

## EL SESGO DEL OPERADOR EN LA REPLICABILIDAD DE LOS ESTUDIOS TAFONÓMICOS COMPARATIVOS

MATIAS DO NASCIMENTO RITTER, HEITOR FRANCISCHINI, LIDIA AUMOND KUHN,  
NATHÁLIA CARVALHO DA LUZ, FERNANDO HECK MICHELS, ANDERSON LUIZ  
MARTINS DE MORAIS, PROTÁSIO ANTÔNIO VERVLOET PAIM,  
PEDRO LUIS AMMON XAVIER

Programa de Pós-Graduação em Geociências, Instituto de Geociências, UFRGS, Cx.P. 15001, 91501-970,  
Porto Alegre, Rio Grande do Sul, Brasil. *mnr Ritter@gmail.com, heitorfrancischini@hotmail.com,*  
*lidia.kuhn@yahoo.com.br, ncarvalho.luz@gmail.com, fernandohm9011@hotmail.com,*  
*crescermoraes@hotmail.com, protasio.paim@ufrgs.br, xavier.pedroluis@gmail.com*

CLAUDIO G. DE FRANCESCO

Instituto de Investigaciones Marinas y Costeras (IIMyC), CONICET, UNMdP, Juan B. Justo 2550, 7600  
Mar del Plata, Argentina. *cgdefra@mdp.edu.ar*

**ABSTRACT – OPERATOR BIAS AND REPLICABILITY IN COMPARATIVE TAPHONOMIC STUDIES.** The operator effect is a well-known analytical bias already quantified in some taphonomic studies. However, the influence of operator bias in the replicability on taphonomic studies has still not been considered. Here, we quantified for the first time this bias using different multivariate statistical techniques, testing if the operator effect is related to the replicability. We analyzed the results reported by 15 operators working on the same dataset. Each operator analyzed 30 bioclasts (bivalve shells) by site, from a total of five sites, considering the following taphonomic attributes: shell fragmentation, edge rounding, corrosion, bioerosion, and color alteration. The operator effect followed the same pattern reported in previous studies, characterized by a worse correspondence for those attributes having more than two levels of damage categories. However, the effect did not appear to have relation to replicability, because nearly all operators found differences among sites. The binary attribute bioerosion exhibited 83% of correspondence among operators, but at the same time it was the taphonomic attribute that showed the highest dispersion among operators (28%). Therefore, we concluded that binary attributes, despite indicating a reduction of the operator effect diminishes replicability, resulting in different interpretations of concordant data. We found that a variance value of nearly 8% among operators was enough to generate a different taphonomic interpretation, in a Q-mode cluster analysis. The results reported here showed that the statistical method employed influences the level of replicability and comparability of a study and that the availability of results may be a valid alternative to reduce bias.

**Key words:** taphonomic attributes, bioclasts, analytical bias, statistical approaches.

**RESUMO – O efeito do operador é um viés metodológico conhecido e de certa forma já quantificado em estudos tafonômicos.** Contudo, a influência do viés do operador sobre a replicabilidade ainda não foi estimada em estudos tafonômicos. Aqui, pela primeira vez, nós acessamos e quantificamos este viés empregando diferentes análises estatísticas multivariadas, testando se o efeito do operador tem relação com a replicabilidade. Para tal, utilizamos os resultados reportados por 15 operadores sobre um mesmo conjunto de dados. Cada operador analisou 30 remanescentes biológicos (valvas de bivalves) de cinco distintas localidades, considerando os atributos tafonômicos fragmentação, arredondamento da margem, corrosão, bioerosão e cor secundária. O efeito do operador seguiu o mesmo padrão dos estudos previamente realizados, em que atributos com mais de dois níveis de estados categóricos geralmente possuem uma menor correspondência entre os operadores. Contudo, este efeito não parece ter relação com a replicabilidade, pois praticamente todos os operadores encontraram que as localidades amostrais são diferentes entre si. Em outro nível de comparação, enquanto o atributo binário bioerosão apresentou 83% de correspondência entre os operadores, concomitantemente, foi o atributo tafonômico que apresentou a maior dispersão relativa entre os operadores (28%). Nós também encontramos que um valor de variância entre os operadores, de aproximadamente 8%, foi suficiente para gerar um resultado interpretativo tafonômico distinto, em uma análise de agrupamento em modo Q. Os resultados aqui reportados ilustram que o método estatístico utilizado influencia o nível de replicabilidade e de comparabilidade de um trabalho, e que a disponibilização dos resultados pode ser uma alternativa para evitar vieses relacionados aos diferentes métodos estatísticos empregados.

**Palavras-chave:** atributos tafonômicos, bioclastos, vieses analíticos, métodos estatísticos.

## INTRODUCCIÓN

La cuantificación en Paleontología, de manera general, ha permitido la visualización de diversos patrones y la elaboración de modelos predictivos, cada vez más precisos, sobre diferentes aspectos de la vida en el pasado (Kowalewski & Novack-Gottshall, 2010; Sepkoski, 2012; McMenamin, 2016). De modo similar, el empleo cada vez más sistemático de herramientas estadísticas también en Tafonomía, con el objetivo de cuantificar los sesgos del registro fósil, ha mejorado nuestro conocimiento sobre patrones de preservación (Kosnik *et al.*, 2011), fidelidad composicional (Tomašových & Kidwell, 2009) y magnitudes de *time-averaging* entre taxones (Krause *et al.*, 2010), así como entre ambientes sedimentarios distintos (Kidwell *et al.*, 2005). No obstante, también ha sido posible correlacionar variables ambientales (*e.g.* tipo de sedimento, granulometría, profundidad) con variables tafonómicas categóricas, medidas en bioclastos (Powell *et al.*, 2011). El objetivo de estos estudios tafonómicos ha sido evaluar la utilización de firmas tafonómicas como predictoras de variables ambientales (Fürsich & Oschmann, 1993), en las cuales algunos atributos particulares tomados en su conjunto permitan predecir variables ambientales o ambientes deposicionales (tafacies, Parsons-Hubbard, 2005).

Sin embargo, aún existen numerosos estudios tafonómicos que utilizan aproximaciones cualitativas, lo que dificulta la interpretación numérica, e incrementa la subjetividad de los resultados y, consecuentemente, reduce la replicabilidad de los trabajos. De la misma manera, la falta de una estandarización en la cuantificación ha dificultado la comparación y la generalización de los resultados (Kidwell *et al.*, 2001; Kowalewski & Hoffmeister, 2003; Rothfus, 2004). Se puede reconocer una falta de estandarización en la recolección de las muestras, el esfuerzo de muestreo, la fracción granulométrica y el número mínimo de bioclastos para el análisis, así como en la definición de cuáles atributos tafonómicos son los más adecuados para el problema de interés y de qué manera cuantificarlos, de modo de permitir análisis estadísticos robustos y confiables. Esta falta de métodos objetivos y estandarizados ha limitado enormemente la posibilidad de comparación de los resultados entre diferentes estudios (Kidwell *et al.*, 2001; Kowalewski & Hoffmeister, 2003; Rothfus, 2004; Rodrigues *et al.*, 2009).

Todo estudio científico tiene como premisa la replicabilidad, es decir, cualquier otro investigador debería poder repetir un experimento o un análisis cuantitativo o interpretativo de forma independiente con el mismo conjunto de datos utilizado previamente por otro investigador, obteniendo resultados mínimamente semejantes. El concepto de replicabilidad también puede estar relacionado a la viabilidad de que un determinado método pueda ser reproducido por otro investigador con otro conjunto de datos. Si bien teóricamente todos los estudios siguen esta premisa, los sesgos debidos a errores del operador, a la descripción insuficiente de los métodos, a la forma de presentación de los resultados, o simplemente al método estadístico utilizado, pueden violar parcialmente esta condición (*Open Science Collaboration*,

2015), algo que recientemente comenzó a ser considerado en algunas áreas de la ciencia (McNutt, 2014, Leek & Peng, 2015). De manera general, tanto para las percepciones de replicabilidad como para la utilización de métodos objetivos y numéricamente descriptivos, el acceso libre a los datos a través de su publicación es un primer paso para una futura reproducción. Esto incrementa la probabilidad de que un resultado pueda ser comparado con otros estudios independientes (Nosek *et al.*, 2015).

La replicabilidad de los métodos utilizados en Tafonomía es aún desconocida, aunque la utilización de atributos tafonómicos esté ampliamente reconocida en la actualidad (de forma no estandarizada) para diversos grupos de organismos, como los vertebrados (Araújo-Junior *et al.*, 2015), equinoideos (Mancosu & Nebelsick, 2015), gasterópodos (Pisano *et al.*, 2015), bivalvos (Erthal *et al.*, 2015) y foraminíferos (Berkeley *et al.*, 2009), entre otros.

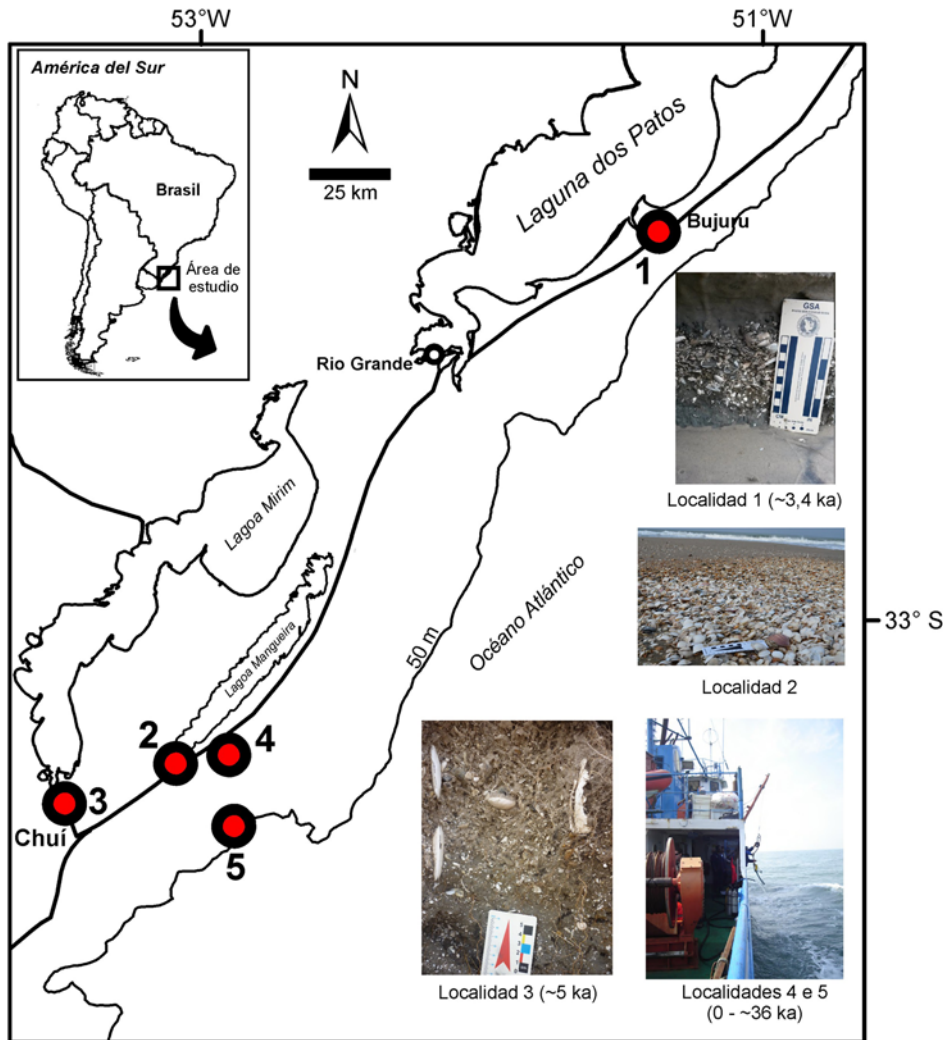
Un sesgo conocido en Tafonomía puede ser atribuido al operador (Rothfus, 2004), es decir, al investigador que analiza las variables tafonómicas en los restos biológicos. Diferentes operadores tienen distintas percepciones y el resultado final puede ser diferente aún para un mismo conjunto de datos. Esto ya fue demostrado tanto para la comparación entre dos operadores (Kidwell *et al.*, 2001), como entre 13 operadores (Rothfus, 2004). Este efecto es lo que denominamos aquí como efecto del operador (también conocido como error inter-observador o error del observador). Es decir, es teóricamente esperable que diferentes operadores atribuyan estados categóricos diferentes a un mismo conjunto de datos. Pero ¿el efecto del operador llevará necesariamente a una interpretación tafonómica diferente? En la presente contribución se cuantifica el efecto del operador y se analiza por primera vez en qué medida este efecto puede interferir en la interpretación de los resultados de un análisis tafonómico comparativo entre diferentes ambientes.

En este sentido, el presente estudio tiene como objetivos evaluar: (i) el sesgo del operador y (ii) la replicabilidad, llevando adelante diferentes análisis independientes sobre un mismo conjunto de datos. Los dos objetivos planteados no actúan como variables dependientes (covariables), es decir que aún asumiendo que el efecto del operador esté presente, la replicabilidad puede ser analizada independientemente.

## MATERIAL Y MÉTODOS

### Origen de las muestras

Las muestras de bioclastos utilizadas en este estudio fueron obtenidas de depósitos sedimentarios marinos y de albufera, a lo largo de la costa del estado de Rio Grande do Sul (RS), sur de Brasil (Figura 1), en los últimos años (2012–2014). Se tomaron muestras sedimentarias de volumen conocido (8 litros) a partir de las cuales se recuperaron los bioclastos (valvas de bivalvos). Para separar los bioclastos del sedimento, se utilizaron tamices de 1 mm de abertura de malla. Los bioclastos fueron posteriormente secados y cuarteados. Específicamente, las muestras utilizadas aquí, numeradas de 1 a 5, fueron obtenidas a partir de una selección



**Figura 1.** Área de estudio y ubicación de las muestras. Edades radiocarbónicas de la localidad 1 están basadas en Dillenburg *et al.* (2004), la localidad 3 en Ritter *et al.* (2014) y las localidades 4 y 5 en Ritter *et al.* (2015).

**Figure 1.** Study area and location map. Radiocarbon ages were based on Dillenburg *et al.* (2004; site 1), Ritter *et al.* (2014; site 3), and Ritter *et al.* (2015; sites 4 and 5).

aleatoria de un cuarto de muestra total en cada localidad (Figura 1). Todos los bioclastos fueron determinados a nivel de género (Kidwell *et al.*, 2001). Las muestras 1 y 3 provienen de depósitos de albufera de edad holocena. Las muestras 2, 4 y 5 representan ensambles de muerte que han sido preservados sobre la interfase agua-sedimento del fondo marino y presentan un amplio rango de edades (por efecto del *time-averaging*; Walker & Bambach, 1971), que va desde 197 a ~36 ka (Pleistoceno) hasta la actualidad (Ritter *et al.*, 2015). La muestra 1 es oriunda de Bujuru, localidad ubicada entre los municipios de Rio Grande y Mostardas; la muestra 2 fue obtenida mediante transectas realizadas en la localidad de los Concheiros, próxima a la playa de Hermenegildo, en el municipio de Santa Vitória do Palmar (Figura 1). La muestra 3 proviene de Arroio Chuí, en el litoral sur del RS. Las muestras 4 y 5 fueron recolectadas por medio de dragas de arrastre (~20 cm de penetración de sedimento) a lo largo de la plataforma interna del sur del estado (Figura 1).

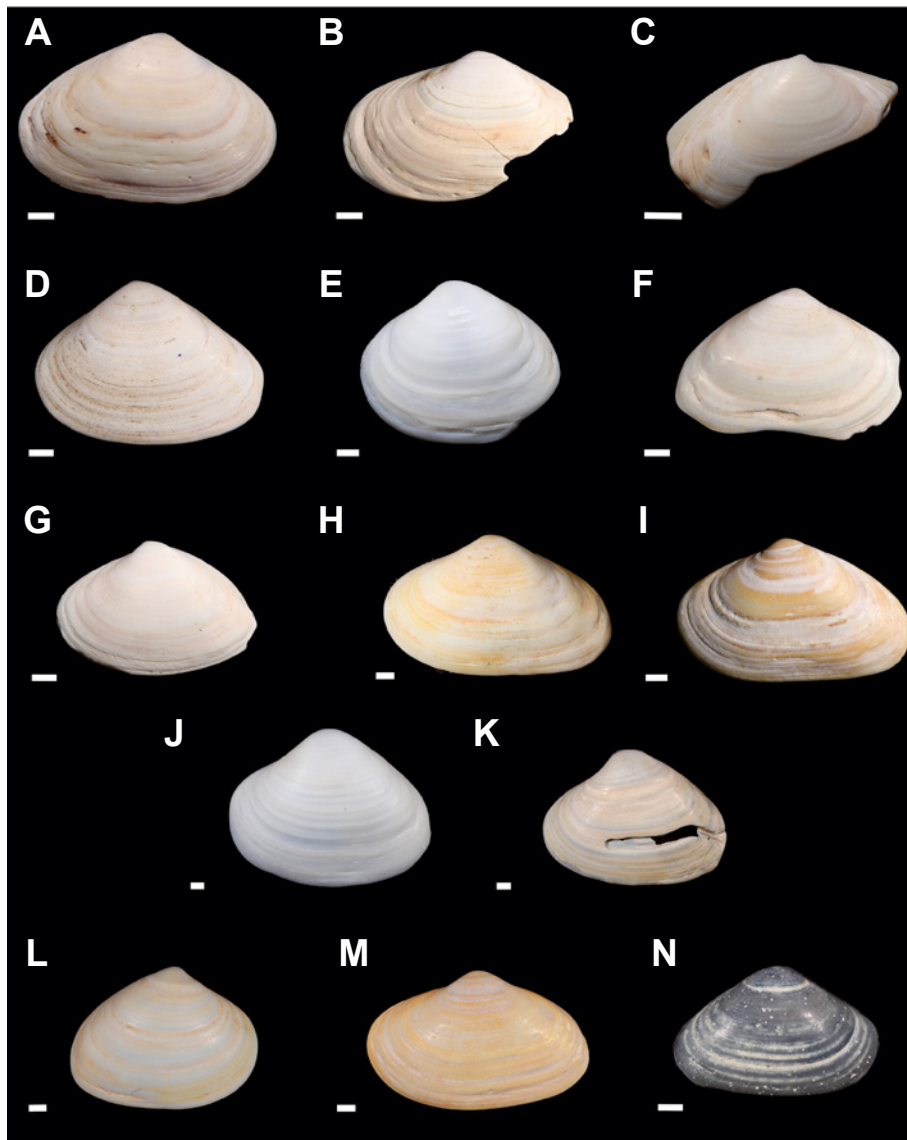
### Análisis tafonómico

Se atribuyeron estados de alteración tafonómica para distintas variables tafonómicas (atributos tafonómicos) utilizando un sistema de categorización semicuantitativa, adaptado de trabajos previos (Tabla 1, Figura 2). Cada atributo tiene un significado exclusivo, aunque su origen puede ser múltiple. Por ejemplo, el atributo corrosión (también conocido como alteración de la superficie a escala fina; Best, 2008) puede estar relacionado a procesos químicos, físicos o biológicos, los que sólo pueden ser distinguidos con observaciones de microscopio electrónico (Best & Kidwell, 2000).

Se analizaron 30 valvas (*Mactra* sp.) por localidad, sumando un total de 150 valvas. Cada espécimen fue observado por los 15 operadores. Las 30 primeras valvas analizadas por el primer observador en cada localidad fueron las que se consideraron como válidas y los restantes 14 observadores analizaron estas mismas muestras. Por ejemplo, el bioclasto número 1 de la localidad 1 fue analizado para todos

**Tabla 1.** Atributos tafonómicos y descripción de las variables categóricas utilizadas en este estudio.**Table 1.** Taphonomic attributes and description of the categorical variables used in this study.

Atributo	VARIABLES CATEGÓRICAS	Información adicional
Fragmentación	ausente (entera) = 0 intermedia = 1 alta = 2	Zuschin <i>et al.</i> (2003) – Figs. 2A, B y C.
Redondeamiento del margen	ausente = 0 intermedia = 1 alta = 2	Ritter <i>et al.</i> (2013) – Figs. 2D, E y F.
Corrosión externa	ausente = 0 intermedia = 1 alta = 2	Ritter <i>et al.</i> (2013) – Figs. 2G, H y I.
Bioerosión (interna y externa)	ausente = 0 presente = 1	Best (2008) – Figs. 2J y K
Coloración secundaria	ausente = 0 tonos cálidos = 1 tonos fríos = 2	Best (2008) – Figs. 2L, M y N.

**Figura 2.** Atributos tafonómicos utilizados en este estudio con sus respectivas categorías (descritas en Tabla 1). Todas las valvas corresponden al género *Mactra*. Escalas = 2 mm.**Figure 2.** Taphonomic attributes used in this study and their categorical variables, which are described in Table 1. All shells belong to *Mactra* genus. Scale bars = 2 mm.

los atributos tafonómicos por la totalidad de los operadores, y así sucesivamente. Kidwell *et al.* (2001) recomendaron analizar un número ideal de 150 bioclastos por localidad. Sin embargo, Ritter *et al.* (2013) no hallaron diferencias estadísticas entre analizar 50 o 200 conchas para estos ambientes estuarinos. Como el objetivo del presente estudio fue cuantificar variaciones entre los operadores y no una cuantificación tafonómica de cada localidad, consideramos que el número de valvas analizado en la presente contribución es suficiente. Estudios similares que estimaron el efecto del operador, inclusive el de Kidwell *et al.* (2001) o el de Rothfus (2004), utilizaron un número menor de bioclastos, así como de observadores (n=2 y n=13, respectivamente) en comparación al presente estudio.

Algunos atributos tafonómicos pueden ser atribuidos a estados (variables tafonómicas) binarios (presencia/ausencia) o presentar diferentes estados, es decir más de dos grados de alteración. En el primer caso, la condición de alteración se considera como 0 (cero) cuando la valva se encuentra en un estado de preservación excelente y 1 (uno) cuando presenta alteración tafonómica. En el caso de atributos con diferentes estados de alteración, se atribuye un valor entero creciente para cada estado (*i.e.* se trata de variables ordinales). Por ejemplo, 0 representa la condición natural (*i.e.* sin alteración), 1 la condición de menor alteración, 2 la condición de alteración mayor que 1 (aunque no necesariamente el doble), y así sucesivamente. En el presente estudio, se analizaron cinco atributos tafonómicos: (1) fragmentación (rotura de la valva, cuantificada como 0=valva intacta, sin alteración), 1=alteración menor (>80% de la valva preservada) y 2=alteración mayor (≤50 de la valva preservada), (2) redondeamiento del margen (cuantificado como 0=margen recto (sin redondeamiento), 1=margen parcialmente redondeado y 2=margen totalmente redondeado), (3) corrosión (desgaste de la superficie de la valva, cuantificada como 0=valva intacta, sin alteración, 1=alteración menor, presencia de pequeñas marcas de desgaste y 2=alteración mayor, importantes marcas de desgaste con acentuación en la ornamentación externa), (4) bioerosión (trazas producidas en las valvas por actividad biogénica, cuantificada como 0=ausente, sin alteración y 1=con trazas de actividad biológica de cualquier tipo) y (5) coloración secundaria (coloración de la valva, cuantificada como 0=blanco, 1=rojizo y 2=gris). Se cuantificó como cero la ausencia de alguno de estos atributos tafonómicos. Estos atributos fueron seleccionados luego de una detallada observación preliminar de la totalidad de las muestras por parte de los 15 operadores, debido a que presentaban variabilidad de estados y eran de fácil observación. Los estados fueron consensuados entre los operadores, a los fines de que todos tuvieran la misma escala de medida.

### Análisis estadísticos

Debido a la naturaleza de los datos (categóricos, con diversos descriptores y con variados objetos), se optó por análisis estadísticos no paramétricos. Cada operador fue considerado una réplica, ya que el análisis fue independiente

entre ellos: es decir que un operador no influenció al otro al momento de cuantificar los atributos tafonómicos. De esta manera, los análisis se focalizaron en las diferencias y las relaciones entre los operadores y las variables tafonómicas, en las diferentes localidades de muestro. El resultado tafonómico interpretativo, por lo tanto, no es el objetivo del presente trabajo, sino comprender de qué manera diversos operadores pueden interferir sobre un análisis tafonómico (el efecto del operador), y cómo esto puede influenciar la replicabilidad de este tipo de análisis en futuros trabajos.

El primer análisis consistió en evaluar el porcentaje de correspondencia entre los operadores para cada atributo tafonómico. Es decir, el porcentaje promedio de operadores que estuvieron de acuerdo en asignar un mismo valor a un determinado espécimen para cada categoría de alteración. Para ello, primero se calculó el número de operadores que coincidieron en asignar un mismo estado tafonómico a cada uno de los 150 especímenes, eligiéndose en cada caso el valor más frecuente (*e.g.* si el espécimen 1 fue valorado como “0” por nueve operadores, como “1” por cinco operadores y como “2” por un operador, el valor final de correspondencia asignado a este espécimen es 9). Luego se calculó el valor promedio de los 150 valores obtenidos, se lo dividió por 15 y se lo multiplicó por 100 (ver Rothfus, 2004). Se utilizó la siguiente fórmula (Fórmula 1):

$$\text{Concordancia (\%)} = \sum \frac{\text{media de concordancia entre los operadores en cada espécimen}}{\text{Número total de operadores}}$$

También se calculó el porcentaje de especímenes que alcanzó un determinado nivel de concordancia entre los operadores (por ejemplo 15 de 15 operadores, 14 de 15, 13 de 15), mediante la siguiente fórmula (Fórmula 2):

$$\text{Especímenes en concordancia (\%)} = \sum \frac{\text{número de especímenes en concordancia}}{\text{Número total de especímenes}} \times 100$$

Para graficar la variación de la concordancia entre los operadores se utilizaron diagramas de caja (*box plots*) con los valores de concordancia media entre los operadores para cada espécimen, por cada variable tafonómica (n=150), en los que fue posible verificar los intervalos intercuartiles y la mediana. Para evaluar si la concordancia fue estadísticamente diferente entre las variables tafonómicas, se realizó una prueba de Kruskal-Wallis (equivalente no paramétrico del ANOVA; Legendre & Legendre, 2012).

Las diferencias entre los operadores para cada localidad y entre las variables tafonómicas fueron evaluadas a través de pruebas de Kruskal-Wallis sobre la media de cada operador por variable tafonómica (alteración tafonómica media), al igual que el resultado individual de cada operador entre las distintas localidades para cada variable tafonómica. Las comparaciones de a pares entre las distintas localidades (por cada atributo tafonómico y por cada operador) fueron realizados con la prueba de Mann-Whitney.

Para evaluar la similitud entre las localidades, así como la dispersión entre los operadores, se realizó un Escalonamiento Multidimensional no Métrico (nMDS, de la sigla en inglés) por cada atributo tafonómico, utilizando la totalidad de las valvas medidas, con base en una matriz con distancia de Bray-Curtis. El nMDS redimensiona las variables en al menos dos ejes ortogonales, con el objetivo de ordenar espacialmente los objetos, colocando aquellos que son más similares entre sí, próximos el uno del otro, preservando las distancias originales entre los objetos, pero en rankings (Gotelli & Ellison, 2012). El valor del *stress* (de la sigla en inglés *STandard RESiduals Sum of Squares*) (Kruskal, 1964) representa cuánto de la información original se perdió al redimensionar las variables en estas dos dimensiones. Convencionalmente, se recomienda un valor de *stress* inferior a 0,20 (Kruskal, 1964). En el presente caso, aquellos atributos tafonómicos que presentan grandes variaciones entre operadores podrían producir valores de *stress* superiores a 0,20, debido a la presencia de *outliers*, lo que no significa que el nMDS sea necesariamente inválido. Las diferencias entre localidades y entre pares de localidades del nMDS fueron calculadas a través de un PERMANOVA (análisis multivariado no paramétrico de la varianza) utilizando Bray-Curtis como índice de distancia, con 9999 permutaciones.

Para analizar la replicabilidad se realizó un Análisis de Componentes Principales (ACP) con el objetivo de establecer una correlación bidimensional entre los operadores, es decir, la varianza entre ellos basada en los ejes ortogonales generados por el ACP. A través de la tabla de *loadings*, generada en el ACP, es posible establecer la varianza máxima y mínima (medida de dispersión) entre los observadores, porque el cuadrado de los *loadings* indica la proporción de la varianza explicada por los componentes, una vez que en un ACP las variables son normalizadas y la varianza de cada eje es igual a 1 (Abdi & Williams, 2010). En este sentido fue posible medir la varianza máxima entre los 15 operadores por atributo tafonómico. En otras palabras, la diferencia entre los valores de *loadings* (al cuadrado) más discrepantes entre sí refleja la varianza máxima entre los observadores, en tanto que la menor distancia representa la varianza mínima entre dos observadores por atributo. Gráficamente, es posible observar en el ACP que la varianza máxima refleja la mayor distancia entre dos observadores, así como la menor varianza refleja también la menor distancia, entre los observadores más próximos, en relación al eje que contiene el autovalor significativo (que explica la mayor varianza entre los datos; véase más abajo sobre *screen plot*, Legendre & Legendre, 2012). La varianza media refleja la media de los *loadings* generados en el ACP para cada atributo.

Por lo tanto, si solamente el eje 1 es significativo y explica el 90% de la varianza de un determinado atributo, la varianza máxima y mínima entre los observadores (*square loadings*) está limitada a la varianza del eje principal en el cual los datos están relacionados. Es importante destacar que el cálculo de la varianza entre los observadores presentado aquí tiene sus limitaciones, porque una vez que el eje significativo explica apenas un 60% de la varianza de los datos, por ejemplo, la

varianza entre los observadores sufre una reducción de un 40% en su resolución total.

Para generar los *scores* y *loadings* del ACP se utilizó una matriz con la media de cada operador por variable tafonómica, considerando cada localidad. Para el ACP se utilizó una matriz de distancia de Manhattan. Para determinar cuántos ejes del ACP son necesarios para representar un alto porcentaje de la variabilidad de los datos, se realizó un gráfico de visualización (*screen plot*), en el que se grafican los autovalores (en %) de cada eje versus una curva generada aleatoriamente. Como criterio, se utilizó el número de ejes que presentó un autovalor superior a los valores generados por una curva aleatoria. Para visualizar la varianza entre los operadores, se realizaron gráficos *box plots* y *jitter plots* entre variables tafonómicas considerando todas las localidades con los valores de los *loadings* (al cuadrado) del eje significativo, en el que es posible visualizar la varianza máxima (cuartil superior) y mínima (cuartil inferior) y la mediana. Para evaluar de qué manera una varianza entre operadores mayor influencia en el resultado final del análisis tafonómico (simulación de una interpretación tafonómica), se utilizó un análisis de agrupamiento basado en una matriz de distancia de Bray-Curtis con 9999 permutaciones sobre el conjunto de datos de cada variable tafonómica, en modo Q, es decir, enfatizando el agrupamiento de las localidades en función de cada variable tafonómica analizada.

Todos los análisis estadísticos fueron realizados con el software PAST v. 3.08 (Hammer *et al.*, 2001). Los datos, figuras y resultados suplementarios utilizados en este artículo son de libre acceso y están disponibles para descarga (y para que cualquiera pueda realizar los análisis *post-hoc*) en el repositorio digital PANGAEA (<https://doi.pangaea.de/10.1594/PANGAEA.859093>).

## RESULTADOS

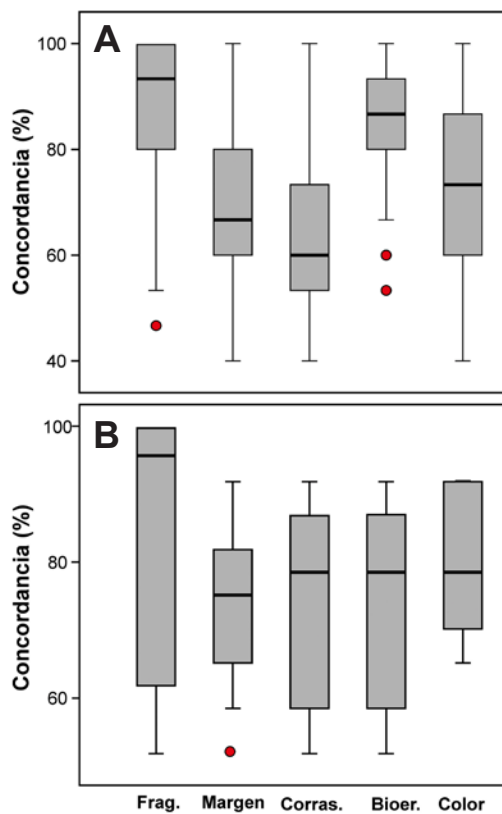
Los operadores presentaron diferentes valores de concordancia según el atributo tafonómico analizado (Tabla 2). La fragmentación y la bioerosión tuvieron los valores más altos, seguidos por la coloración secundaria y el redondeamiento del margen. El menor porcentaje de concordancia correspondió a la corrosión (Tabla 2, Figura 3A). En los atributos fragmentación y bioerosión algunos operadores presentaron valores atípicos de concordancia en relación a los demás (Figura 3A). Los valores de concordancia entre los operadores fueron significativamente diferentes entre los atributos tafonómicos analizados (Kruskal-Wallis,  $p < 0,001$ ; Figura 3A).

La comparación del estado de preservación (evaluado independientemente para cada atributo tafonómico) entre las distintas localidades fue similar entre los 15 operadores (Tabla 3). En particular, para la fragmentación, la corrosión y la coloración secundaria, el 100% de los operadores coincidieron en sus conclusiones, respecto a la existencia de diferencias entre las localidades. Mientras que para el redondeamiento del margen, solamente hubo un 80% de concordancia. La bioerosión fue el atributo que más dificultad presentó, ya que sólo hubo un 60% de concordancia (Tabla 3).

**Tabla 2.** Valores de concordancia (%) entre los 15 operadores para cada atributo tafonómico. Véase también la Figura 3A. **Margen:** redondeamiento del margen; **Color:** coloración secundaria.

**Table 2.** Correspondence scores (%) among the 15 operators for each taphonomic attribute. See also, Figura 3A. **Margen:** edge rounding; **Color:** secondary color.

Concordancia (%)	Atributo tafonómico				
	Fragmentación	Margen	Corrasión	Bioerosión	Color
Media	87,11	69,11	63,73	83,82	73,87
Desvío estándar	14,85	14,84	12,99	14,11	15,01
Mediana	93,33	66,67	60,00	86,67	73,33
Máximo	100,00	100,00	100,00	100,00	100,00
Mínimo	46,67	40,00	40,00	53,33	40,00



**Figura 3.** Box plots indicando el porcentaje de concordancia entre los operadores para cada atributo tafonómico. Los círculos rellenos representan valores de concordancia discrepantes. Datos numéricos en Tablas 2, 4. **A,** concordancia entre los operadores considerando los estados tafonómicos (variables tafonómicas). **B,** concordancia entre los operadores al realizar una prueba estadística comparativa entre cada sitio, por atributo. **Abreviaturas:** **Frag.**, fragmentación; **Margen**, redondeamiento del margen; **Corras.**, corrasión; **Bioer.**, bioerosión; **Color**, coloración secundaria.

**Figure 3.** Box plots illustrating the percentage of agreement among operators for each taphonomic attribute. Filled spots represent discrepant agreement values. Numerical data are available in Tables 2, 4. **A,** correlation between the operators considering the taphonomic states (taphonomic variables). **B,** the correlation between operators when performed a comparative statistical test between each site by taphonomic attribute. **Abbreviations:** **Frag.**, fragmentation.; **Margen**, edge rounding; **Corras.**, corrasion; **Bioer.**, bioerosion; **Color**, secondary color.

**Tabla 3.** Valores de significancia estadística (*p*) de la comparación del estado tafonómico entre las cinco localidades (véase Figura 1) por cada operador individual para cada atributo tafonómico. Los valores fueron obtenidos a partir de pruebas de Kruskal-Wallis. En negrita se destacan aquellos operadores que no hallaron diferencias significativas entre las localidades para dicho atributo tafonómico. **Margen:** redondeamiento del margen; **Color:** coloración secundaria.

**Table 3.** Significance values of statistical probability (*p*) of the taphonomic comparison among the five sample sites (see Figure 1) by each individual operator for each taphonomic attribute. Values were obtained from Kruskal-Wallis tests. Operators who found no significant differences among sites are highlighted in bold. **Margen:** edge rounding; **Color:** secondary color.

Operador	Atributo tafonómico				
	Fragmentación	Margen	Corrasión	Bioerosión	Color
1	<0,001	<0,001	<0,001	<b>0,118</b>	<0,001
2	<0,001	<b>0,109</b>	0,014	0,001	<0,001
3	0,006	0,002	0,022	0,034	<0,001
4	<0,001	0,007	<0,001	<b>0,076</b>	<0,001
5	<0,001	<0,001	<0,001	<b>0,261</b>	<0,001
6	<0,002	<0,001	<0,001	0,003	0,001
7	<0,001	<0,001	<0,001	0,001	<0,001
8	<0,001	0,002	<0,001	0,014	<0,001
9	<0,001	<0,001	<0,001	0,003	<0,001
10	<0,001	<0,001	<0,001	<b>0,119</b>	<0,001
11	<0,001	0,021	<0,001	0,017	<0,001
12	<0,001	<b>&gt;0,500</b>	<0,001	0,004	<0,001
13	<0,001	<b>0,186</b>	0,004	<0,001	<0,001
14	<0,002	<0,001	<0,001	<b>0,238</b>	<0,001
15	<0,001	<0,001	<0,001	<b>0,658</b>	<0,001

Al comparar los resultados obtenidos para cada atributo tafonómico entre los distintos operadores, con base en las comparaciones de a pares entre localidades (prueba de Mann-Whitney), se pudo reconocer que para la fragmentación, en promedio, el 85,33% de los operadores concordaron, mientras que para el redondeamiento del margen, lo hicieron un 76 %. Para la corrasión y la bioerosión, coincidieron el 77,33% y el 74,67% de los operadores, respectivamente, mientras que el 80,67% presentó valores de correspondencia para la coloración secundaria (Figura 3B; Tabla 4). Los valores de cada comparación de a pares pueden consultarse en las Tablas 1S–5S (datos suplementarios).

Los datos tomados por cada operador fueron prácticamente diferentes para todos los atributos tafonómicos y para todas las localidades, con excepción de la fragmentación (Tabla 5, Figura 5; Figuras 1S–4S, datos suplementarios). En otras palabras, en la mayoría de los atributos tafonómicos existieron diferencias estadísticas significativas entre los operadores (Tabla 5).

El porcentaje de especímenes que alcanzó un determinado nivel de concordancia entre los operadores varió entre los diferentes atributos tafonómicos, siendo más alto para la fragmentación (Tabla 6). En efecto, el 80% de las muestras tuvieron un alto porcentaje de acuerdo (entre 80 y 100%) entre los operadores. En segundo lugar, se ubicó la bioerosión con un 77% de las muestras con un acuerdo entre 80 y 100%.

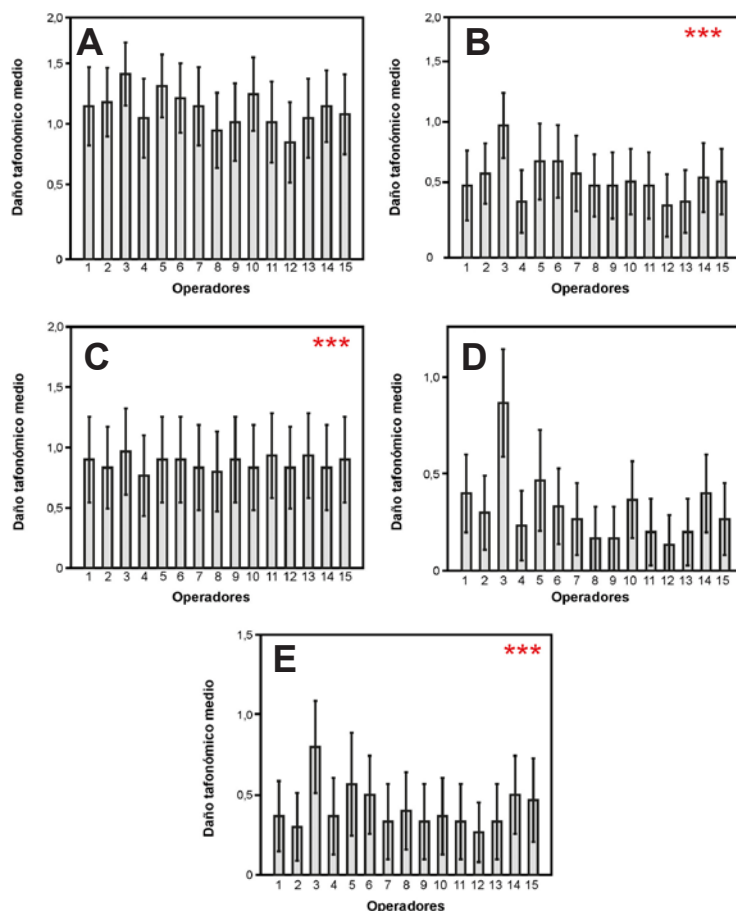
**Tabla 4.** Valores de correspondencia (%) entre los operadores, tomadas de a pares entre todas las localidades. Los valores fueron obtenidos utilizando los mayores valores de coincidencia de los datos disponibles en las Tablas 1S-5S (datos suplementarios). **Margen:** redondeamiento del margen; **Color:** coloración secundaria.

**Table 4.** Percentage (%) of operators in agreement, for paired localities. The values were obtained using the highest values matching the available data in the Tables 1S-5S (supplementary data). **Margen:** edge rounding; **Color:** secondary color.

Comparación	Atributo tafonómico				
	Fragmentación	Margen	Corrosión	Bioerosión	Color
1 x 2	100,00	93,33	80,00	80,00	80,00
1 x 3	80,00	53,33	80,00	73,33	66,67
1 x 4	66,67	80,00	80,00	80,00	93,33
1 x 5	53,33	73,33	93,33	86,67	73,33
2 x 3	100,00	66,67	86,67	86,87	93,33
2 x 4	60,00	73,33	60,00	60,00	66,67
2 x 5	93,33	60,00	86,67	80,00	80,00
3 x 4	100,00	80,00	60,00	53,33	93,33
3 x 5	100,00	86,67	93,33	93,33	93,33
4 x 5	100,00	93,33	53,33	53,33	66,67
Media	85,33	76,00	77,33	74,67	80,67
Desvío estándar	18,80	13,40	14,47	14,35	11,94
Mediana	96,67	76,67	80,00	80,00	80,00

En este mismo rango de acuerdo entre operadores (80-100%) solamente un 47,3%, 33,3% y 15,3% de los especímenes estuvo representado para la coloración secundaria, el redondeamiento del margen y la corrosión, respectivamente.

El análisis de la dispersión de los operadores entre las localidades en un espacio bidimensional (nMDS) mostró que para la fragmentación todos presentaron resultados próximos entre sí (en al menos tres de las cinco localidades que no presentaron diferencias estadísticas; Tabla 5, Figura 5A), y que los grupos formados (correspondientes a cada localidad) fueron significativamente diferentes entre ellos (PERMANOVA,  $F=68,27$ ,  $p<0,001$ ). Para el redondeamiento del margen, este patrón también se mantuvo, formando cinco grupos (localidades) bien distintos, con poca superposición, también con diferencias significativas dentro cada grupo (entre los operadores) (Tabla 5) y entre los grupos (PERMANOVA,  $F=16,08$ ,  $p<0,001$ , Figura 5B). Mientras que, para la corrosión, la bioerosión y la coloración secundaria, los grupos formados estuvieron superpuestos, indicando una mayor dispersión entre los operadores (Figura 6), siendo posible



**Figura 4.** Variación de la alteración tafonómica media entre los 15 operadores para el atributo fragmentación, único atributo que no presentó diferencias significativas entre los distintos operadores para algunas localidades (\*\*\*) indica valor de  $p>0,05$ ). **A,** localidad 1. **B,** localidad 2. **C,** localidad 3. **D,** localidad 4. **E,** localidad 5. Para los demás atributos, véase Figuras 1S-4S (datos suplementarios). La barra de error comprende un intervalo de confianza de 95%.

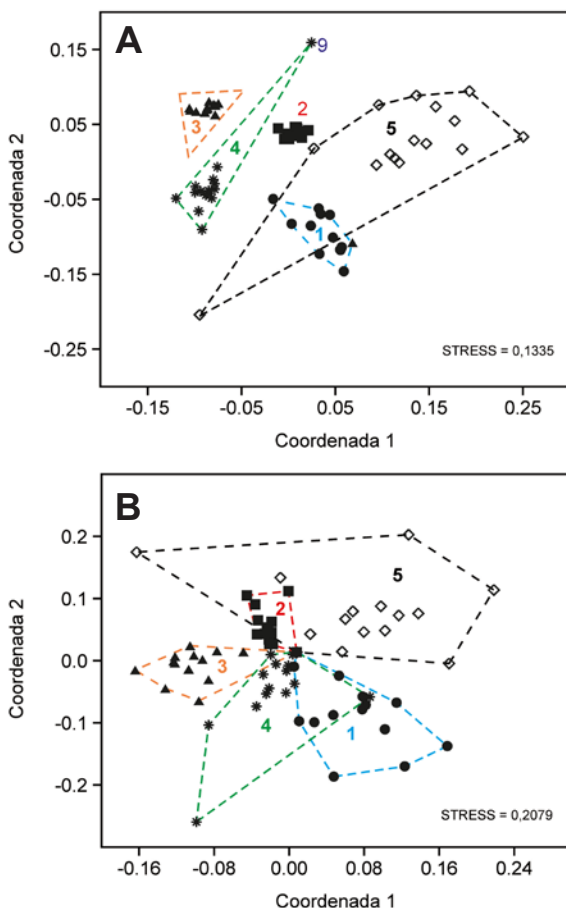
**Figure 4.** Bar graph in which the variation of the average taphonomic damage between 15 operators for the attribute fragmentation is represented, the one in which there were no significant differences between operators for sites (\*\*\*) indicates  $p$ -value  $>0.05$ . **A,** site 1. **B,** site 2. **C,** site 3. **D,** site 4. **E,** site 5. For the remaining attributes, see Figures 1S-4S (supplementary data). Error bar comprises a 95% confidence interval.



**Tabla 5.** Valores de significancia estadística ( $p$ ) de la comparación del estado tafonómico de las muestras de una determinada localidad entre los distintos operadores. Los valores fueron obtenidos a partir de pruebas de Kruskal-Wallis. En negrita se destacan aquellas localidades que no presentaron diferencias significativas entre los operadores para dicho atributo tafonómico. **Margen:** redondeamiento del margen; **Color:** coloración secundaria.

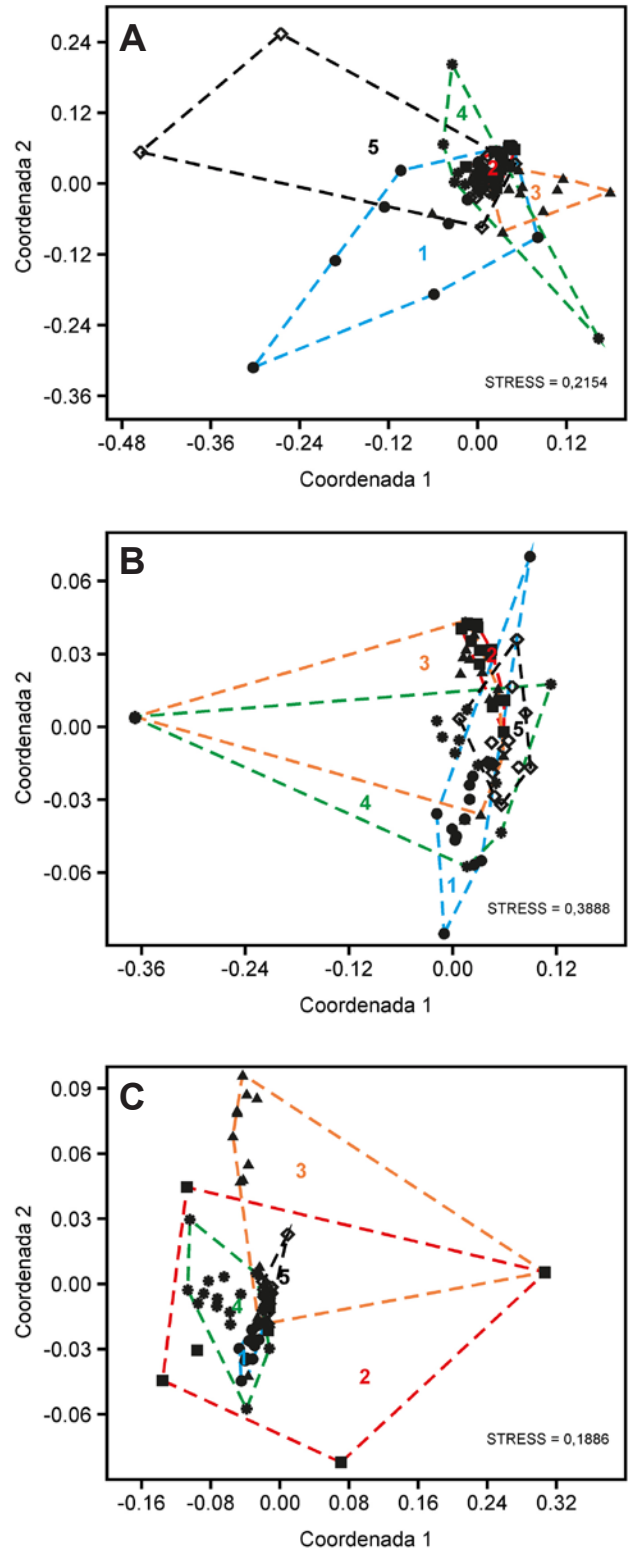
**Table 5.** Significance values of statistical probability ( $p$ ) of the comparison of the taphonomic grade of samples at a particular locality among operators. Values were obtained from Kruskal-Wallis tests. Localities, where no significant differences among operators were found, are highlighted in bold. **Margen:** edge rounding; **Color:** secondary color.

Localidad	Atributo tafonómico				
	Fragmentación	Margen	Corrasión	Bioerosión	Color
1	<b>1,000</b>	0,017	<0,001	<0,001	<0,001
2	<b>0,139</b>	<0,001	<0,001	<0,001	<0,001
3	<b>0,690</b>	<0,001	<0,001	<0,001	<0,001
4	<0,001	<0,001	<0,001	<0,001	<0,001
5	<b>0,261</b>	<0,001	<0,001	<0,001	<0,001



**Figura 5.** Gráfico de Escalamiento multidimensional no métrico (nMDS) de los operadores y las localidades para los diferentes atributos tafonómicos. **A,** fragmentación; **B,** redondeamiento del margen.

**Figure 5.** Non-metric multidimensional scaling (nMDS) illustrating operators and groups (locations) formed by taphonomic attribute on a two-dimensional space. **A,** fragmentation; **B,** edge rounding.



**Figura 6.** Gráfico de Escalamiento multidimensional no métrico (nMDS) de los operadores y las localidades para los diferentes atributos tafonómicos. **A,** corrasión; **B,** bioerosión; **C,** coloración secundaria.

**Figure 6.** Non-metric multidimensional scaling (nMDS) illustrating operators and groups (locations) formed by taphonomic attribute on a two-dimensional space. **A,** corrasion; **B,** bioerosion; **C,** secondary color.

incluso observar algunos que presentaron resultados bien diferentes en relación a los demás. Para estos tres atributos, los análisis tafonómicos también fueron diferentes dentro (Tabla 5) y entre los grupos formados (PERMANOVA,  $F=5,66, p<0,001$ , Fig. 6A; PERMANOVA,  $F=13,93, p<0,001$ , Fig. 6B) y (PERMANOVA,  $F=13,17, p<0,001$ , Figura 6C). Para todos existió diferencia entre los grupos cuando se los analizó de a pares (Tablas 6S–11S, datos suplementarios).

El ACP permitió visualizar las variaciones entre los operadores por cada atributo tafonómico en un gráfico bidimensional (Figuras 5S, 6S, datos suplementarios). Aunque para la mayoría de los atributos tafonómicos (la única excepción fue la fragmentación) hubo diferencias, la varianza entre los operadores no alcanzó valores superiores al 30%. La fragmentación presentó la menor varianza entre los operadores, seguida del redondeamiento del margen, la bioerosión, la corrosión y la coloración secundaria (Tabla 7, Figura 7). No obstante, la varianza de los operadores no presentó diferencias significativas entre los atributos tafonómicos (Kruskal-Wallis,  $p=0,8699$ , Figura 7).

**Tabla 6.** Porcentaje de especímenes (valvas) que alcanzan un determinado nivel de concordancia entre los operadores. **Margen:** redondeamiento del margen; **Color:** coloración secundaria.

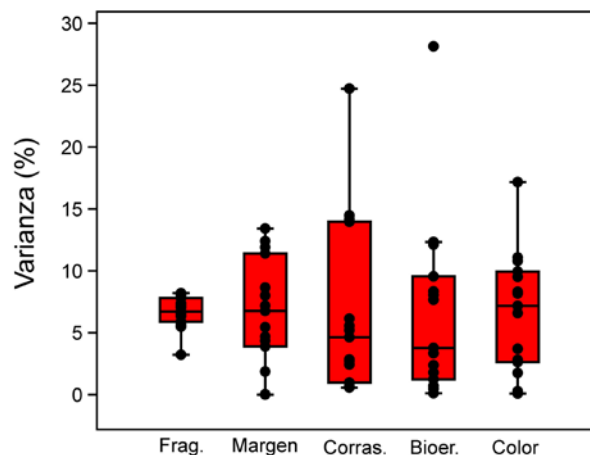
**Table 6.** The percentage of specimens is satisfying a given level of agreement among operators. **Margen:** edge rounding; **Color:** secondary color.

	Fragmentación	Margen	Corrosión	Bioerosión	Color
15 de 15	33,3	1,3	0,7	18,7	4,0
14 de 15	26,0	6,0	2,0	20,7	11,3
13 de 15	12,0	15,3	4,7	22,7	14,0
12 de 15	8,0	10,7	8,0	14,7	18,0
11 de 15	4,7	12,7	18,7	6,0	16,7
10 de 15	4,0	17,3	14,7	3,3	10,0
9 de 15	4,0	12,7	18,0	6,0	8,7
8 de 15	7,3	14,0	16,7	8,0	10,0
7 de 15	0,7	6,0	12,7	-	6,0
6 de 15	-	4,0	4,0	-	1,3

**Tabla 7.** Valores relativos de varianza obtenidos a través de las cargas del Análisis de Componentes Principales (ACP) entre los operadores (considerando solamente el eje 1, en los que su respectivo autovalor no sobrepasó valores generados aleatoriamente; véase también Figuras 5S–7S, y Tabla 12S; datos complementarios). **Margen:** redondeamiento del margen; **Color:** coloración secundaria.

**Table 7.** Relative values of variance obtained from the loadings of the Principal Component Analysis (PCA) between the operators (considering only the axis 1, in which their respective self-worth did not exceed values randomly generated; see also Figures 5S–7S, and Tabla 12S; supplementary data). **Margen:** edge rounding; **Color:** secondary color.

Varianza (%)	Atributo tafonómico				
	Fragmentación	Margen	Corrosión	Bioerosión	Color
Máxima	8,20	13,42	24,73	28,14	17,18
Media	6,70	6,76	4,64	0,11	7,17
Mínima	3,22	0,03	0,57	3,77	0,09



**Figura 7.** Box plots y jitter plots en donde se muestra la varianza entre los operadores para cada atributo tafonómico considerando todas las localidades. **Abreviaturas:** **Frag.**, fragmentación; **Margen**, redondeamiento del margen; **Corras.**, corrosión; **Bioer.**, bioerosión; **Color**, coloración secundaria.

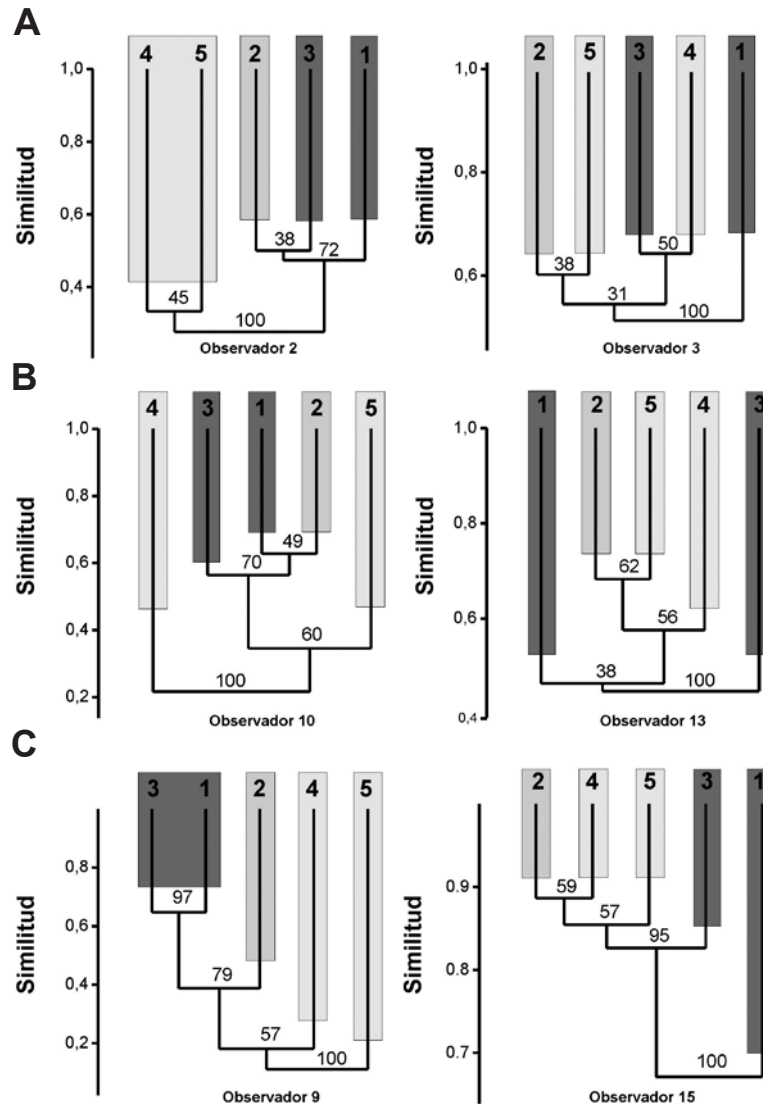
**Figure 7.** Box plots and jitter plots showing the variance between operators on each taphonomic attribute at all sites. **Abbreviations:** **Frag.**, fragmentation.; **Margen**, edge rounding; **Corras.**, corrosion; **Bioer.**, bioerosion; **Color**, secondary color.

A pesar de que la varianza fue menor a un 30%, una simulación de un resultado tafonómico interpretativo en modo Q (comparando las localidades) indicó que los resultados fueron muy diferentes entre los dos operadores que presentaron la mayor varianza (Figuras 8-9). Una variación inferior a un 8% entre los operadores, registrada para el atributo fragmentación fue suficiente para generar un cluster en modo Q con resultados diferentes (Figura 8A).

## DISCUSIÓN

### El efecto del operador

La fragmentación presentó el mayor valor de concordancia entre los operadores, al igual que en el trabajo de Rothfus (2004). Postulamos que estos resultados se deben a que la fragmentación es el atributo más fácilmente identificable y con menor nivel de subjetividad, porque resulta teóricamente sencillo categorizar una fragmentación intensa, moderada o ausente (esta afirmación se puede testear en la Figuras 2A–C). Exceptuando este atributo, el efecto del operador presentó una relación con la cantidad de las variables categóricas medidas. Así, el atributo bioerosión, que fue el único de carácter binario, presentó la segunda mayor correspondencia entre los operadores, concordando con los resultados obtenidos previamente por Kidwell *et al.* (2001) y Rothfus (2004). Por último, los atributos coloración secundaria, redondeamiento del margen y corrosión, que poseen tres categorías de alteración tafonómica, presentaron los menores valores de concordancia. De estos, los dos últimos atributos también presentaron los menores índices en el estudio de Rothfus (2004), quien utilizó el atributo alteración de la superficie externa como un análogo a la corrosión utilizada aquí. Los atributos tafonómicos con más de dos grados aumentan la



**Figura 8.** Análisis de agrupamiento (*clusters*) entre los operadores que presentaron valores más discrepantes de la varianza (véase Tabla 7). Los diferentes colores indican diferentes ambientes sedimentarios. Las localidades 4-5 son de origen marino, mientras que las localidades 1, 3 son de albufera; la localidad 2 está relacionada al ambiente marino, pero es distinta de las localidades 4-5. **A**, fragmentación. **B**, redondeamiento del margen. **C**, corrosión.

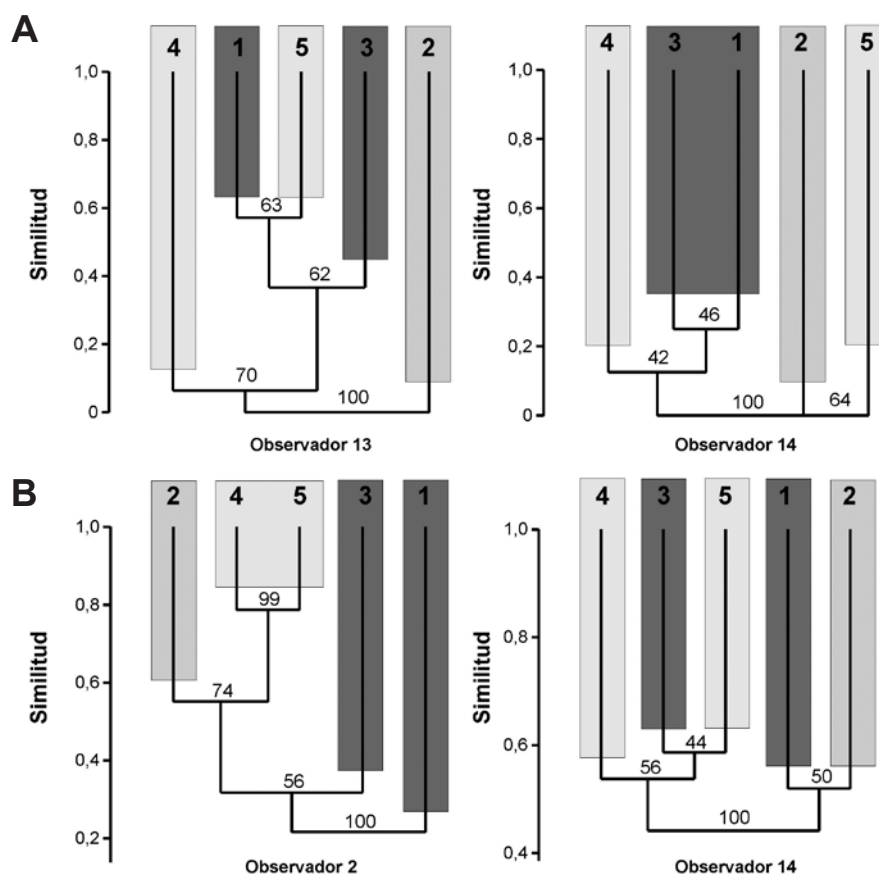
**Figure 8.** Cluster analysis of the operators with the highest differing values of variance (see Table 7). The sites were highlighted according to their sedimentary environment: sites 1-3 are lagoonal while sites 4-5 are marine; site 2 is enclosed to 4-5, but quite differently. **A**, fragmentation. **B**, edge rounding. **C**, corrosion.

subjetividad en los análisis de multi-operadores (Kidwell *et al.*, 2001), dado que cuantas más categorías se definen, mayor es la probabilidad relativa de encontrar discordancia entre los operadores. Por esta razón, Rothfus (2004) concluyó que convertir las categorías multi-estado (0, 1, 2 etc.) en categorías binarias incrementaría la proporción de especímenes para cada nivel de categoría de alteración tafonómica.

Kidwell *et al.* (2001) verificaron el efecto del operador entre observadores ( $n=44$ , bioclastos  $>8$  mm) analizando valvas de bivalvos marinos tropicales, y percibieron que el error era mayor cuando los atributos tafonómicos presentaban más de dos estados. Rothfus (2004) también cuantificó este efecto entre 13 operadores que analizaban

conchas de bivalvos, hallando que, por ejemplo, la fragmentación presentaba una correspondencia entre 96,67% de los operadores, 72,82% para redondeamiento del margen, y 91,41% para bioerosión (un promedio entre interna y externa). Este mismo patrón, aunque con diferentes valores, fue obtenido en el presente estudio (Tabla 2, Figura 3A).

Otro factor, que no fue medido en el presente estudio, es el conocimiento previo que cada operador tiene de su objeto de estudio. En este sentido, aunque la fragmentación presente 3 categorías, aparentemente resultó fácil distinguir entre ellas, lo que no pareció ser el caso de la corrosión, el redondeamiento del margen y la coloración secundaria. De hecho, la distinción



**Figura 9.** Análisis de agrupamiento (*clusters*) entre los operadores que presentaron los valores más atípicos de la varianza (véase Tabla 7). Los diferentes colores indican diferentes ambientes sedimentarios. Las localidades 4-5 son de origen marino, mientras que las localidades 1, 3 son de albufera; la localidad 2 está relacionada al ambiente marino, pero es distinta de las localidades 4-5. **A**, bioerosión. **B**, coloración secundaria.

**Figure 9.** Cluster analysis between operators with the highest differing values of variance (see Table 7). The sites were highlighted according to their sedimentary environment. The sites were highlighted according to their sedimentary environment: sites 1, 3 are lagoonal while sites 4-5 are marine; site 2 is enclosed to 4-5, but quite differently. **A**, bioerosion. **B**, secondary color.

entre colores, por ejemplo, es algo muy subjetivo entre los seres humanos (*e.g.* Lafer-Souza *et al.*, 2015), contribuyendo a esta discordancia. No obstante, tal vez el principal factor sea que los operadores no conocían la coloración original de los bioclastos analizados aquí, lo que les dificultaba atribuir los colores a una determinada categoría.

Kidwell *et al.* (2001) y Rodrigues *et al.* (2009) citan también que la falta de estandarización de las variables es un problema importante que disminuye la comparabilidad de los trabajos tafonómicos. Actualmente aún se pueden hallar diferencias entre distintos estudios tafonómicos. Por ejemplo, Ritter *et al.* (2013) consideraron al atributo “corrosión” como un sinónimo del atributo “alteración de la superficie a escala fina” de Best (2008). Sin embargo, a diferencia de Best (2008), Ritter *et al.* (2013) utilizaron diferentes categorías para cuantificar este atributo. Algo similar ocurrió para este mismo atributo en el estudio de Pisano *et al.* (2015): las categorías utilizadas fueron diferentes de las de Best (2008) y también de las de Ritter *et al.* (2013). Este ejemplo ilustra claramente la forma en que la comparabilidad se ve reducida debido a una variación en la metodología empleada, aún en trabajos con objetivos similares.

A la vez que se recomienda escoger estados categóricos bien distintos (como presencia/ausencia) en los estudios que necesitan de multi-operadores, un único operador es idealmente la mejor opción para eliminar el efecto del operador (Kidwell *et al.*, 2001; Rothfus, 2004). Así, la mayoría de los estudios tafonómicos han venido utilizando un único operador en todos los análisis tafonómicos (*e.g.* Best, 2008; Rodrigues *et al.*, 2009; Ritter *et al.*, 2013). Esto permite que pueda incrementarse considerablemente el número de estados tafonómicos (para un mismo atributo) que en el caso de multi-operadores, dado que no existe el efecto del operador, lo que deviene en un mayor grado de detalle e información. Por ejemplo, aquí la bioerosión tenía dos niveles, mientras que en Best (2008) el mismo atributo presentaba 12 niveles. En este sentido, existen dos alternativas: (i) cuanto más categorías tafonómicas se erijan, mayor será el nivel de información y de detalles que un estudio puede proveer, permitiendo correlacionarlo con otras variables independientes, pero a la vez (ii) el nivel de replicabilidad del estudio se reducirá drásticamente, es decir, la posibilidad de que otro operador repita de manera independiente el estudio con el mismo conjunto de datos y encuentre resultados mínimamente

semejantes, se verá reducido por el efecto del operador. Este efecto será mayor a medida que un determinado atributo tafonómico presente más categorías (como en el caso de la coloración secundaria), y como fuera reportado por Kidwell *et al.* (2001), Rothfus (2004) y en el presente estudio.

### Replicabilidad

Otra forma de verificar el efecto del operador es analizando de qué manera las diferencias entre los datos recogidos por cada operador interfieren sobre el resultado de un análisis tafonómico. Al comparar el resultado estadístico individual de cada operador evaluando si existen diferencias significativas entre las localidades para cada atributo, es posible observar que el resultado final es semejante, con raras excepciones (Tabla 3). En otras palabras, en este nivel de análisis el efecto del operador tiene un efecto reducido sobre un análisis interpretativo cuantitativo, dado que para los atributos fragmentación, corrosión y coloración secundaria todos los operadores obtuvieron el mismo resultado: la existencia de diferencias significativas entre las localidades. Para redondeamiento del margen y bioerosión, por su parte, el 20% (3 de 15 operadores) y el 40% (6 de 15 operadores) de los resultados, respectivamente, no presentaron diferencias significativas, a diferencia de los restantes. Y, aunque si bien se reportó previamente que los atributos tafonómicos con apenas dos categorías distintas teóricamente reducen el efecto del operador (Kidwell *et al.*, 2001; Rothfus, 2004), su efecto sobre la interpretación tafonómica (replicabilidad) parece ser el opuesto, dado que, en este nivel de análisis estadístico, la bioerosión presentó la menor replicabilidad relativa (Tabla 7, Figura 7). Además, una mayor correspondencia entre los datos obtenidos por los operadores (Tabla 2, Figura 3A) para el atributo fragmentación produjo también una mayor correspondencia en los resultados (Tabla 4, Figura 3B). Por el contrario, este patrón no fue observado para el caso de la bioerosión, que presentó, conjuntamente con el redondeamiento del margen, los menores valores medios de correspondencia (Tabla 4, Figura 3B). Es decir que estos resultados sugieren que si bien los atributos con categorías binarias reducen el efecto del operador (como ya fuera demostrado por Rothfus, 2004), esto no garantiza que los resultados obtenidos por los distintos operadores sean semejantes tafonómicamente.

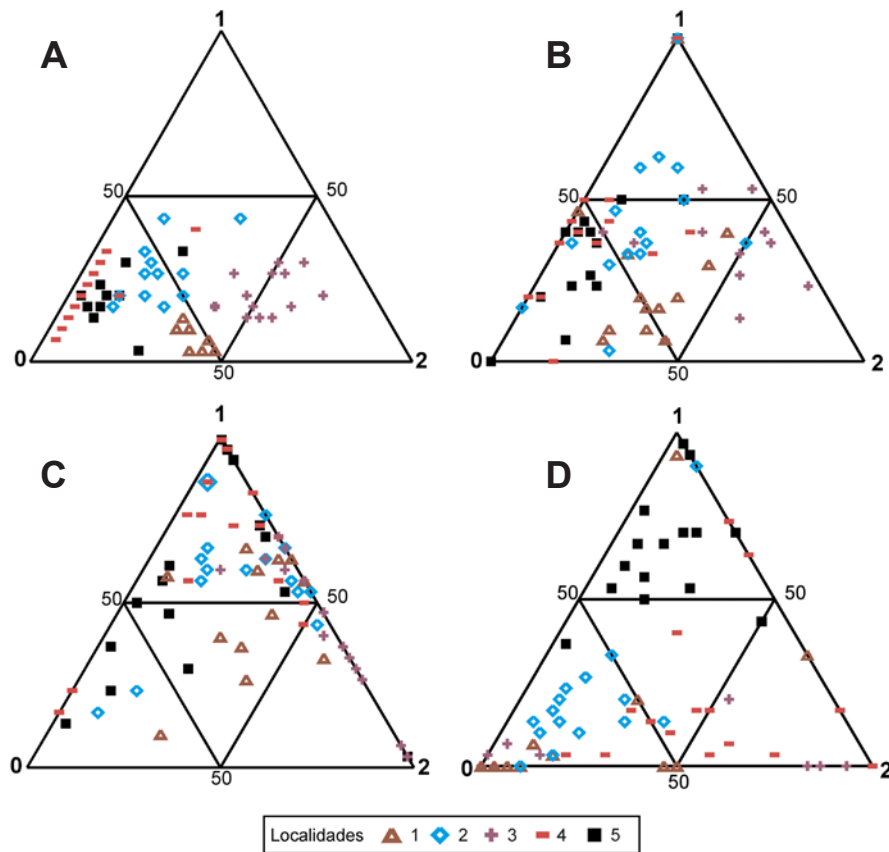
La distribución de los operadores y de las localidades en un espacio bidimensional permite conocer la dispersión de los operadores en relación a las localidades (grupos), para cada atributo tafonómico (Figuras 5-6). Cada operador podría graficar independientemente sus resultados en este tipo de análisis, observando cómo los distintos atributos son escalonados de acuerdo con la localidad, permitiendo verificar, de esta manera, relaciones de dispersión entre atributos y localidades. Aquí, debido al objetivo del trabajo, se graficó el resultado individual de los 15 operadores para cada atributo tafonómico entre las localidades. En este sentido, fue posible comprobar gráficamente los datos presentados en las Tablas 3 y 5. La fragmentación, por ejemplo, que fue semejante entre los operadores en cuatro localidades (Tabla 5), presentó gráficamente grupos bien

distintos, con todos los operadores ubicados muy próximos entre sí. Cuando se compararon las localidades entre sí (sin considerar individualmente a los operadores), para todos los atributos, existieron diferencias significativas entre ellos, al igual que al ser comparados de a pares (Figuras 5-6; Tablas 6S–11S, datos suplementarios). El significado interpretativo ya fue reportado anteriormente: en prácticamente todos los atributos, para todos los operadores, existen diferencias entre las localidades. Por lo tanto, el efecto del operador, si bien es un sesgo presente, parece no tener gran influencia sobre el resultado tafonómico comparativo entre las localidades o facies. O sea, la replicabilidad de este método de análisis es válida científicamente, pues en la mayoría de los casos, diferentes operadores llegaron a los mismos resultados, aunque existan diferencias entre ellos en casi todos los atributos (Tablas 3 y 5).

Un aspecto interesante para destacar es que la bioerosión presentó el mayor valor de *stress* (Figura 6B), lo que se debe a la presencia de valores atípicos en algunos operadores en relación a los demás, aunque aún los grupos continúen siendo diferentes entre sí, lo que aumenta la pérdida de información en este tipo de gráfico bidimensional, aún siendo un atributo binario, teóricamente con los efectos entre los operadores minimizados (Kidwell *et al.*, 2001; Rothfus, 2004). Aunque la segunda mayor concordancia entre los operadores (efecto del operador) es para la bioerosión (Tabla 2), cuando se testea la replicabilidad, se observa que la relación no es necesariamente positiva. Esto permite concluir que la reducción del efecto del operador no garantiza valores interpretativos más semejantes y, consecuentemente, una mayor o menor replicabilidad.

Los gráficos de ordenamiento bidimensional (nMDS), en síntesis, sugieren que a pesar de no estar visualmente agrupados, todos los grupos formados (localidades) son diferentes entre sí (Figuras 5-6) para casi todos los operadores (Tabla 3, Kruskal-Wallis para cada operador entre las localidades). O sea, en este tipo de interpretación tafonómica, independientemente del efecto del operador, el resultado final será semejante para todos los operadores (la excepción lo constituyen la bioerosión y el redondeamiento del margen), sugiriendo una alta replicabilidad de esta metodología cuantitativa.

Alternativamente, los diagramas ternarios (tres estados tafonómicos) en un espacio también bidimensional (Kowalewski *et al.*, 1995; Figura 10) representan una buena forma de observar estas diferencias mencionadas. En esta representación, cada vértice del triángulo representa una condición tafonómica categórica (Tabla 1), y es posible graficar conjuntamente los operadores por localidad. Las diferencias entre los 15 operadores por atributo en cada localidad (Tabla 3) y entre cada operador por localidad (Tabla 5) y, así como las pruebas realizadas para testear relaciones entre grupos, son los mismos utilizados en las Figuras 5-6. En estos dos tipos de representación gráfica se observa claramente que la fragmentación formó los grupos más cohesivos, mientras que el redondeamiento del margen resultó el grupo con mayor dispersión. La bioerosión, por ser binaria, no fue posible representarla en un tafograma, lo que



**Figura 10.** Diagramas ternarios (tafogramas) donde se muestra la dispersión de los operadores por localidad, en aquellos atributos tafonómicos de categorías ternarias. Compárese la dispersión con aquella obtenida en las Figuras 5–7 **A**, fragmentación. **B**, redondeamiento del margen. **C**, corrosión. **D**, coloración secundaria.

**Figure 10.** Taphograms illustrating the dispersion of operators by location. Compare the dispersion to Figures 5–7. **A**, fragmentation. **B**, edge rounding. **C**, corrosion. **D**, secondary color.

confiere una desventaja en relación al nMDS. En síntesis, si todos los atributos fueran definidos *a priori* de manera ternaria, como en Pisano *et al.* (2015), el tafograma resulta un método rápido y eficaz de representación. En otros casos, el nMDS es considerado un mejor método para representar estas variaciones y más fidedigno porque, además de brindar una visualización rápida de los patrones, preserva la distancia entre las variables (Tyler & Kowalewski, 2014).

Además de la concordancia entre los operadores, es posible también estimar la varianza entre los operadores por medio de un ACP (Figuras 5S–6S, datos suplementarios). En este sentido, con base en el eje 1 del ACP (Tabla 12S, datos suplementarios), se generó un gráfico de varianza (una medida de la dispersión) entre los operadores (Figura 7), que permitió afirmar que el 50% de los operadores estuvieron contenidos en una franja de dispersión en torno al 6% para la fragmentación, mientras que la máxima dispersión fue de 8,20% (Tabla 7). Así, en orden creciente, la fragmentación y la alteración del margen presentaron la menor varianza entre los operadores. En el extremo opuesto, la bioerosión, la corrosión y la coloración secundaria presentaron la mayor varianza (en orden decreciente, respectivamente) (Tabla 7, Figura 7). Es decir, la

medida de la varianza resulta diferente de la de concordancia, dado que, por ejemplo, para el caso de la bioerosión, aunque aproximadamente el 84% de los operadores concordaron en la atribución de los estados tafonómicos, la dispersión entre ellos fue la mayor, en torno a un 28% (Tabla 7).

De esta manera, fue posible analizar la replicabilidad a otro nivel de análisis estadístico, no solamente observando los valores relacionados a las diferencias entre localidades por operadores, sino agrupando localidades más próximas en relación a un atributo tafonómico en particular. Así, con base en los valores de varianza, es posible analizar si los valores máximos entre los operadores aquí reportados (Tabla 7) son suficientes para generar resultados interpretativos diferentes. Utilizando los dos operadores más distantes (Figuras 5S–6S, datos suplementarios) considerando apenas el eje 1 del ACP (Figura 7S, datos suplementarios), se realizó un análisis de agrupamiento (*cluster*) en modo Q (considerando las localidades) para cada par de operadores por atributo tafonómico y se compararon luego los resultados entre ambientes sedimentarios (Figuras 8–9). Con este análisis fue posible percibir que tanto una varianza relativamente pequeña, de 8% (en el atributo fragmentación), entre los

operadores fue suficiente para que el resultado interpretativo fuese diferente entre ellos. O sea, para este tipo de análisis, la replicabilidad del estudio tafonómico puede ser baja, en relación a comparaciones de análisis estadísticos mas generalistas.

### CONSIDERACIONES FINALES

El efecto del operador, como ya fuera reportado en estudios previos (e.g. Kidwell *et al.*, 2001; Rothfus, 2004), es un hecho presente en cualquier estudio tafonómico que intente cuantificar variables categóricas sobre un mismo conjunto de datos. Este efecto puede ser reducido al utilizar variables distintas entre sí, como, por ejemplo, simplemente la presencia o ausencia de un atributo tafonómico. No obstante, las variables binarias, a pesar de presentar una reducción del efecto del operador, disminuyen la replicabilidad, lo que resulta en interpretaciones distintas para datos concordantes. Es decir que en aquellos estudios en los que se necesita de más de un operador, no resulta recomendable utilizar atributos tafonómicos binarios, dado que aparentemente es el tipo de atributo que más negativamente interfiere en la replicabilidad del estudio tafonómico.

La replicabilidad, aquí cuantificada por primera vez para variables tafonómicas, mostró que la varianza entre los operadores podría ser un indicador de la replicabilidad de un estudio. Basados en los resultados obtenidos aquí, podemos concluir que la replicabilidad se vio restringida en variaciones superiores a un 8%. Este hecho, sin embargo, debe ser analizado en relación al análisis estadístico elegido por el operador, una vez que los métodos estadísticos utilizados influyen el nivel de replicabilidad y de comparabilidad de un trabajo.

En este sentido, proponemos algunas recomendaciones para incrementar la replicabilidad y la comparabilidad de futuros estudios tafonómicos: (i) utilizar variables categóricas, abandonando siempre que sea posible los análisis cualitativos, ya que el significado de “buena preservación” para un operador puede no ser el mismo para otro, y esta subjetividad es inconmensurable; (ii) describir detalladamente cada variable dentro de un determinado atributo tafonómico, e incluir fotografías de los ejemplares correspondientes; (iii) utilizar métodos estadísticos adecuados para variables categóricas, recomendando como punto de partida los análisis aquí empleados, aunque siempre se debe realizar previamente un análisis crítico; (iv) brindar una mayor transparencia, disponiendo la información suplementaria de los análisis auxiliares, así como los datos sobre los que se realizaron dichos análisis. Esto permitiría que un estudio pueda ser comparado fácilmente por otro investigador, el cual podría identificar variaciones derivadas del tipo de método estadístico utilizado o variaciones no sesgadas por la naturaleza de los datos o del ambiente sedimentario previamente estudiado. De este modo, se garantiza que los estudios puedan ser comparados de una mejor forma, lo que aumentaría tanto su replicabilidad como su comparabilidad.

### AGRADECIMIENTOS

MNR agradece a FAPERGS (1982-2551/13-7) y a CNPq (140568/2014-0). CGDF agradece a CNPq (APV, 451359/2015-1) y a J.C. Coimbra, que le permitió financiar su visita a Brasil. HF, LAK, NCL, FHM, ALMM y PLAX agradecen al CNPq por la beca de estudios. Un agradecimiento especial a R.P. Lopes, E.R. Beltram, y S.M. Petró por su ayuda en la recolección de material de la localidad 3; a Jair Weschenfelder por la donación de muestras de la localidad 1 y a F. Caron por la donación de muestras de las localidades 4 y 5 y por las fotos de los muestreos en estos sitios. Las muestras de la localidad 2 fueron recolectadas por Vanessa Agostini, a quien también se extiende nuestro agradecimiento. Los autores agradecen a todos los participantes del curso de “Tafonomía actualista y su importancia en Paleontología” que tomaron participación en esta experiencia y produjeron parcialmente estos datos: A.G. Martinelli, A.C.B. Brust, E.R. Beltram, E. Mayer, F.H. Veiga, M. Guignard, P.R.R.V. Martinez y R. Casati. Los autores agradecen también los valiosos comentarios de S. Martínez, F. Erthal, F. Archuby y de un revisor anónimo.

### REFERENCIAS

- Abdi, H. & Williams, L.J. 2010. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, **2**:433–459. doi:10.1002/wics.101
- Araújo-Júnior, H.I.; Porpino, K.O. & Bergqvist, L.P. 2015. Vertebrate taphonomy and paleoecology in an Upper Pleistocene tank deposit of Paraíba, Brazil: taphonomic modes, evidence of temporal and spatial resolutions and paleoecological patterns of the Brazilian Intertropical Region. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **437**:1–17. doi: 10.1016/j.palaeo.2015.07.009
- Berkeley, A.; Perrya, C.T. & Smithers, S.G. 2009. Taphonomic signatures and patterns of test degradation on tropical, intertidal benthic foraminifera. *Marine Micropaleontology*, **73**:148–163. doi:10.1016/j.marmicro.2009.08.002
- Best, M.M.R. 2008. Contrast in preservation of bivalve death assemblages in siliciclastic and carbonate tropical shelf settings. *Palaios*, **23**:796–809. doi:10.2110/palo.2005.p05-076r
- Best, M.M.R. & Kidwell, S.M. 2000. Bivalve taphonomy in tropical mixed siliciclastic-carbonate settings. I. Environmental variation in shell condition. *Paleobiology*, **26**:80–102. doi:10.1666/0094-8373(2000)026<0080:BTITMS>2.0.CO;2
- Dillenburg, S.R.; Tomazelli, L.J. & Barboza, E.G. 2004. Barrier evolution and placer formation at Bujuru, Southern Brazil. *Marine Geology*, **203**:43–56. doi:10.1016/S0025-3227(03)00330-X
- Erthal, F.; Kotzian, C.B. & Simões, M.G. 2015. Multi-step alteration in fluvial mollusk shells: the case of the Touro Passo Formation (Pleistocene-Holocene), Southern Brazil. *Palaios*, **30**:388–402. doi:10.2110/palo.2013.104
- Fürsich, F.T. & Oschmann, W. 1993. Storm shell beds as tools in basin analysis: the Jurassic of Kachchh, Western India. *Journal of the Geological Society*, **150**:169–185. doi:10.1144/gsjgs.150.1.0169
- Gotelli, N. & Ellison, A.M. 2012. *A primer of ecological statistics*. 2<sup>nd</sup> ed. Sunderland, Sinauer, 614 p.

- Hammer, Ø.; Harper, D.A.T. & Ryan, P.D. 2001. PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica*, **4**. Disponível em [http://palaeo-electronica.org/2001\\_1/past/issue1\\_01.htm](http://palaeo-electronica.org/2001_1/past/issue1_01.htm); acesso 12/09/2015.
- Kidwell, S.M.; Best, M.M.R & Kaufman, D.S. 2005. Taphonomic trade-offs in tropical marine death assemblages: differential time averaging, shell loss, and probable bias in siliciclastic vs. carbonate facies. *Geology*, **33**:729–732. doi:10.1130/G21607.1
- Kidwell, S.M.; Rothfus, T.A. & Best, M.M.R. 2001. Sensitivity of taphonomic signatures to sample size, sieve size, damage scoring system, and target taxa. *Palaios*, **16**:26–52. doi:10.2307/3515551
- Kosnik, M.A.; Behrensmeier, A.K.; Fürsich, F.T.; Gastaldo, R.A.; Kidwell, S.M.; Kowalewski, M.; Plotnick, R.E.; Rogers, R.R.; Wagner, P.J. & Alroy, J. 2011. Changes in the shell durability of common marine taxa through the Phanerozoic: evidence for biological rather than taphonomic drivers. *Paleobiology*, **37**:303–331. doi:10.1666/10022.1
- Kowalewski, M.; Flessa, K.W. & Hallman, D.P. 1995. Ternary taphograms: triangular diagrams applied to taphonomic analysis. *Palaios*, **10**:478–483. doi: 10.2307/3515049
- Kowalewski, M. & Hoffmeister, A.P. 2003. Sieves and fossils: effects of mesh size on paleontological patterns. *Palaios*, **18**:460–469. doi:10.1669/0883-1351(2003)018<0460:SAFEOM>2.0.CO;2
- Kowalewski, M. & Novack-Gottshall, P. 2010. Resampling methods in Paleontology. In: J. Alroy & G. Hunt (eds.) *Quantitative Methods in Paleobiology*, Boulder, The Paleontological Society, p. 19–54 (Papers 16).
- Krause, R.A.; Barbour, S.L.; Kowalewski, M.; Kaufman, D.S.; Romanek, C.S.; Simões, M.G. & Wehmiller, J.F. 2010. Quantitative estimates and modeling of time averaging in bivalve and brachiopod shell accumulations. *Paleobiology*, **63**:428–452. doi:10.1666/08072.1
- Kruskal, J.B. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**:1–27. doi:10.1007/BF02289565
- Lafer-Souza, R.; Hermann, K.L. & Conway, B.R. 2015. Striking individual differences in color perception uncovered by ‘the dress’ photograph. *Current Biology*, **25**:R545–R546. doi:10.1016/j.cub.2015.04.053
- Leek, J.T. & Peng, R.D. 2015. Reproducible research can still be wrong: adopting a prevention approach. *Proceedings of the National Academy of Sciences of the United States of America*, **112**:1645–1646. doi: 10.1073/pnas.1421412111
- Legendre, P. & Legendre, L. 2012. *Numerical Ecology*. 3<sup>rd</sup> ed. Amsterdam, Elsevier Science BV, 1006 p.
- Mancosu, A. & Nebelsick, J.H. 2015. The origin and paleoecology of clypeasteroid assemblages from different shelf settings of the Miocene of Sardinia, Italy. *Palaios*, **30**:373–387. doi:10.2110/palo.2014.087
- McMenamin, M.A.S. 2016. *Dynamic Paleontology using quantification and other tools to decipher the history of life*. Switzerland, Springer, 251 p. doi:10.1007/978-3-319-22777-1
- McNutt, M. 2014. Reproducibility. *Science*, **343**:229. doi:10.1126/science.1250475
- Nosek, B.A. et al. 2015. Promoting an open research culture. *Science*, **348**:1422–1425. doi:10.1126/science.aab2374
- Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science*, **349**:aac4716. doi:10.1126/science.aac4716
- Parsons-Hubbard, K. 2005. Molluscan taphofacies in recent carbonate reef/lagoon systems and their application to sub-fossil samples from reef cores. *Palaios*, **20**:175–191. doi:10.2110/palo.2003.p03-105
- Pisano, M.F.; De Francesco, C.G. & Fucks, E.E. 2015. Taphonomic signatures in concentrations of *Heleobia* Stimpson, 1865 from Holocene deposits of the Salado river basin, Buenos Aires, Argentina: their utility in paleoenvironmental reconstructions. *Palaios*, **30**:248–257. doi:10.2110/palo.2014.034
- Powell, E.N.; Staff, G.M.; Callender, W.R.; Ashton-Alcox, K.A.; Brett, C.E.; Parsons-Hubbard, K.M.; Walker, S.E. & Raymond, A. 2011. The influence of molluscan taxon on taphofacies development over a broad range of environments of preservation: the SSETI experience. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **312**:233–264. doi:10.1016/j.palaeo.2011.01.008
- Ritter, M.N.; Erthal, F. & Coimbra, J.C. 2013. Taphonomic signatures in molluscan fossil assemblages from the Holocene lagoon system in the northern part of the coastal plain, Rio Grande do Sul State, Brazil. *Quaternary International*, **305**:5–14. doi:10.1016/j.quaint.2013.03.013
- Ritter, M.N.; Erthal, F. & Coimbra, J.C. 2014. Time-averaging remarks of molluscan assemblages in the Coastal Plain of South Brazil. In: INTERNATIONAL PALAEONTOLOGICAL CONGRESS, 4, 2014. *Abstracts*, Mendoza, IPA, p. 102. doi:10.13140/2.1.3921.2802
- Ritter, M.N.; Erthal, F. & Coimbra, J.C. 2015. Time-averaging of shell-rich accumulations on Brazilian Southern shelf. In: CONGRESSO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS DO QUATERNÁRIO, 15, 2015. *Resumos expandidos*, Imbé, ABEQUA, p. 116–117. doi:10.13140/RG.2.1.1236.8089
- Rodrigues, S.C.; Simões, M.G. & Pires-Domingues, R.A. 2009. Viés analítico em estudos tafonômicos com macroinvertebrados: implicações (paleo)ambientais e (paleo)ecológicas. *Geologia USP, Série Científica*, **9**:101–114. doi:10.5327/Z1519-874X2009000100007
- Rothfus, T.A. 2004. How many taphonomists spoil the data? Multiple operators in taphofacies studies. *Palaios*, **19**:514–519. doi:10.1669/0883-1351(2004)019<0514:HMTSTD>2.0.CO;2
- Sepkoski, D. 2012. *Rereading the fossil record. The growth of Paleobiology as an evolutionary discipline*. Chicago, The University of Chicago Press, 440 p.
- Tomašových, A. & Kidwell, S.M. 2009. Preservation of spatial and environmental gradients by death assemblages. *Paleobiology*, **35**:122–148. doi:10.1666/07081.1
- Tyler, C.L. & Kowalewski, M. 2014. Utility of marine benthic associations as a multivariate proxy of paleobathymetry: a direct test from recent coastal ecosystems of North Carolina. *Plos One*, **9**:e95711. doi:10.1371/journal.pone.0095711
- Walker, K.R. & R.K. Bambach. 1971. The significance of fossil assemblages from fine-grained sediments: time-averaged communities. *Geological Society of America Abstracts with Programs*, **3**:783–784.
- Zuschin, M.; Stachowitsch, M. & Stanton Jr., R.J. 2003. Patterns and processes of shell fragmentation in modern and ancient marine environments. *Earth-Science Reviews*, **63**:33–82. doi:10.1016/S0012-8252(03)00014-X

Received in March, 2016; accepted in August, 2016.