol como taxónomo?

Ejercicio 9

Los coronavirus (HCoV) causan epidemias y pandemias anuales, y difieren en cuanto a hospedadores, infectividad, capacidad de propagación y sintomatología. Diferentes investigaciones involucran estudios de hospedadores originales (grupos animales) y su transmisión a humanos (zoonosis). Con el fin de explorar algunas similitudes entre sus características, se examinaron las secuencias de ARN del genoma de diferentes subclases de HCoV, mediante un PCA (Figura 4.21).

Figura 4.21. PCoA de diferentes subclases de HCoV, con sus hospedadores.



Nota. Adaptado de *Classes of coronaviruses separated into PC1 and 2*, de Konishi, 2020. Alpha: alfacoronavirus (murciélagos); delta: coronavirus delta (aves); Duvinaco: Duvinacovirus (humano) resfrío común; embeco (Embecovirus (mamíferos)); merbeco (Merbecovirus (murciélagos)); nobeco (Nobecovirus (murciélagos)); sarbeco (sarbecovirus (murciélagos y otros mamíferos)); Azul: muestras humanas representadas con distintos códigos.

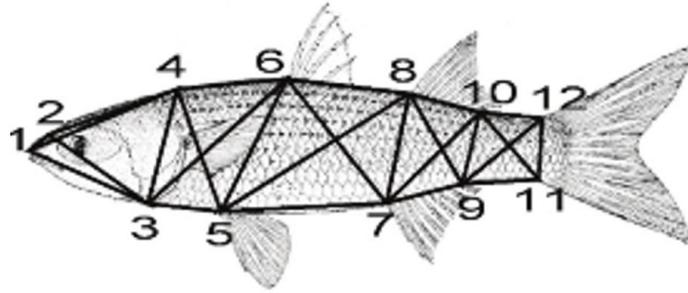
- ¿Cuántos grupos de HCoV puede distinguir en la Figura 4.21?
- ¿Cuántos de esos grupos presentan coronavirus como posibles causantes de zoonosis?
- Considerando que algunas de las cepas de coronavirus pueden tener una alta infectividad y causar brotes, como el síndrome respiratorio agudo severo (SARS) -CoV, el síndrome respiratorio de Oriente Medio (MERS) y el SARS-CoV-2 (SCoV2), ¿Desde cuales posibles hospedadores sugeriría sus orígenes a partir de los resultados del PCA?

Ejercicio 10

El estatus taxonómico de la lisa del Atlántico Sudoccidental, *Mugil cephalus* (Linnaeus 1758), se encuentra actualmente en discusión, ya que es considerada por Thomson (1997) y Harrison (2002) como sinónimo de la lisa, *Mugil liza* Valenciennes, 1836.

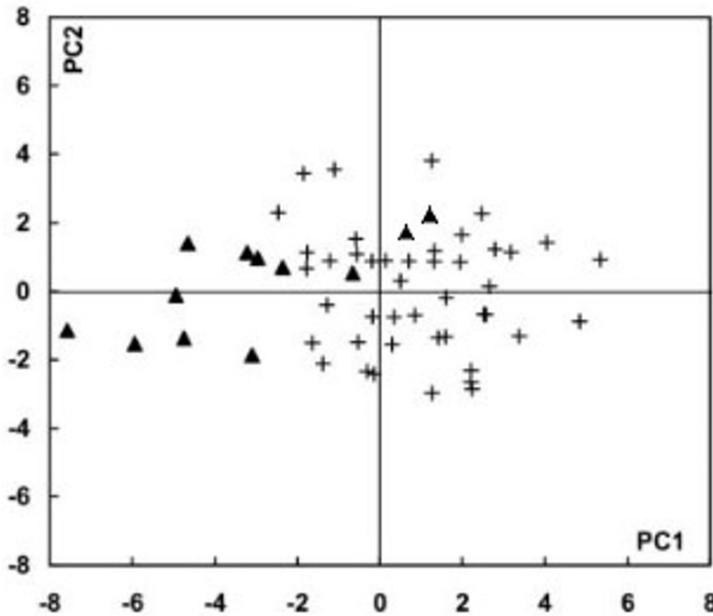
A fin de obtener más evidencia y resolver el conflicto se realizó un análisis multidisciplinario (taxonomía integrativa), empleando secuencias del gen mitocondrial Citocromo b, datos merísticos y morfometría basada en *landmarks* (Figura 4.22) de ambas especies.

Figura 4.22. Landmarks obtenidos para los distintos análisis morfométricos de especies de la familia Mugilidae.



Nota. Adaptado de Landmarks obtained using truss network for morphometric analysis, de González Castro et al., 2008.

Figura 4.23. PCA basado en datos morfométricos.



Nota. Adaptado de Principal components analysis based on morphometric data, de González Castro et al., 2008. Triángulos: *Mugil cephalus*, Cruces: *M. liza*.

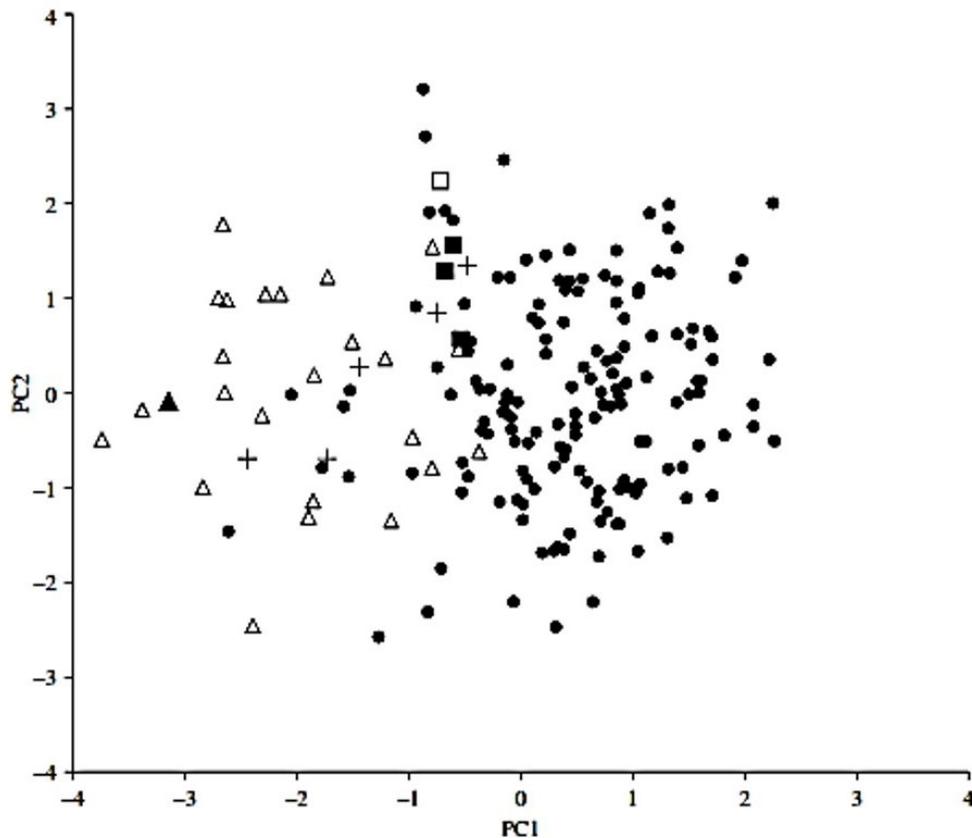
- A partir del gráfico de PCA (PC I y II acumularon el 60% de la varianza) basado en datos morfométricos (Figura 4.23), ¿cuántos grupos distingue?
- Dado que las variables obtenidas para el análisis de landmarks mostraron que *M. liza* presenta mayor robustez en los segmentos medios y caudales del cuerpo, con relación a *M. cephalus*, sumados a las diferencias en los rangos de las escamas de la serie lateral (merística) entre las dos especies, y que las distancias genéticas revelaron una clara separación de ambas taxa. ¿Qué decisión taxonómica adoptaría? ¿Coincide con lo concluido en el punto a? Explique.

- c. Considerando las distribuciones geográficas de *Mugil cephalus* y *Mugil liza* (disponibles en la base de datos Fishbase, <https://www.fishbase.se/search.php>) ¿Qué tipos de especies representan?

Ejercicio 11

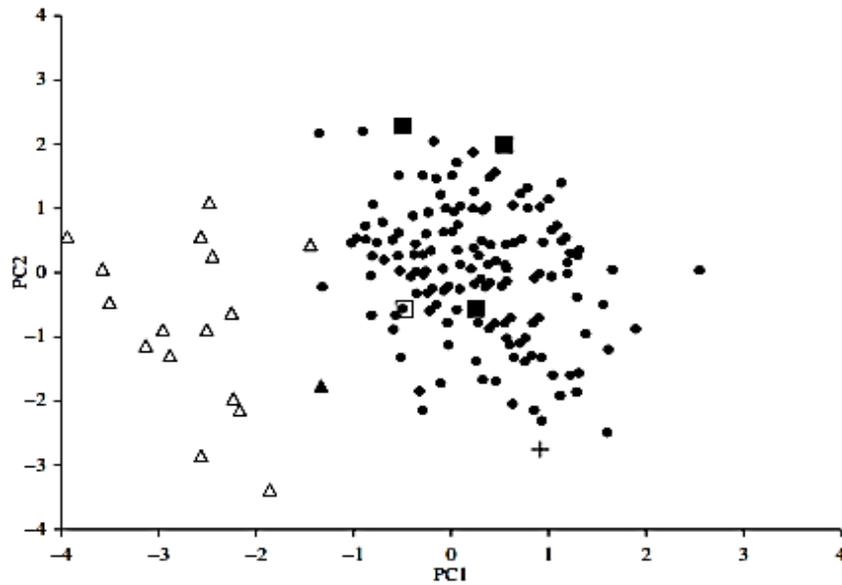
Con el fin de evaluar el estatus taxonómico de 3 especies del género *Merluccius* (*M. hubbsi* Marini 1933, *M. australis* (Hutton 1872) y *M. patagonicus* Lloris y Matallanas 2003), se realizó un análisis de variables morfométricas lineales (Figura 4.24), datos merísticos (Figura 4.25) y distancias entre puntos homólogos (*landmarks*) (Figura 4.26) de 241 especímenes colectados en diferentes regiones, junto a la reexaminación de los ejemplares tipo disponibles.

Figura 4.24. PCA basado en datos morfométricos.



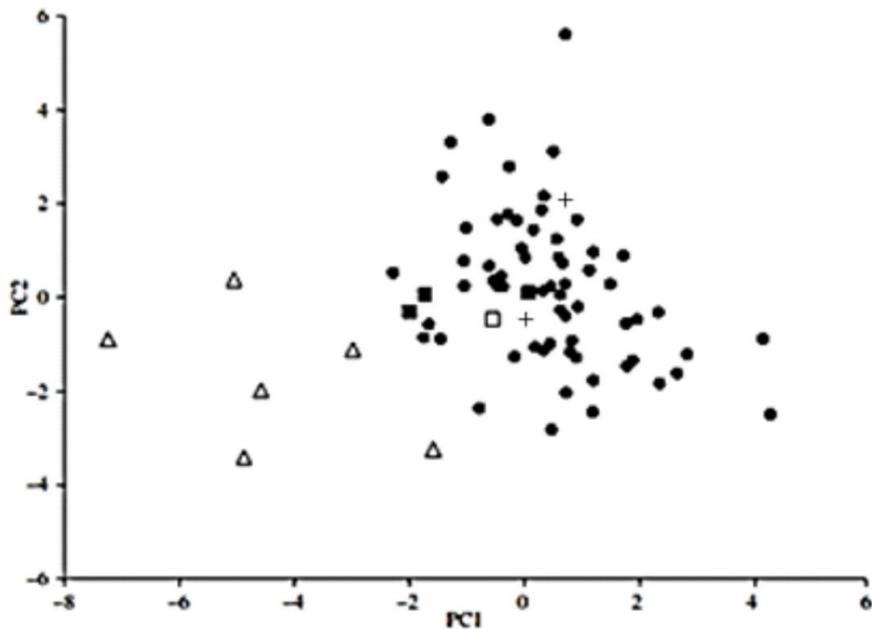
Nota. Adaptado de *Morphological, morphometric, meristic and osteological evidence for two species of hake (Actinopterygii: Gadiformes: Merluccius) in Argentinean Waters*, de Díaz de Astarloa et al., 2011. *Merluccius hubbsi* (circulo negro), *M. hubbsi* sintipos (cruz), *M. australis* (triangulo blanco), *M. australis* sintipo (triangulo negro), *M. patagonicus* holotipo (cuadrado negro) y paratipos (cuadrado negro).

Figura 4.25. PCA basado en datos merísticos.



Nota. Adaptado de Principal component (PC) analysis based on meristic data, de Díaz de Astarloa et al., 2011. *Merluccius hubbsi* (circulo negro), *M. hubbsi* sintipo (cruz), *M. australis* (triangulo blanco), *M. australis* sintipo (triangulo negro), *M. patagonicus* holotipo (cuadrado blanco) y paratipos (cuadrado negro).

Figura 4.26. PCA basado en landmarks.



Nota. Adaptado de Principal components (PC) analysis based on interlandmark distance data, de Díaz de Astarloa et al., 2011. *Merluccius hubbsi* (circulo negro), *M. hubbsi* sintipos (cruz), *M. australis* (triangulo blanco), *M. patagonicus* holotipo (cuadrado blanco) y paratipo (cuadrado negro).

A partir de los gráficos obtenidos de cada una de las fuentes de datos empleadas (Figuras 4.24–26)

- ¿Cuántas especies considera que se reconocen desde los agrupamientos de especímenes en cada caso?
- ¿Tomaría alguna decisión taxonómica-nomenclatural diferente a la planteada en el enunciado? Explique.
- ¿Cuál sería su decisión taxonómica y nomenclatural si se concluye que los datos basados en morfometría (Figura 4.24) resultaran más relevantes y de valor diagnóstico para delimitar especies que el resto de los datos considerados?

Ejercicio 12

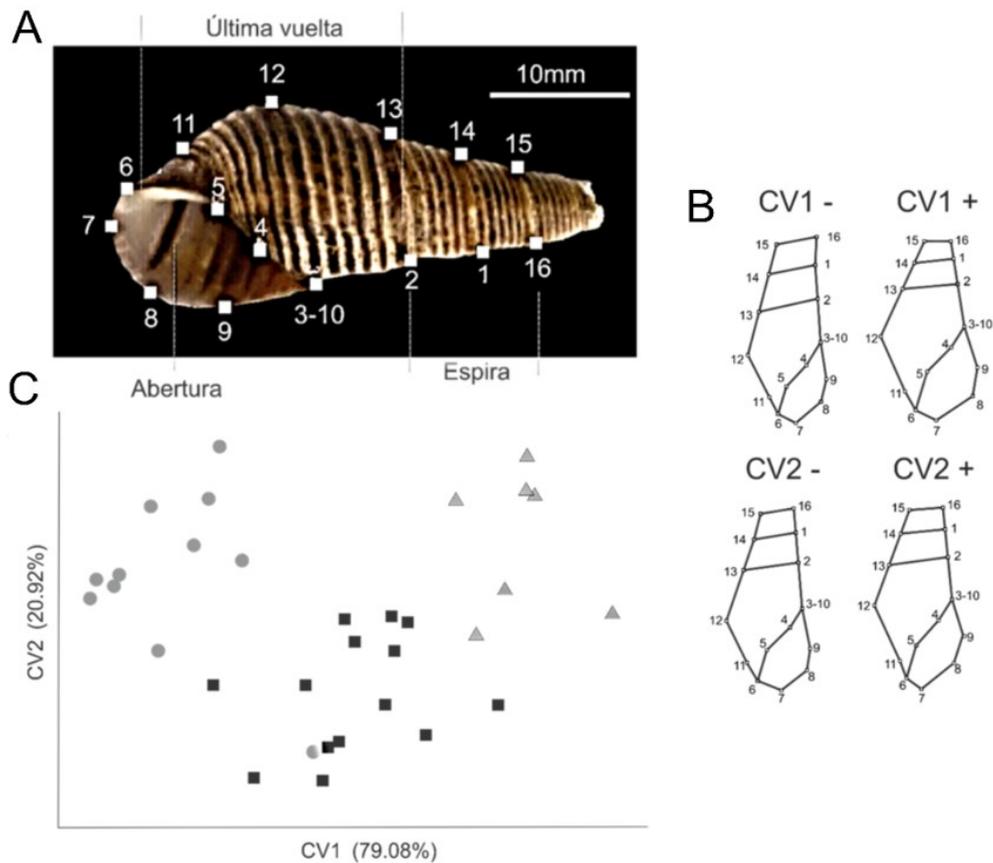
Las especies del género *Aylacostoma* (Mollusca: Gastropoda) son gasterópodos de agua dulce, vivíparos, que se reproducen por partenogénesis. Se conocen tres especies endémicas de Paraguay y Argentina: *A. guaraniticum* (Hylton-Scott 1953), *A. chloroticum* Hylton-Scott 1954 y *A. stigmaticum* Hylton Scott 1954. Las tres especies fueron descritas de ambientes altamente oxigenados (rápidos de Yacyretá-Apipé). A partir de 1978, el hábitat de estas especies sufrió importantes modificaciones debido a la construcción de la Represa de Yacyretá. Al llenarse el lago del embalse los rápidos desaparecieron. Actualmente sólo *A. chloroticum* puede encontrarse en la naturaleza, habiéndose extinguido las otras dos especies. Teniendo en cuenta que la morfología de la conchilla puede reflejar adaptaciones a las características ambientales, Vogler *et al.* (2012) realizaron un análisis de morfometría geométrica a fin de caracterizar la forma de la conchilla de cada una de las especies.

En la Figura 4.27 se ilustran los gráficos resultantes del análisis realizado a partir de 16 *landmarks*.

- ¿Considera que las tres especies de *Aylacostoma* pueden ser discriminadas sobre la base de la morfología de la conchilla? Explique por qué.
- ¿Cuál es la variabilidad total expresada en el gráfico del análisis canónico de la varianza?
- Considerando la Figura 4.27 A y B, analice las retículas extremas de deformación a lo largo de los dos ejes canónicos. Indique cuáles son las deformaciones más importantes para el eje 1 y para el eje 2 considerando la abertura (*landmarks* 3-10), espira (*landmarks* 1, 2, 13-16) y última vuelta (*landmarks* 11-12).
- El desarrollo de aberturas grandes en gasterópodos puede indicar la presencia de un pie –estructura que se utiliza para el desplazamiento del animal y su sujeción al sustrato– de mayor tamaño, brindando una mejor capacidad de mantener al espécimen adherido en

ambientes de corriente. A partir de los resultados de este análisis (Figura 4.27C) y considerando las características de la única especie viviente, ¿considera que las variaciones morfológicas en las especies pueden explicar diferencias en el hábitat preferido por ellas? Explique.

Figura 4. 27. Resultado del análisis de morfometría geométrica de tres especies endémicas de *Aylacostoma*.

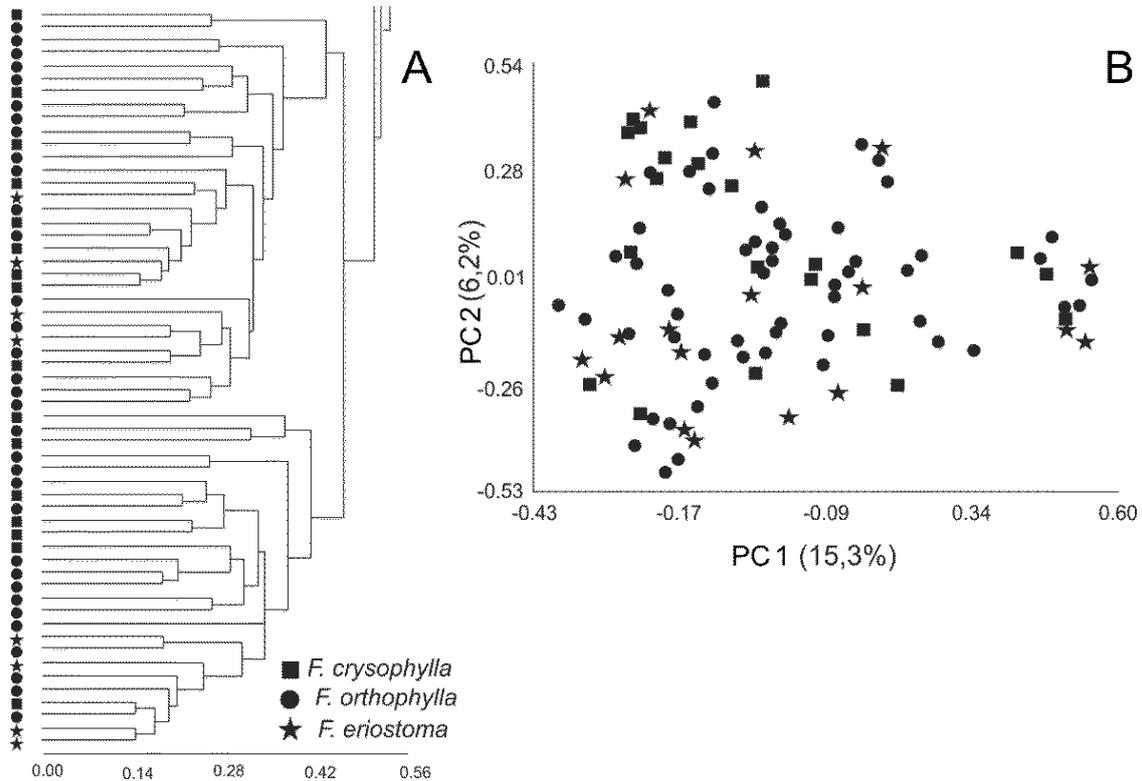


Nota. Adaptado de Figuras 2-4, de Vogler et al., 2012. A) posición de los landmarks seleccionados en las conchillas. B) retículas de deformación a lo largo de dos ejes canónicos. C) Diagrama de dispersión de los puntos individuales (Análisis Canónico de Varianza) de las tres especies, *A. choloticum* (círculo gris), *A. guaraniticum* (cuadrado negro) y *A. stigmaticum* (triángulo gris).

Ejercicio 13

Ospina González et al. (2013) analizaron la variación morfológica de tres especies de poáceas o gramíneas, morfológicamente muy similares *Festuca orthophylla* Pilg. 1898, *F. chrysophylla* Phil. 1891 y *F. eriostoma* Hack. 1903. Mediante análisis multivariado, se estudiaron 100 ejemplares y 10 caracteres morfológicos de valor diagnóstico para la identificación de las unidades taxonómicas. Los resultados obtenidos a partir del análisis se muestran en la Figura 28.

Figura 4.28. Análisis multivariado de 100 ejemplares basado en datos morfológicos de tres especies de *Festuca*.



Nota. Adaptado de Análisis estadísticos multivariados, de Ospina González et al., 2013. A) sección del Fenograma obtenido. B) PCoA.

Sobre la base de la figura 4.28A y B responda:

- ¿Qué técnicas numéricas se aplicaron en cada caso?
- ¿Las agrupaciones resultantes en los análisis se corresponden con las tres especies consideradas previamente? Justifique.
- ¿Existe congruencia entre los resultados de las dos técnicas? Explique.
- ¿Qué indica la escala en la Figura 4.28A?
- Explique que representan los ejes del gráfico en la figura 4.28B. ¿Cuál es la variabilidad total expresada en el gráfico?
- En virtud de los resultados, ¿qué decisión taxonómica adoptaría? ¿debería tomar alguna decisión nomenclatural? Justifique.

Ejercicio 14

En un estudio se abordó la problemática taxonómica de *Opuntia anacantha* Speg. (Opuntioideae, Cactaceae), en la Región Chaqueña Argentina, ya que su identificación es difícil.

Tomando en cuenta la variación morfológica, tanto vegetativa como reproductiva se analizaron 80 ejemplares (*OTUs*) y 70 caracteres, de los cuales diez son cualitativos (tres vegetativos, siete reproductivos), 17 son cuantitativos discretos (doce vegetativos, cinco reproductivos) y 43 cuantitativos continuos (21 vegetativos, 22 reproductivos). La tabla 4.11 muestra algunos ejemplos de dichos caracteres. Se realizó un *PCA* sobre 80 individuos con caracteres morfológicos cuantitativos vegetativas solamente y con cuantitativos vegetativos y reproductivos en conjunto (Figuras 4.29 y 4.30). Además, sobre 80 individuos, se realizó un análisis de agrupamiento (empleando el algoritmo *UPGMA*) utilizando caracteres morfológicos vegetativos y reproductivos cualitativos, y otro análisis con caracteres cualitativos y cuantitativos (Figuras 4.31 y 4.32).

Tabla 4.11. Valores de los eigen-vectores en las *PC1* y *PC2* de 14 caracteres morfológicos vegetativos y reproductivos cuantitativos, usados en el *PCA* (considerando 80 *OTUs*)

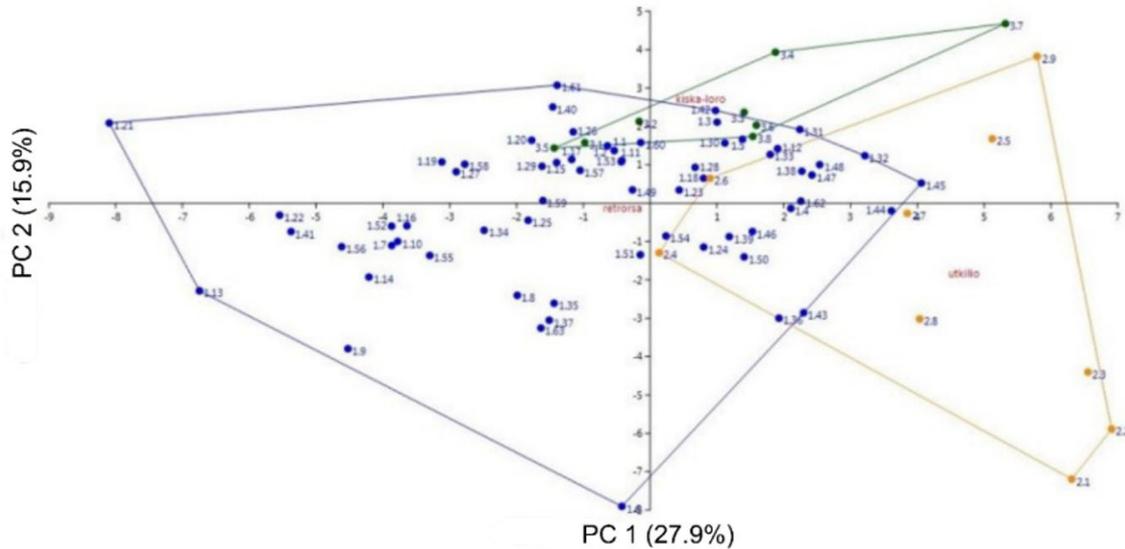
Variable o carácter	CP 1	CP 2
Número de areolas con espinas en artejo proximal (NaAPe)	0,60528	-0,57579
Número mínimo de espinas por areola en artejo proximal (NmeaAP)	0,28592	-0,63057
Número máximo de espinas por areola en artejo proximal (NMeaAP)	0,27715	-0,78644
Número de areolas con espinas en artejo medio (NaAMe)	0,5608	-0,54072
Número mínimo de espinas por areola en artejo medio (NmeaAM)	0,20594	-0,66633
Número máximo de espinas por areola en artejo medio (NMeaAM)	0,21384	-0,75578
Número mínimo de espinas por areola en artejo distal (NmeaAD)	0,25644	-0,5785
Número máximo de espinas por areola en artejo distal (NMeaAD)	0,2304	-0,70039
Número de semillas por fruto (NSF)	0,031492	0,55569
Ancho del artejo proximal (cm) AAP	0,83256	0,17474
Ancho del artejo medio (cm) AAM	0,82704	0,18139
Ancho del artejo distal (cm) AAD	0,78287	0,10856
Longitud del pericarpelo (mm) LP	0,32471	0,51698
Longitud del fruto (mm) LF	0,31826	0,57862

Sobre la base de la tabla 4.11 y figuras 4.29-32, responda:

- ¿Qué caracteres aportan más al *PC1* y al *PC2*?
- Observando las figuras 4.29 y 4.30 ¿Qué tipo de caracteres aportan a una mejor separación de los grupos, los que contemplan caracteres cuantitativos vegetativos solamente o caracteres cuantitativos vegetativos y reproductivos en conjunto? Explique.

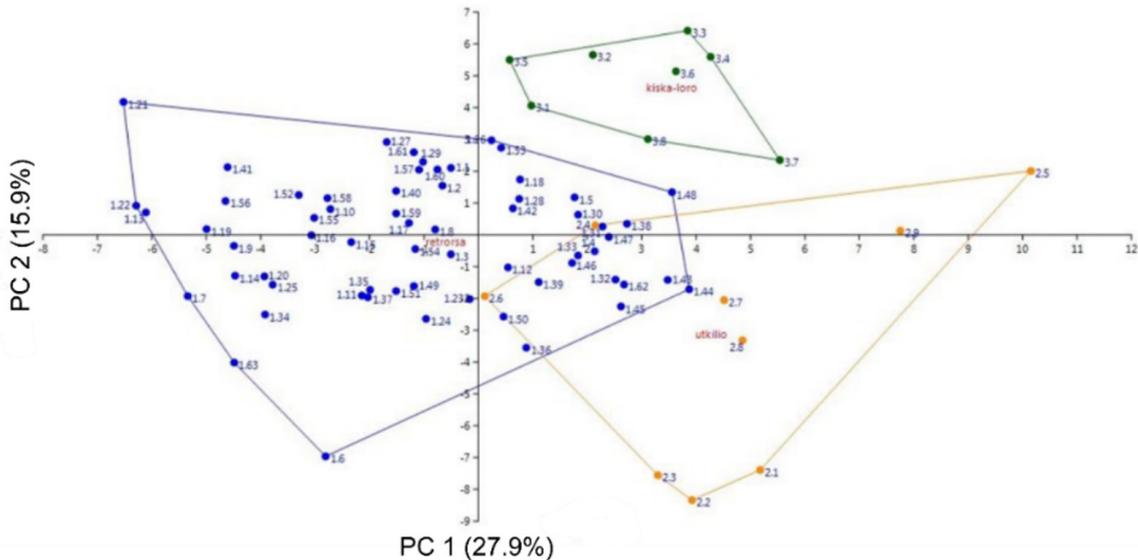
- c. ¿Las agrupaciones resultantes en las Figuras 4.31 y 4.32 se corresponden con las tres especies consideradas hasta el momento?
- d. ¿Existe congruencia entre los resultados de las dos técnicas empleadas? Explique.

Figura 4.29. ACP de 33 variables vegetativas cuantitativas.



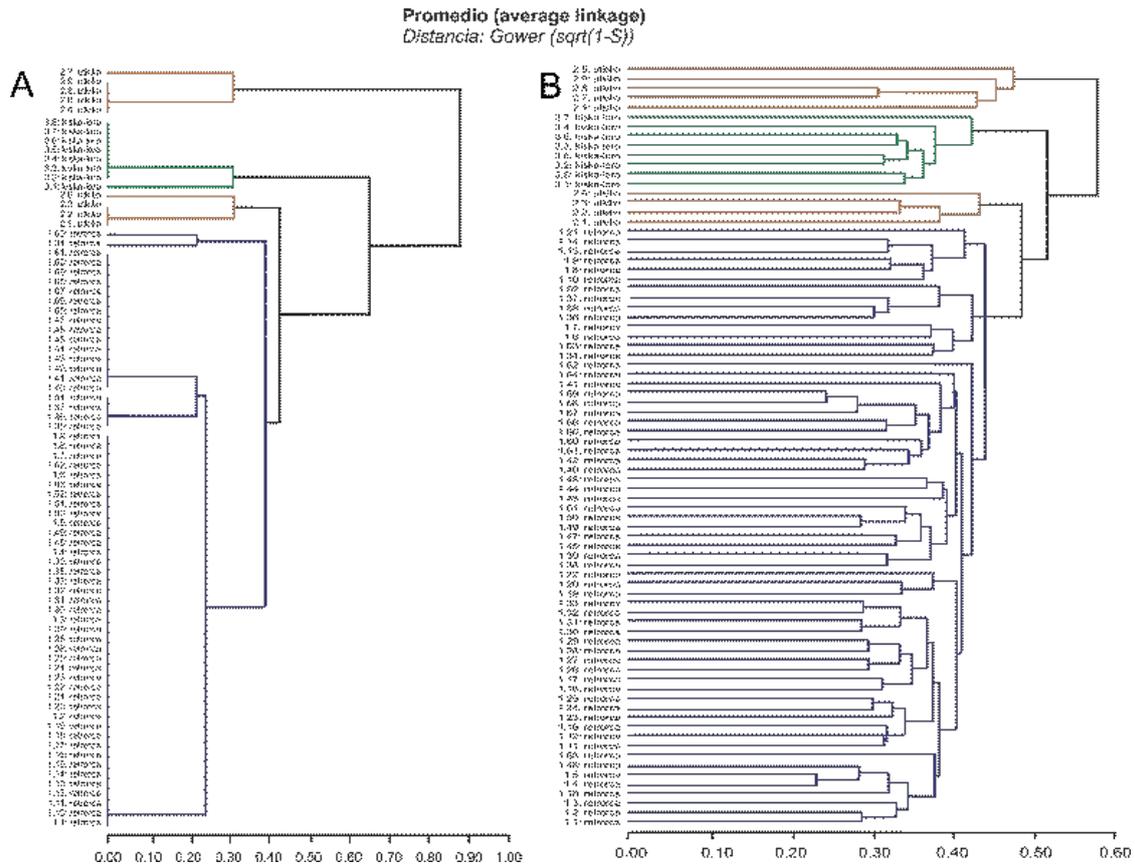
Nota. Adaptado de *Proyección de los individuos estudiados en relación a las componentes principales 1 y 2, resultado del ACP de 33 variables vegetativas cuantitativas sobre 80 individuos, de Oakley, 2019.* En azul: *Opuntia anacantha* var. *retrorsa*; naranja: *O. anacantha* var. *utkilio*; verde: *O. anacantha* var. *kiska-loro*, según la identificación previa de los especímenes. Las observaciones (individuos) se indican con puntos y número de OTU.

Figura 4.30. ACP de 60 variables morfológicas vegetativas y reproductivas cuantitativas.



Nota. Adaptado de *Proyección de los individuos estudiados en relación a las componentes principales 1 y 2, resultado del ACP de 60 variables morfológicas vegetativas y reproductivas cuantitativas sobre 80 individuos, de Oakley, 2019.* En azul: *Opuntia anacantha* var. *retrorsa*; naranja: *O. anacantha* var. *utkilio*; verde: *O. anacantha* var. *kiska-loro*, según la identificación previa de los especímenes. Las observaciones (individuos) se indican con puntos y número de OTU.

Figura 4.31. Análisis con caracteres cualitativos y cuantitativos.



Nota. Adaptado de UPGMA, de Oakley, 2019. Colores representan las especies y las observaciones (individuos) se indican con número de OTU e identificación previa de los especímenes: *Opuntia anacantha* var. *retrorsa* (azul), *O. anacantha* var. *utkilio* (naranja) y *O. anacantha* var. *kiska-loro* (verde). A) Análisis con caracteres morfológicos vegetativos y reproductivos cualitativos. B) Análisis con caracteres cualitativos y cuantitativos.

Sobre la base de la tabla 4.11 y figuras 4.29-31, responda:

- ¿Qué caracteres aportan más al PC1 y al PC2?
- Observando las figuras 4.29 y 4.30 ¿Qué tipo de caracteres aportan a una mejor separación de los grupos, los que contemplan caracteres cuantitativos vegetativos solamente o caracteres cuantitativos vegetativos y reproductivos en conjunto? Explique.
- ¿Las agrupaciones resultantes en la figura 4.31 (A y B) se corresponden con las tres especies consideradas hasta el momento?
- ¿Existe congruencia entre los resultados de las dos técnicas empleadas? Explique.

Ejercicio 15

Las macroalgas *Crassiphycus corneus* (J. Agardh) Gurgel, J.N. Norris y Fredericq y *C. usneoides* (Mertens ex C. Agardh) Gurgel, J.N. Norris y Fredericq, se encuentran dentro de las Rhodophyta y presentan una considerable similitud morfológica y áreas de distribución solapadas, que hacen que la separación entre ambas especies resulte complicada. En este marco, se consideró hacer una evaluación morfométrica de algunos caracteres anatómicos para determinar una

posible discriminación entre las especies. Los resultados obtenidos se muestran a continuación en las Figuras 4.32 y 4.33 y Tabla 4.12.

Figura 4.32. Dendrograma

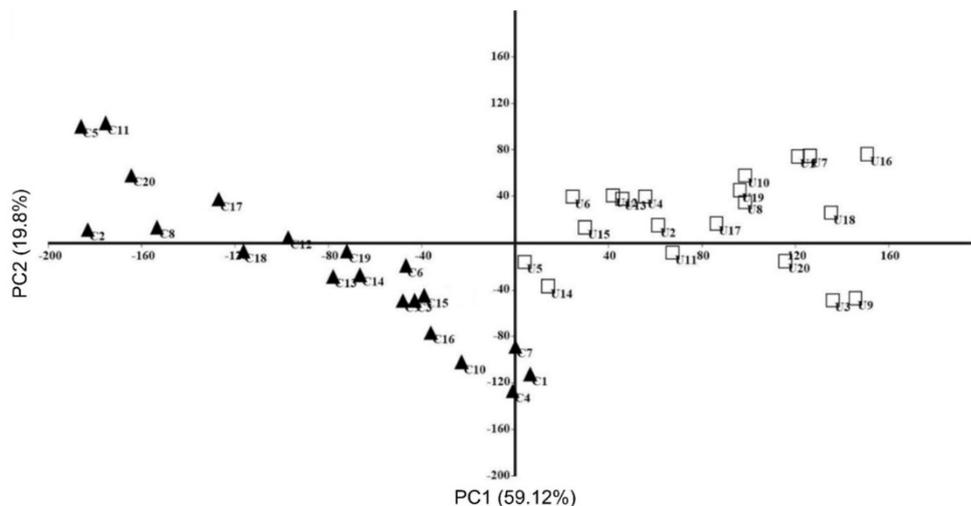


Nota. Adaptado de Dendrograma resultado del análisis de agrupamiento bajo el índice de disimilitud de distancia euclidiana y el algoritmo UPGMA, de Vilchis et al., 2020. Basado en 24 caracteres tomados de 40 ejemplares de: Crassiphycus corneus (OTUs representadas con la letra c + número); y C. usneoides (OTUs representadas con la letra u + número).

A partir de los resultados observados en dicho análisis (Figuras 4.32 y 4.33 y Tabla 4.12) y teniendo en cuenta conceptos abordados en trabajos prácticos anteriores responde.

- ¿En qué categoría taxonómica se ubica el grupo Rhodophyta?
- ¿Cómo denominaría a las especies considerando el tipo de distribución geográfica?
- A partir de los resultados obtenidos en las figuras 4.32 y 4.33. ¿Considera que el análisis realizado fue de utilidad para discriminar a las especies estudiadas? Justifique
- ¿Alguno/s de los caracteres analizados podrían resultar más relevantes? Justifique.

Figura 4.33. PCA.



Nota. Adaptado Diagrama de los dos primeros componentes principales, obtenidos a partir de 24 variables morfométricas, de Vilchis et al., 2020. Crassiphycus corneus (□) y C. usneoides (▲).

Tabla 4.12. Caracteres, su ubicación en el talo del alga y los eigen valores obtenidos del PCA.

VARIABLES O CARACTERES DE LOS FRAGMENTOS DEL TALO MEDIDOS	Porción del talo	CP1	CP2	CP3
Diámetro mayor del talo	Apical	0,335	0,407	0,545
Diámetro menor del talo	Apical	0,337	-0,723	0,340
Diámetro mayor del talo	Media	0,615	0,090	-0,349
Diámetro menor del talo	Media	0,370	0,004	-0,355
Diámetro mayor del talo	Basal	0,256	0,422	0,342
Diámetro menor del talo	Basal	0,436	-0,350	0,194
Largo células corticales	Apical	0,006	0,012	0,004
Ancho células corticales	Apical	0,006	0,011	0,005
Largo células subcorticales	Apical	0,009	0,016	0,009
Ancho células subcorticales	Apical	0,008	0,013	0,018
Largo células medulares	Apical	-0,001	-0,013	0,040
Ancho células subcorticales	Media	0,017	0,023	0,014
Largo células medulares	Media	0,005	0,003	-0,038
Largo células corticales	Basal	0,018	0,027	0,012
Ancho células corticales	Basal	0,013	0,021	0,012
Largo células subcorticales	Basal	0,011	-0,008	-0,005
Ancho células subcorticales	Basal	0,007	0,000	0,000
Largo células medulares	Basal	0,013	0,006	-0,025
Ancho células medulares	Basal	0,023	0,009	-0,001

Nota. Adaptado de Eigenvalores obtenidos del análisis de componentes principales (CP) a partir de 24 medidas lineales de las especies *Crassiphycus corneus* y *C.*, de Vilchis et al., 2020.

Referencias

- Adams, D. C., Rohlf, F. J. y Slice, D. E. (2004). Geometric morphometrics: ten years of progress following the “revolution”. *The Italian Journal of Zoology*, 71, 5-16.
- Álvarez-Presas, M., Carbayo, F., Rozas, J. y Riutort, M. (2011). Land planarians (Platyhelminthes) as a model organism for fine-scale phylogeographic studies: understanding patterns of biodiversity in the Brazilian Atlantic Forest hotspot. *Journal of Evolutionary Biology*, 24, 887-896.
- Arce, O. E., De Marco, N. E. y Santillán, M. R. (2009). Componentes principales y coordenadas principales: estudio comparativo basado en una aplicación a la taxonomía numérica. *Lilloa*, 46, 10-33.
- Bandelt, H. J., Forster, P. y Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16(1), 37-48.
- Bookstein, F. L. (1991). *Morphometric tools for landmark data: geometry and biology*. Cambridge: Cambridge University Press.
- Bookstein, F. L. (1997). Landmark methods for forms without landmarks: Localizing group differences in outline shape. *Medical Image Anal*, 1, 225-243.

- Campbell, N. A. y Mahon, R. J. (1974). A multivariate study of variation in two species of rock crab of genus *Leptograpsus*. *Australian Journal of Zoology*, 22, 417-425.
- Cardini, A. (2019). Integration and modularity in procrustes shape data: Is there a risk of spurious results? *Evolutionary Biology*, 303, 2747-2765.
- Cattell, R. B. (1952). *Factor analysis: an introduction and manual for the psychologist and social scientist*. Harper: Oxford.
- Cavalli-Sforza L. L. y Edwards, A. W. (1967). Phylogenetic analysis: models and estimation procedures. *American journal of human genetics*, 19, 233-257.
- Cheema, J. y Dicks, J. (2009). Computational approaches and software tools for genetic linkage map estimation in plants. *Briefings in Bioinformatics*, 10(6), 595-608.
- Crisci, J. V. y López Armengol, M. E. (1983). *Introducción a la teoría y práctica de la taxonomía numérica*. Washington: Organización de los Estados Americanos.
- Díaz de Astarloa, J. M., Bezzi, S. I., González Castro, M., Mabrugaña, E., Hernández, D., Delpiani, S. M., Figueroa, D. E., Cousseau, M. B., Deli Antoni, M. Y. y Tringali, L. (2011). Morphological, morphometric, meristic and osteological evidence for two species of hake (Actinopterygii: Gadiformes: *Merluccius*) in Argentinean Waters. *Journal of Fish Biology*, 78(5), 1336-58.
- González Castro, M., Heras, S., Cousseau, M. B. y Roldán, M. I. (2008). Assessing species validity of *Mugil platanus* Günther, 1880 in relation to *Mugil cephalus* Linnaeus, 1758 (Actinopterygii). *Italian Journal of Zoology*, 75, 319-325.
- Guy, F., Brunet, M., Schmittbuhl, M. y Viriot, L. (2003). New approaches in Hominid Taxonomy: Morphometrics. *American Journal of Physical Anthropology*, 21, 198-218.
- Harrison, I. J. (2002). Mugilidae: mullets. En: K. Carpenter (Ed)., *The living marine resources of the Western Central Atlantic. Vol. 2. Bony fishes part 1 (Acipenseridae to Grammatidae)*. FAO Species identification Guide for Fisheries purposes (pp. 1071-1085). Rome: FAO.
- Hernández, T., Canales, M., Caballero, J., Durán, A. y Lira, R. (2005) Análisis cuantitativo del conocimiento tradicional sobre plantas utilizadas para el tratamiento de enfermedades gastrointestinales en Zapotitlán de las Salinas, Puebla, México. Inci [online], 30(9), 17-27.
- Husson, F, Lê, S. y Pagès, J. (2017). *Exploratory multivariate analysis by example using R. and Hall/CRC. Boca Raton: CRC press*.
- James, F. C y McCulloch, C. E. (1990). Multivariate analysis in ecology and systematics: panacea or Pandora's box? *Annual Review of Ecology and Systematics*, 21(1), 129-166.
- Jansky S. H., Dempewolf H., Camadro E. L., Simon R., Zimnoch-Guzowska E., Bisognin D. A. y Bonierbale M. (2012). A Case for Crop Wild Relative Preservation and Utilization in Potato (*S. tuberosum* L.). *Crop Science*, 53, 746-754.
- Jolliffe. I. T. y Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Kassambara, A. (2017). *Practical guide to principal component methods*. Marsella: STHDA.

- Konishi, T. (2020) Principal component analysis of coronaviruses reveals their diversity and seasonal and pandemic potential. *PLoS ONE*, 15(12), e0242954. <https://doi.org/10.1371/journal.pone.0242954>
- Konishi, T., Matsukuma, S., Fuji, H., Nakamura, D., Satou, N. y Okano, K. (2019). Principal Component Analysis applied directly to Sequence Matrix. *Scientific Reports*, 9(1), 1-13.
- Lanteri, A. A. y Cigliano, M. M. (eds). (2006). *Sistemática Biológica: Fundamentos teóricos y ejercitaciones*. (3ª Edición). La Plata: Editorial de la Universidad de la Plata.
- Legendre, P. y Legendre, L. (1998). *Numerical ecology: developments in environmental modeling*. Amsterdam: Elsevier.
- Lemey, P., Salemi, M. y Vandamme, A. M. (2009). *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge: Cambridge University Press.
- Marcus, L. F., Corti, M., Loy, A., Naylor, G. J. P. y Slice, D. E. (1996). *Advances in morphometrics*. New York: Springer Science & Business Media.
- McCune, J. M. (1991). Epstein-Barr virus associated lymphoproliferative disease in mice and men. *Laboratory investigation; a journal of technical methods and pathology*, 65(4), 377-380.
- Mercado-Salas, N. y Suárez-Morales, E. (2012). Morfología, diversidad y distribución de los Cyclopoida (Copepoda) de las zonas áridas del centro-norte de México. II. Eucyclopiniae y análisis biogeográfico. *Hidrobiológica: revista del Departamento de Hidrobiología*, 22, 99-124.
- Moss, W. (1968). Experiments with various techniques of numerical taxonomy. *Systematic Zoology*, 17(1), 31-47.
- Nei, M. (1972). Genetic Distance between Populations. *American Naturalist*, 106, 283-292.
- Nei, M. y Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences*, 76(10), 5269-5273.
- Nei, M., Tajima, F. y Tatenno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of molecular evolution*, 19, 153-170.
- Núñez-Colín, C. A. y Escobedo-López, D. (2011). Uso correcto del análisis clúster en la caracterización de germoplasma vegetal. *Agronomía Mesoamericana*, 22(2), 415-427.
- Oakley, L. (2019). Análisis de la variabilidad intraespecífica y de distribución de *Opuntia anacantha* Speg. (Cactaceae-Opuntioideae) en la Región Chaqueña Argentina (Tesis de maestría) UNR.
- Ospina González, J. C., Aliscioni, S. S. y Denham, S. S. (2013). Estudios taxonómicos en el género *Festuca* L. (Poaceae) de Argentina y Chile. Universidad de Concepción. *Gayana Botánica*, 70(1), 01-15.
- Palacio, F. X., Apodaca, M. J. y Crisci, J. V. (2020). *Análisis multivariado para datos biológicos. Teoría y su aplicación utilizando el lenguaje R*. Buenos Aires: Fundación de Historia Natural Félix de Azara.
- Peña Cervera, M. A. (2020) Caracterización Morfológica y Comportamiento Reproductivo de Poblaciones de *Solanum brevicaule* del Parque Nacional Los Cardones, Salta, Argentina (Tesina de Grado). Universidad Nacional de Cuyo.

- Roch, S. (2010). Toward extracting all phylogenetic information from matrices of evolutionary distances. *Science*, 327(5971), 1376-1379.
- Rogers, J. S. (1972). Measures of genetic similarity and genetic distances. *Studies in Genet VII* 7213, 145-153.
- Rohlf, F. J. (1990). *Numerical taxonomy system of multivariate statistical programs*. Version 1.8. New York: State University at Stony Brook.
- Rohlf, F. J. (1996). Morphometric spaces, shape components, and the effects of linear transformations. En L. F. Marcus, M., A. Corti, A. Loy y G. J. Naylor (Eds), *Advances in morphometrics* (pp. 117-129). New York: Plenum.
- Rohlf, F. J. y Marcus, L. F. (1993). A revolution morphometrics. *Trends in Ecology and Evolution*, 8(4), 129-132.
- Saitou, N. y Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406-425.
- Sneath, P. H. y Sokal, R. R. (1973). *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. (1st Edition). San Francisco: W. H. Freeman.
- Sokal, R. R. y Rohlf, F. J. (1986). *Introducción a la bioestadística*. Barcelona: Reverte.
- Takezaki, N. y Nei, M. (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics*, 144, 389-399.
- Tamura, K. y Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular biology and evolution*, 10, 512-526.
- Tamura, K., Nei, M. y Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences (USA)* 101, 11030-11035.
- Tamura, K., Dudley, J., Nei, M. y Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24, 1596-1599.
- Testoni, D., y Villamil, C. (2014). Estudios en el género *Cortaderia* (Poaceae). I. Sistemática y nomenclatura de la sect. *Cortaderia*. *Darwiniana, Nueva Serie*, 2(2), 260-276.
- Thomson J. M. (1997). The Mugilidae of the world. *Memoirs of the Queensland Museum*, 41, 457-562.
- Toro Ibacache, M. V., Manriquez Soto, G. y Suazo Galdames, I. (2010). Morfometría Geométrica y el Estudio de las Formas Biológicas: De la Morfología Descriptiva a la Morfología Cuantitativa. *International Journal of Morphology*, 28(4), 977-990.
- Vogler, R. E., Beltramino, A. A., Gutiérrez-Gregoric, D. E., Peso, J. G., Griffin, M. y Rumi, A. (2012). Threatened Neotropical mollusks: analysis of shape differences in three endemic snails from High Paraná River by geometric morphometrics. *Revista Mexicana de Biodiversidad*, 83(4), 1045-1052.

- Vilchis, M. I., Dreckmann, K. M., Quintanar, A. y Senties, A. (2020). Análisis morfométrico de las especies simpátricas del complejo *Crassiphycus corneus*-*C. usneoides* (Gracilariaceae, Rhodophyta) en las costas mexicanas del Golfo de México y Mar Caribe. *Acta Botánica Mexicana*, 127, e1596.
- Yang, Z. y Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, 13(5), 303-314.