

# Generando Pseudorréplicas, a Partir de Variables Morfométricas Reales, para Contrastar Hipótesis Taxonómicas y Evolutivas: El Caso de un Pez Ampliamente Distribuido en las Tierras Bajas de la Orinoquia Venezolana

Generating Pseudo-Replicates, from Real Morphometric Variables, for Contrast Taxonomic and Evolutionary Hypotheses: The Case of a Widely Distributed Fish in the Lowlands of the Venezuelan Orinoquia

Juan Elías García-Pérez<sup>1,2</sup>

**GARCÍA-PÉREZ, J. E.** Generando pseudo réplicas, a partir de variables morfométricas reales, para contrastar hipótesis taxonómicas y evolutivas: el caso de un pez ampliamente distribuido en las tierras bajas de la Orinoquia venezolana. *Int. J. Morphol.*, 38(4):1212-1223, 2021.

**RESUMEN:** Con el propósito de evaluar la distintividad morfométrica de cinco poblaciones de *Mikrogeophagus ramirezi* (Pisces: Cichlidae) de las tierras bajas de la Orinoquia venezolana, se tomaron 14 hitos homólogos a 83 individuos, y se transformaron a variables Procrustes. Se realizó una prueba MANOVA/CVA, pero ésta no funcionó bien, por el bajo número de ejemplares usados, y se decidió ampliar el tamaño de la muestra. Se determinó la distribución probabilística de las variables, que fue normal, después se detectó que las varianzas eran similares. Luego se procedió a generar 100 valores de cada variable, bajo el modelo  $Y = X + \mathcal{E}$ , donde  $Y$  es el valor a estimar de la variable,  $X$  es el valor promedio de esa variable, y  $\mathcal{E}$  es el producto de la desviación estándar real multiplicado por valores aleatorios de una distribución normal, para cada variable y población. Esta simulación, se realizó para tres tipos de desviaciones estándar: La de las variables reales, la de valores ponderados simple y con media calculada con bootstrap. Los datos generados fueron validados comparando con los valores reales, y a través de algunas propiedades de ellos, tanto antes, como después de aplicar la prueba MANOVA/CVA. En todas las pruebas, los valores generados no fueron distintos a los valores reales. Usando los valores generados, como sustitutos válidos, se determinó, estadísticamente, que las cinco poblaciones eran morfométricamente diferentes, y candidatas a ser consideradas especies plenas. Por otra parte, se realizó una regresión entre las coordenadas de los centroides de las muestras reales y se comparó con las de los generados, mientras las coordenadas reales se distribuyen sobre una línea recta casi perfecta, y altísimamente correlacionada, las de los valores generados están muy dispersas. Se infiere que esta línea, llamada Línea de Estasis Centroidal, sea una señal filogenética surgida de los datos reales.

**PALABRAS CLAVE:** Coordenadas Procrustes; Línea de Estasis Centroidal; MANOVA/CVA; *Mikrogeophagus ramirezi*; Técnicas Monte Carlo.

## INTRODUCCIÓN

La Morfometría Geométrica, es una herramienta que usa análisis numérico y estadístico, sin ninguna hipótesis a priori sobre la distribución, importancia o polaridad de los caracteres, que en su caso, consisten en coordenadas cartesianas en el plano o el espacio, que son tomadas como puntos, sobre imágenes, que son identificables en todos los ejemplares, llamados hitos (Bookstein, 1991). Esta técnica permite separar especies crípticas (Soria-Barreto *et al.*, 2011; Ruane, 2015; Zúñiga-Reinoso & Benítez, 2015).

Como hay diferencias en tamaño, forma, disposición, de las imágenes de cada individuo, entre otras variables que pueden ocasionar ruidos o perturbaciones en el sistema de coordenadas, existe una transformación de los datos, que permite eliminar los efectos del tamaño, orientación y la disposición, llamada Procrustes (Bookstein), que implica que los valores correspondientes de cada variable, para todos los individuos, presentarán una mínima distancia hacia el centroide de cada variable, entonces las diferencias encon-

<sup>1</sup>Museo de Zoología, Programa CAM, Universidad Nacional Experimental de Los Llanos Occidentales Ezequiel Zamora, UNELLEZ, Guanare, Portuguesa, Venezuela.

<sup>2</sup>Doctorado de Biodiversidad, PRESAV, Universidad Nacional Experimental de Los Llanos Occidentales Ezequiel Zamora, UNELLEZ, Guanare, Portuguesa, Venezuela.

tradas, entre los ejemplares, son atribuibles, estrictamente, a la forma (Bookstein; Zelditch *et al.*, 2004), y la magnitud o tamaño del centroide para cada conjunto de valores tomados de un individuo, cuando han sido transformados a Procrustes, es igual a 1 (Zelditch *et al.*).

Strauss (2010) sugiere aplicar, para discriminar grupos con base a caracteres morfométricos, las técnicas de Análisis de Variación Canónica (CVA), las distancias de Mahalanobis y Análisis Multivariado de Varianza (MANOVA). Pero hay un inconveniente, técnicas tan poderosas, como el MANOVA, son sensibles al tamaño muestral y al número de variables a analizar. Por ejemplo, Strauss señala un número mínimo de muestras de, al menos, cinco o más veces la cantidad de variables. Como a veces es difícil encontrar un número adecuado de muestras, y dado que el problema puede agravarse si alguna de estas poblaciones está en declinación, o resulta muy costoso ir a coleccionar nuevas muestras, y se necesita evaluar el estado taxonómico de dichas poblaciones. Una opción es la reducción del número de hitos, pero con ello, se puede perder información valiosa.

Entonces, se hace necesario buscar la manera de aumentar el tamaño de la muestra, sin intervenir nuevamente las poblaciones naturales. Para ello, se sugiere realizar un aumento artificial del tamaño muestral, a través de una simulación de las variables de cada conjunto de individuos, que corresponden con un grupo de interés a ser analizado, generando las variables simuladas, a partir de los datos reales.

En el presente trabajo se propone una metodología para evaluar la biodiversidad en poblaciones naturales de *Mikrogeophagus ramirezi* (Myers & Harry, 1948) (Teleostei:

Cichlidae), una especie considerada endémica de los llanos de la Orinoquia Colombiana y Venezolana, su distribución abarca desde los ríos Inírida y Guaviare, en los Llanos Orientales de Colombia (Lasso *et al.*, 2009) hasta ríos que desembocan en el Golfo de Paria en Venezuela (Lasso *et al.*, 2010).

*Mikrogeophagus ramirezi*, tal como se conoce hasta ahora, pertenece a un género, cuyo otro integrante habita en Bolivia (Kullander, 2011). Estudios morfológicos y genéticos ubican a *M. altispinosus*, como su pariente más cercano (López-Fernández *et al.*, 2005a,b), la cual habita en cuerpos de agua en el sur de la Amazonía. Esto representa un hiato bastante extenso entre la distribución de ambas especies.

Es de suponer que las poblaciones de *M. ramirezi*, actualmente aisladas, en cuerpos de agua con diferentes orientaciones de su drenaje (Fig. 1), pueden estar evolucionando independientemente y presentar algún tipo de diferenciación morfológica. Por ello, se realizó un análisis de morfometría geométrica para evaluar esa posible diferenciación de las poblaciones.

## MATERIAL Y MÉTODO

Se estudiaron muestras de cinco localidades, número de ejemplares examinados entre paréntesis, cuatro en la región de los llanos de Venezuela codificadas como PI (12), FC (23), GU (16) y AN (9) y una de la región de Guayana BOL (23), representadas espacialmente en la Figura 1, para un total 83 ejemplares, pertenecientes a la colección de Ictiología del Museo de Zoología de la UNELLEZ-Guanare.

Se tomaron fotografías digitales de los 83 ejemplares de *Mikrogeophagus ramirezi*. Para cada individuo se digitalizaron, con el programa tpsDig2,17 (Soria-Barreto *et al.*; Rohlf, 2013), 14 hitos homólogos de dos dimensiones (Fig. 2), que fueron leídos y procesados con el programa PAST 2,71 (Hammer, *et al.*, 2001), y se aplicó una transformación a coordenadas Procrustes, para eliminar los efectos de tamaño, traslación y rotación (Bookstein; Zelditch *et al.*).

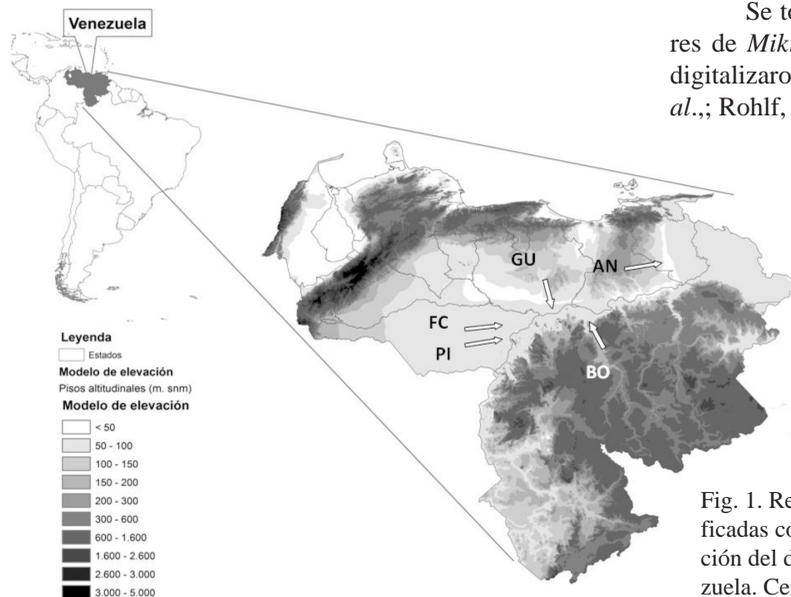


Fig. 1. Representación gráfica de las cinco localidades, identificadas como en el texto, con una flecha que indica la orientación del drenaje de las aguas, sobre un modelo DEM de Venezuela. Centro de Cartografía, Biocentro-UNELLEZ.

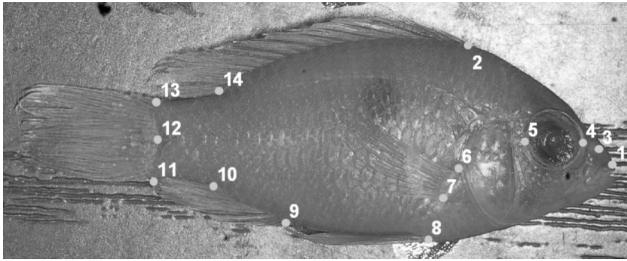


Fig. 2. Ejemplar de *Mikrogeophagus ramirezi*, donde se muestran los 14 hitos homólogos tomados para el análisis morfométrico. La definición de los hitos, algunos modificados de Young *et al.* (2009), es la siguiente: (1) Extremo anterior de la cara, (2) Inserción anterior de la aleta dorsal, (3) Narina, (4) Extremo anterior horizontal de la órbita ocular, (5) Extremo posterior horizontal de la órbita ocular, (6) Inserción dorsal de la aleta pectoral, (7) Inserción ventral de la aleta pectoral, (8) Inserción dorso-lateral de la aleta pélvica (9) Inserción anterior de la aleta anal, (10) Inserción posterior de la aleta anal, (11) Inserción ventral de la aleta caudal, (12) Borde caudal de la placa hipural, a nivel de la línea lateral, (13) Inserción dorsal de la aleta caudal, (14) Inserción posterior de la aleta dorsal.

Claramente, con estos 83 individuos no se cumplen los requerimientos de número de muestras por variable, pues son 28 variables, y el cociente entre el tamaño de la muestra y las variables, es de 2,964, que están algo lejos del valor de, al menos 5 veces, señalado por Strauss.

Se realizó un Análisis Multivariado de Varianza (MANOVA), y se determinó su significancia estadística con las pruebas Lambda de Wilks y Traza de Pillai, luego se graficaron los puntajes de un Análisis de Coeficiente de Variación (CVA) entre grupos.

Teniendo en cuenta los resultados con estos primeros datos, y sabiendo que se tenían problemas de falta de datos, se llevaron a cabo una serie de pasos para amplificar, a través de pseudorréplicas, los tamaños muestrales, bajo un modelo  $Y = X + \varepsilon$  (Roff, 2006), donde  $X$  es el valor promedio de la variable, y  $\varepsilon$  es la desviación estándar de la variable, multiplicada por un valor aleatorio de una distribución normal con media= 0 y Varianza= 1,

Para ello, primero se sometieron a consideración dos hipótesis:

**1) Las coordenadas, cuando se transforman a Procrustes, presentan una distribución normal para cada variable X y cada variable Y.** Para ello a cada población analizada se le realizó la prueba de Jarque-Bera, por cada variable, para lo cual se usó el programa PAST®, y los valores de probabilidad obtenidos, fueron sometidos a una prueba condicional en Excel®, en la que se le daba valores de ceros y unos, donde cero equivalía a una distribución normal para

esa variable, definida por aquellos valores de la prueba de Jarque-Bera que eran de una significancia  $> 0,05$  y uno, si el valor de probabilidad de la prueba de Jarque-Bera era  $< 0,05$  y, por tanto, la distribución no era Normal. Luego se sumaron todos los valores, y esta sumatoria fue contrastada a través de una prueba Binomial, en Excel®, donde la probabilidad de éxito era 0,05 y los ensayos eran el número de variables consideradas, en este caso 28 por población, y 140 en total para las cinco poblaciones.

**2) Las poblaciones relacionadas entre sí, presentan varianzas similares.** Esta es una hipótesis seguida, casi sin verificar, por muchos autores, asumiendo que las diferencias, en grupos relacionados, en los valores de las distintas variables se encuentran entre las medias, y no en las varianzas (Day, 2005). Así, para cada variable, se realizó un análisis de Homogeneidad de Varianzas, con el programa MINITAB®, éste presenta dos pruebas, la de Bartlett y la de Levene, siendo esta última la recomendada para este caso, con menos de 50 muestras por grupo (Van Valen, 2005). En caso de existir varianzas diferentes, se remueve la(s) población(es) discordante(s) y se procede a calcular nuevamente la prueba, con las poblaciones restantes. Se suman las varianzas discordantes, y se someten a una prueba binomial, similar a como se procedió con la hipótesis de Normalidad.

Esta prueba sirve para justificar el uso una varianza ponderada, para cada variable, y usarla para todas las poblaciones. En caso de que las variables tengan distribución normal, y tengan varianzas similares, se procede a desarrollar el modelo propuesto.

**Generación de las pseudorréplicas:** Si las variables tienen una distribución normal, y si tienen varianzas similares, entonces es posible generar una serie de pseudorréplicas, para aumentar los tamaños muestrales. Para ello se genera, en Excel®, una distribución normal con Media = 0, y Desviación Estándar = 1, con la función de distribución normal inversa y la función de números aleatorios como insumo. Para este procedimiento, se usó Excel®, por su versatilidad y rapidez, como hoja de cálculo. Luego se procedió a realizar la simulación de los cálculos de los valores de cada variable, aplicando la fórmula señalada anteriormente del modelo  $Y = X + \varepsilon$  (Roff). Se usaron tres valores para la desviación estándar: a) el valor obtenido a partir de los datos reales, b) un valor ponderado, calculado a partir del cociente entre el número de individuos de cada población y el número total de individuos examinados, multiplicado por la desviación estándar real de la variable para cada una de las cinco poblaciones y c) con las probabilidades usadas en el cálculo de la desviación estándar ponderada, de cada una de las variables, que, ahora, para cada población, fueron some-

tidas a un Bootstrap, de 2000 repeticiones, con el programa Minitab® siguiendo la metodología de Brase & Brase (2012) y calculada la media de los valores. Para cada una de estas tres desviaciones estándar, se generaron 100 muestras, por cada población, para un total de 1500 muestras usadas para aplicar las diferentes pruebas MANOVA.

Se realizaron pruebas de consistencia interna de los datos generados, que se basaron en la hipótesis, de que con una probabilidad de  $\alpha < 0,05$ , los datos generados, se distribuyeran fuera del ámbito del 95 % de confianza estimado para una distribución normal, con la media real, y las desviaciones estándar calculadas.

Estos datos generados, entonces, deben cumplir con las siguientes condiciones:

**A) Son consistentes con el nivel  $\alpha$  de variación, con el que se generó la muestra.** Buscando la consistencia interna de los datos producidos, se evaluó la probabilidad de ocurrencia de los valores, a partir de una distribución normal, con la media de los valores reales, y las desviaciones estándar reales, ponderadas y ponderadas con Bootstrap. Para cada población se generaron 100 muestras virtuales, y en la verificación de la consistencia de los datos generados, se sometieron los valores de cada celda a una prueba a ver si su probabilidad de ocurrencia de  $\alpha < 0,05$ , bajo una distribución normal, usando la media real y la desviación estándar de cada variable (bajo cada una de las tres presentaciones), estaba dentro del 95 % esperado, (valor 0), o si estaba en el 5 % restante (valor 1). Luego se sumaron todos los datos con valor 1 y se dividieron entre 2800, que es el número total de celdas generadas, 100 muestras, con 28 variables. Este paso se repitió 50 veces, para un total de 140000 réplicas independientes, a través de una hoja de cálculo Excel®, para cada tratamiento, tres por población, para un total de 15, con 750 valores de probabilidad calculados. En total, para cumplir este paso, se generaron 2100000 valores de celdas.

Esos 750 valores de probabilidad, fueron sometidos a ocho pruebas ANOVA con Minitab®, una prueba para cada población con los tres tratamientos diferentes, y tres pruebas con los datos de todas las poblaciones para cada desviación estándar, para detectar posibles diferencias entre los tratamientos.

**B) Los datos generados se comparan con los datos reales.** Para contrastar eso, se realizaron pruebas para comparar las diferencias entre los valores reales de cada variable, y los valores generados, en un principio, la prueba t de diferencia entre dos medias, entre valores reales y generados, y para cada variable. Luego se compararon los valores con el

nivel de significancia, a través de una función condicional, si los valores eran menores que el  $\alpha$  establecido de 0,05, se le daba un valor de 1, si no, un valor de 0. Se aplicó después una prueba binomial con el programa Excel®. Esto para determinar si el número de variables con diferencias estadísticamente significativas era, a su vez, estadísticamente significativo.

**C) Los datos generados deben cumplir la condición de que para cada individuo, el tamaño del centroide debe tener un valor de 1.** Esta condición, surge porque los datos fueron transformados a coordenadas Procrustes, y una característica de cada muestra, es que el centroide tiene un valor de 1. El tamaño de los centroides se obtiene a partir de la raíz cuadrada de la suma de los 28 valores de las variables, elevados al cuadrado, por cada individuo o pseudorréplica. Acá se evaluaron los centroides de todas las poblaciones replicadas, en total 500 centroides para cada tratamiento, y se sometieron a una prueba t de diferencia entre dos medias, con los centroides de los valores reales de todas las poblaciones. Además, para evaluar indirectamente la idoneidad de los valores generados, se graficó la distribución de frecuencias para todos los centroides generados.

Con los centroides calculados, se procedió a realizar un análisis de regresión, de los componentes X y Y para los 83 valores reales, y se compararon con 249 valores generados, 83 por cada tratamiento, guardando las mismas cantidades de muestras por localidad.

**D) Pruebas Post Hoc:** Para los datos reales y generados de cada grupo, se realizó un análisis con la función MANOVA/CVA, de PAST®, para cuatro restricciones, por ser datos de dos dimensiones con transformación a variables Procrustes, donde se comparan los valores reales con los generados, usando la desviación estándar real, y las dos ponderadas, para ver si se observaban diferencias apreciables entre las muestras producidas y las reales. Aquí, se aplican varias pruebas, primero está la Lambda de Wilks y la Traza de Pillai, para probar si hay diferencias en las muestras, luego, se sigue con el Test de Hottelling, para comparaciones pareadas de cada par de tratamientos dentro de cada grupo, con la corrección de Bonferroni, que sirve para darle más robustez a las pruebas, disminuyendo la posibilidad de errores de Tipo II, que son los errores potencialmente más peligrosos (Dytham, 2011). Estas pruebas, en este caso intragrupal, se realizaron para determinar si se podía considerar a los datos generados, como equivalentes a los datos reales.

Se compararon, ahora, conjuntamente, los cuatro tratamientos, de las cinco poblaciones, a través de un MANOVA/CVA, siguiendo la misma técnica anterior, pero esta vez además de las pruebas pareadas de Hotelling, allí

también se compararon los valores de las distancias de Mahalanobis tomadas de los valores reales, con las distancias de los valores generados, a través de tres pruebas de análisis de correlación de matrices de distancias de Mantel, que fueron realizadas con el programa BioEstat5®, para ver si en las distancias de Mahalanobis, había una estructuración diferente entre los valores reales y los valores construidos. Con los valores de cada tratamiento, se construyeron las matrices simétricas de cinco filas por cinco columnas, para ser sometidas a cada prueba.

Por último, se graficaron los puntajes de los dos primeros ejes, de cada población, para cada uno de los conjuntos de muestras generadas, para compararlos con los valores reales.

## RESULTADOS

El análisis MANOVA/CVA, para los datos reales, produjo unos valores altamente significativos, para el Lambda de Wilks (0,001956,  $p \ll 0,0001$ ) y la Traza de Pillai (3,043,  $p \ll 0,0001$ ) sin embargo, en las pruebas pareadas de Hottelling, con corrección de Bonferroni, sólo las poblaciones FC y BO, presentaron diferencias significativas entre ellas, mientras los demás pares de poblaciones no presentaron diferencias significativas en sus valores, y otros, simplemente, no presentaron ningún tipo de valor, sin embargo los puntajes del CVA ubican las muestras, de cada población, agrupadas en diferentes posiciones dentro del Morfoespacio (Fig. 3). Otro de los problemas, fue que las elipses de 95 % de confianza se solapaban grandemente, y estas figuras geométricas son importantes para visualmente observar el grado de discriminación de poblaciones anima-

les, tal como fueron usada, con simulación de Bootstrap incluida, para fines taxonómicos por Von Zuben *et al.* (1998).

Por tanto, como no tenía el número mínimo, de muestras, recomendado de, al menos, cinco veces el número de variables que el número de variables para llevar a cabo el MANOVA, se procedió a la realización de todo el protocolo para la generación de las muestras:

Los resultados de las 140 pruebas de Jarque-Bera con PAST®, mostraron sólo 4 variables con distribución diferente a la Normal, para una proporción de 0,0286, luego se aplicó una prueba Binomial con Excel® (Prueba Binomial,  $\alpha=0,05$ ,  $p=0,834$ , n.s.). Por tanto, se asume que las variables tienen distribución normal.

El próximo paso fué la verificación de que las cinco poblaciones tienen varianzas similares, para cada una de las variables, aplicando la prueba de Levene, con el programa Minitab®, se encontraron 2 varianzas diferentes, del total de 140 varianzas analizadas, con una proporción de 0,0143, (Prueba Binomial,  $\alpha=0,05$ ,  $p=0,973$ , n.s.). Por tanto se asume que las varianzas son iguales para cada una las variables analizadas en las cinco poblaciones.

Dado que se puede inferir, con alto grado de certeza, que los valores de las variables se distribuyen normalmente, y que las varianzas son iguales, entonces se procede a generar los tres conjuntos de datos para cada población: simulados con la varianza real de cada variable, simulados con varianza ponderada, y simulados con Bootstrap de la varianza ponderada.

Una vez generados los datos, se encontró que no existieron diferencias significativas entre la prueba t de diferencia entre dos medias, cuando se compararon los valores reales, para cada variable y población, con los valores generados: simulados con la varianza real de cada variable, simulados con varianza ponderada, y simulados con bootstrap de la varianza ponderada. En otras palabras, ninguna diferencia significativa para 420 pruebas, desglosadas en 28 pruebas para cada una de las cinco poblaciones, realizado en tres oportunidades distintas, por existir tres grupos de valores generados (Prueba Binomial,  $\alpha=0,05$ ,  $p=1$ , n.s.). Los datos generados no se diferencian de los valores reales.

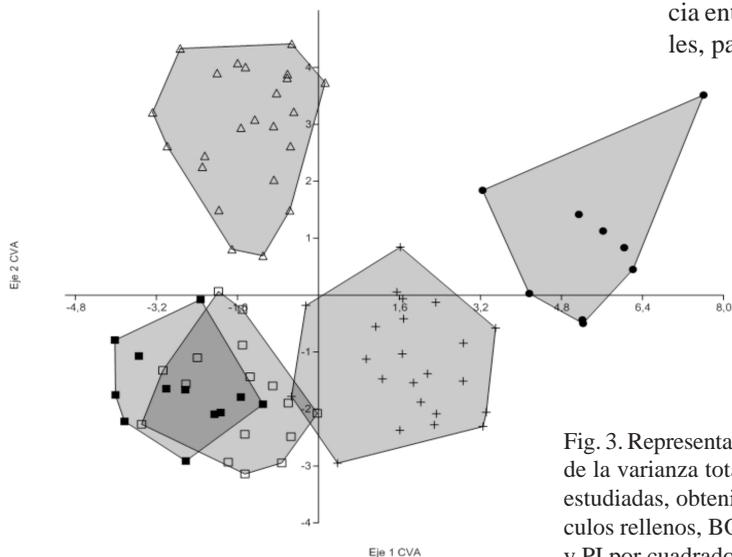


Fig. 3. Representación gráfica de los puntajes del CVA, que explican el 76.61% de la varianza total de la muestra, para los valores reales de las poblaciones estudiadas, obtenida con el programa PAST®. AN está representada por círculos rellenos, BO por triángulos, FC por cruces, GU por cuadrados rellenos y PI por cuadrados.

Tabla I. Resultados de las pruebas de normalidad y para los cocientes fuera del ámbito del 95% de confianza, para cada una de las poblaciones, de la suma de los tres tipos de datos generados con las desviaciones estándar real, ponderada y ponderada con Bootstrap.

Población (n=150)	Prueba de normalidad de Anderson-Darling	Media con intervalos de confianza de 95 %	Mediana con intervalos de confianza de 95 %	ANOVA
AN	AD=0,34 p=0,495 (n.s.)	0,04991 (0,04921-0,05061)	0,0500 (0,04893-0,05071)	p=0,157 (n.s.)
BO	AD=0,38 p=0,403 (n.s.)	0,04898 (0,04832-0,0506)	0,04857 (0,04805-0,04929)	p=0,484 (n.s.)
FC	AD=0,26 p=0,707 (n.s.)	0,05071 (0,04918-0,05095)	0,0500 (0,04948-0,05107)	p=0,071 (n.s.)
GU	AD=0,28 p=0,634 (n.s.)	0,04924 (0,04859-0,04989)	0,04929 (0,04937-0,0500)	p=0,678 (n.s.)
PI	AD=0,32 p=0,528 (n.s.)	0,05008 (0,04936-0,05081)	0,0500 (0,04948-0,05107)	p=0,600 (n.s.)

AD es el valor de la prueba de Anderson-Darling; p es la probabilidad; n.s. es no significativo.

Las pruebas de consistencia interna de los datos generados, se basaron en la hipótesis, de que con una probabilidad de  $\alpha < 0,05$ , existen datos generados que se distribuyeran fuera del ámbito del 95 % de confianza estimado para una distribución normal, con la media real, y las desviaciones estándar calculadas.

Comparando los cocientes generados bajo cada tratamiento, para todas las poblaciones, se tiene que, a) presentaron una distribución normal, b) los valores estuvieron alrededor de 0,05, y c) no se encontraron diferencias significativas en los resultados de las pruebas ANOVA donde se comparaban los cocientes, dentro de cada población, para los datos generados con los tres tipos de desviación estándar (Tabla I).

Los valores de los cocientes, agrupando conjuntamente las cinco poblaciones, para cada una de las tres desviaciones estándar, presentan un comportamiento parecido a las pruebas de la sección anterior, los valores nuevamente se agrupan en distribuciones normales, las medias y medianas alrededor de 0,05 (Fig. 4), y las pruebas ANOVA presentan resultados no significativos (Tabla II). Sin embargo, cuando se suman todos los 750 valores, y se aplica una última prueba ANOVA, ésta resulta ser significativa a un nivel de  $\alpha < 0,05$ . Cuando se realiza una prueba t entre dos medias, se tiene que las diferencias se deben a la comparación entre los valores simulados y los ponderados con bootstrap ( $t= 2,57, 495$  gl,  $p= 0,011$ ).

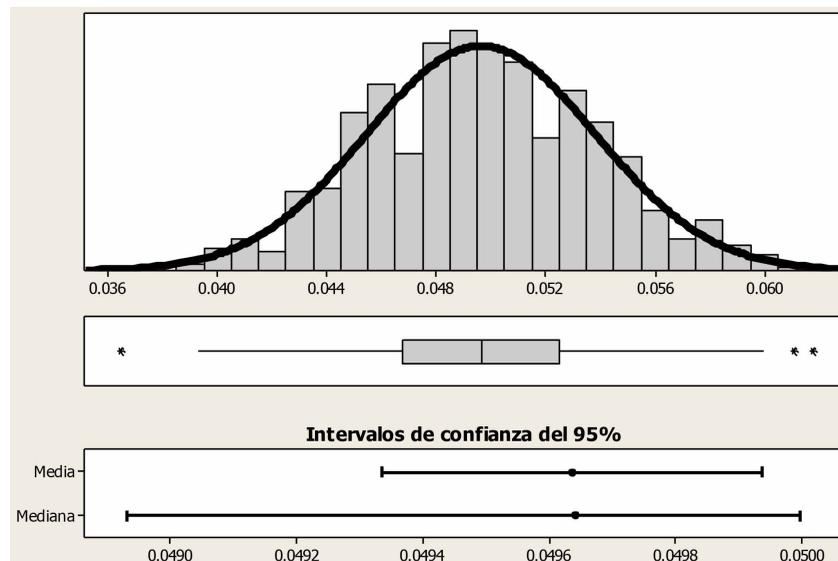


Fig. 4. Distribución de las probabilidades de error, entre los valores generados, para todos los tratamientos combinados.

La próxima prueba, que fué de carácter deductivo pues, partiendo de la premisa de que los tamaños de los centroides para cada muestra transformada a coordenadas Procrustes, tienen un valor de 1, se deduce que los centroides de los datos generados, estarían distribuidos alrededor de ese valor. Los resultados muestran una distribución normal para los valores de los centroides (prueba de Anderson-Darling= 0,24,  $p= 0,790$ , n.s.), con unos valores similares para la media y mediana (1,0001), e igual intervalo de valores con 95 % de confianza (0,9998-1,0004) (Fig. 5). Además, como el resultado de la prueba t de diferencia entre dos medias, entre los centroides reales y todos los generados ( $t=1,31, 1553$  g.l.,  $p= 0,191$ , n.s.), se puede decir, con certeza, que estos conjuntos de valores no difieren entre sí.

Tabla II. Resultados de las pruebas para los cocientes fuera del ámbito del 95% de confianza, de los tres tipos de datos generados con las desviaciones estándar real, ponderada y ponderada con Bootstrap. AD es el valor de la prueba de Anderson-Darling; p es la probabilidad;

Muestra	Prueba de normalidad de Anderson-Darling	Media con intervalos de confianza de 95 %	Mediana con intervalos de confianza de 95 %	ANOVA
Simulada (n=250)	AD=0,32, p=0,528 (n.s.)	0,05011 (0,04957-0,05071)	0,0500 (0,04929-0,05071)	p=0,600 (n.s.)
Ponderada (n=250)	AD=0,38, p=0,402 (n.s.)	0,04965 (0,04912-0,05017)	0,04964 (0,04857-0,05036)	p=0,067 (n.s.)
Ponderada Bootstrap (n=250)	AD=0,25, p=0,739 (n.s.)	0,04915 (0,04865-0,04965)	0,04893 (0,04857-0,04964)	P=0,835 (n.s.)
Todos Generados (n=750)	AD=0,51, p=0,195 (n.s.)	0,04964 (0,04934-0,04999)	0,04964 (0,04893-0,05000)	P=0,039*

n.s. es no significativo, \* es significativo a nivel de  $\alpha < 0,05$ .

Una vez determinada la consistencia interna de los datos, se procedió a la aplicación de la prueba MANOVA, para las comparaciones entre los valores reales y los tres tratamientos de los valores generados, para cada una de las cinco poblaciones, y no se encontraron diferencias los valores producidos y los reales, incluso la probabilidad de todas las muestras pareadas, fue de 1, y el poder de explicación de la varianza de los dos primeros ejes del CVA, fue de mínimo 96,82 % (Tabla III).

Un ejemplo de cómo se representan, en el morfoespacio, los puntajes de los cuatro tratamientos de una población, es presentada en la Figura 6, donde el centro de cada elipse de 95 % de confianza, está ubicado muy cerca de las coordenadas del punto de origen (0,0), y la elipse de los valores reales, abarca a las otras tres elipses.

Luego de llevar a cabo la prueba MANOVA, entre las cinco muestras reales, y las 15 muestras simuladas, las distancias de Mahalanobis, para los seis pares de cada población, dieron valores menores a 1, Las pruebas de Mantel, fueron realizadas a partir de las distancias de Mahalanobis calculadas, para los pares interpoblacionales y todas resultaron tener correlaciones altamente significativas (Tabla IV). Las distancias de Mahalanobis, calculadas en Minitab®, con la función Análisis Discriminante, a partir de los valores reales de las cinco poblaciones, también dieron correlaciones altamente significativas con las matrices de las distancias de las muestras generadas (Tabla IV), y también estuvo altamente correlacionada con la matriz real ( $r=0,9030$ , 8 g.l.,  $p < 0,001$ ).

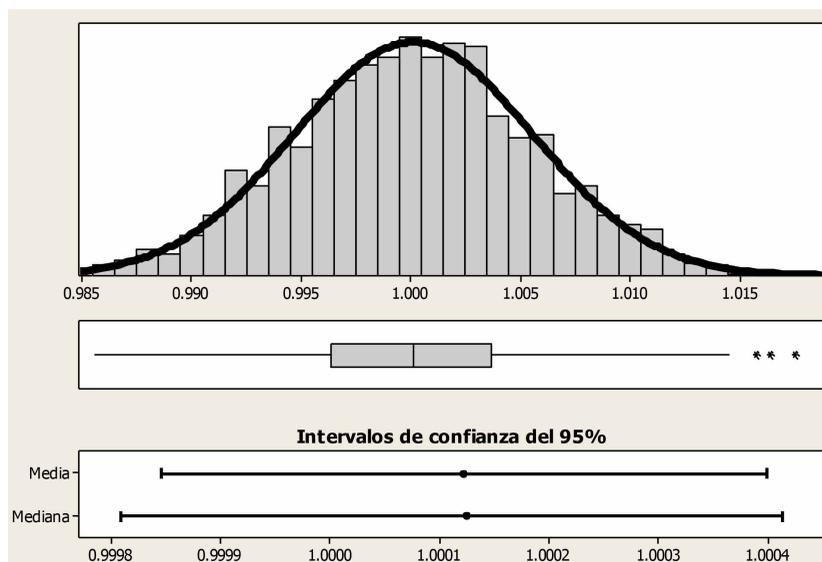


Fig. 5. Distribución de los valores de los centroides para todas las muestras generadas.

Tabla III, Valores de probabilidades para las pruebas de Manova, para cada población, entre los cuatro tratamientos realizados, Lambda de Wilks y Traza de Pillai muestran las probabilidades generales de la prueba; Hotelling-Bonferroni se refiere a el promedio de la probabilidad de similitud de las muestras pareadas de los valores reales vs los tres conjuntos de datos generados; Porcentaje Explicado, se refiere al porcentaje de la varianza explicado por los dos primeros ejes del Morfoespacio.

Población	Lambda de Wilks	Traza de Pillai	Hotelling- Bonferroni	Porcentaje Explicado
AN	0,2262	0,2646	1	99,18
BO	0,9735	0,9762	1	98,37
FC	0,9467	0,9511	1	96,89
GU	0,7758	0,7978	1	98,59
PI	0,9863	0,9877	1	97,71

Tabla IV, valores del coeficiente de correlación, con 8 g.l., para la prueba de las matrices de Mantel, entre las distancias de Mahalanobis, calculadas en el MANOVA de todas las poblaciones, para los cuatro tratamientos, más un promedio de los tres tratamientos generados. Reales Minitab, se refiere a los valores de la muestra original, que no se pudieron calcular en PAST®, y se calcularon con la función Análisis Discriminante de Minitab®. \*\*\* es significativo a nivel de  $\alpha < 0.001$

Datos	Simulados	Ponderados	Bootstrap	Promedio
Reales	0,9952***	0,9954***	0,9961***	0,9985***
Simulados		0,9897***	0,9963***	
Ponderados			0,9880***	
Reales Minitab	0,9063***	0,8734***	0,9208***	0,9010***

Las tres pruebas de MANOVA, para todas las poblaciones, con las muestras generadas bajo las tres desviaciones estándar, dieron valores de probabilidades para las pruebas Lambda de Wilks y Traza de Pillai de  $p = 0$ , y todas las comparaciones de pares de poblaciones tuvieron diferencias, como mínimo de  $p < 10^{-51}$ ,

Los puntajes del CVA, para los dos primeros ejes de variación, para cada conjunto de datos generados, se muestran en las Figuras 7, 8 y 9. Cuando estos gráficos se comparan con el gráfico de los valores reales (Fig. 3), se encuentra que los puntajes, con la excepción de los de PI, mantienen su posición en el plano, éstos últimos parecen haber migrado desde las cercanías a GU hasta proximidades de BO, sin embargo, las probabilidades de pertenecer ambas a la misma muestra, están por el orden de  $p < 10^{-51}$ , lo que es muy improbable, que sean de la misma población.

Aún cuando PI y BO, tuvieron la mayor proximidad aparente, de entre todos los pares de especies en el Morfoespacio, en los tres gráficos de las muestras generadas (Figs. 7, 8 y 9). Sin embargo en el tercer eje, la separación ocurre y ambas poblaciones se segregan muy bien en el morfoespacio (prueba  $t = 19,20, 197$  g.l.,  $p < 0,0001$ ), así en el gráfico de los valores Ponderados (Fig. 8), que es donde se ven mas cercanos, los valores de la prueba Hotelling es del orden de  $10^{-75}$ , y la distancia de Mahalanobis es más del doble de la distancia entre BO y GU, que aparecen muy separados en el morfoespacio, para los dos primeros ejes, en ese mismo gráfico. También en el promedio de los datos

Tabla V. Valores promediados de las distancias de Mahalanobis, para los tres tipos de datos generados, entre los diferentes pares de poblaciones.

	AN	BO	FC	GU	PI
AN	0	59,1098	25,6483	59,9242	53,7414
BO	59,1098	0	36,1641	17,1391	33,2311
FC	28,8444	39,5528	0	23,4597	34,2331
GU	64,9179	17,3996	23,4597	0	28,4070
PI	58,9040	34,3069	34,2331	28,4070	0

de las tres distancias (Tabla V), se observa que el par BO-GU es el que tiene la menor distancia entre sí. Con las distancias reales calculadas con el Análisis Discriminante, la más pequeña, entre dos pares de especies, correspondió al par PI-GU lo cual se corresponde con lo observado en la Figura 3.

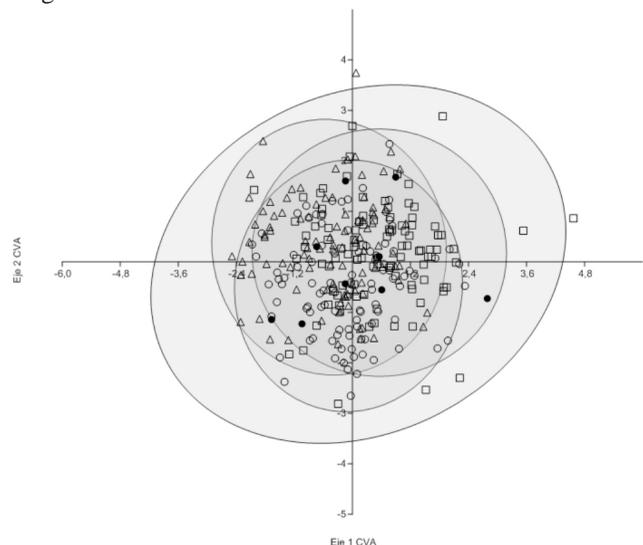


Fig. 6. Representación gráfica, en los dos primeros ejes del Morfoespacio, de la población de AN, que explican el 99.18 % de la varianza para el Coeficiente de Variación Canónica (CVA) de los valores reales (círculo relleno), simulados (cuadrado), ponderados (círculo) y ponderados con Bootstrap (triángulo), usando la función MANOVA del programa PAST®, las elipses representan la región de 95% de confianza.

La Figura 10, muestra la ubicación de las coordenadas de los centroides en el plano cartesiano, en ella se puede apreciar que los valores reales se alinean alrededor de una recta, mientras los valores generados, se encuentran dispersos. La recta de regresión para los valores reales presenta una ecuación de  $Y = 3,5734 - 3,4309X$  ( $r = 0,9974$ ,  $F = 15384,27$ ,  $p < 0,0001$ ), y la de todos los valores generados es de  $Y = 0,7476 - 0,4880X$  ( $r = 0,3170$ ,  $F = 27,60$ ,  $p < 0,0001$ ), con una diferencia entre las pendientes ( $t = -9,3913$ , 328 gl,  $p$

< 0,0001). Además, si se comparan los ámbitos de variación del 95 % de confianza, con la amplitud de la variación, para la pendiente de la recta para los valores reales ( -3,486 a -3,376, 0,09) y los generados (-0,670 a -0,306, 0,364), y para

los interceptos es (3,520 a 3,626, 0,106) y (0,573 a 0,922, 0,347) respectivamente. Los ámbitos para cada ítem no se solapan, y la amplitud de la variación es, al menos, tres veces mayor para los datos generados.

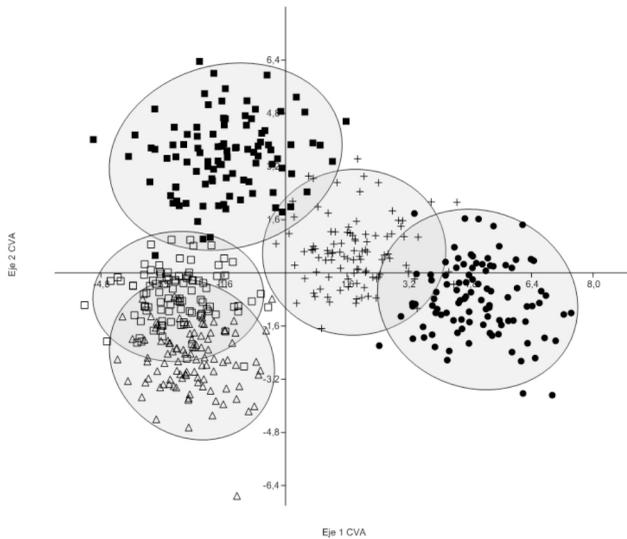


Fig. 7. Representación gráfica de los puntajes del CVA, que explican el 74.81% de la varianza total de la muestra, con las elipses del 95% de confianza, para los valores simulados de las poblaciones estudiadas, obtenida con el programa PAST®. AN está representada por círculos rellenos, BO por triángulos, FC por cruces, GU por cuadrados rellenos y PI por cuadrados.

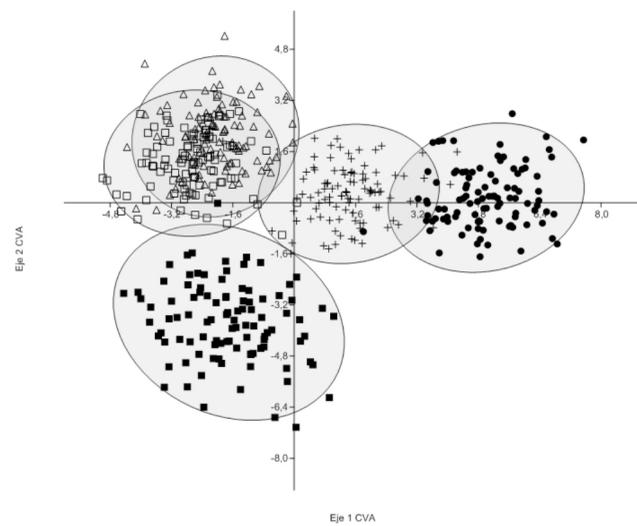


Fig. 9. Representación gráfica de los puntajes del CVA, de los dos primeros ejes, que explican el 75.48% de la varianza total de la muestra, con las elipses del 95% de confianza, para los valores ponderados con Bootstrap de las poblaciones estudiadas, obtenida con el programa PAST®. AN está representada por círculos rellenos, BO por triángulos, FC por cruces, GU por cuadrados rellenos y PI por cuadrados.

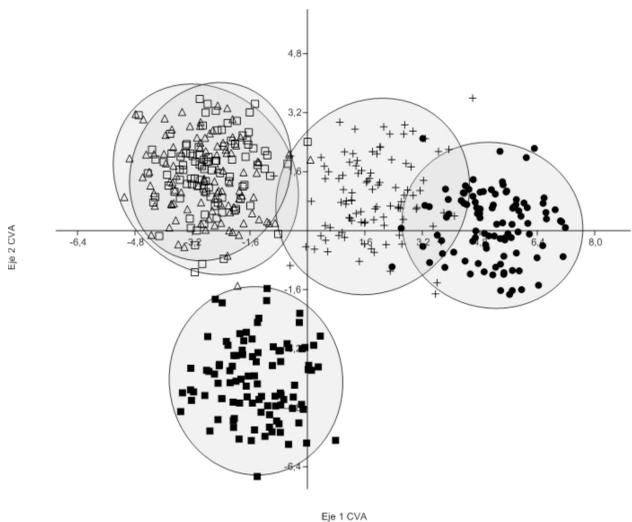


Fig. 8. Representación gráfica de los puntajes del CVA, que explican el 77.92% de la varianza total de la muestra, con las elipses del 95% de confianza, para los valores ponderados de las poblaciones estudiadas, obtenida con el programa PAST®. AN está representada por círculos rellenos, BO por triángulos, FC por cruces, GU por cuadrados rellenos y PI por cuadrados.

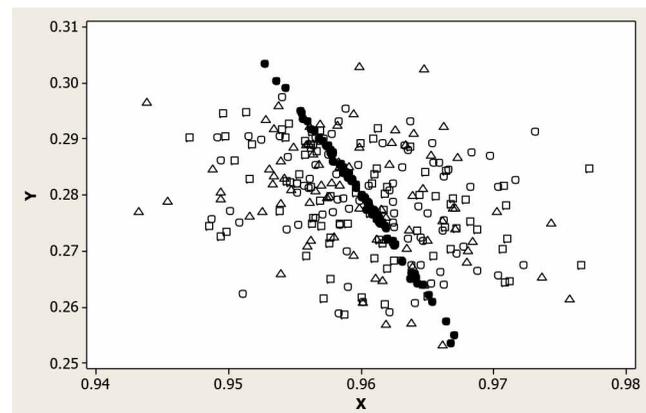


Fig. 10. Gráfico de dispersión de las coordenadas de los centroides, los círculos rellenos representan las muestras reales; los cuadrados vacíos, datos simulados; los círculos vacíos, datos ponderados; y los triángulos vacíos, datos ponderados con Bootstrap.

## DISCUSIÓN

Teniendo en cuenta los siguientes resultados, que sirven para confirmar, o para no refutar, las hipótesis propuestas: a) Las coordenadas morfométricas, cuando son transformadas a coordenadas Procrustes, presentan una distribución Normal; b) las varianzas de las variables Procrustes, son similares entre poblaciones relacionadas. Por tanto, se pueden generar toda una serie de pseudorréplicas de las muestras reales, para aumentar los tamaños muestrales, a través de una distribución Normal.

Luego de generados estos nuevos datos, y ser sometidos a diversas pruebas directas contrastables con los datos reales tales como: a) prueba t de diferencia entre dos medias; b) prueba Binomial; c) Lambda de Wilks y Traza de Pillai, d) pruebas pareadas de Hotelling con corrección de Bonferroni. No se encontraron diferencias entre los valores reales y los valores generados con varianza real, varianza ponderada simple y varianza ponderada con Bootstrap.

Los datos generados también fueron sometidos a pruebas indirectas como: a) la distribución de los cocientes de errores; b) el promedio de esos errores; c) la distribución de los centroides y el valor promedio de éstos; d) la prueba t de diferencia entre los centroides de los datos reales y los de los datos generados; e) las matrices de Mantel para evaluar correlaciones de las distancias de Mahalanobis entre las distancias de los datos reales y las de los datos generados. Y no se encontraron diferencias apreciables que permitieran sugerir que la metodología no fue acorde, y que los datos generados se comportan de manera distinta, a los datos reales, para las situaciones a las que fueron generados.

Los resultados obtenidos tanto de las pruebas directas, como de las indirectas, apoyan la hipótesis de similitud, entre los valores reales y los generados. Con lo cual se puede tener la certeza de que son equivalentes ambos valores y, por tanto, se pueden usar los valores generados, con esta metodología, para acompañar, o sustituir a los valores reales.

Como no hubo diferencias entre los valores reales y los generados, se recomienda, dada su sencillez de cálculo, usar los valores simulados con desviación estándar real, si las varianzas de las poblaciones reales no son iguales, y usar los valores generados con desviación estándar ponderada simple, si las varianzas son iguales. No vale la pena usar la desviación estándar ponderada con Bootstrap, porque no mejora sustancialmente los valores obtenidos con la desviación estándar ponderada simple, y su cálculo es un poco más complicado.

Por la cantidad de evidencia acumulada, a favor de la aplicación de estas pseudorréplicas, se infiere que es completamente factible el empleo de la metodología acá sugerida.

Esta metodología de simulación de muestras, fue diseñada con fines netamente taxonómicos y comparativos, como un intento para resolver el problema ocasionado por tamaños muestrales relativamente pequeños, y donde existe el dilema de reducir las variables usadas, o coleccionar mas muestras, si se tiende a reducir las variables usadas, puede perderse información morfométrica, por no decir biológica, importante, y aumentar el número de individuos coleccionados, o estudiados, puede implicar enormes inversiones de tiempo y dinero, además de que puede ser difícil encontrar nuevas muestras, sobre todo si son poblaciones en declive.

La técnica, acá desarrollada, que sirve para discriminar poblaciones con poca diferenciación morfológica aparente, debe ser usada como complemento a otras técnicas o metodologías de caracterización y evaluación taxonómica de organismos, y no debe ser usada para otros fines. Si entre las variables existen correlaciones entonces, este modelo, sólo puede servir para definir la hipótesis nula, de que entre las variables no existe ninguna correlación, pues los valores fueron generados considerando a cada variable como independiente de las demás.

Los resultados del Análisis de Regresión entre las coordenadas de los centroides, son realmente sorprendentes, para los valores reales se presenta un coeficiente de determinación ajustado es de 0,9947, lo que equivale a decir que el 0,53 % de la varianza mostrada, es producto del azar, mientras que los valores para los datos generados, el coeficiente de determinación ajustado es de 0,0969, lo que equivale a decir que el 90,3 % de la varianza observada es producto del azar.

Existe una disparidad evidente en el valor del estadístico F, entre las rectas de regresión de las coordenadas de los valores reales, y las de los valores simulados, donde existen diferencias para las pendientes, de hasta tres órdenes de magnitud, y de un orden de magnitud para los interceptos. Se puede inferir, entonces, que la recta de regresión de las coordenadas de los centroides, para los datos reales, indica una señal filogenética, puesto que esa alineación de los valores es demasiado precisa, mientras que la distribución de las coordenadas de los centroides de los datos generados acá, con valores estimados para cada variable, independientemente de las demás, representaría, en este caso, la hipótesis nula o hipótesis a contrastar.

Esta señal, que pudiera denominarse Línea de Estasis Centroidal, pues aún cuando se observaron diferencias morfométricas apreciables entre las cinco poblaciones, las coordenadas de los centroides para los valores reales, caen todas en una línea recta muy bien definida. Esta recta, se hipotetiza como una propiedad única para cada taxa, o grupos de taxa relacionados, y sería función de la relación entre los hitos seleccionados y las formas mantenidas por grupos emparentados. Por tanto, se infiere que diferencias significativas en la pendiente y/o el intercepto con el eje Y, usando los mismos hitos, implicaría la existencia de taxa evolutivamente separados. Esto puede ser usado como una evidencia de distintividad taxonómica, tanto a nivel específico, como supraespecífico.

Ahora, regresando a la especie estudiada, se puede decir con un alto grado de certeza, que las poblaciones están evolucionando independientemente, y que tienen una diferenciación morfométrica, definida con los 14 hitos digitalizados, bastante pronunciada, la cual se infiere por las probabilidades obtenidas, tanto en las pruebas generales de Lambda de Wilks y Traza de Pillai, como las pruebas pareadas de Hotelling con corrección de Bonferroni, y las distancias de Mahalanobis.

La Línea de Estasis Centroidal, sugiere una estrecha relación entre las cinco poblaciones, y los datos morfométricos, que indican divergencias estadísticamente apreciables en la forma, son indicio de que se estaría en presencia de un complejo de especies crípticas.

Como la distribución de las poblaciones de este género, se presenta en extremos opuestos y periféricos de la Amazonía, éstas parecieran constituir parte de un linaje muy antiguo, posiblemente originado antes de la formación de la actual cuenca amazónica, y que su separación se corresponde con una elevación de la Cordillera Oriental de los Andes Colombianos en el Mioceno tardío, al menos hace 11 millones de años (Mora *et al.*, 2010).

Entonces, de ser cierta esta hipótesis, las poblaciones de *M. ramirezi*, tienen, al menos, 11 millones de años evolucionando independientemente de su especie hermana. Los resultados obtenidos en este estudio, muestran probabilidades de diferencias interpopulacionales tan pronunciadas que se hace necesario realizar estudios sobre otras dimensiones morfológicas, merísticas, genéticas, ecológicas y biogeográficas, las cuales permitirían apoyar o refutar la hipótesis, surgida de este trabajo, de la existencia de, al menos, cinco poblaciones diferenciadas, que pudieran ser consideradas especies plenas, y ser descritas cuatro de ellas, como especies nuevas, si se acumulan más evidencias.

En caso de demostrarse la singularidad taxonómica de cada una de éstas poblaciones, la especie *M. ramirezi*, sería restringida a la población denominada GU en este trabajo, por estar muy cercana a la localidad típica de la especie, sur de Palenque, Venezuela (Kullander).

---

**GARCÍA-PÉREZ, J. E.** Generating pseudo-replicates, from real morphometric variables, for contrast taxonomic and evolutionary hypotheses: The case of a widely distributed fish in the lowlands of the Venezuelan Orinoquia. *Int. J. Morphol.*, 39(3):1212-1223, 2021.

**SUMMARY:** In order to evaluate the morphometric differences of five populations of *Mikrogeophagus ramirezi* (Pisces: Cichlidae) from the lowlands of the Venezuelan Orinoquia, 14 homologous landmarks from 83 specimens were taken and transformed to Procrustes variates. A MANOVA/CVA test was performed, however, this test did not perform well due to the low number of specimens used, and it was decided to increase the sample size. The probabilistic normal distribution for the variables was determined, and variances of these variables were similar. Subsequently, 100 values for each variable were generated, from the model  $Y = X + \varepsilon$ , being Y the variable value to estimate, X is the real mean value of that variable, and  $\varepsilon$  is the product between real standard deviate and random values of a normal distribution. This simulation was made for three types of standard deviates: the real variables SD, the prorated values from real variables, and mean bootstrap prorated values. The data generated values were validated with the real data, and some inner properties, both before and after applying the MANOVA/CVA test. In all tests, the values generated were no different from the real ones. Using the generated values as a valid surrogate, it was statistically determined that five populations were different morphometrically, and candidate species. On other hand, a regression between real data coordinates was performed, and compared with the generated data. Whilst the real data coordinates arranged on an almost perfect straight line, and were highly correlated, the generated data ones were dispersedly arranged. This fact, did permit infer that this line, the Stasis Centroid Line, is a phylogenetic signal emerging from the real data.

**KEY WORDS:** MANOVA/CVA; *Mikrogeophagus ramirezi*; Monte Carlo Techniques; Procrustes coordinates; Stasis centroid line.

## REFERENCIAS BIBLIOGRÁFICAS

- Bookstein, F. L. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge, Cambridge University Press, 1991.
- Brase, C. H. & Brase, C. P. *Understandable Statistics: Concepts and Methods*. 10a ed. Boston, Books/Cole, Cengage Learning, 2012.
- Day, T. *Modelling the Ecological Context of Evolutionary Change: Deja Vu or Something New?*. En Cuddington, K. & Beisner, B. (Eds.). *Ecological Paradigms Lost: Routes of Theory Change*. San Diego, Elsevier, Academic Press, 2005. pp.273-309.
- Dyham, C. *Choosing and Using Statistics: A Biologist's Guide*. Hoboken, Wiley-Blackwell, 2011.

- Hammer, Ø.; Harper, D. A. T. & Ryan, P. D. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontol. Electron.*, 4(1):1-9, 2001.
- Kullander, S. O. Nomenclatural availability of putative scientific generic names applied to the South American cichlid fish *Apistogramma ramirezi* Myers & Harry, 1948 (Teleostei: Cichlidae). *Zootaxa*, 3131(1):35-51, 2011.
- Lasso, C. A.; Provenzano, F. & Lasso-Alcalá, O. M. Ictiofauna dulceacuñola y estuarina de la cuenca del golfo de Paria, Venezuela: composición y relaciones biogeográficas con la cuenca del Orinoco. *Biota Colomb.*, 11:53-73, 2010.
- Lasso, C. A.; Usma Oviedo, J. S.; Villa, F.; Sierra-Quintero, M. T.; Ortega-Lara, A.; Mesa, L. M.; Patiño, M. A.; Lasso-Alcalá, O. M.; Morales-Betancourt, M. A.; González-Oropeza, K.; *et al.*, Peces de la estrella fluvial inñrida: ríos Guaviare, Inñrida, Atabapo y Orinoco (Orinoquia colombiana). *Biota Colomb.*, 10:90-122, 2009.
- López-Fernández, H.; Honeycutt, R. L. & Winemiller, K. O. Molecular phylogeny and evidence for an adaptive radiation of geophagine cichlids from South America (Perciformes: Labroidae). *Mol. Phylogenet. Evol.*, 34(1):227-44, 2005a.
- López-Fernández, H.; Honeycutt, R. L.; Stiassny, M. L. J. & Winemiller, K. O. Morphology, molecules, and character congruence in the phylogeny of South American geophagine cichlids (Perciformes, Labroidae). *Zool. Scr.*, 34(6):627-51, 2005b.
- Mora, A.; Baby, P.; Roddaz, M.; Parra, M.; Brusset, S.; Hermoza, W. & Espurt, N. *Tectonic History of the Andes and Sub-Andean Zones: Implications for the Development of the Amazon Drainage Basin*. En: Hoorn, C. & Wesseling, F. P. (Eds.). *Amazonia, Landscape and Species Evolution: A Look Into the Past*. London, Wiley-Blackwell, 2010, pp.38-60.
- Roff, D. A. *Introduction to Computer-Intensive Methods of Data Analysis in Biology*. Cambridge, Cambridge University Press, 2006.
- Rohlf, F. *Tps series*. New York, State University of New York, Stony Brook, 2013, Disponible en: <http://life.bio.sunysb.edu/morph/index.html>
- Ruane, S. Using geometric morphometrics for integrative taxonomy: an examination of head shapes of milksnakes (genus *Lampropeltis*). *Zool. J. Linn. Soc.*, 174(2):394-413, 2015.
- Soria-Barreto, M.; Rodiles-Hernández, R. & González-Díaz, A. Morfometría de las especies de Vieja (Cichlidae) en ríos de la cuenca del Usumacinta, Chiapas, México. *Rev. Mex. Biodivers.*, 82:569-79, 2011.
- Strauss, R. E. *Discriminating Groups of Organisms*. En: Elewa, A. M. T. (Ed.). *Morphometrics for Nonmorphometricians. Lecture Notes in Earth Sciences 124*. Heidelberg, Springer-Verlag, 2010, pp.73-91.
- Van Valen, L. *The Statistics of Variation*. En: Hallgrímsson, B. & Hall, B. K. (Eds.). *Variation: A Central Concept in Biology*. Amsterdam, Elsevier, Academic Press. 2005. pp.29-47.
- von Zuben, F. J.; Duarte, L. C.; Stangenhuis, G.; Pessôa, L. M. & dos Reis, S. F. Bootstrap confidence regions for canonical variates: application to studies of evolutionary differentiation. *Biom. J.*, 40:327-39, 1998.
- Young, K. A.; Snoeks, J. & Seehausen, O. Morphological Diversity and the Roles of Contingency, Chance and Determinism in African Cichlid Radiations. *Plos One*, 4(3):e4740, 2009.
- Zelditch, M. L.; Swiderski, D. L.; Sheets, H. D. & Fink, W. L. *Geometric Morphometrics for Biologists: A Primer*. Amsterdam, Elsevier, Academic Press, 2004.
- Zúñiga-Reinoso, A. & Benítez, H. A. The overrated use of the morphological cryptic species concept: An example with *Nyctelia darkbeetles* (Coleoptera: Tenebrionidae) using geometric morphometrics. *Zool. Anz.*, 255:47-53, 2015.

Dirección para correspondencia:

Juan Elías García-Pérez  
Doctorado de Biodiversidad, PRESAV  
Universidad Nacional Experimental  
de Los Llanos Occidentales Ezequiel Zamora  
UNELLEZ, Guanare  
Portuguesa  
VENEZUELA

E-mail: [ecologia2unellez@gmail.com](mailto:ecologia2unellez@gmail.com)

Recibido : 25-03-2021

Aceptado: 18-05-2021