USING ELLIPTICAL FOURIER ANALYSIS TO COMPARE SIZE OF MORPHOSPACE OCCUPATION BETWEEN MODERN EDENTULOUS FRESHWATER UNIONOID MUSSELS AND THE FOSSILS AT L6516 (SLOPE COUNTY, NORTH DAKOTA, U.S.A.), WITH REMARKS ON PRESERVATION

by

Matthew E. Burton-Kelly
Bachelor of Science, St. Lawrence University, 2005

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota
December
2008

This thesis, submitted by Matthew E. Burton-Kelly in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

_____

Chairperson

_____

_____

This thesis meets the standards for appearance, conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

_____

Dean of the Graduate School

_____

Date

PERMISSION

Title	Using Elliptical Fourier Analysis to Compare Size of Morphospace Occupation Between Modern Edentulous Freshwater Unionoid Mussels and the Fossils at L6516 (Slope County, North Dakota, U.S.A.), with Remarks on Preservation
Department	Geology
Degree	Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the chairperson of the department or the dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Signature _____

Date _____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Financial assistance to present various aspects of this research was provided by the North Dakota EPSCoR Program (NSF grant #EPS-0447697) through UND Research Development and Compliance for travel to Antwerp, Belgium, and by the Geological Society of America for travel to Houston, Texas. Thanks to Alan Cvancara, who was collecting modern clams in North Dakota before I was even born, and to Kevin Cummings, who sent me many wonderful photographs. Thanks to Art Sweet, who collected pollen samples at Das Goods before we knew that they would be useful for this project. Thanks to Henning Scholz, who graciously took the time to explain his methods. Thanks to F. James Rohlf, who was helpful in his explanation of what data I could and could not use in my multivariate analyses. Thanks to Art Bogan, who supplied me with modern photographs and some very interesting stories. Thanks to Dean Pearson and the Pioneer Trails Regional Museum for use of lab space and access to fossil collections, and to the Horse Creek Grazing Association for letting us collect on their land. Thanks to Matt Borths for assisting in the field and Red Gierach for keeping things from getting too serious. Thanks to Marron Bingle-Davis, Kristyn Voegele, and the Fall 2007 Introduction to Paleontology class, who collected a large amount of the material I used. Thanks to Tanya Justham for giving up some of her own field time to help collect even more specimens at Das Goods. Thanks to my parents, who have always been supportive. Thanks to J. Mark Erickson for eternally questioning what I know. Thanks to my committee members, Rich LeFever and Ron Matheney, who have been waiting with bated breath for longer than I would have liked. Finally, thank you especially to Alison Finstad, who kept reminding me that I

really did want to graduate, and to my advisor, Joseph Hartman who probably had no idea what he was getting into when he accepted me as a student.

ABSTRACT

The purpose of this project was to use outline morphometric techniques to determine the number of freshwater mussel taxa present at locality L6516 ("Das Goods"), which is located in Cretaceous strata in the basal Ludlow Member of the Fort Union Formation in Slope County, North Dakota, U.S.A. Secondary goals were to improve the outline morphometric procedures that have been previously used to characterize fossil and extant freshwater mussels, and to describe in more detail the preservation of the edentulous unionoid specimens from the Das Goods area.

Outlines of individual mussel valves were used for morphometric analyses due to lack of preserved taxonomic characters. Outlines were digitized as a series of (x,y) coordinates, after which elliptical Fourier analysis was used to standardize the valves according to position, size, and rotation. A series of coefficients was produced for each valve that was used to characterize the freshwater mussels from L6516 as well as a selection of modern edentulous genera and species. The size of morphospace occupation was calculated and compared between the fossil and modern groups to determine the probability of multiple taxa at L6516. Different regions of preservation on the best-preserved specimens from locality L6516 were mapped to characterize preservation trends at this site. Traces attributed to epizoic organisms were also examined for position, size, and relief.

The fossil edentulous freshwater mussel assemblage from locality L6516 does not possess statistically significantly more or less morphological variation than the modern genera and species with which it was compared, suggesting a single taxon is preserved. Some of the effects of procedural changes on the outcome of the statistical tests were

described in order to enhance the efficacy of the morphometric techniques: standardization of both digitization size and increased smoothing during elliptical Fourier analysis lead to a reduction in size of morphospace occupation. Detailed descriptions of the specimens led to new insight into the epifaunal association of other organisms with the unionoid assemblage. Suggestions were made for the improvement of morphometric methods in the natural sciences, including making standardized datasets and detailed documentation more available to new researchers.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Mussels of the family Unionoidea (or Unioniformes) are freshwater, benthic organisms with bivalve shells made of calcium carbonate (Dunca et al., 2005; Bogan, 2008). They are most notable for utilizing a parasitic larval stage for reproduction, allowing colonization of flowing water with the assistance of a fish host to move larvae upstream against the current (Cvancara, 1983; Bauer, 2001; Wächtler et al., 2001; Scholz, 2003). This family is represented by 180 genera worldwide and 800–900 species, most of those being part of the subfamily Unioninae (Graf and Cummings, 2006; Bogan, 2008). Fifty-three genera and 302 species of unionoids exist in the Nearctic (North America and Greenland), making this region the most diverse on the planet (Bogan, 2008).

Fossil unionoids from Cretaceous strata of the Western Interior are also quite diverse leading up to the Cretaceous-Paleogene (K-Pg) extinction at about 65.95 Ma (Kuiper et al., 2008) and can be tied biostratigraphically into the North American Land Mammal "Ages" both above and below (Lancian, Puercan, Torrejonian, and Tiffanian) (Hartman, 1998, 1992). This fauna underwent a loss of diversity just prior to the K-Pg extinction event, resulting in a reduction of the number of sculptured taxa (Hartman and Butler, 1995; Hartman, 1996a,b; Scholz and Hartman, 2007a,b). In southwestern North Dakota and eastern Montana, U.S.A., unionoid taxa with relatively featureless exteriors and a common elliptical shape appear to have survived the extinction event and have remained the primary freshwater mussel ever since (Hart-

man, 1996a,b). Sculptured forms returned to the fossil record during the Pleistocene Epoch (Hartman, UND, pers. comm., 2008).

The Late Cretaceous freshwater mollusks in this area have been studied for well over a century, and continue to be better understood. Problematically, however, the molluscan fossils in the Hell Creek Formation and the Ludlow Member of the Fort Union Formation in North Dakota are relatively poorly preserved in comparison with those of Montana (west of the Miles City Arch) (Justham, 2008). Unionoids and gastropods alike are typically present as "ironstone" (siderite, quartz, and goethite) steinkerns and replaced original material, resulting in a taphonomic loss of taxonomic information (Justham, 2008). The combination of locally poor preservation, proximity to the K-Pg extinction event, and a morphologically undiagnostic, depauperate fauna provides impetus for continued study of the freshwater mussels in southwestern North Dakota. Understanding the taxa present in this interval will help determine whether the K-Pg extinction was the result of one or many causes.

Over the past decade a small number of sites exhibiting a wholly different type of preservation have been discovered in the base of the Ludlow Member of the Fort Union Formation northwest of Marmarth, Slope County, North Dakota. The first of these localities (L6516) was named "Das Goods" for its leaf flora; this name is used herein to refer to the type of preservation exhibited there. These fossil localities preserve gastropods and unionoids, the latter as molds and casts of the interior and exterior of the valves in mudstone. Pollen analysis has previously shown L6516 to be of latest Cretaceous age (Hartman et al., 2001); the location of the K-Pg boundary is discussed herein (Chapter 1.2.2). Although the Das Goods unionoids are better preserved than other North Dakota freshwater mussels of similar age, they are unsculptured and generally elliptical, and retain no taxonomically useful characters to aid in identification aside from a lack of hinge teeth (preservation, including varia-

tion within individual specimens, is discussed in Chapter 2.2). Diagnosis of these specimens is an apt problem for morphometric analysis to address.

"Traditional" morphometrics has long been a process of measuring various discrete distances, manipulating those data, and attempting to determine the taxonomic usefulness of sets of measurements through multivariate statistical methods (Marcus, 1990). This approach has had general success and acceptance in the literature, but does not allow for shape (a description of the organism independent of scaling, rotation, or translation) to be addressed independently of size (Marcus, 1990; Zelditch et al., 2004). The more recent field of "geometric" morphometrics seeks to improve mathematical representation of shape by comparing the physical relationship between biologically homologous "landmarks" that have been standardized (Bookstein, 1991; Zelditch et al., 2004). Similarly, outline analysis techniques that standardize data by converting outlines into mathematical functions can be used on shapes that have few or no biologically homologous landmarks (Kuhl and Giardina, 1982; Ferson et al., 1985; McLellan and Endler, 1998; Haines and Crampton, 2000; Lestrel et al., 2004; Scholz and Hartman, 2007b; Scholz and Scholz, 2007). A method of the latter type (Elliptical Fourier Analysis) was applied to the unionoids from L6516.

Although computers and statistics will never be able to replace a trained human eye when identifying important biological characters and how they relate to an organism in a taxonomic system, analytical techniques provide useful forms of communication and added rigor. By definition, if a computer could "see" all the morphological characters a specimen possesses, it should be able to diagnose that specimen to a known taxon or identify it as an unknown; otherwise, taxon diagnoses need to be rewritten without ambiguity (Winston, 1999). Questions need to be carefully constructed in order to utilize the growing mathematical and statistical power at our disposal yet allow the researcher to have the final word in assessing the accuracy

of the results based on his or her own knowledge (Zelditch et al., 2004).

Three major questions are addressed in this thesis that cover data capture and morphometrics methods, applying these methods to modern taxa of known identification to test their efficacy, and attempting to translate them to the unionoid fossils from L6516. These questions are described in more detail in Chapter 3.4:

- How can Elliptical Fourier Analysis be optimized for modern and fossil unionoid clam specimens?
- Can Elliptical Fourier Analysis describe the shapes of modern unionoid taxa sufficiently to allow discrimination between those taxa?
- Can Elliptical Fourier Anaylsis be used to determine how many fossil taxa may be present at the Das Goods localities?

1.2 Abbreviations

1.2.1 Institutions

CC — Concordia College, Moorhead, Minnesota, U.S.A.

DMNS — Denver Museum of Nature and Science, Denver, Colorado, U.S.A.

GSC — Geological Survey of Canada, Calgary, Alberta, Canada

INHS — Illinois Natural History Survey

MfN — Museum für Naturkunde, Berlin, Germany

NCSM — North Carolina Museum of Natural Sciences

OSU — Ohio State University, Columbus, Ohio, U.S.A.

PTRM — Pioneer Trails Regional Museum, Bowman, North Dakota, U.S.A.

UND — University of North Dakota, Grand Forks, North Dakota, U.S.A.

UND-PC — UND Paleontology Collections

1.2.2 Symbols

AD — Among-groups dispersion

ANOVA — Analysis of Variance

EFA — Elliptical Fourier Analysis

L-number — Hartman locality number (Hartman, 1998).

LSD — Least significant difference [test]

MANOVA — Multivariate Analysis of Variance

S-number — Hartman specimen number

ΣV — Sum of variance

T-number — Burton-Kelly temporary specimen number

WD — Within-group dispersion

CHAPTER 2

REGIONAL STRATIGRAPHY AND LOCALITY INFORMATION

2.1 Regional Stratigraphy

The lithostratigraphy of the Hell Creek Formation, especially in southwestern North Dakota and eastern Montana, has been exceptionally documented, specifically with regard to the upper contact with the overlying Ludlow Member of the Fort Union Formation. The close stratigraphic proximity of the K-Pg boundary to the Hell Creek-Ludlow contact, as well as continued interest in dinosaurs near the end of the Cretaceous (Pearson et al., 2001) and the evolution of the mammalian fauna before (Hunter and Archibald, 2002), during (Hunter et al., 1997; Clemens, 2002), and after the K-Pg extinction event, have provided impetus for tighter lithostratigraphic, magnetostratigraphic (Hicks et al., 2002), biostratigraphic, and palynostratigraphic (Kroeger, 2002; Sweet, 2006; Nichols, 2007) control in order to tie thousands of plant, invertebrate, and vertebrate fossil localities together in time and space.

Both the Hell Creek-Ludlow lithostratigraphic contact and the K-Pg chronostratigraphic boundary are used for approximating local positions in the Hell Creek Formation and the Ludlow Member, which are located within a short distance below or above these markers. The lithostratigraphic contact is more obvious, being the transition from the drab, "somber" gray, "popcorn" weathering Hell Creek Formation, to the continuous light browns and yellows of the variegated siltstone and claystone beds of the Fort Union (Murphy et al., 2002). The most evident small-scale marker is the base of the lowermost laterally extensive lignite coal unit, the beginning of many such thin units in the Fort Union Formation. In North Dakota,

the chronostratigraphic boundary is typically designated by pollen and leaf horizon sampling bracketed by magnetostratigraphic intervals (Hicks et al., 2002; Johnson, 2002; Nichols, 2002), and in two places (Mud Buttes and Pyramid Butte) is known to be concurrent with the iridium and ejecta spherule layer representing the globally recognized K-Pg boundary (Nichols and Johnson, 2002).

2.2 Das Goods Localities

The Das Goods mollusk localities are located in western Slope County, North Dakota on land owned by the Horse Creek Grazing Association (sec. 9, T. 134 N., R. 106 W.) (Figure 2.1). The area is accessible via Old Route 16 about 20.5 km north of the turnoff from Marmarth. Six numbered localities produce Das Goods "type" molluscan fossils in this area (Table 2.1); five on a NW/SE trending butte, approximately halfway between two bends in the road on the southeast side, and one on a parallel butte to the northeast where the road bends to the north (Figure 2.2).

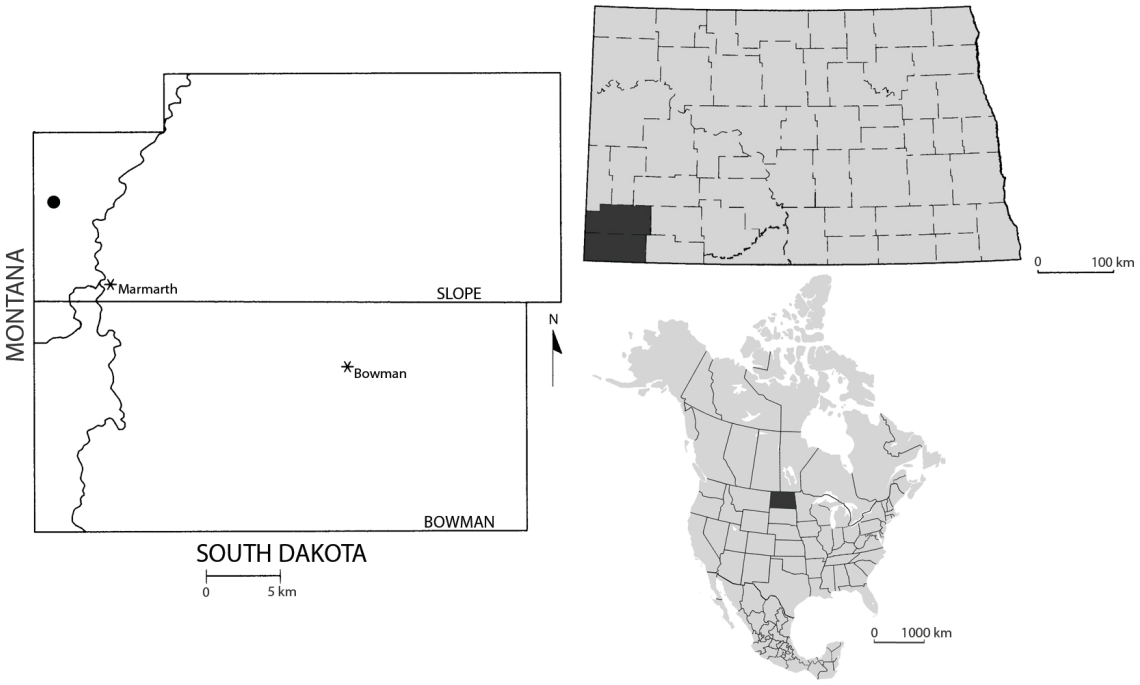

Figure 2.1: Location of Das Goods locality area in western Slope County, North Dakota, U.S.A.

Table 2.1: Locality L6516-style mollusk localities in the Das Goods area.

Locality Number	**Name**	**Latitude / Longitude (WGS 84)**	**UTM Coordinates (NAD 27)**	**Date Discovered**	**Discovered by**	**Section, Township, and Range**
L6516	Das Goods	46.4378 -104.006 El. 868 m	13T 0576440N 5142964E ±8.2 m	1998	K. R. Johnson and T. Farnham	SE NE NW NW sec. 9, T. 134 N., R. 106 W.
L6803	Das Goods III	46.43951 -104.00348 El. 867 m	13T 0576603N 5142752E ±5.7 m	06 September 2007	M. Burton-Kelly	SE SE NE NW sec. 9, T. 134 N., R. 106 W.
L6804	Das Goods IV	46.43593 -104.00373 El. 868 m.	13T 0576584N 5142755E ±6.9 m	1998	M. Burton-Kelly	SE SW NE NW sec. 9, T. 134 N., R. 106 W.
L6805	Das Goods V	46.43729 -104.00511 El. 884 m	13T 0576476N 5142904E ±10 m	07 September 2007	M. Burton-Kelly	NW SE NE NW sec. 9, T. 134 N., R. 106 W.
L6806	Das Goods II	46.43759 -104.00519 El. 867 m	13T 0576470N 5142937E ±8 m	06 September 2007	M. Burton-Kelly	SW NW NE NW sec. 9, T. 134 N., R. 106 W.
L6807	Das Goods VI	46.43898 -104.0016 El. 862 m	13T 0576743N 5143095E ±5.7 m	07 September 2007	M. Burton-Kelly	NW NE NE NW sec. 9, T. 134 N., R. 106 W.

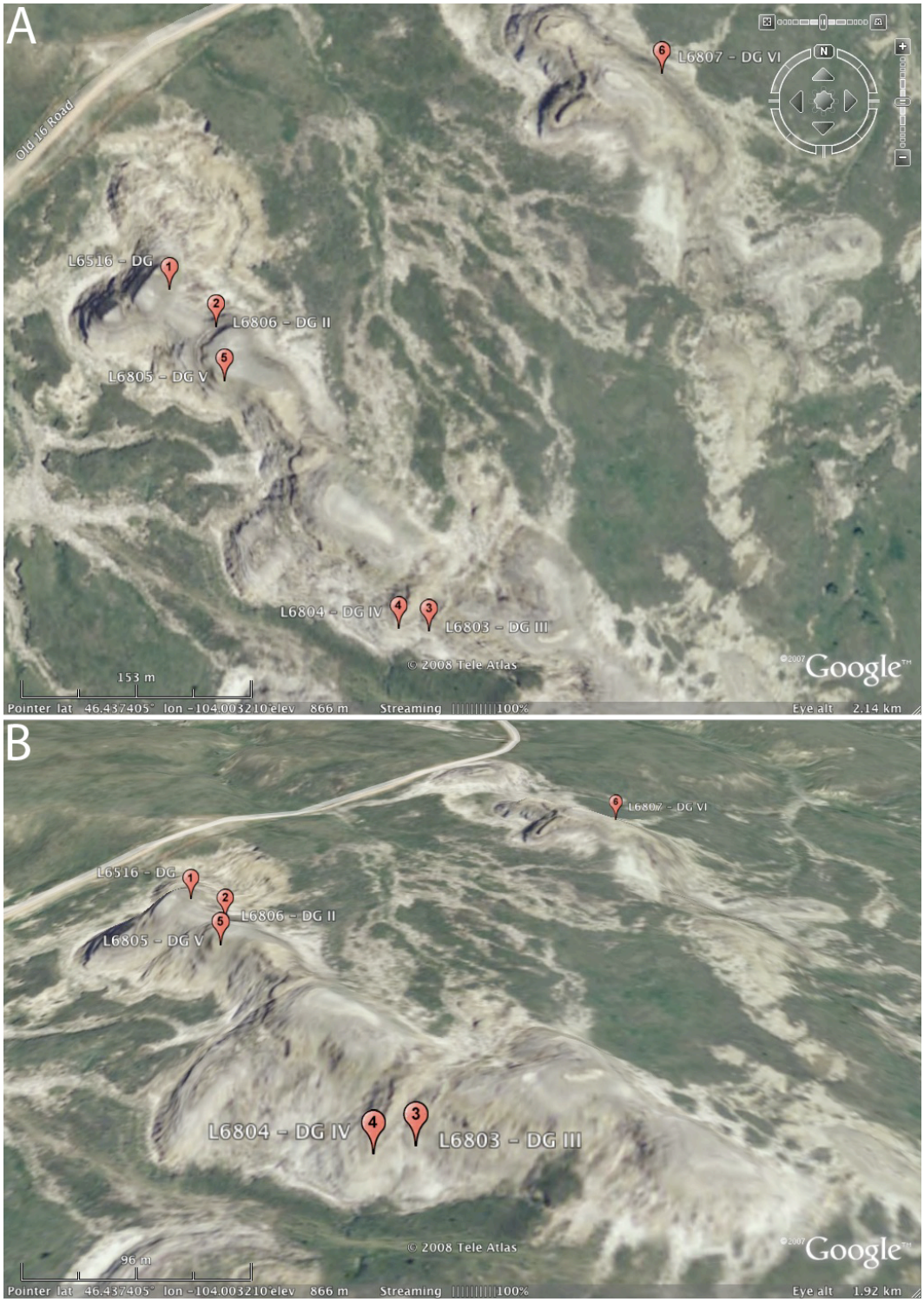

Figure 2.2: L6516-type mollusk localities in the Das Goods area. A–Google Earth^{TM} view of the Das Goods area. Access to the sites is via Old Route 16 (dirt road to northwest). B–Oblique north-facing Google Earth^{TM} view of the Das Goods area. Sites are numbered according to order of discovery.

The mollusk-producing horizon at Das Goods locality L6516 is 83.5-90.5 cm above the base of the lowermost coal (the base of the Ludlow Member). This contact lignite delimits the uppermost Hell Creek Formation and the lowermost Ludlow Member of the Fort Union Formation in this area (Murphy et al., 2002) (Figure 2.3). The molluscan assemblage at L6516 was discovered while measuring the stratigraphy for Johnson's (2002) fossil leaf locality 9858 (2217) and was first reported by Hartman et al. (2001). The leaf flora was summarized by Johnson (2002). Hartman et al. (2001) placed the top of the mollusk-producing horizon from 0-50 cm below the K-Pg boundary based on pollen. Later analysis of independently sampled pollen shifted and narrowed this range to 63-73 cm below (Sweet, 2006; Hartman et al., 2007). The Das Goods leaf locality 9858 (2217) located 2.39 m above the Hell Creek-Fort Union contact contains leaves from Johnson's (2002) FUI megafloral zone. The megafloral zone of the leaf layer directly underlying the mollusk-producing horizon at L6516 is difficult to estimate based on stratigraphy alone; no in-depth study of these leaves was undertaken, so no megafloral zone placement can be made at this time. Currently 176 numbered mussel or gastropod specimens have been recovered; additional unnumbered specimens exist in the UND-PC collections (Appendix E.2).

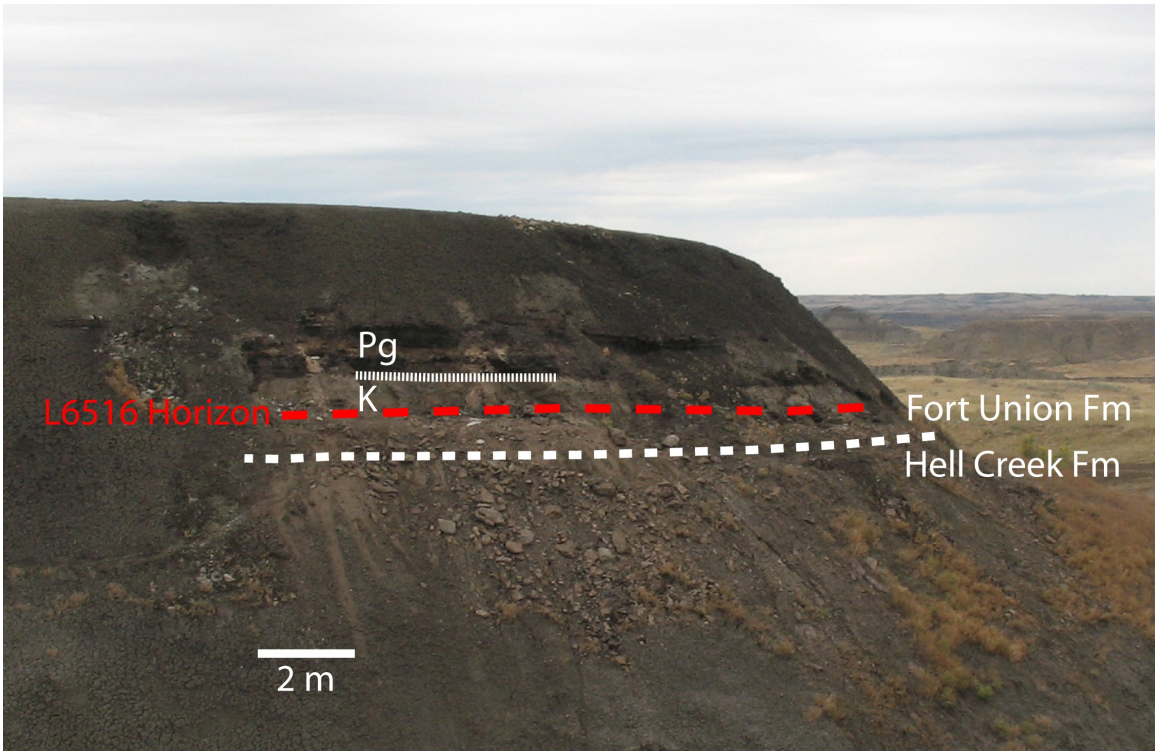

Figure 2.3: Original Das Goods mollusk locality (L6516). View is to the northwest. Thick dashed line approximates location of mollusk-producing horizon, thin dashed line approximates top of Hell Creek-Fort Union formational contact coal, dotted line approximates K-Pg boundary based on pollen (Sweet, 2006; Hartman et al., 2007). The Das Goods plant locality 9858 (2217) is located approximately 2.39 m above the formational contact and 1.49 m above the L6516 horizon. Old Route 16 is visible in the background.

Figure 2.4: Stratigraphic column around L6516, between Hell Creek-Ludlow contact coal and second Ludlow coal. Chronostratigraphy is based on pollen analysis (Sweet 2006). Lithologic units are numbered according to relative position to contact coal. Figure integrates information from Hartman et al. (2001, 2007). Description is as follows:

Unit **22**: Light grey fissile plant-rich shale with abundant roots. 35 cm thick.

Unit **21**: Plant-producing layer (9858 (2217)). Light grey fissile plant-rich shale with abundant roots. 30 cm thick.

Unit **20**: Brown carbon-rich fissile shale, coally at top. 22 cm thick.

Unit **19**: Very dark grey carbon-rich shale. 3.5 cm thick.

Unit **18**: Pink tonstein. 3 cm thick.

Unit **17**: Vitreous to dull black coal. 19 cm thick.

Unit **16**: Light grey to light brown coarse siltstone, rooted at top. 12 cm thick.

Unit **15**: Vitreous to dull black coal. 10 cm thick.

Unit **14**: Moderate brown platey claystone with coalified horizons and gypsum crystals. 7 cm thick.

Unit **13**: Grey-brown claystone with coalified horizons and gypsum crystals. 7 cm thick.

Unit **12**: Olive black platy to papery claystone with some plant debris. 3 cm thick.

Unit **11**: Light yellow grey homogeneous massive fine-grained sandstone with minor plant debris. 24 cm thick.

Unit **10**: Light olive-green to yellow-grey fine-grained sandstone with small nodules. 4 cm thick.

Unit **9**: Yellow-orange to grey-yellow coarse speckly very fine- to fine-grained sandstone with lenses of clay or silty clay. 4 cm thick.

Unit **8**: Pale yellow-brown massive fine-grained sandstone. Some organic debris. 12 cm thick.

Unit **7**: Pale brown slightly silty/clayey fine-grained sandstone with organic debris along bedding planes. 10 cm thick.

Unit **6**: Speckly fine-grained sandstone with plant debris and small nodules. 9 cm thick.

Unit **5**: Mollusk-producing layer (L6516). Pale yellow-brown silty very fine-grained sandstone with plant debris. Splits preferentially along clams or other debris. 7 cm thick.

Unit **4**: Yellow-grey or -brown siltstone with small nodules and plant stems. 16 cm thick. Olive grey when damp.

Unit **3**: Slightly clayey light olive grey siltstone with small nodules. 20 cm thick.

Unit **2**: Yellow-grey to light grey silty claystone with organic matting. 20 cm thick.

Unit **1**: Vitreous to dull black coal. 25 cm thick. Base is the Hell Creek Formation-LudlowMember of the Fort Union Formation contact.

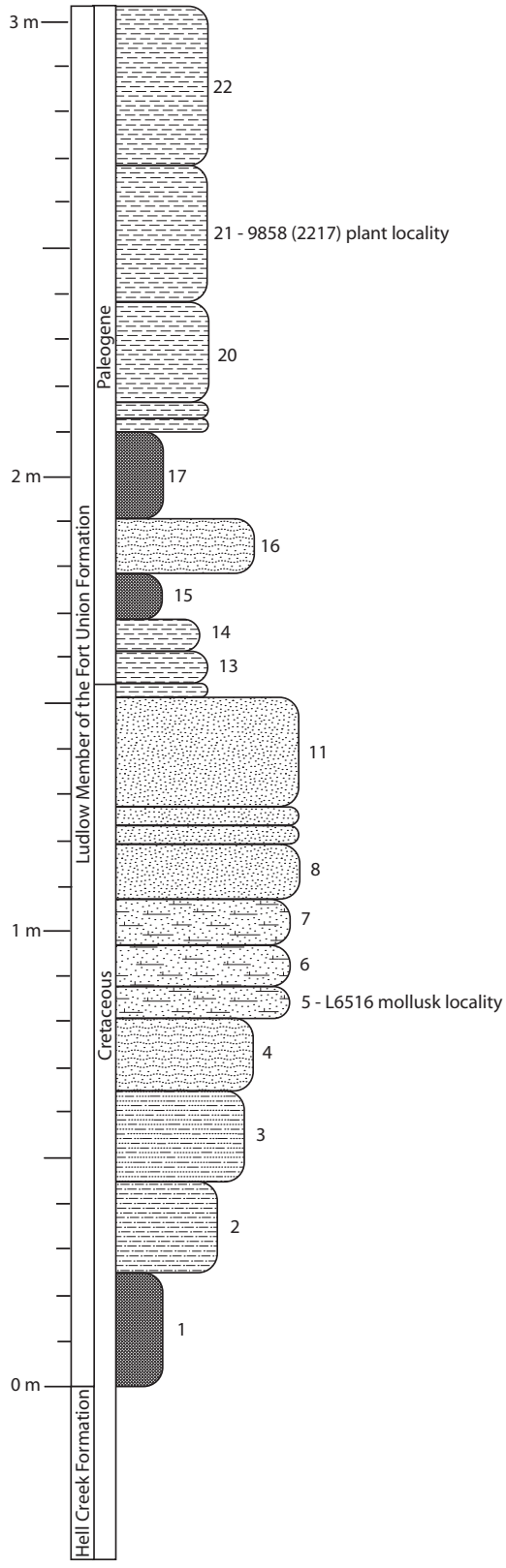

Figure 2.4: See opposite page for explanation.

Less stratigraphic information is known about the other Das Goods localities. The other localities were not tied to a stratigraphic datum, but all are found between the basal coal and the second coal (out of up to four coals locally) of the Ludlow Member, a thickness of approximately 1.46 m based on measurements at locality L6516, but varying throughout the immediate area. Of the six localities, Das Goods VI (L6807) is closest to the second coal (Figure 5).

Das Goods II (L6806) is located to the southeast of L6516, across a low saddle (Figure 2.2). This locality produced a single valve fragment and snail impression (specimen S2950), but was not extensively quarried because of the amount of overburden that would need to be removed in order to expose the producing horizon and the visibility of the locality from the road (attempts were made to minimize the visual impact of fossil collecting activities in this area).

Das Goods III (L6803) is located approximately 260 m southeast of L6516, on the southwest slope of the same long butte (Figure 2.5). Again, the amount of overburden present because of the steep slope prevented a large-scale quarrying operation, but five molluscan fragments were produced (specimens S2951 through S2955).

Das Goods IV (L6804) is located on the tip of a southwest-trending spur off the main butte, approximately 20 m west of L6803 (Figure 2.5). There is very little overburden and the potential to excavate more fossils along the producing horizon (Figure 2.6). This site was quarried for an afternoon by members the UND Introduction to Paleontology class and produced 41 specimens (S2956 through S2996). A stratigraphic section was measured nearby (D. Pearson, PTRM, pers. comm., 2008), but correlation between the measured section and L6804 has been difficult.

Das Goods V (L6805) is located approximately 35 m south of L6806, in conditions very similar to L6804, resulting in little excavation after the initial discovery (3 specimens, S3247-S3249).

Figure 2.5: Das Goods III (L6803; right dashed line) and IV (L6804; left dashed line) mollusk localities. View is to the north. Stratigraphic section measured by D. Pearson, J. Hunter, M. Borths, N. Burgei, D. Weinstein, and J. Wood on 9 September 2007.

Das Goods VI (L6807) is located on the northeast side of the butte parallel to and northeast of the butte containing the other localities (Figure 2.2). Twenty mollusk fragments were quarried over the course of a few days (S2997-2999, S3230-3246). L6807 is estimated to be the closest of all known localities in the area to the second Ludlow coal (the first coal above the Hell Creek-Ludlow contact coal), but no stratigraphic column was measured (Figure 2.7).

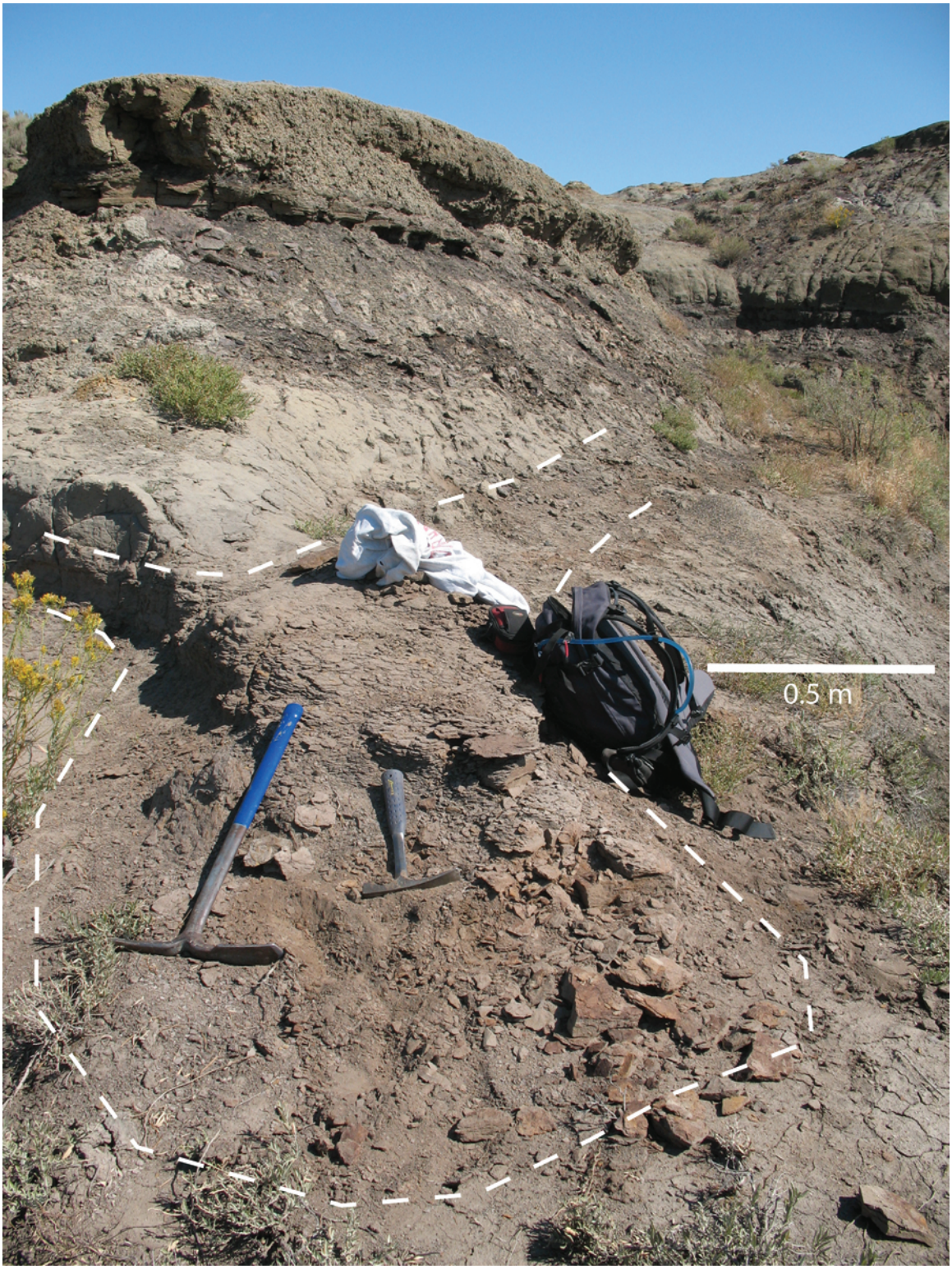

Figure 2.6: Das Goods mollusk locality IV (L6804). Dashed line approximates upper and lower boundaries of producing horizon in foreground. View is to the north.

Figure 2.7: Das Goods VI mollusk locality (L6807). View is to the southwest. Dashed line approximates location of mollusk-producing horizon. A: Oblique view, before excavation. B: Normal view, overlying layers exposed. Scales are equal.

CHAPTER 3

PRESERVATION OF SELECTED L6516 UNIONOID SPECIMENS

3.1 Purpose

The aim of this chapter is to describe and illustrate the preservation of the most well-preserved unionoid specimens from locality L6516.

3.2 Introduction

Twenty-seven valves of specimens from L6516 were available and preserved in enough detail to allow description of the variation in preservation over the surface of a single valve. Descriptions were done with respect to 1) the quality and position of preserved growth lines, 2) whether the valve was preserved as an internal or external mold or cast, and 3) the character (concave or convex epirelief), and 4) position of trace fossils on the valve surface (Osgood Jr., 1987). For these descriptions, concave or convex epirelief refers to the apparent relationship between traces and the surface of the specimen as it appears, not to the relationship between the traces and the original shell material or periostracum. Trace fossils were not measured for size classes. All valves are figured at the same scale.

Of the 27 specimens examined (14 left and 13 right), 14 are preserved as concave forms (e.g. S2928, S2781aL) and 13 as convex (e.g., S2800, S2835). Of these, 2 are complete casts of the exterior of the valve (S2924, S2883), 13 are complete molds of the exterior of the valve (e.g., S2922, S2942), and 12 are a combination of mold and cast of the interior and exterior (e.g., S2754, S2781aR). All retain at least some growth lines; 7 specimens are more than about 75% covered (e.g., S2929, S2946), 8 are approximately between 50% and 75% covered (e.g., S2773, S2879), and 12 are less

than approximately 50% covered with growth lines (e.g., S2865, S2876). Trace fossils are preserved on the surface of all examined valves; 4 specimens possess traces in only concave epirelief (e.g., S2886, S2923) and 4 in only convex epirelief (e.g., S2882, S2948), 6 have an approximately equal covering of concave and convex epirelief (e.g., S2920), 5 have more concave epirelief traces than convex (e.g., S2921, S2944, S2947), and 7 have more convex epirelief than concave (e.g., S2881, S2930).

Apart from a few specimens (S2754, Figure 3.2; S2947, Figure 3.27), there seems to be no relationship between the preservation of the valve surface and the character of the trace fossils. Growth lines are almost always more apparent near the margins of the valves.

Descriptions of the preservation of each valve follow in the figure captions below.

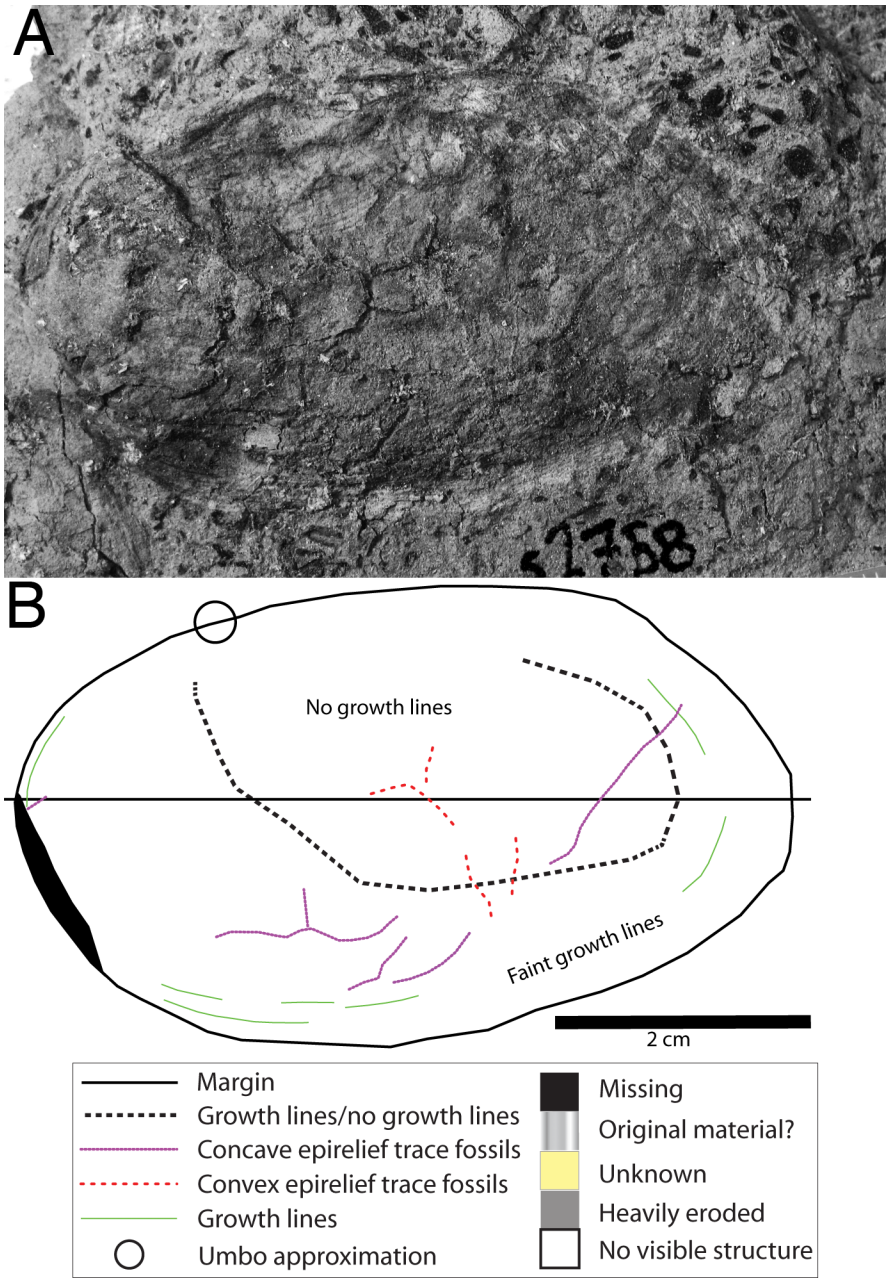

Figure 3.1: Preservation of specimen S2758, a slightly convex left valve. Growth lines are conspicuous in anteroventral region and faint on posterodorsal. There are few trace fossils, most in concave epirelief (the larger traces). Poor preservation; surface is not smooth, valve seems compacted posterior to umbo. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

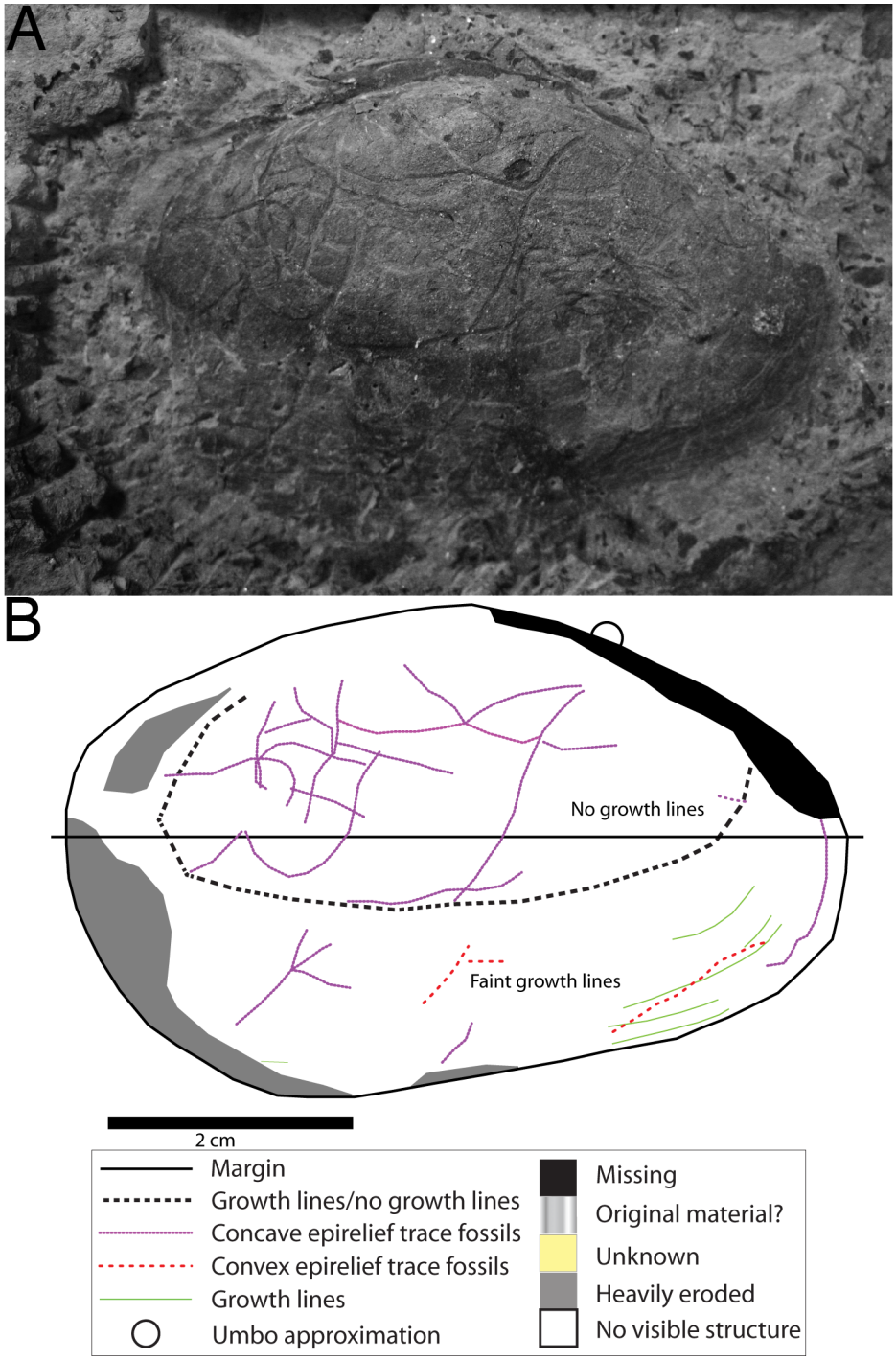

Figure 3.2: Preservation of specimen S2754. Convex right valve has poor preservation of posterior and faint growth lines on the ventral margin. Umbonal cavity (mold of the interior) and marginal growth lines (cast of the exterior) are preserved. Most traces in central part of valve are larger and in concave epirelief and smaller and in convex epirelief near margins. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

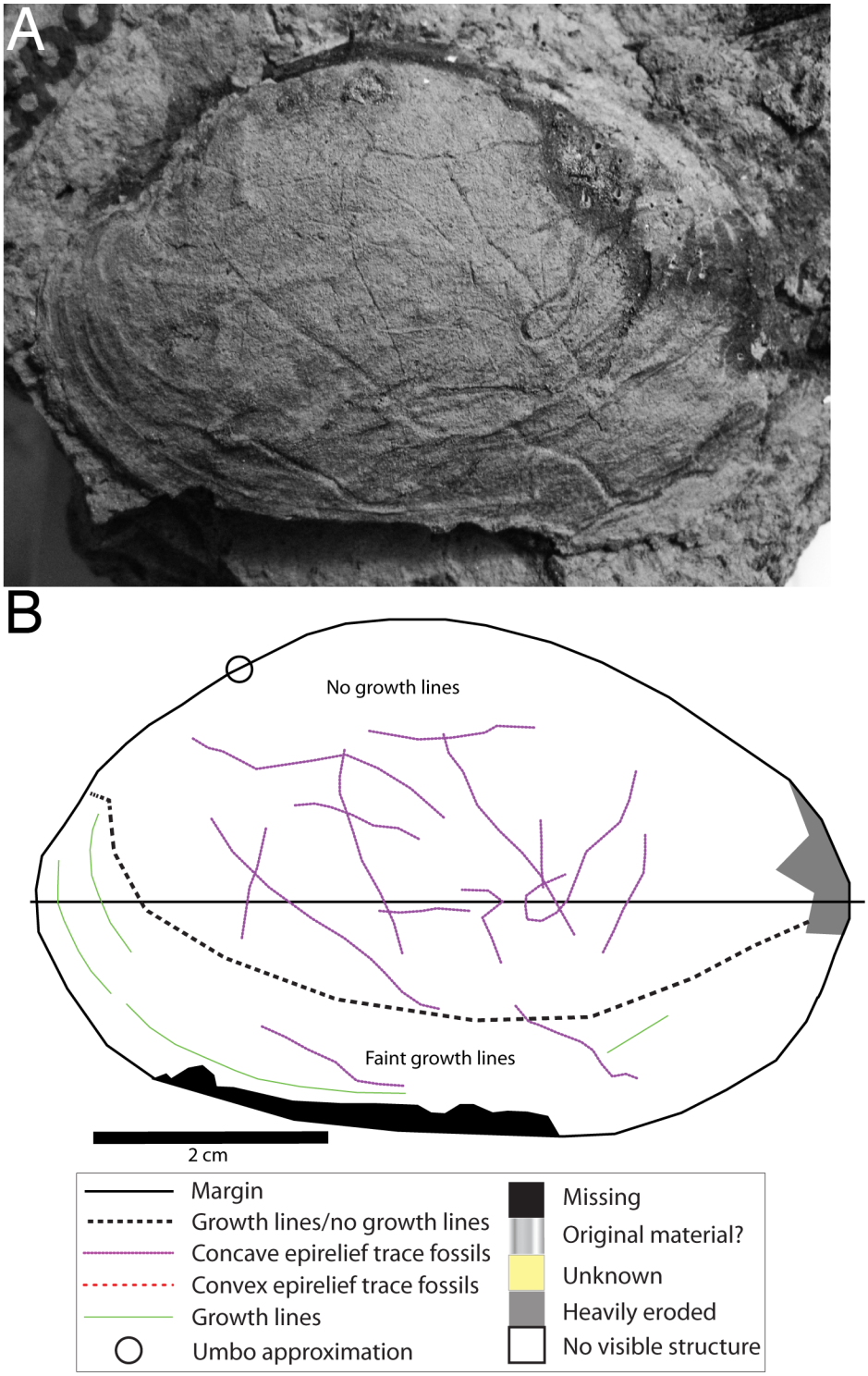

Figure 3.3: Preservation of specimen S2800, a convex left valve. A mold of the interior (center) and cast of the exterior (edges) are preserved. Interior umbonal cavity is preserved as well as growth lines. Most traces are in concave epirelief, some in convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

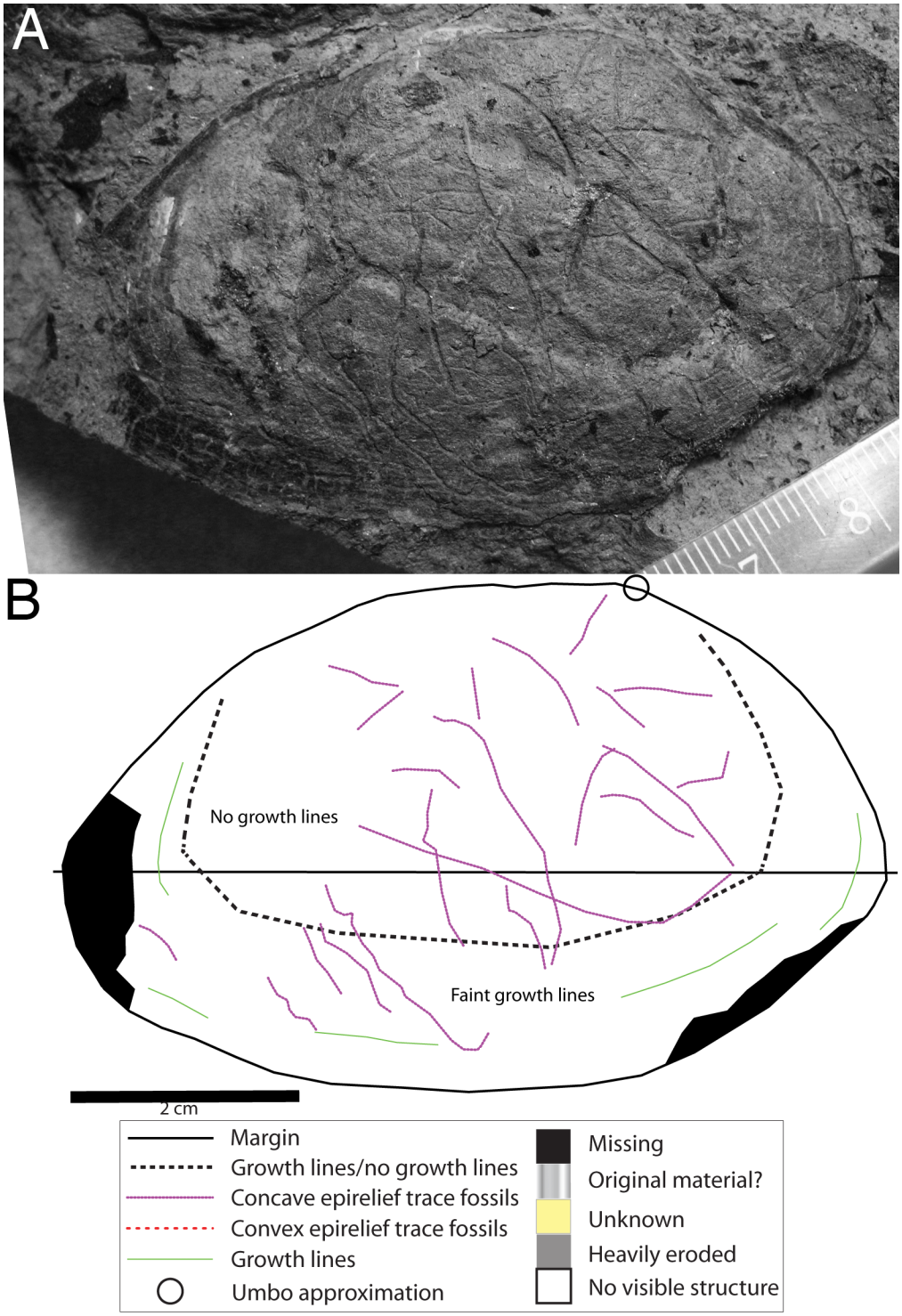

Figure 3.4: Preservation of specimen S2835, a convex right valve. A mold of the interior (center) and cast of the exterior exterior (edges) are preserved. The interior umbonal cavity is preserved as well as growth lines. Most traces are in concave epirelief, some in convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

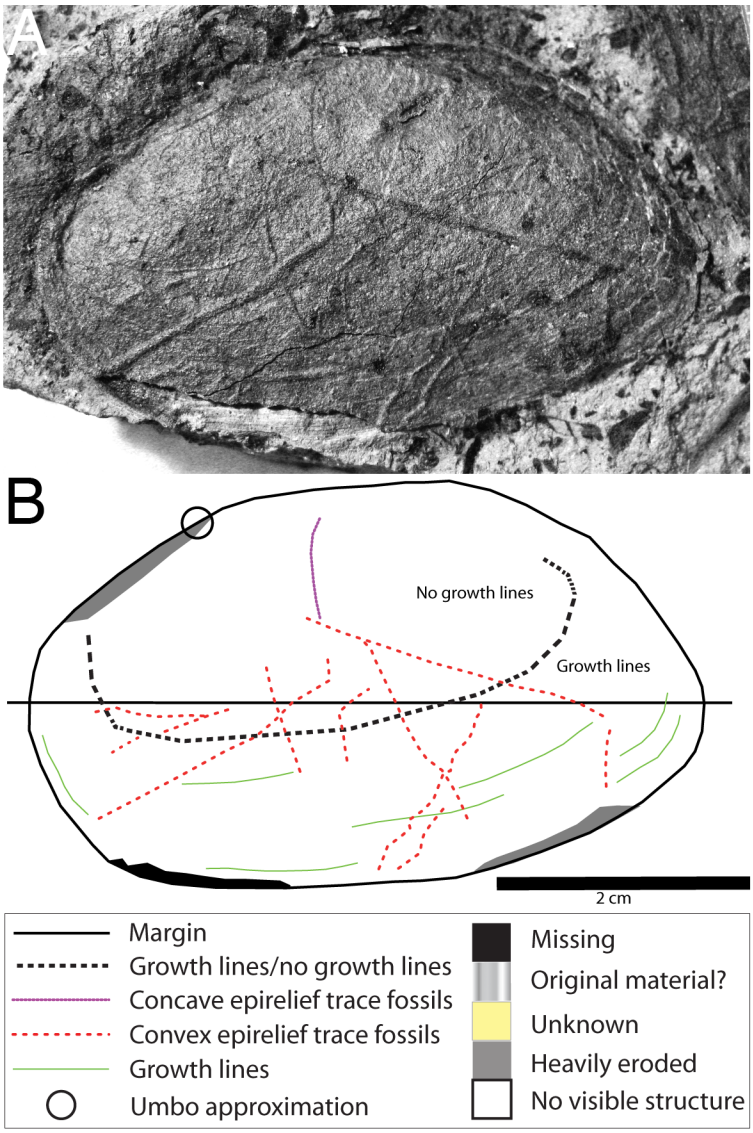

Figure 3.5: Preservation of specimen S2881, a convex left valve. Growth lines preserved on margin (cast of exterior). The middle is a mold of interior (the umbonal cavity is vaguely preserved). Trace fossils are mostly preserved in convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

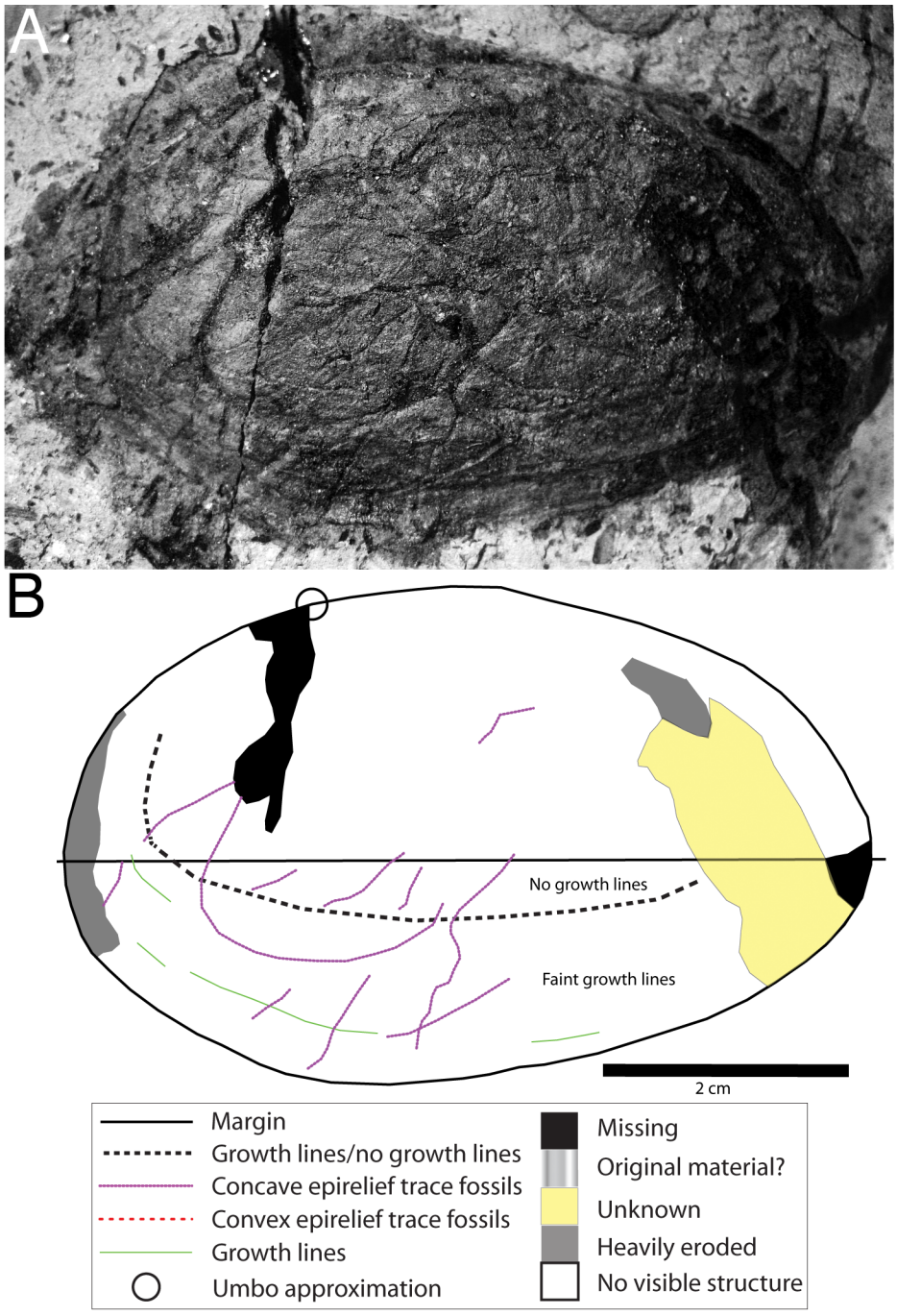

Figure 3.6: Preservation of specimen S2882, a convex left valve. Faint growth lines on the ventral margin (a cast of exterior) are preserved. The central area is a mold of the interior and lacks a defined umbonal cavity. All trace fossils in are preserved in convex epirelief. The area marked on the posterior appears to be plant debris cutting through the shell. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

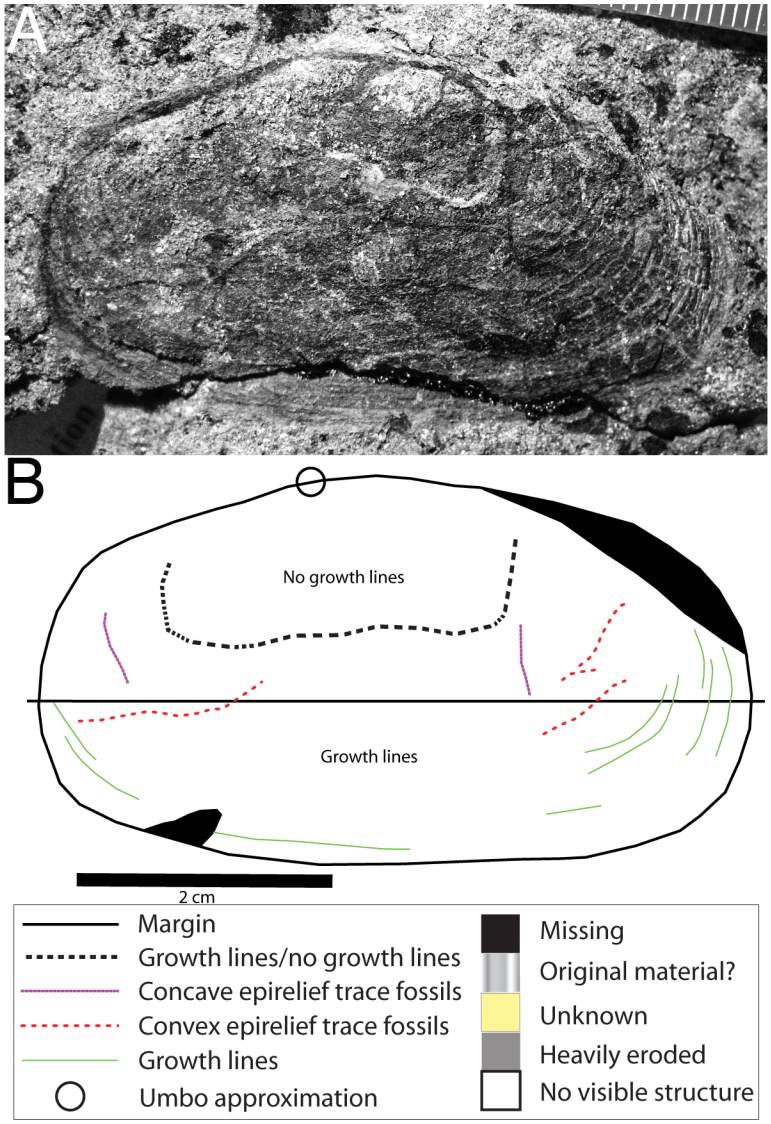

Figure 3.7: Preservation of specimen S2883, a slightly convex left valve. Growth lines are preserved over most of the valve (a cast of the exterior). The few trace fossils are preserved in both concave and convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

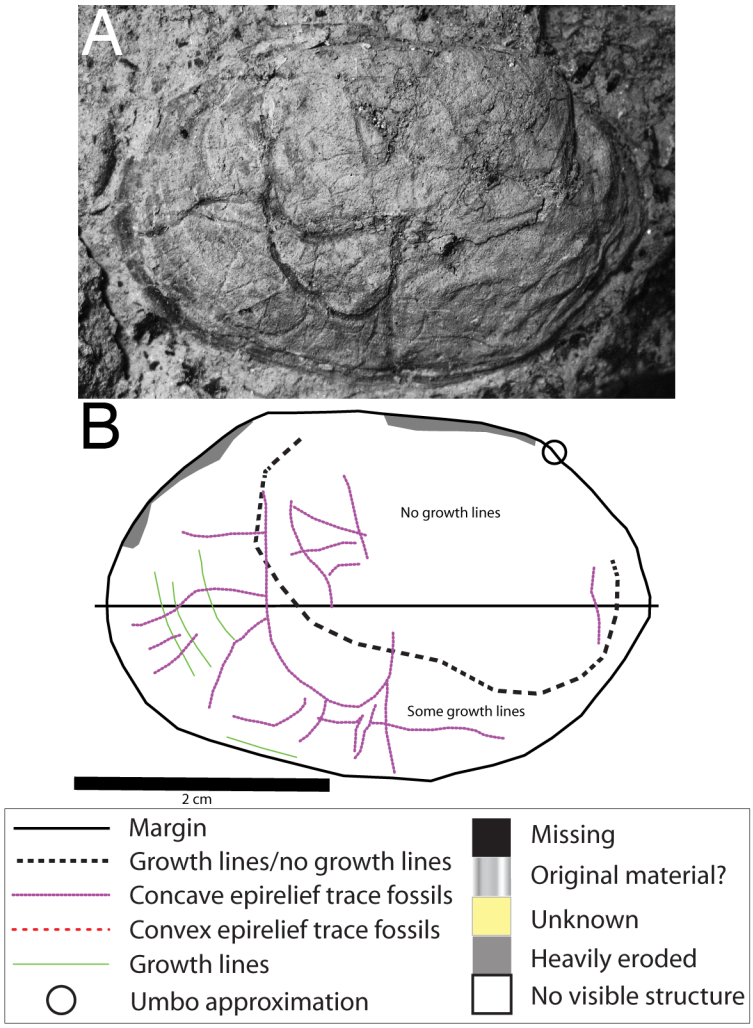

Figure 3.8: Preservation of specimen S2886, a convex right valve. Growth lines are preserved on the ventral and posterior margins (this is a cast of exterior). Center of valve is a mold of the interior (the umbonal cavity is visible). Trace fossils are all in concave epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

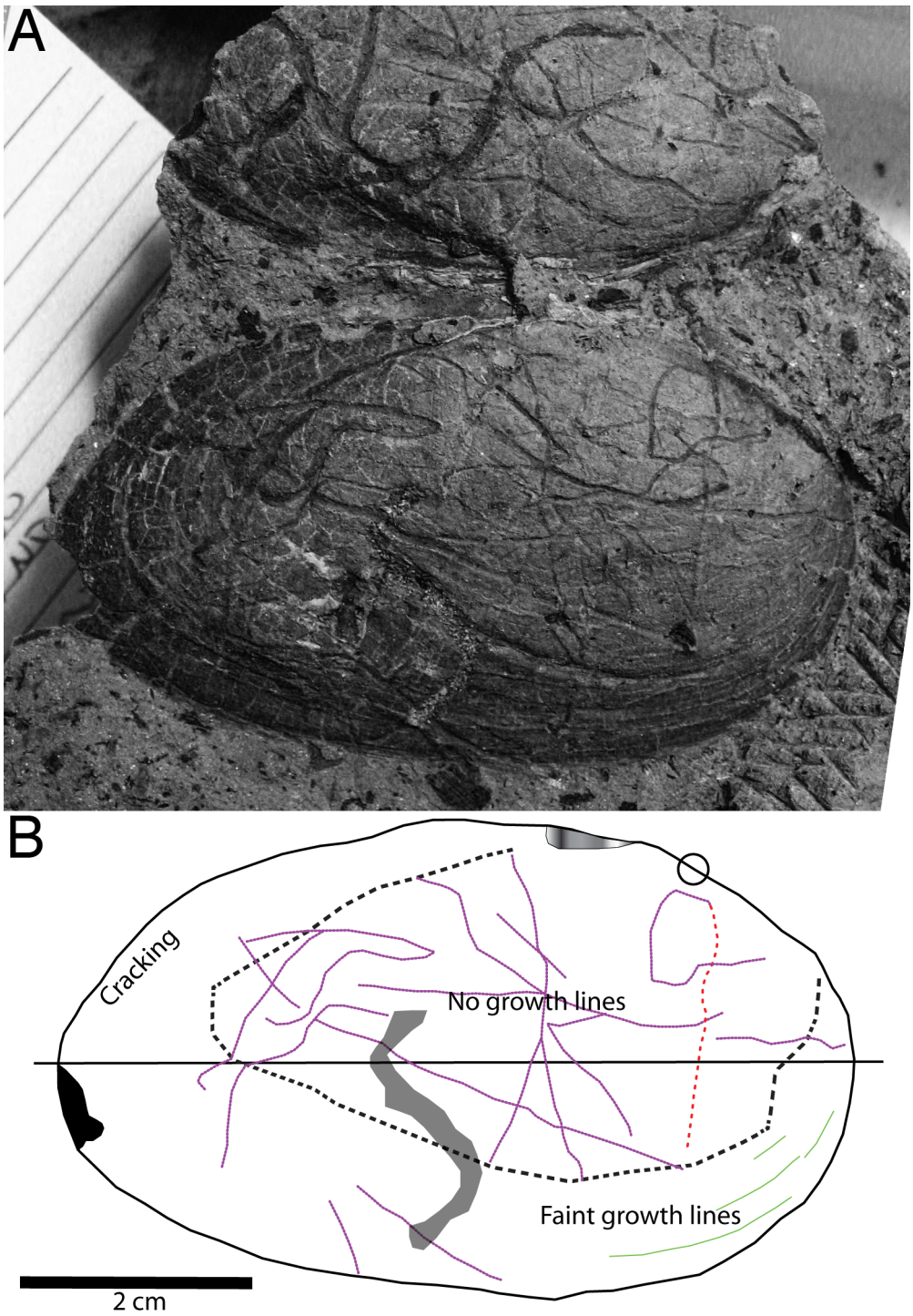

Figure 3.9: Preservation of specimen S2921, a convex right valve. Posterior margin is cracked. Faint growth lines are preserved only on the ventral anterior. There are many traces, almost all in concave epirelief. The posterior and ventral margins preserve a cast of the exterior; the anterior and dorsal margin and umbo preserve a mold of the interior of the valve. The central area is stained and shell evidence is removed; this could be a root trace but age is undetermined. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

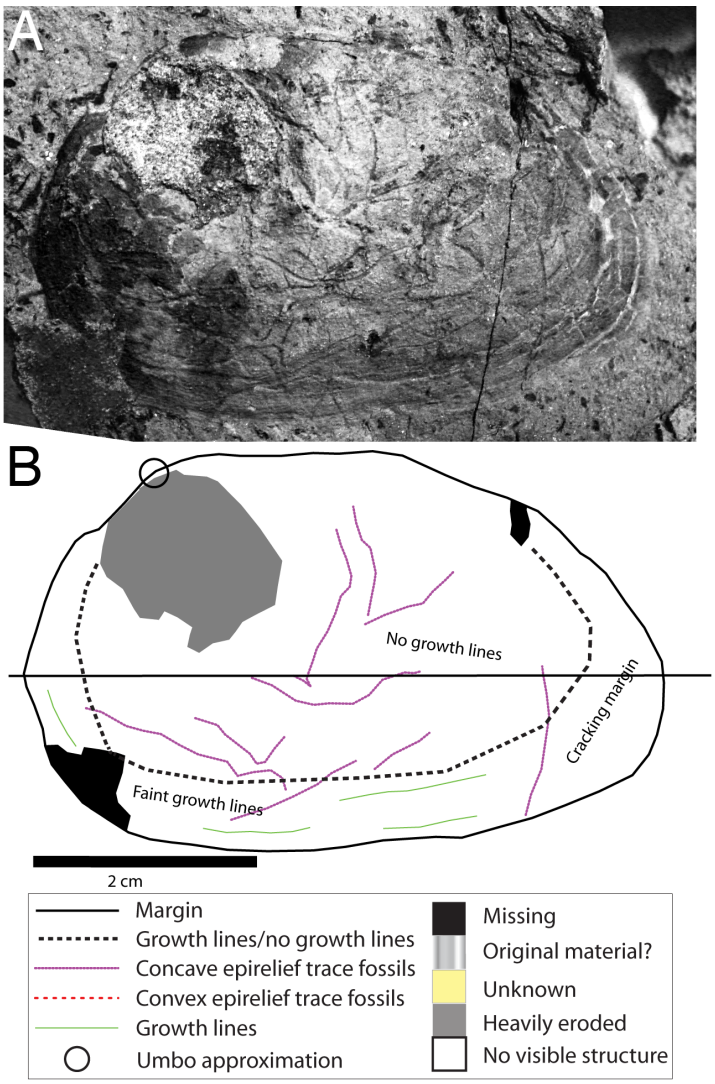

Figure 3.10: Preservation of specimen S2923, a convex left valve. Posterior margin is slightly cracked. The posterior and ventral margins preserve a cast of the exterior; the anterior and dorsal margin and umbo preserve a mold of the interior of the valve. Trace fossils are preserved in concave epirelief. There is a possible root trace through umbonal area. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

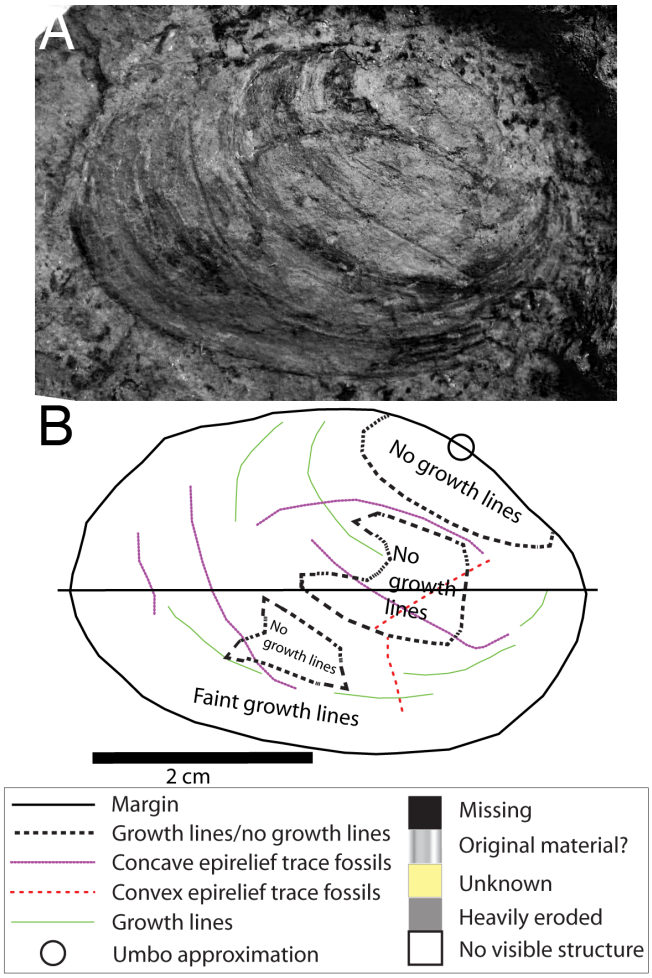

Figure 3.11: Preservation of specimen S2924, a convex right valve. Edges are probably incomplete. Valve is not cracked and preserves few growth lines. There is little evidence of the shape of the umbonal cavity. This specimen is likely a cast of the exterior. Trace fossils are preserved in both concave and convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

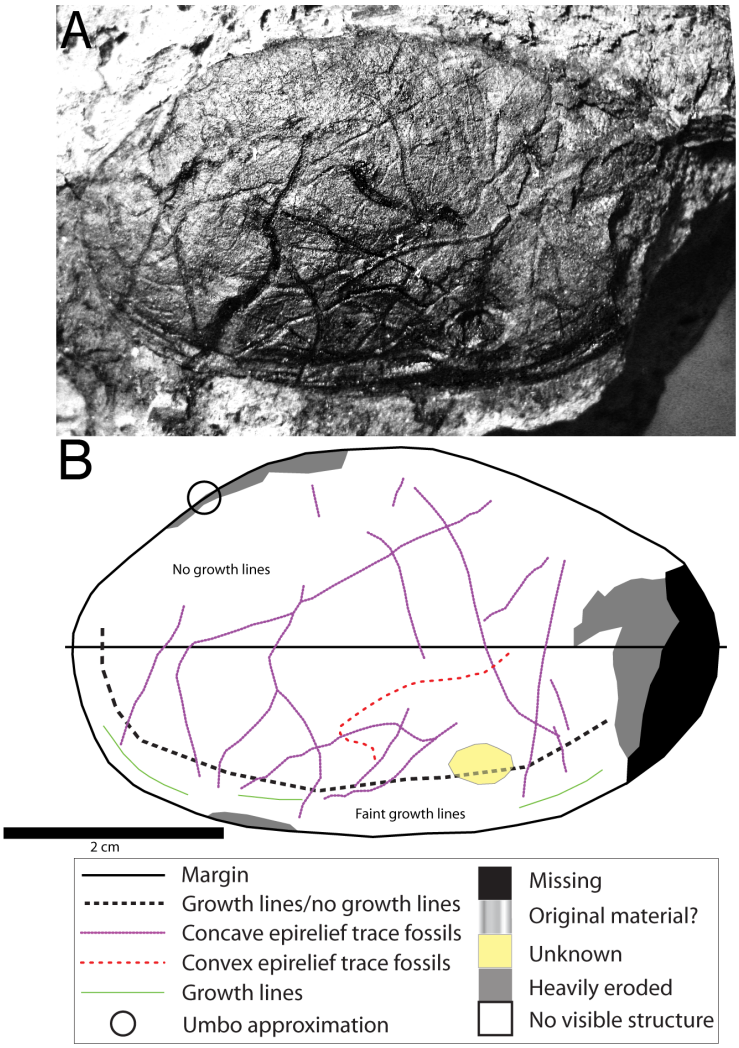

Figure 3.12: Preservation of specimen S2944, a convex left valve. Growth lines are preserved on the extreme ventral margin. Preserved as a mold of the interior in the center, a cast of the exterior on the margin. Trace fossils are mostly preserved in concave epirelief. Marked area near ventral posterior margin indicates a hole through the surface of the valve. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

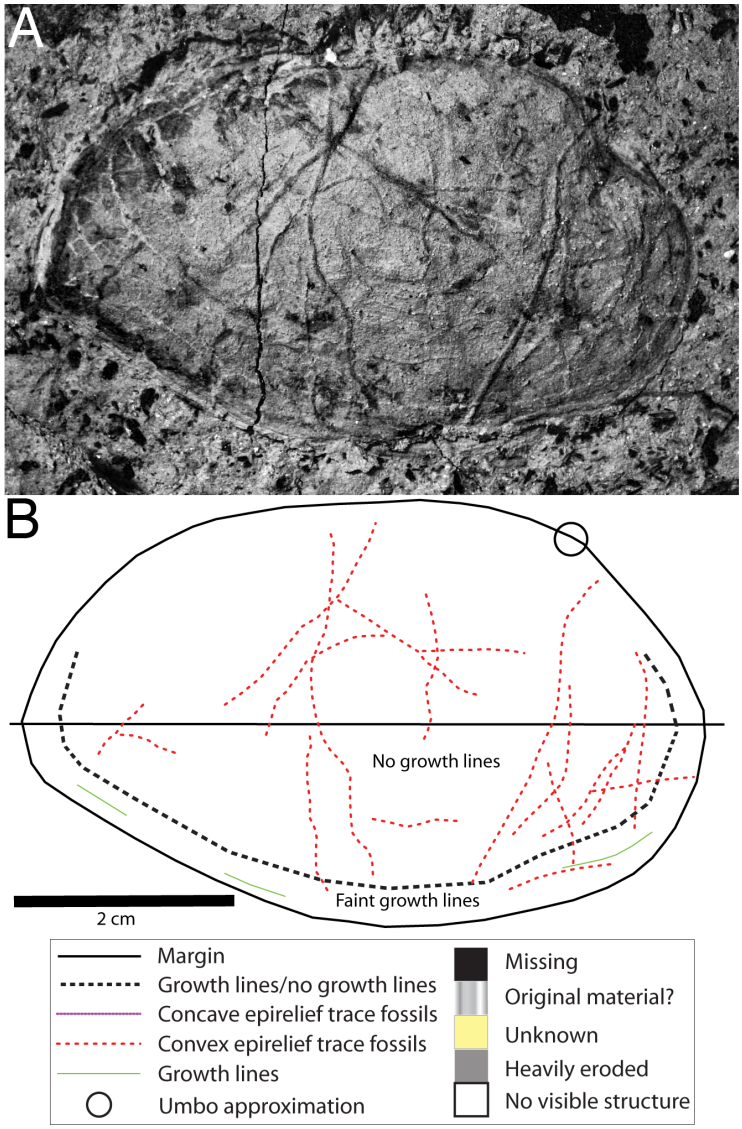

Figure 3.13: Preservation of specimen S2948, a convex right valve. The mold of the interior (umbonal cavity) is clear, faint growth lines (cast of exterior) are preserved near the margins. All traces are preserved in convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

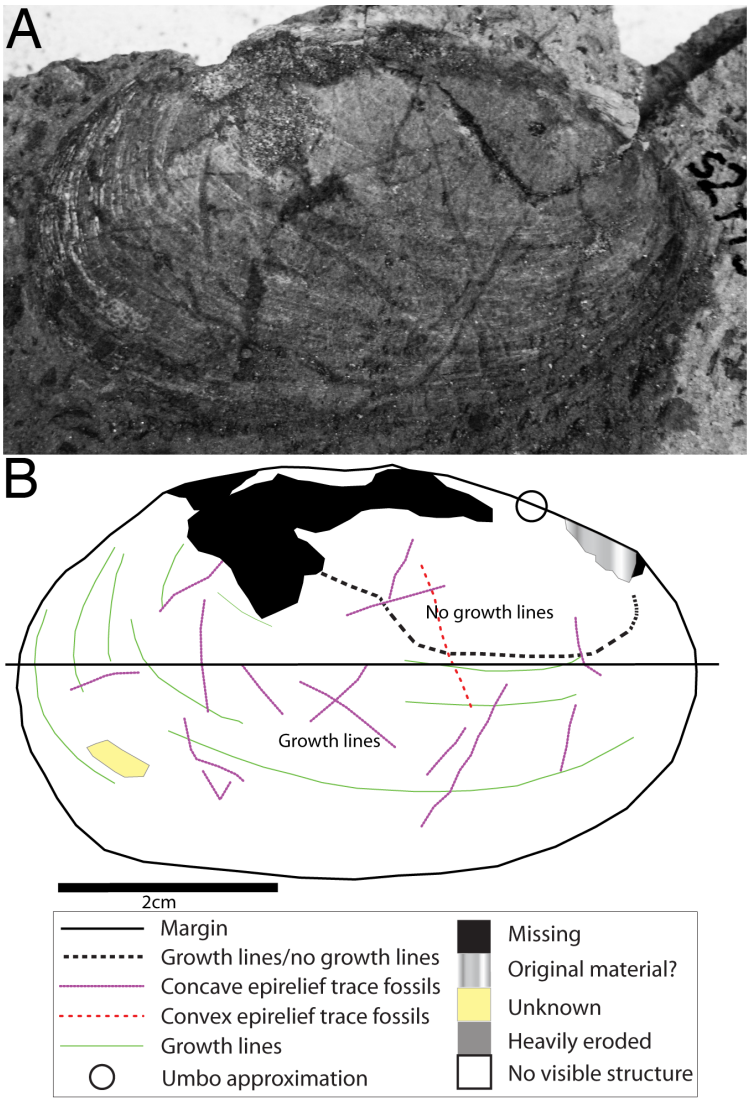

Figure 3.14: Preservation of specimen S2773, which is flat or slightly concave and probably a mold of the exterior of the left valve. Traces appear as both concave and convex epirelief on the surface of the specimen. Convex epirelief traces are much more common. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

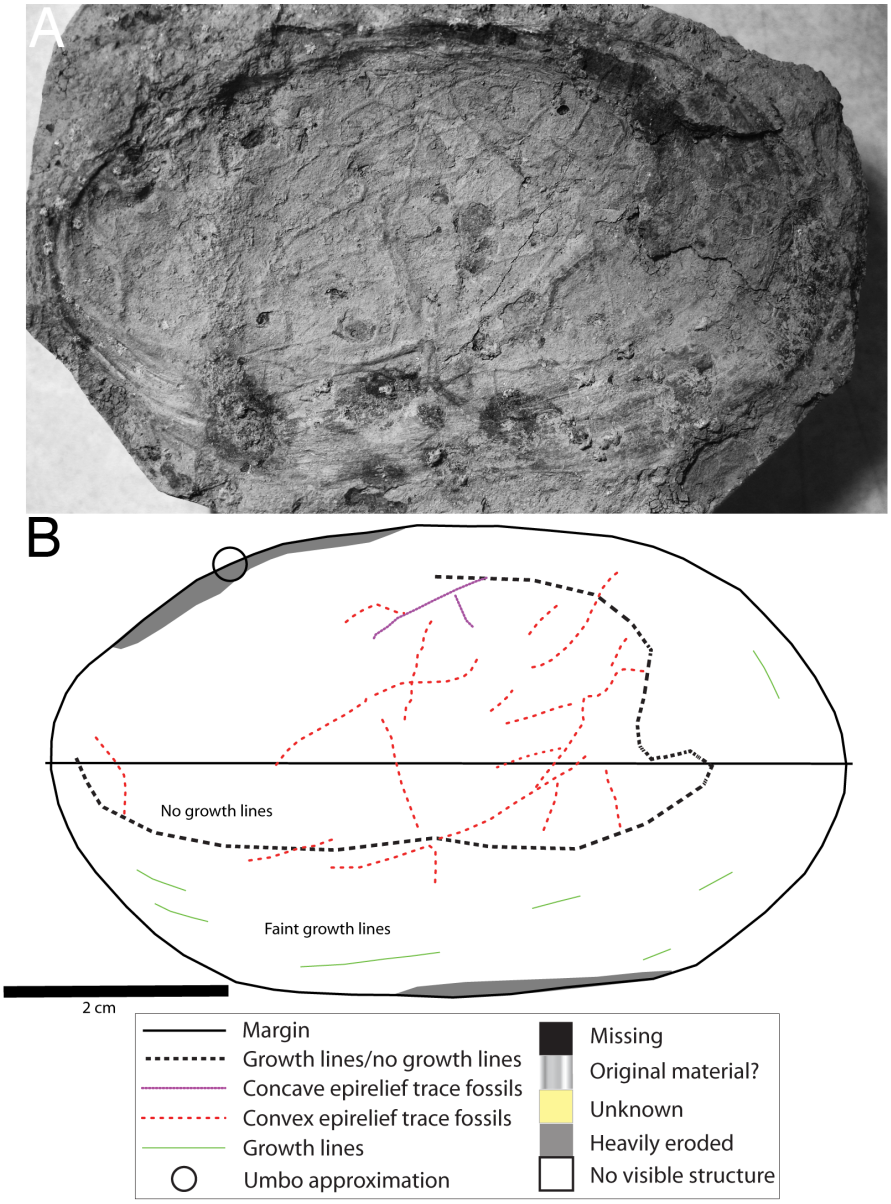

Figure 3.15: Preservation of specimen S2865, a concave right valve. Growth lines are preserved around the margins (this is a mold of the exterior). Most traces in are in convex epirelief, at least one in concave epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

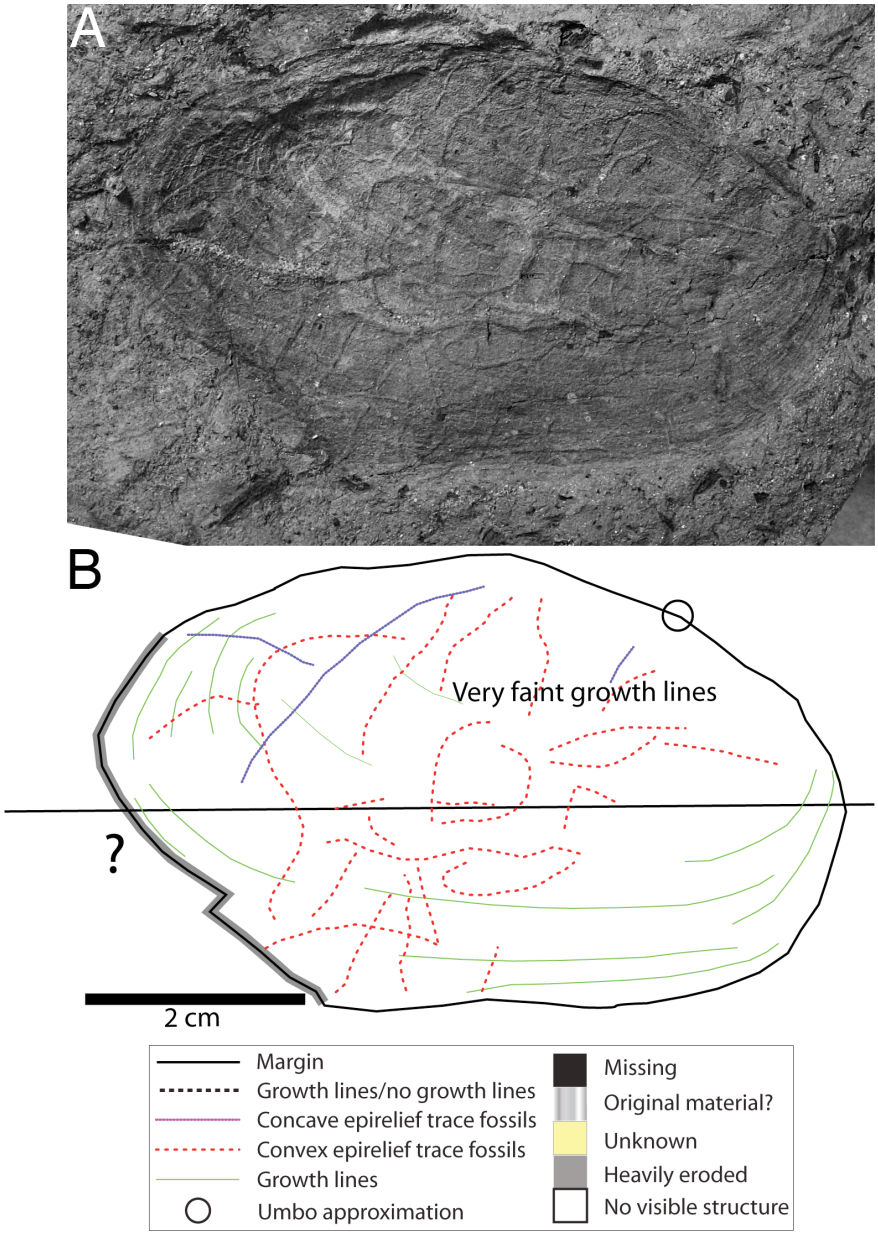

Figure 3.16: Preservation of specimen S2871aL, a concave left valve preserved as a mold of the exterior. Faint growth lines are preserved over the entire surface. Most trace fossils are preserved in concave epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

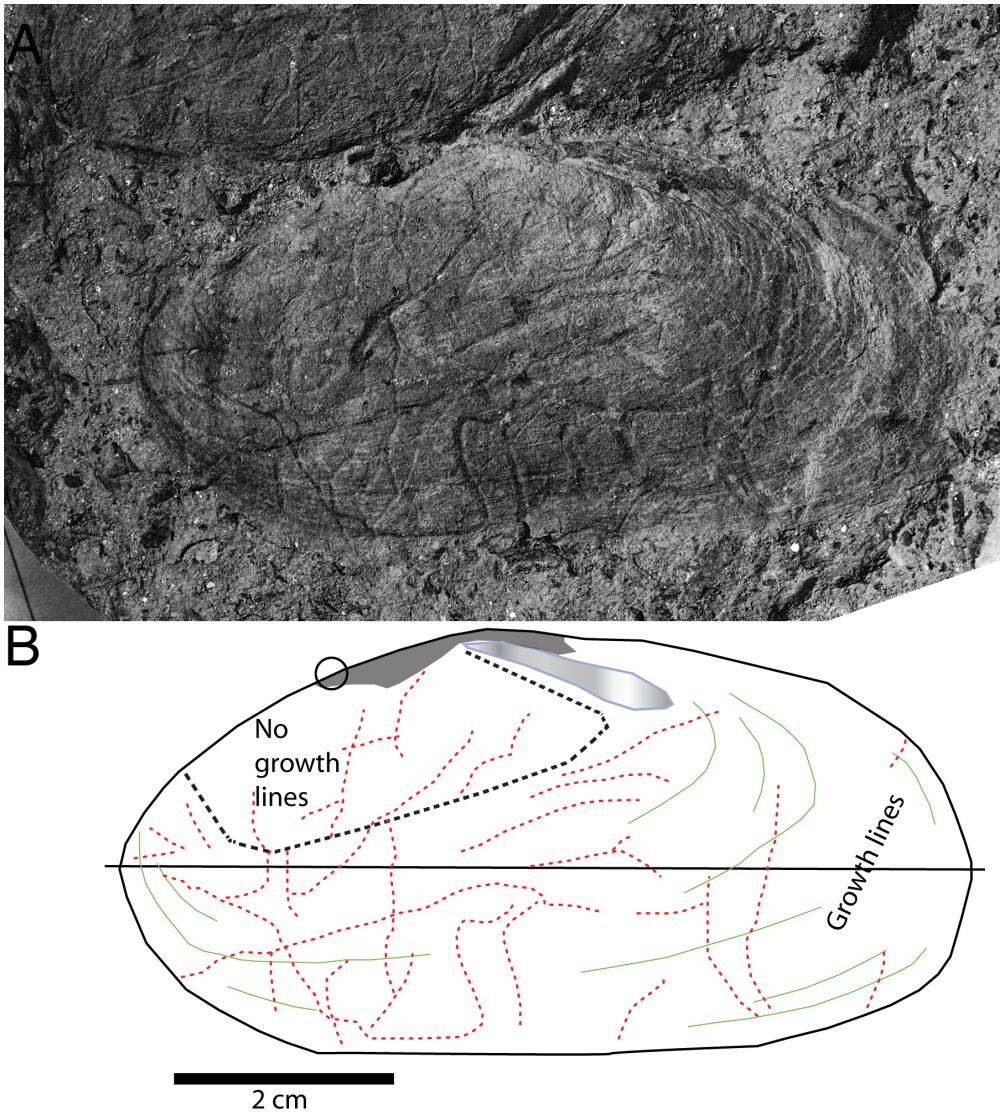

Figure 3.17: Preservation of specimen S2871aR, a concave right valve. Specimen is preserved as a mold of exterior (most) and cast of the interior (center). All traces are in convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

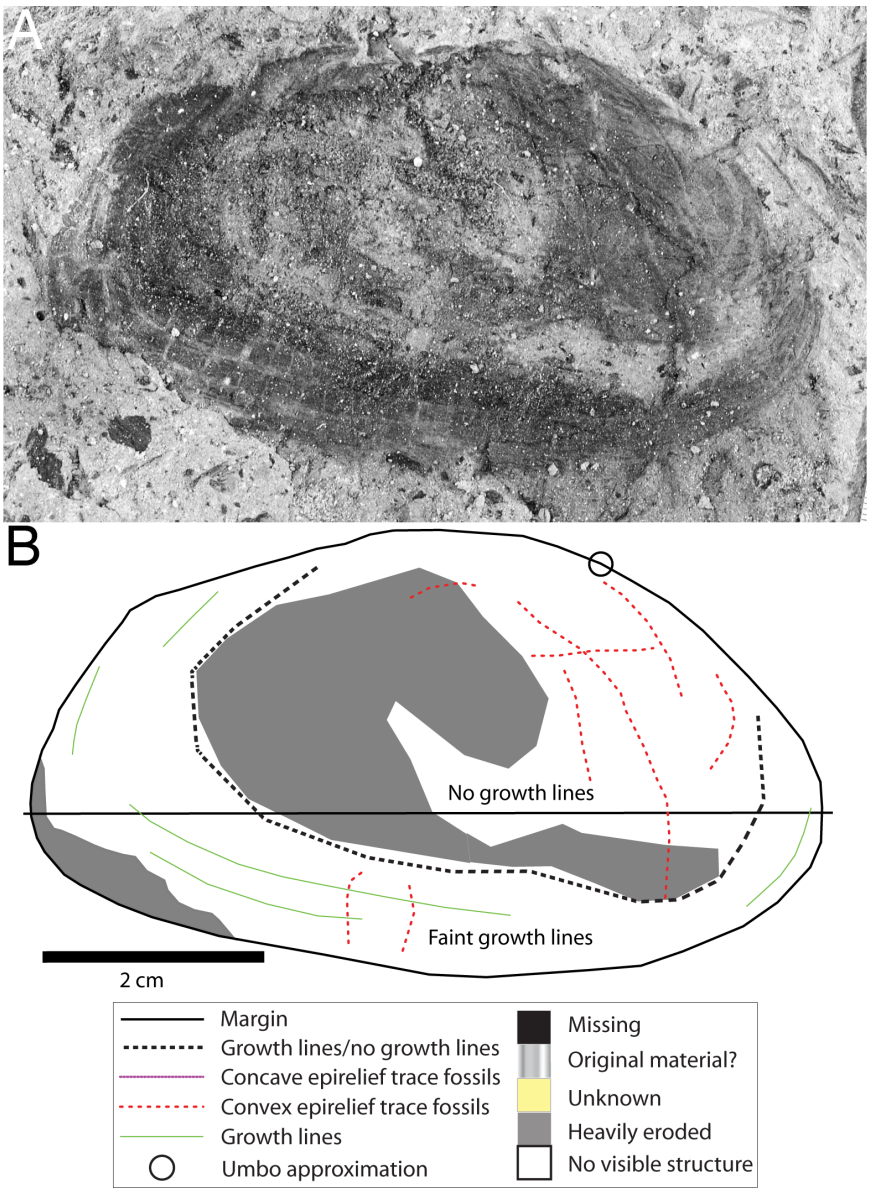

Figure 3.18: Preservation of specimen S2876, a concave left valve. Growth lines are preserved near the margin (a mold of the exterior); the center may be a cast of the interior. All trace fossils are preserved in convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

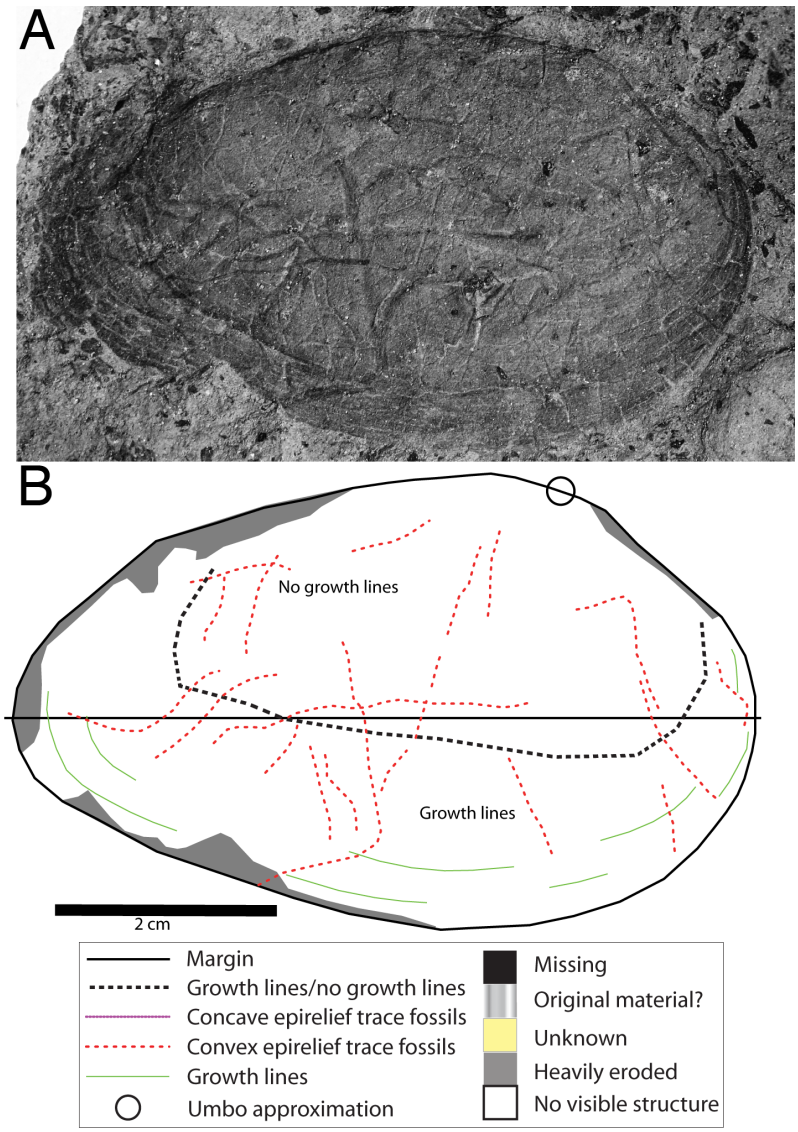

Figure 3.19: Preservation of specimen S2879, a concave left valve. The posterior margin is ragged. Growth lines are preserved on the ventral margin (a mold of the exterior); the center may be a cast of the interior. Trace fossils are preserved in convex epirelief throughout, varying in size. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

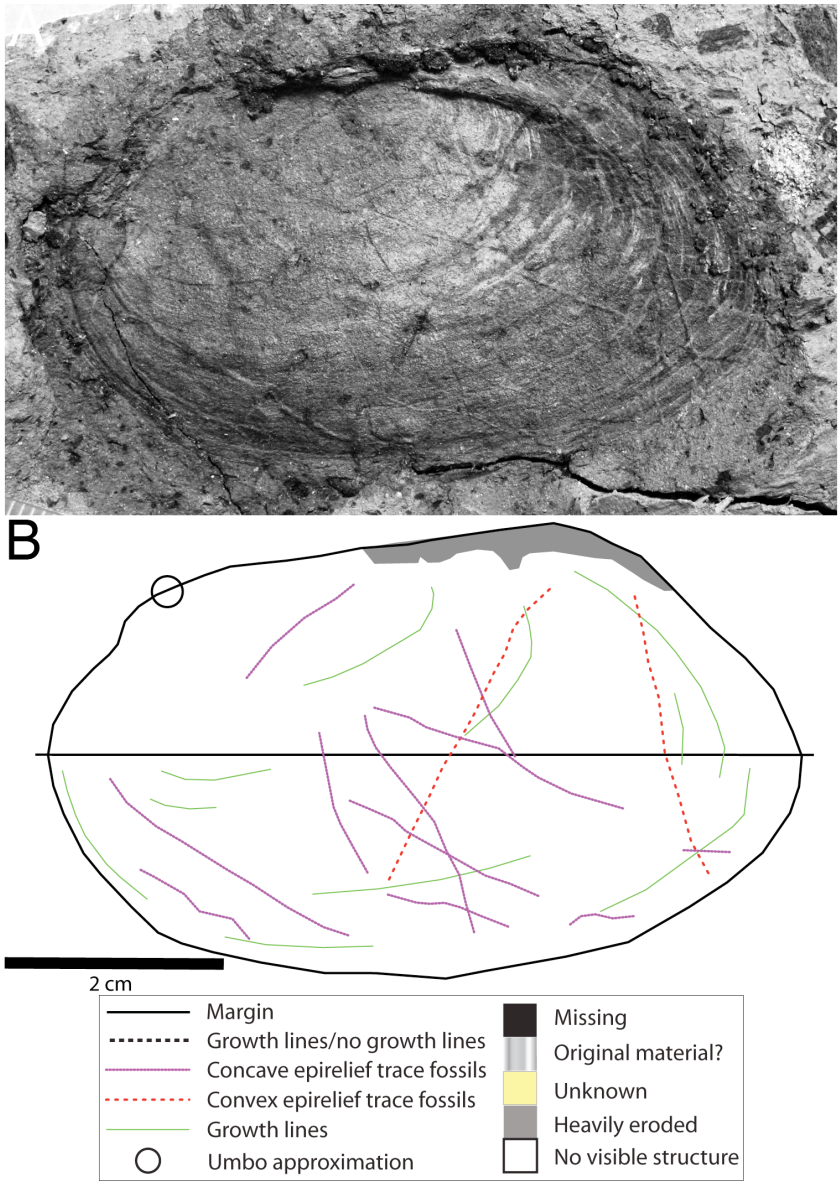

Figure 3.20: Preservation of specimen S2920, a concave right valve preserved as a mold of the exterior. Growth lines are visible over the entire surface. Trace fossils are preserved in both concave and convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

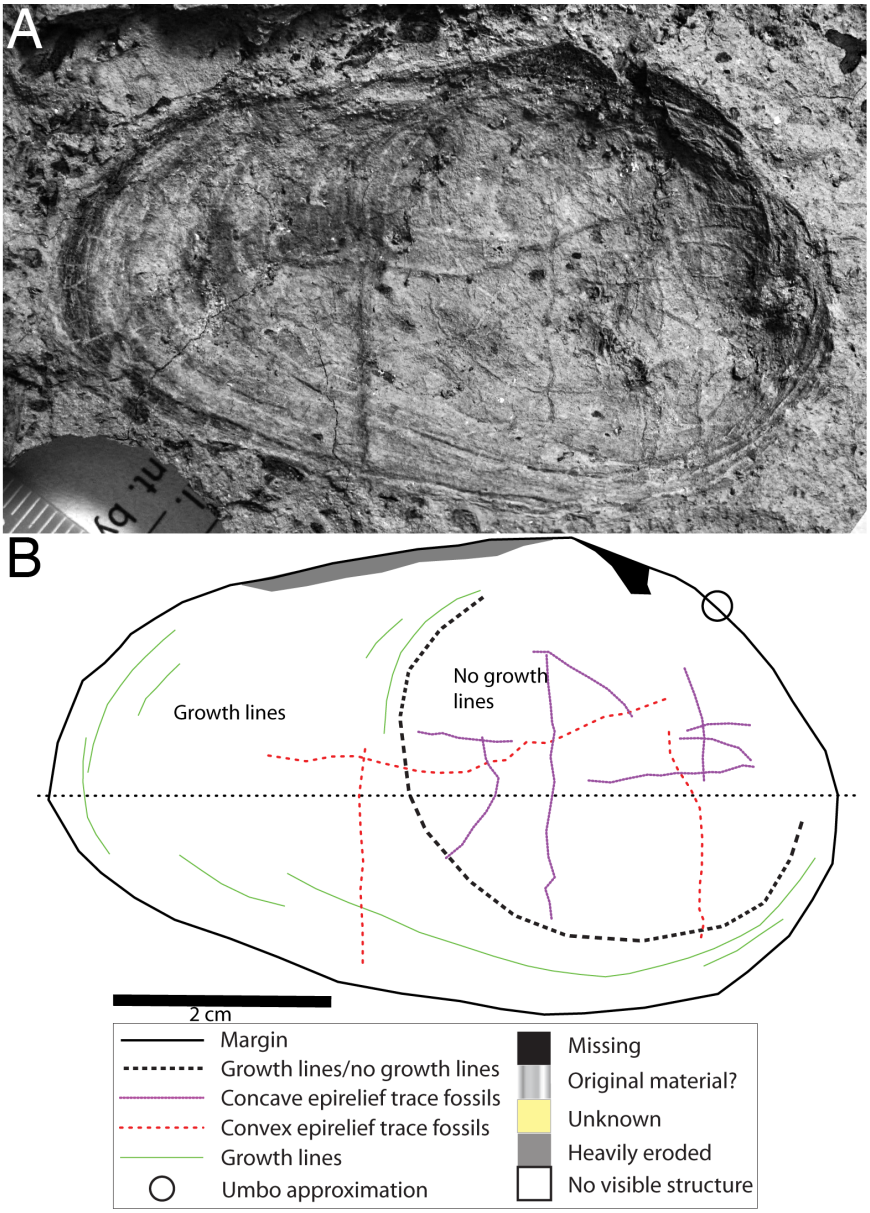

Figure 3.21: Preservation of specimen S2922, a concave left valve preserved as a mold of the exterior. Growth are visible lines around the margins. Most trace fossils are preserved in convex epirelief, some in concave epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

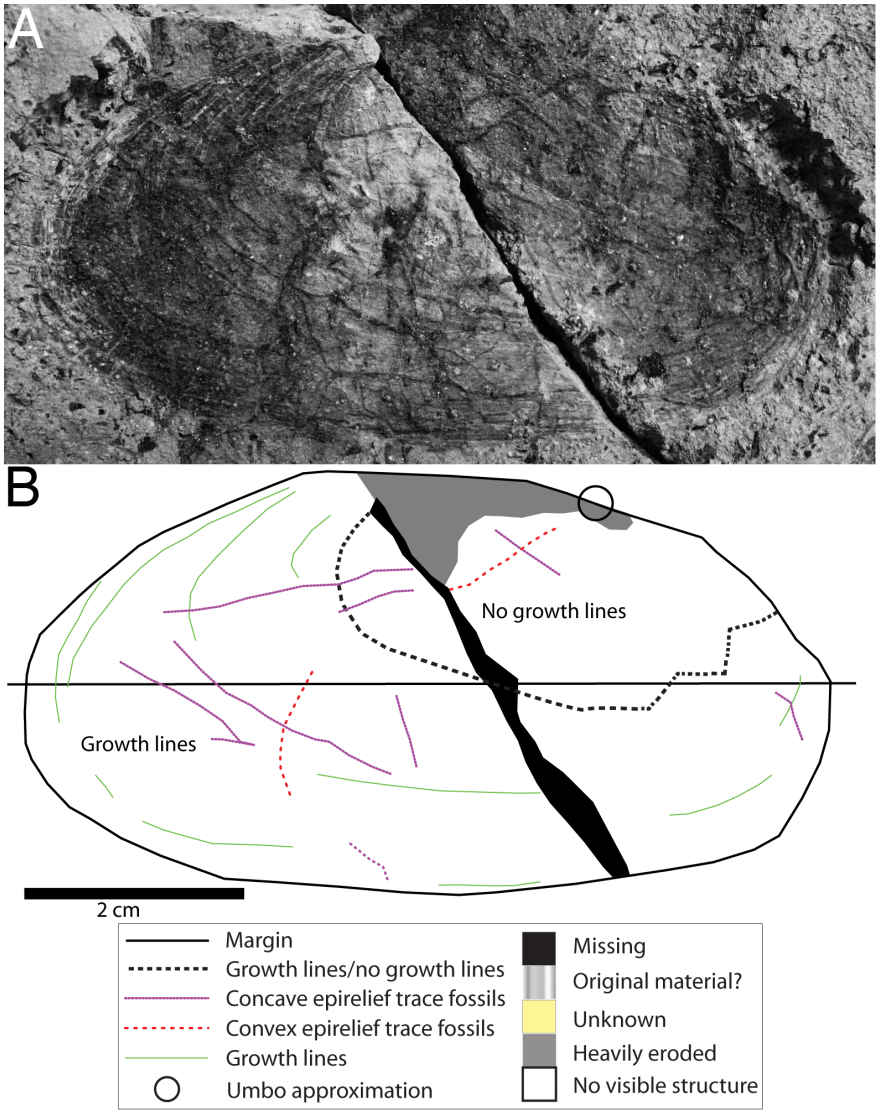

Figure 3.22: Preservation of specimen S2928 is similar to that of S2773. Specimen is a concave right valve preserved as a mold of the exterior. Traces are preserved in both concave and convex epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

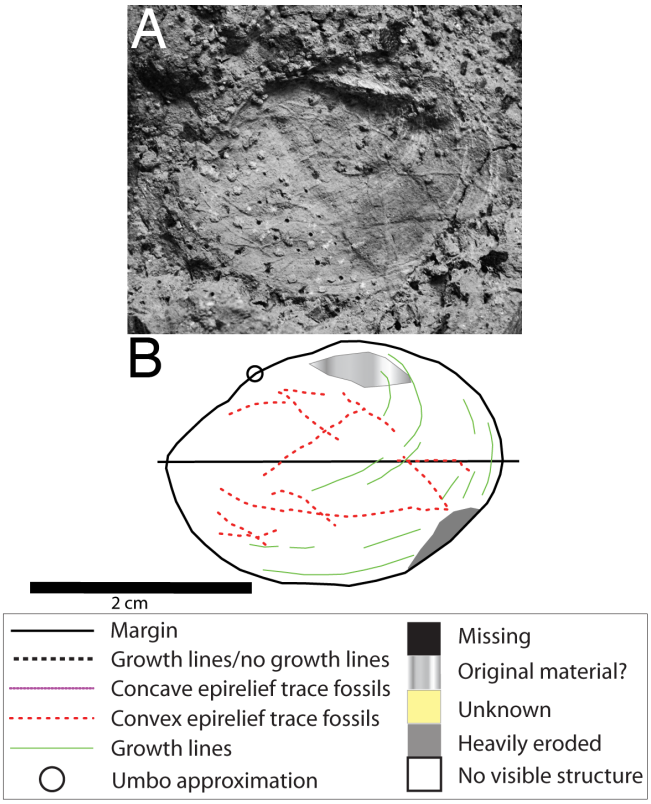

Figure 3.23: Preservation of specimen S2929, a concave right valve preserved as a mold of the exterior. Growth lines are mostly visible. Trace fossils are preserved in convex epirelief. Marked area near dorsal posterior margin indicates possible shell material (hinge plate?). A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

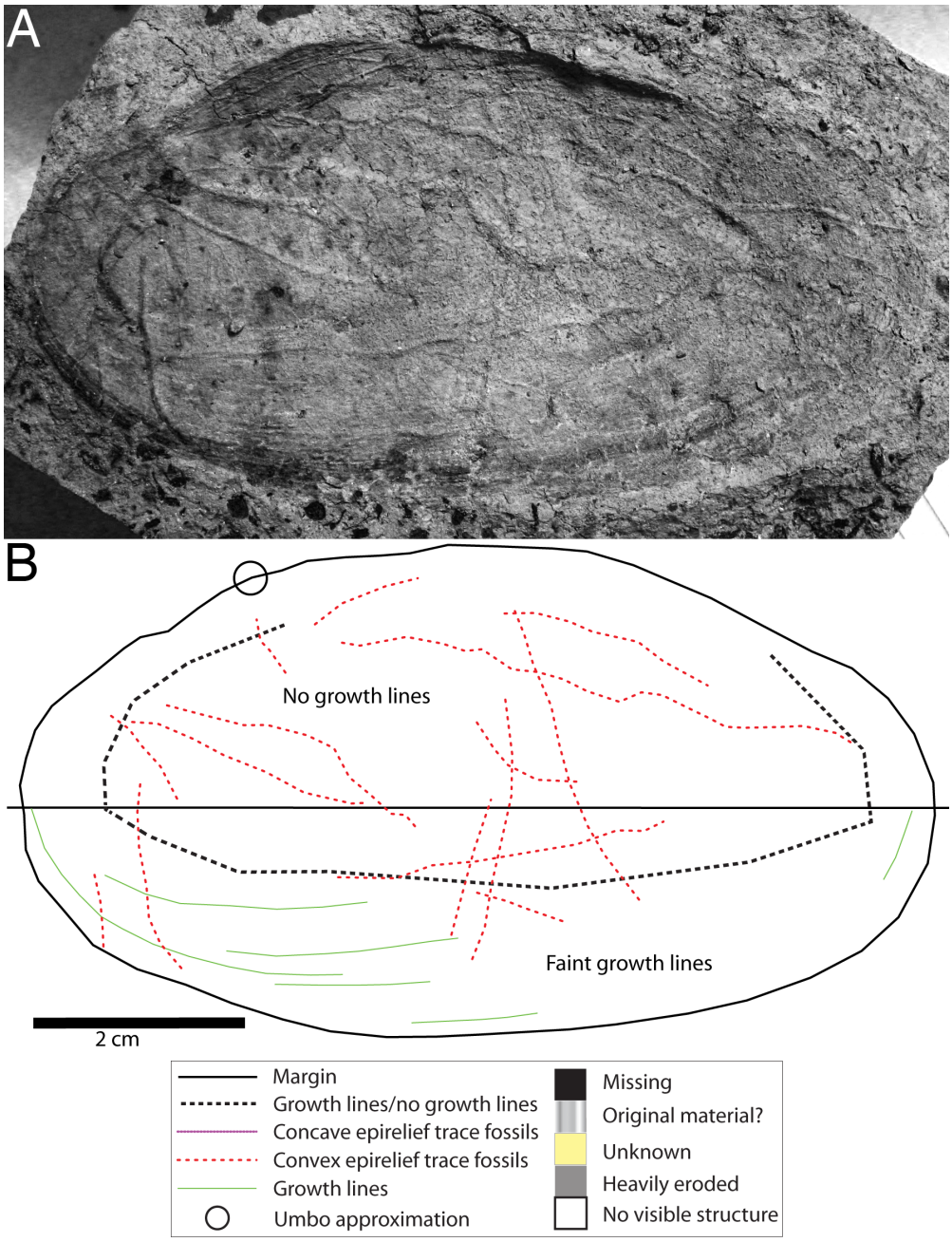

Figure 3.24: Preservation of specimen S2930, a concave right valve preserved as a mold of the exterior. Growth lines are visible around the margins. Most trace fossils are preserved in convex epirelief, some in concave epirelief. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

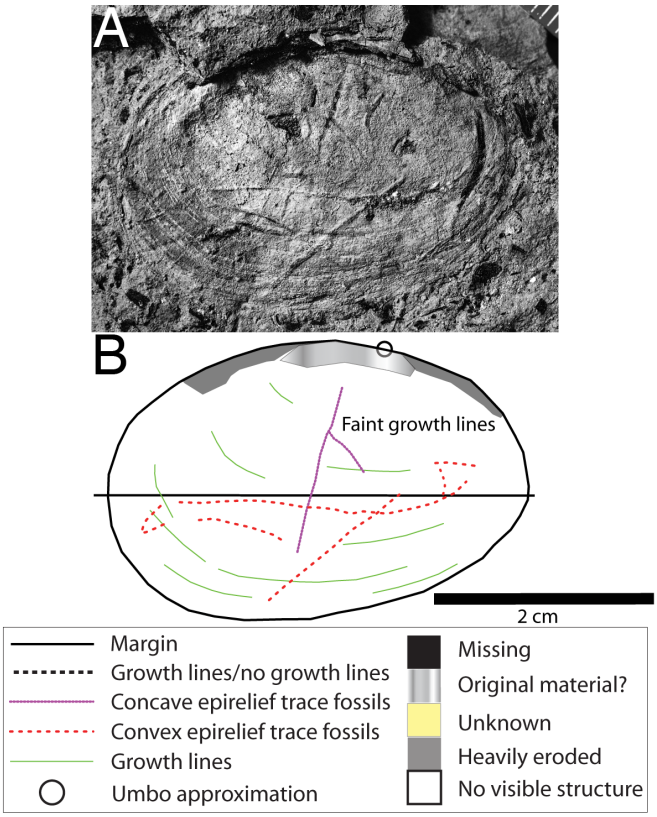

Figure 3.25: Preservation of specimen S2942, a concave left valve. Growth lines are visible overall (specimen is a mold of exterior). Most trace fossils are preserved in convex epirelief, at least one is in concave epirelief. Marked area on dorsal margin indicates possible shell material. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

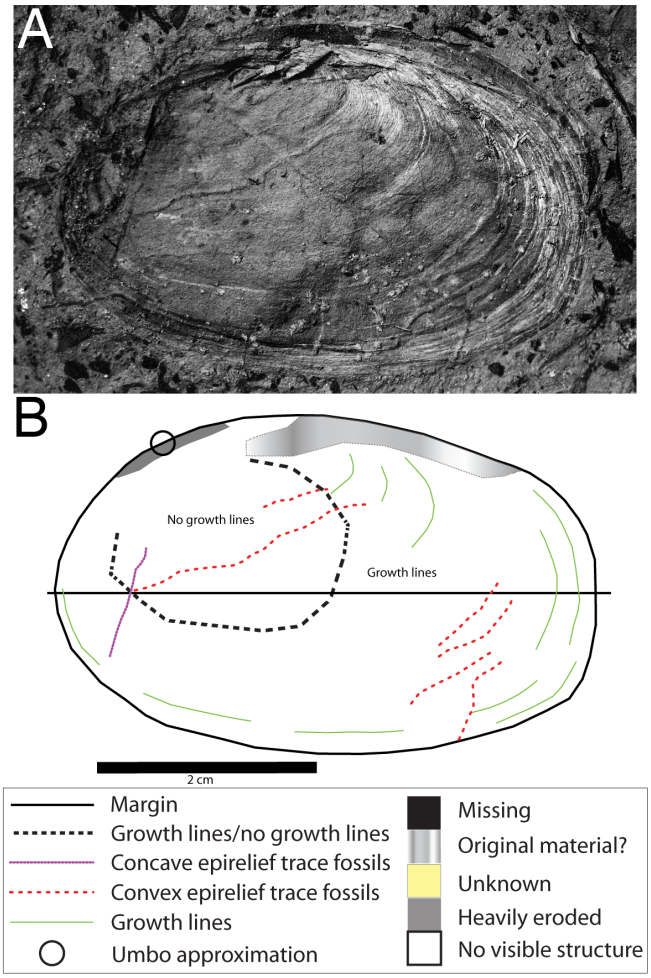

Figure 3.26: Preservation of specimen S2946, a concave right valve preserved as a mold of the exterior and mostly covered with growth lines. Few traces are visible, mostly in convex epirelief. Marked area on dorsal margin indicates possible shell material. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

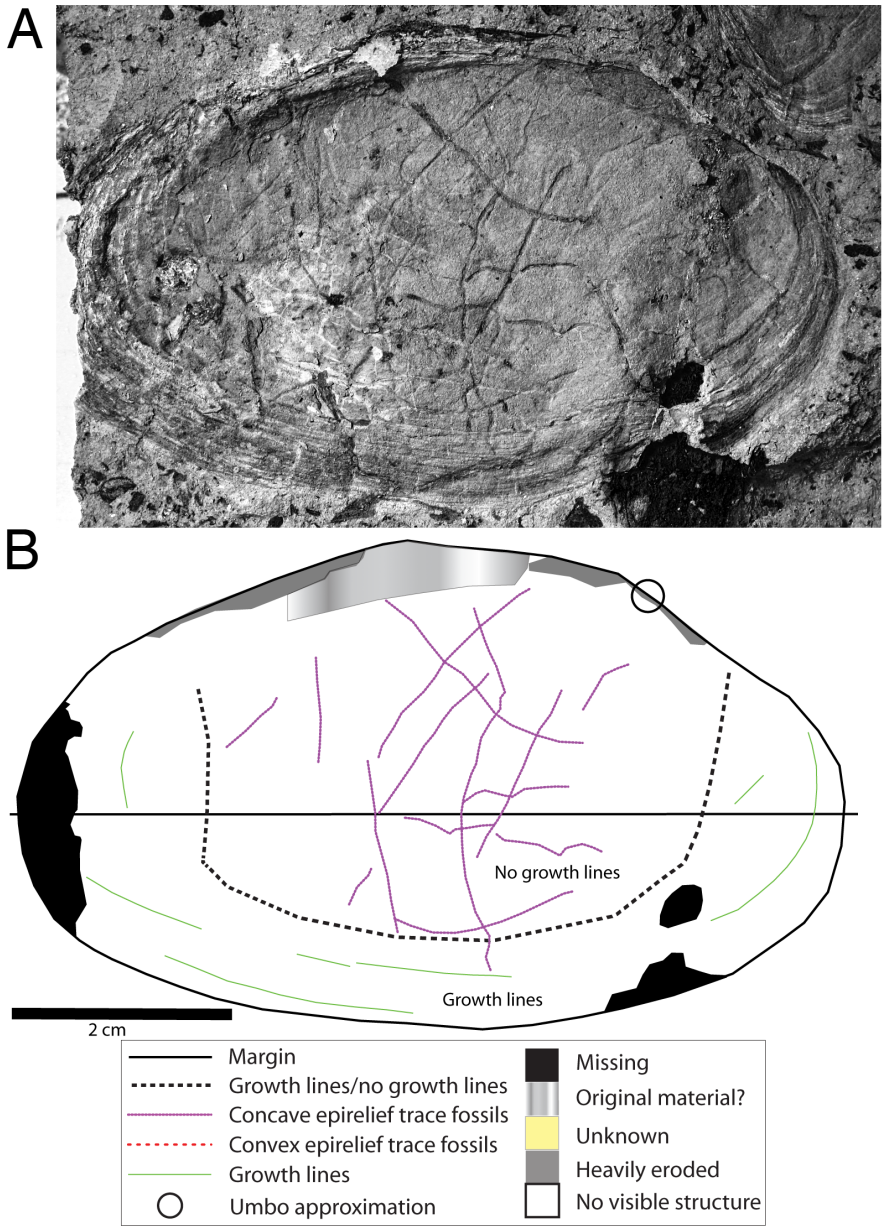

Figure 3.27: Preservation of specimen S2947, a concave left valve. Growth lines are preserved around the margins (specimen is a mold of the exterior). Trace fossils are preserved in concave epirelief, localized to non-growth-line area (center is a cast of the interior). Circular area near anterior is a hole through the valve, possibly deformation from compaction. A–Photograph of specimen. B–Schematic diagram of specimen showing preservation and trace fossils.

CHAPTER 4

QUANTITATIVE SHAPE ANALYSIS OF POORLY PRESERVED FOSSIL UNIONOID MUSSELS BASED ON MORPHOSPACE OCCUPATION OF MODERN FORMS

4.1 Purpose

This chapter describes the different methods used in the field, lab preparation, and statistical procedures.

4.2 Introduction

Quantitative study of fossil material has always been an important part of paleontology. The type and number of direct measurements that can be made necessarily vary according to the taxa being studied and their preservation. The measurements that can be made on unionoid mussels similar to the Das Goods fossils is limited by their relatively featureless exterior; many studies of better preserved specimens have chosen to investigate only length, height, thickness (inflation), umbonal position, and various derived ratios (Eager, 1948, 1974, 1977, 1978; Aldridge, 1999; Scholz and Scholz, 2007). Although such "traditional" morphometric measurements have led to a better understanding of the relation of shell shape to habitat and life habits, such as burrowing depth and rate (Eager, 1948, 1974, 1978; Innes and Bates, 1999), they can be used for identification and classification only in the roughest sense. Recently, popular geometric morphometric methods such as Elliptical Fourier Analysis (EFA) (Kuhl and Giardina, 1982; Rohlf and Archie, 1984; Ferson et al., 1985; Foote, 1989; Crampton and Haines, 1996; Haines and Crampton, 2000; Schmittbuhl et al., 2003) and Landmark methods (Bookstein, 1991; Zelditch et al., 2004) are an attempt to mathematically capture as much quantitative shape information as possible for use

in multivariate statistical tests, with varying results.

Confounding the idea of using any quantitative shape measure for these purposes is the noted morphological plasticity of unionoids (Balla and Walker, 1991; Eager, 1948, 1974, 1977; Hinch and Bailey, 1988) with regard to habitat, leading to repeated convergence in shape (Watters, 1993). Unfortunately, geometric morphometric methods cannot on their own distinguish convergent forms or solve problems of homology. Instead, they can be used as another way to visualize an organism (or part of an organism) in addition to qualitative or presence/absence characters so that specific questions can be answered about shape. These data can then be used in support of an argument for or against homology or convergence with other related forms.

The majority of the unionoid fossils in this study preserve little more than an outline of the valve or valves and incomplete growth line traces, leaving almost nothing that can be treated as an identifiable character. In order to differentiate between subjectively obvious morphological groups (akin to fossil leaf morphotypes after Johnson, 2002) in these assemblages and to try to identify possible affinities of these morphotypes to modern genera or groups of genera, EFA was chosen as a method in order to capture outline data rather than point (landmark) data. The EFA method used herein was created by Ferson et al. (1985) and improved upon in MS DOS program format by Crampton and Haines (1996). EFA produces a series of scores (termed "Fourier Coefficients") that define the shape of a closed curve (in this case, a unionoid valve outline); these scores can then be used in multivariate analyses to examine similarity or difference between individuals or groups of specimen outlines. A thorough explanation of the theory behind EFA is given by Crampton and Haines (1996); Haines and Crampton (2000); Scholz (2003).

One goal of this study is to improve EFA of shell outline for the analysis of the size of morphospace occupation of the fossils from L6516 in comparison to similar

modern forms. Optimization of these methods is based on the metrics of sum of variance (ΣV) and within-group dispersion (WD). The variables of smoothing, number of Fourier harmonics, and normalization to a certain Fourier harmonic can all be tested for with a synthetic group to determine the combination that results in the highest discreteness (AD/WD), the ratio of among-group dispersion (AD) to within-group dispersion (Foote, 1989). A model system made of specimens that fall into easily identifiable morphological groups would have a low within-group dispersion (variation) and a high morphological disparity. Optimization is limited to the current data set, and other data sets would most likely be optimized with different input values during EFA. Optimized methods for specimens of modern genera can be applied to the mussel specimens from the Das Good assemblage, as long as the modern taxa possess similar morphologies. The maximum WD and ΣV of the modern genera contain the morphospace of the modern genera used within general upper and lower limits. These limits are postulated to be one guide to determining the number of taxa present in an assemblage.

4.3 Material: Fossil Specimens

Specimens specific to this project were collected over a period of two summers, composed of two incomplete field days in August 2006 (L6516) and three complete days in August and September 2007 (L6516, L6803, L6804, L6805, L6806, and L6807) by the author with field assistance from Joseph Hartman (UND), Arthur Sweet (CGS), Matthew Borths (OSU), Marron Bingle (UND), Tanya Justham (UND), Kristyn Voegele (CC), and the UND Introduction to Paleontology class of Fall 2007. Material was previously collected in Summer 2000 by Joseph Hartman, S. Bowman, and David Lamb, and in Summer 1999 by K. Johnson (DMNS), R. Barclay, S. Bowman, and G. Knauss. The site was first recorded in 1998 by Johnson and Tim Farnham. Extraction methods were similar to those outlined by Johnson (2002) for

the leaves at this site. Fossils were removed by quarrying large blocks with hoe picks and then splitting these blocks parallel to bedding planes with rock hammers. Due to the narrowness of the bed producing fossils of interest at these sites, care was taken to minimize the amount of overburden removed and to focus on this single producing horizon (for notes on specific locality information and correlation between sites, see Chapter 2.2). Fossil recovery was difficult at all but one site (L6804) because the Das Goods localities are all situated on steep slopes, which causes the amount of overburden to increase rapidly as the bed is followed back into the hill.

After fossil-producing sites were located, the methodology followed that of fossil leaf workers (Johnson, 2002). When specimens were exposed on the surface of blocks or fragments, these blocks (typically several to tens of centimeters across) were wrapped for transport in toilet paper, labeled, and stored in plastic zip-top bags. Well-exposed (full shell outline visible) or fragile specimens were treated in the field with super-fine-thin Star Bond cyanoacrylate to preserve them. Adhesive was not used on the majority of specimens that needed further preparation. A small proportion of blocks were oriented for up-direction on extraction; these were marked with an arrow pointing "up" on a side perpendicular to the bedding plane or by writing "top" or "bottom" on the appropriate surface. After transport from site to vehicle, specimens were packed in cardboard boxes for storage and to prevent further erosion of the fossils or fracture of the blocks themselves. These methods sufficed to keep specimens in good condition even when, in some cases, boxes were left in the vehicle for a period of days to over a week of subsequent field work.

Upon return to UND, specimens were removed from boxes, unwrapped and placed in cardboard specimen trays in cabinets. Specimens were allowed to dry for varying lengths of time before further preparation or study was undertaken. Some blocks became moldy from being in moisture-bearing zipped bags. Moldy surfaces were

sprayed with alcohol and allowed to dry. Neither mold nor alcohol application had a noticeable effect on preservation quality of these specimens. Specimens seem to survive the process of drying quite well (including those collected in 1999), with little cracking of fossil surfaces. Specimens were arranged by locality in cabinets.

Physical lab preparation primarily consisted of cutting parent blocks into smaller pieces in order to reduce storage space needed and allow further splitting of excess material. This was accomplished using a rock saw (Hillquist coolant diluted with water) to minimize damage to specimens. Cut specimens were stored in smaller specimen trays; relationships between specimens previously associated on the same bedding plane or the same block were not retained unless biological relationships were obvious. Excess material was combined and later split, producing a fair number of new specimens. Individual specimens were further cleared of overlying sediment, where necessary, to expose the entire shell surface where possible using a dissecting pin. This process also produced rare new specimens.

4.4 Material: Modern Specimens

Modern specimens of freshwater mussels of known identification were needed for optimizing the quantitative methods used below and for comparison of the shapes of modern genera with the fossils from L6516. Modern specimens were chosen for comparison based on 1) an edentulous or nearly edentulous hinge; 2) lack of surface sculpture; 3) lack of extraneous dorso-posterior "wings" (cf. *Cristaria*) that are dissimilar to the fossil material; and 4) preference for a silty or muddy substrate. Little information was available about substrate preference, which was weighted considerably less than the first three criteria. Full specimen data can be found in Appendix E.2. Modern specimens were identified to the species level according to collection labels and word of the source; no attempt was made to check identification or to deal with possible synonymies.

4.5 Specimen Imaging

Most fossil specimens were photographed with a digital FujiFilm FinePix S1 Pro camera, which produced images of 5 megapixel resolution (17 MB TIFF files). Some fossil and modern specimens were scanned with an HP Scanjet 4070 Photosmart scanner at 400 dpi resolution. Modern specimens were photographed with a variety of camera models depending on home institution and photographer. Some photographs of modern specimens from institutional websites were of low resolution; these photographs were not originally intended for morphometric analysis.

Specimens (both fossil and modern) were oriented so that the commissural plane was parallel to the plane of focus. Scanned specimens were laid flat on the scanning bed. As the outline was considered the most important feature of the specimens, scans were performed only on the interior of the valves. In many cases the specimen label was included in the photograph or scan so that errors would not occur from poor transcription of the specimen identification number to the filename. All fossil specimens from Das Goods localities were given a Hartman specimen number (S). Modern specimens and figures from Cretaceous examples of *Anodonta* were numbered internally to this study with a prefix (T). Appendix E.2 matches (T) numbers with institutional numbering systems.

Specimen photos were stored in computer directories according to locality and (S) number or, in the case of modern specimens and voucher figures, by (T) number. Subsequent versions of these images and primary geometric morphometric data (outline data) produced by software were stored in the same directory as the original photographs. Valves from extant genera belonging to the same individual were given the same specimen number and differentiated by an appended "L" (for left valve) or "R" (for right valve). These valves were treated as individual specimens during subsequent analysis in order to increase the number of examples of each genus. All images

were oriented with the anterior to the left and the umbo on the upper margin whether the image was of the interior or exterior surface; images of exterior right or interior left valves were reflected horizontally to match this pattern. Early exploratory work suggested that there is no statistically significant shape difference between left and right valves in similar extant genera (Burton-Kelly and Hartman, 2007).

4.6 Computer Software

Many software packages were used for this project. Since a number of programs were used together to manipulate data and perform operations and statistical tests, a number of different file types were produced, each with its own specific internal organization. See Appendix 6.8 for a description of software, file formats, organization, and uses.

4.7 Specimen Outline Digitization

Before digitization, specimen images were oriented in Adobe® Photoshop® with the longest axis of the specimen generally parallel to the horizontal (any deviation from this was adjusted for by rotation of the outline during EFA). Specimen outlines were manually digitized using the program tpsDig 2.05 (Rohlf, 2008). Outlines were manually digitized using the pencil tool in a clockwise direction, beginning and ending at the umbo or the nearest approximation that could be determined. Manual rather than automated outline tool digitization was chosen due to lack of a defined edge on most fossil specimens. Although interpreted outlines that were traced manually over photographs (first in CorelDraw®, then in Adobe® Illustrator®) could be subsequently digitized automatically, this would still result in a digitized outline based on a manually defined edge. Due to the variety of photography and scanning techniques and the condition of many of the modern specimens, automated outline digitization in tpsDig did not produce an accurate representation of the valve even after repeated testing and image processing. The decision was made to save time by

manually digitizing the modern valves manually without first creating an outline in CorelDraw® or Adobe® Illustrator®, and thereafter to manually digitize the manually traced fossil specimens in order to reduce the disparity in the number of outline points (and possibly the amount of smoothing necessary) between the two groups.

Only fossil specimens from L6516 were digitized. Digitization was performed with accuracy to the valve outline in mind. Small irregularities in outlines were included where possible in order to capture as much "natural" variation as possible (e.g., small-scale variations in shell shape due to life history of the individual are probably representative of a general weakness in the shell and therefore would be a repeatable occurrence. This has not yet been statistically tested, however). Obvious breaks in the shell due to handling or drying were digitized over as smoothly as possible. Digitization of fossil specimens was more problematic than that of modern specimens due to the larger portions of outline missing and the incomplete preservation of the "real" margins of the valve.

The perceived size of the specimens during the manual digitization process has been hypothesized to have an effect on the repeatability of the shape produced by digitization. Variation between repeated manual digitizations of the same valve at the same perceived size (size of the specimen on the screen when plotting the outline) is here defined as "digitization error." Results (Chapter 5.1) suggest that specimens measured at approximately 25 cm on the screen during digitization will have lower error than those digitized at other screen lengths. Most specimens were digitized at an arbitrary size dependent on the resolution of the original image and the size of the monitor (in this case, 15 inches (38.1 cm) diagonal, resolution 1280 by 1024 pixels).

One file consisting of paired (x,y) coordinates was produced by tpsDig for each digitized specimen. These were given a *.tps filename. Appendix E.2 lists the number of points that make up each outline.

4.8 EFA-based Analyses

Multiple statistical analyses were run to quantify the variation found in the Das Goods area fossils and the sample of modern specimens. These tests assume that a) modern mussel genera occupy the same or greater amount of morphospace based on shell shape than fossil mussel genera, b) each modern specimen used is representative of its assigned genus and species, c) fossil specimens represent the complete ecological assemblage. Each statistical test has additional underlying assumptions, detailed below, that contribute to the power of that test.

4.8.1 Minimizing Digitization Error

"Digitization error," for the purposes of this project, was defined as the total size of morphospace occupation as defined by within-group dispersion (WD) or sum of variance (ΣV) of Fourier scores of a single specimen over a series of repeated manual digitizations using tpsDig (Rohlf, 2008). WD was calculated as "the mean of all pairwise distances between samples within a group" (Foote, 1989), which is the mean Euclidean distance of the pairwise Euclidean distances between the trials (represented as Euclidean centroids in n-dimensional space) of a single specimen (Figure 4.1). For example, given five manually digitized outlines with 11 retained Fourier coefficients (22 variables), five centroids can be calculated with C being a synthetic shape formed by the unweighted average of each Fourier score (variable). Treating each Fourier score as a variable, and each variable as an axis in multivariate space, this centroid also exists in n-dimensional space. The mean Euclidean distance was calculated from the pairwise Euclidean distance between these centroids in each group. This unitless value represents the size of the morphospace occupation of the repeated digitizations, and hence acts as a measure of error resulting from the skill or ability of the digitizer (e.g., eyesight, mouse coordination) and from the interpolation of the selected point on the screen to the pixel on the image.

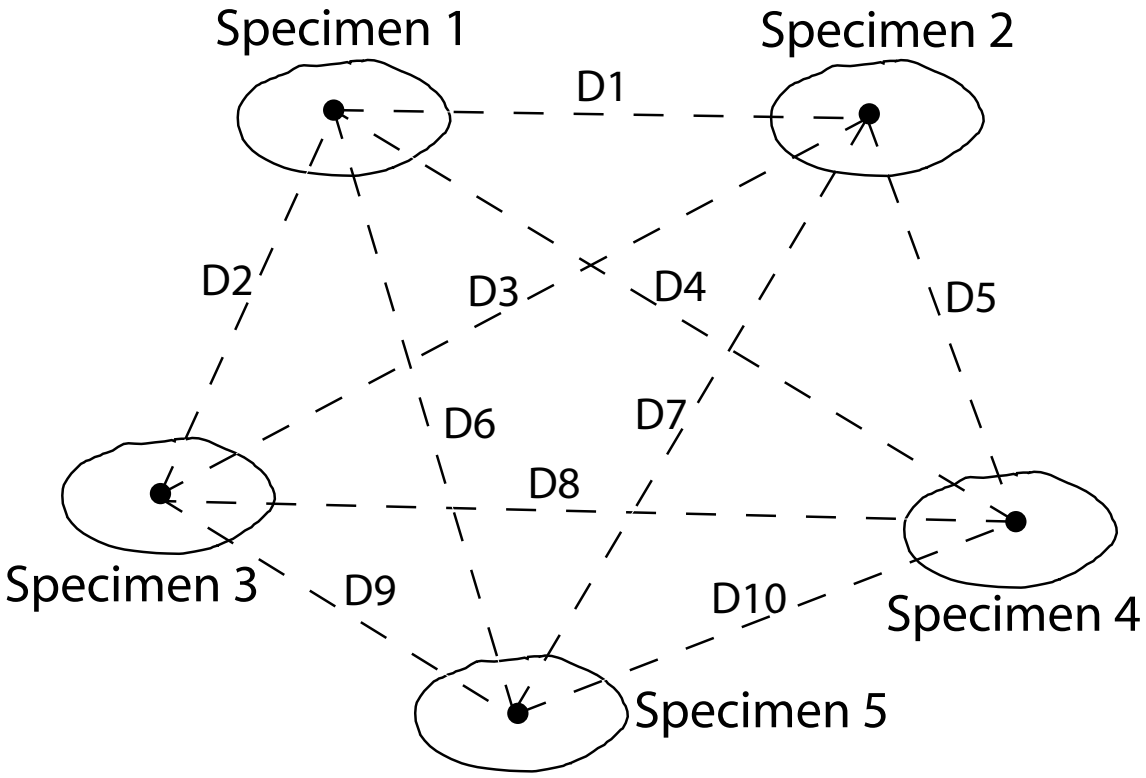

Figure 4.1: The within-group dispersion of a group of digitized outlines is calculated by averaging the Euclidean distances between all pairs in that group, WD = mean($D_1 \ldots D_n$). This figure is a simplified version projected from n-dimensional space (n equal to the number of Fourier coefficients representing the shape of each specimen) into two dimensions.

The sum of the variances of Fourier scores across specimens (ΣV) was calculated as the trace (the sum of the diagonal elements) of the covariance matrix. It was chosen as a secondary metric to assess the usefulness of WD in determining morphospace occupation and was calculated as a bootstrapped value (Rodgers, 1999; Zelditch et al., 2004; Hesterberg et al., 2005) in order to overcome the non-normality of the Fourier score data when comparing groups of specimens. Bootstrapping the variance (or other statistic) begins by resampling a sample (with replacement) multiple times (in most cases, 1000 or more) and calculating the variance for each of these resamples. If these variances are plotted in a histogram, they will show a near-normal distribution.

The variance of the 1000 or more calculated variances is then calculated, which is the bootstrapped variance. Hesterberg et al. (2005) describe this procedure in detail.

One source of digitization error initially was variation in the perceived ("screen") size (measured with a physical ruler in the screen surface) of the specimens in the tpsDig window during manual digitization. Standardization of this size was used to remove some error. To discover the most appropriate digitization size (that yielded the smallest WD), a sample set of nine synthetic outlines based on fossil specimens was manually digitized five times each for specimens with screen sizes (lengths) of 5, 10, 15, 20, 25, and 30 cm. This allowed an average WD to be calculated over the Fourier scores of nine specimens for each digitization size (see Appendix A.3 for details).

Manual digitization of outlines was chosen over automated digitization (an option available in tpsDig) due to the preservation of the fossil material from L6516. Because these specimens exist as molds and casts, are inseparable from the rock, and are incomplete, automated digitization (even with preprocessing in tpsDig or Adobe® Photoshop®) could not locate a suitable edge to work with. Furthermore, tracing the outline by hand in another program (such a CorelDraw® or Adobe® Illustrator®) and then using automated digitization on the drawn outline does not create a more accurate mathematical representation of the real shape of the valve than digitizing that same valve by hand; the digitized outline is still the result of one person's interpretation of the shape of the valve. Valves were manually digitized to standardize both the interpretive nature of the outlines (valves of modern and fossil specimens could be corrected if missing small edge fragments) and the approximate number of points per outline. Automated digitization, while possibly faster than manual digitization, results in over a thousand points per outline (the actual number dependent on the resolution of the image being digitized) while manual digitization typically

results in around 100 points per outline. Automated digitization does not depend on the size of the image as it appears to the user.

On-screen digitization lengths greater than 30 cm could not be easily tested due to monitor constraints, but the WD and ΣV per specimen will only decrease as digitization size increases to a certain point, where a greater number of manually digitized outline points would cause more variation in the repeated outlines and an increase in WD and ΣV per specimen would occur. The digitization size resulting in the lowest WD and ΣV among specimens was then used as a guideline to standardize future digitizations for to minimize digitization error.

Data Capture and EFA

Outlines of nine fossil specimens (S2780, S2800, S2919, S2920, S2921, S2922, S2923, S2924, and S2925) from L6516 were hand-traced on transparent sheets and then scanned to produce "synthetic" (because tracing was accurate at a small scale) outlines representing different shell morphologies. These synthetic outlines were scaled to standardized on-screen lengths of 5, 10, 15, 20, 25, and 30 cm. Each specimen was manually digitized five times at each on-screen length, resulting in 270 TPS files representing one outline each. Individual specimens used for this analysis are marked in Appendix E.2. Elliptical Fourier Analysis was performed on the TPS files using the program HAngle with the following settings: File extension = TPR, number of header lines = 3, amount of smoothing = 1, reverse outlines = 0, number of Fourier harmonics = 12, normalize to which harmonic = 2. See Appendix A.3 for detailed workflow.

Statistical Tests

Within-group dispersion was calculated for each specimen at each digitization length (Appendix A.3). A repeated-measures analysis of variance (ANOVA) was performed on each specimen to compare the WD among the digitization lengths

Table 4.1: Statistical tests performed to determine the on-screen length of specimens during digitization that results in the lowest within-group dispersion of repeated digitizations of the same specimen (and hence the lowest digitization error).

Statistical Test	**Assumptions / Met?**	**Comments**
Repeated-measures ANOVA and LSD tests on pairwise Euclidean distances.	Normal distribution of paired Euclidean distances / Yes for all specimen/digitization length combinations except for S2921 at 15 cm (Table D.1 on page 153). Circular variance-covariance matrix / True for S2920, S2923 and S2924.	Mauchly's W is very sensitive to departures from sphericity even with large sample sizes (SPSS). Performed in SPSS.
WD with 95% confidence intervals.	Normal distribution of paired Euclidean distances / Yes for all specimen/digitization length combinations except for S2921 at 15 cm.	Confidence interval based on the t distribution (Davis, 2002). Performed in R.
Bootstrapped ΣV with 95% confidence intervals.	Unknown.	Performed in R.

(Table 4.1). Post hoc pairwise tests of equality of WD among digitization lengths were accomplished with the least significant difference (LSD) test. 95% confidence intervals calculated with the t distribution were compared to examine the change in WD as specimen digitization length increased.

Sum of variances and 95% confidence intervals were calculated for each specimen at each digitization length (Appendix A.3). 95% confidence intervals were compared to examine the change in ΣV as specimen digitization length increased (Table 4.1). Results can be found in Chapter 5.1.

4.8.2 Arbitrary Digitization Lengths versus those Standardized to 25 cm in modern *Anodonta* Specimens

The equivalence of the standardized length of 25 cm and the arbitrary length

(which was not the same for all specimens) was tested because not all specimens were digitized at a standard length. The goal of these analyses was to examine whether WD and ΣV are each statistically significantly different between the two populations and whether the shape difference between specimens digitized at 25 cm and specimens digitized at an arbitrary length were statistically distinguishable. This test is a consequence of having digitized a large number of specimens at an arbitrary length.

Data Capture and EFA

Outlines of 66 modern *Anodonta* specimens were manually digitized, first with an arbitrary scaling factor dependent on size of the original specimen and the screen size, second with a standard on-screen length of 25 cm. Each specimen was therefore digitized twice, resulting in two groups of 66 outlines, a total of 132 TPS files. Individual specimens used for this analysis are designated in Appendix E.2. Elliptical Fourier Analysis was run on the TPS files using the program HAngle with the following settings: File extension = TPR, number of header lines = 3, amount of smoothing = 2, reverse outlines = 0, number of Fourier harmonics = 12, normalize to which harmonic = 2. See Appendix A.3 for detailed workflow.

Statistical Tests

A Students paired-groups t test was performed on the pairwise Euclidean distances to compare WD of specimens digitized at 25 cm with those digitized at arbitrary lengths (Table 4.2). A permutation t test was performed on the same data to double-check the results of the Student's t test. Confidence intervals were calculated for both WD and ΣV to test the equality of these values in specimens digitized at 25 cm and arbitrary lengths.

Discriminant Analysis (DA) was performed on the EFA scores to test for equality between the shapes of specimens digitized at 25 cm and arbitrary lengths. A permuta-

Table 4.2: Statistical tests performed to compare the within-group disparity between modern specimens of *Anodonta* digitized at an arbitrary screen length with that of specimens digitized at a screen length of 25 cm.

Statistical Test	**Assumptions / Met?**	**Comments**
Student's paired-groups t test on pairwise Euclidean distances.	Normal distribution of differences between trials / No (Shapiro-Wilk W = 0.888, p(normal) = 0.000).	A paired-groups test was chosen over an independent-groups test because the same specimens were digitized at both 25 cm and an arbitrary screen length. Performed in SPSS.
Permutation t test on pairwise Euclidean distances.	None / N.A.	Equivalent to Student's t test but without needing normally distributed data. Performed in PAST.
WD with 95% confidence intervals.	Normal distribution / No (25 cm Shapiro-Wilk W = 0.991, p(normal) = 0.000; arbitrary Shapiro-Wilk W = 0.987, p(normal) = 0.000).	Confidence interval based on the t distribution (Davis, 2002). Performed in R.
Bootstrapped ΣV with 95% confidence intervals.	Unknown.	Performed in R.
Discriminant Analysis with Hotelling's T^2 test on EFA scores.	Multivariate normality / No (Table D.3 on page 155). Similar covariance matrices / Yes (Box's M = 124.61, p(same) = 1).	Performed in PAST.
Permutation test for two multivariate groups on EFA scores.	Similar distributions (variances) between paired multivariate groups / Unknown.	Equal to Hotelling's T^2 test but with fewer assumptions (Hammer, 2002; Hammer et al., 2008). Performed in PAST.

tion test for two multivariate groups was performed on the same data to double-check the DA results (Table 4.2). Results can be found in Chapter 4.8.1.

4.8.3 Examination of Smoothing Effects During Elliptical Fourier Analysis

Previous studies have cited empirical tests of the effects of smoothing (during EFA

using HAngle) on the resulting approximation of shape by the Fourier coefficients (Scholz, 2003; Scholz and Hartman, 2007a,b), but no data have been provided. Ten to 15 retained Fourier coefficients are commonly used across disciplines (Rohlf and Archie, 1984; Ferson et al., 1985; Mehlhop and Cifelli, 1997; Daegling and Jungers, 2000; Stransky, 2002; Schmittbuhl et al., 2001, 2003; Scholz, 2003; Tort, 2003; Sengupta et al., 2005; Bond and Beamer, 2006; Pothin et al., 2006; Tracey et al., 2006; Scholz and Hartman, 2007a,b), although some studies use greater numbers (Innes and Bates, 1999; Tanaka, 1999; Tangchaitrong et al., 2000; Lestrel et al., 2004; Raveloson et al., 2005; Tort and Finizola, 2005; Athreya, 2006) or fewer (Gauldie and Crampton, 2002; Cardini and Slice, 2004). This study differs in using manually digitized outlines rather than outlines produced by computer software recognizing photographic edges, which has the result of producing between 50 and 150 (x,y) coordinate pairs (points) per each outline. Auto-traced outlines can be produced for any number of points, although HAngle resamples all outlines to 1024 pairs (Crampton and Haines, 1996). Auto-tracing outlines naturally produces edge variations based on the pixels making up the image, which causes a more uneven outline at high magnification. When HAngle resamples curves to 1024 points, the unevenness of these edges can cause changes in the shape of the outline (Crampton and Haines, 1996). The smoothing function repeatedly interpolates the points that make up the curve to make it less jagged; since manual tracing has many fewer points to begin with, this is less important.

Smoothing during EFA results in a loss of information about the shape of the valve biased against the retention of smaller features (Figure 4.2). Since it is unclear whether the size of a feature on the valve outline is positively correlated with importance (or usefulness to identifying that shape), the progressive removal of the smallest departures from an ellipse has an unknown effect on how well the resulting

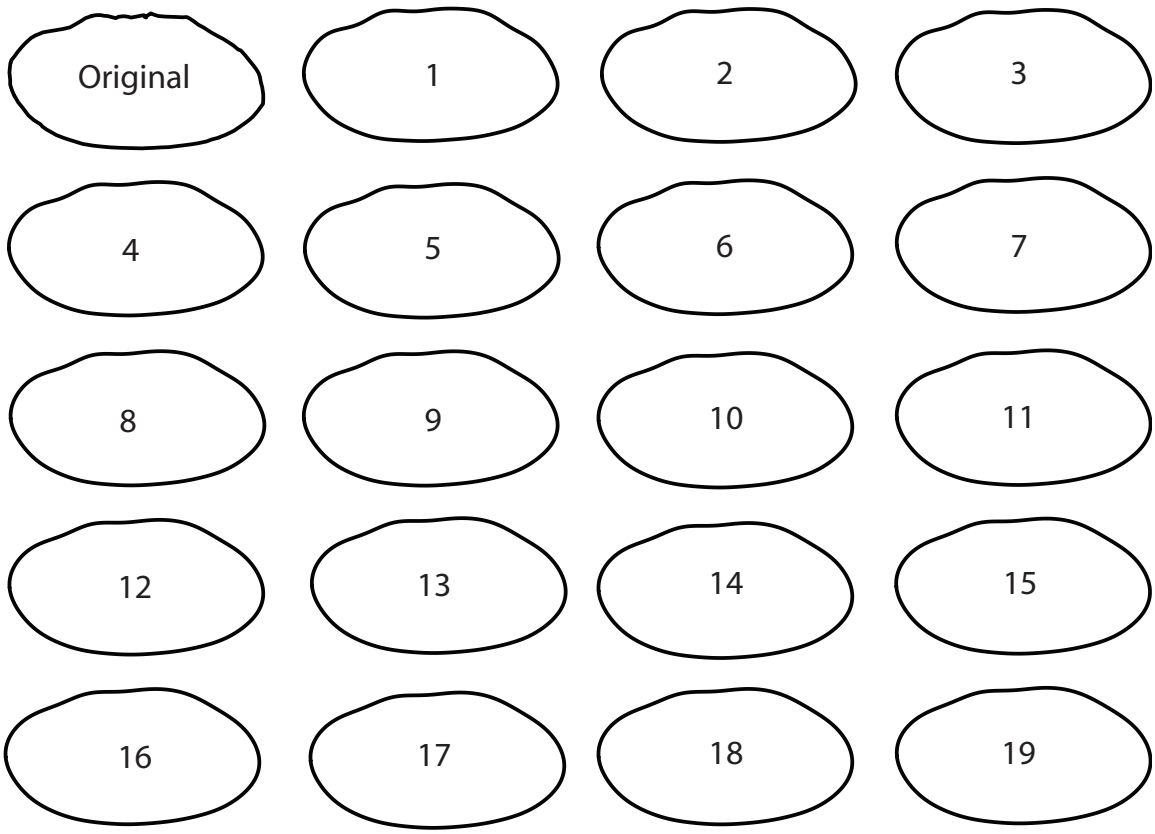

Figure 4.2: Effects of smoothing on an outline of the right valve of an *Anodonta implicata* specimen (Burton-Kelly specimen T0094).

outline matches the morphometric concept of the (generic or specific) group of which the specimen is a part. Too much smoothing adversely affects the accuracy of the outline as described by the Fourier scores, especially when the outlines are manually traced and the smoothing function adjusts the outline in large steps between different smoothing values due to the relatively low number of points.

The effect of smoothing on the variation of manually traced outlines of modern specimens was tested by comparing the size of morphospace occupation (measured by WD) of nine modern edentulous genera with valve shapes similar to L6516 specimens across multiple amounts of smoothing (smoothing = 2, 3, 4, ... 20) with a repeated-measures ANOVA with post hoc paired least significant difference (LSD) tests to

determine significance. Statistical significance of change in sum of variance (ΣV) across smoothing values was determined by bootstraped 95% confidence intervals. EFA was performed with a smoothing value of 1 but the results are not included here because HAngle was not able to include all specimens during analysis and would fail when trying to smooth certain specimens, probably due to irregularities in the digitized outlines.

Data Capture and EFA

Outlines of 384 valves of 9 extant unionid mussel genera (*Anodonta*, *Anodontites*, *Anodontoides*, *Gonidea*, *Pilsbryoconcha*, *Pyganodon*, *Simpsonaias*, *Strophitus* and *Utterbackia*) were manually digitized at either a length of 25 cm or an arbitrary scaling factor dependent on size of the original specimen and the screen size. This procedure resulted in a total of 384 TPS files. Individual specimens used for this analysis are marked in Appendix E.2. Elliptical Fourier Analysis was run on the TPS files using the program HAngle with the following settings: File extension = TPR, number of header lines = 3, amount of smoothing = 1 to 20, reverse outlines = 0, number of Fourier harmonics = 12, normalize to which harmonic = 2. See Appendix A.3 for detailed workflow.

Statistical Tests

A repeated-measures ANOVA was performed on the pairwise Euclidean distances to compare the WD among specimens with different levels of smoothing for each genus. Post hoc comparisons of WD were performed with the least significant difference test. 95% confidence intervals of WD and ΣV were used to examine the change in size of morphospace occupation across smoothing values in each genus (Table 4.3). Results can be found in Chapter 5.2.

Table 4.3: Statistical tests performed to examine the effects of different amounts of smoothing on Elliptical Fourier Analysis.

Statistical Test	**Assumptions / Met?**	**Comments**
Repeated-measures ANOVA with post hoc LSD tests.	Normal distribution of paired Euclidean distances / Yes only for *Gonidea* smoothing 9-20 and *Pilsbryoconcha* smoothing 2-10 (Tables D.14–D.22 on pages 160–164). Circular variance-covariance matrix / No (Mauchly's p(spherical) < 0.001 for each genus).	Mauchly's W is very sensitive to departures from sphericity even with large sample sizes (SPSS).
WD with 95% confidence intervals.	Normal distribution of paired Euclidean distances / Yes only for *Gonidea* smoothing 9-20 and *Pilsbryoconcha* smoothing 2-10 (Tables D.14–D.22 on pages 160–164).	Confidence interval based on the t distribution (Davis, 2002). Performed in R.
Bootstrapped ΣV with 95% confidence intervals.	Unknown.	Performed in R.

4.8.4 Within-Group Dispersion and Sum of Variance in the Fossil Edentulous Freshwater Mussel Assemblage at Locality L6516

Morphospace occupation of unionoid specimens from L6516 was defined by measuring the WD and ΣV based on the scores produced by Elliptical Fourier Analysis of individual valves. This resulted in an envelope of morphospace occupation for the Das Goods specimens that can be used to determine whether theoretical genus- or species-level groups are probable according to modern morphospace occupation.

Data Capture and EFA

Outlines of 27 unionoid valves from L6516 were digitized at a standard on-screen length of 25 cm using tpsDig. Individual specimens used for this analysis are marked in Appendix E.2. Elliptical Fourier Analysis was run on the TPS files using the

program HAngle with the following settings: File extension = TPR, number of header lines = 3, amount of smoothing = 2, reverse outlines = 0, number of Fourier harmonics = 12, normalize to which harmonic = 2. See Appendix A.3 for detailed workflow.

Statistical Tests

A one-way ANOVA was performed on paired Euclidean distances to compare WD among modern genera and the L6516 specimens. Post hoc comparisons of equality were performed with Tukey's HSD tests. The same was done to compare WD among modern species with that of L6516 specimens. 95% confidence intervals of WD and ΣV were also used to compare the size of morphological occupation of the L6516 specimens with that of the modern genera and species (Table 4.4).

4.8.5 Within-Group Dispersion and Sum of Variance in Selected Modern Edentulous Freshwater Mussel Genera and Comparison with L6516 Unionoids

The size of morphospace occupation of modern edentulous freshwater mussel genera was defined by measuring the WD and ΣV based on the Fourier scores produced by Elliptical Fourier Analysis of individual valves from each genus. This resulted in an envelope of morphospace occupation for each genus or group of genera (in this case being those genera that most closely resemble the fossil L6516 specimens in outline) of known size that can be compared with theoretical fossil generic-level groups to determine whether they are probable. Multivariate tests were also used to determine whether modern genera could be identified as different based on Fourier scores; if so, it is reasonable to assume that these same methods can be used to test theoretical fossil generic-level groups.

Data Capture and EFA

Outlines of 384 valves of 9 extant unionid mussel genera (*Anodonta, Anodontites, Anodontoides, Gonidea, Pilsbryoconcha, Pyganodon, Simpsonaias, Strophitus*, and *Utterbackia*) were manually digitized with tpsDig at either a standard on-screen length

Table 4.4: Statistics calculated to define the morphospace occupation of the edentulous freshwater mussel assemblage at L6516.

Statistical Test	**Assumptions / Met?**	**Comments**
One-way ANOVA to compare WD of modern genera and L6516 assemblage. Post hoc tests of equality by Tukey's HSD test.	Normal distribution of pairwise Euclidean distances / L6516: Yes (Shapiro-Wilk $W = 0.9924$, p(same) = 0.06929; modern genera: No except for *Pilsbryoconcha* (Table D.41 on page 174). Similar variances / No (Levene's means p(same) < 0.0001, medians p(same) $<$ 0.0001).	Difference still significant with unequal variances (Hammer et al., 2008).
One-way ANOVA to compare WD of modern species and L6516 assemblage. Post hoc tests of equality by Tukey's HSD test.	Normal distribution of pairwise Euclidean distances / L6516: Yes (Shapiro-Wilk $W = 0.9924$, p(same) = 0.0693; modern species: Yes for 17 out of 24 species (Table D.48 on page 179). Similar variances / No (Levene's means p(same) < 0.0001, medians p(same) $<$ 0.0001).	Difference still significant with unequal variances.
Bootstrapped ΣV with 95% confidence intervals.	Unknown.	Performed in R.
WD with 95% confidence intervals based on the t distribution.	Normal distribution of pairwise Euclidean distances / Yes (Shapiro-Wilks $W =$ 0.9924, p(same) = 0.0693).	Confidence interval based on the t distribution (Davis, 2002). Performed in R.

Table 4.5: On-screen size of photographed valves during digitization and key to specimen prefixes in files used in analyses.

Genus	**Prefix (25 cm)**	**N (25 cm)**	**Prefix (arbitrary)**	**N (arbitrary)**
Anodonta	z	67	e	1
Anodontites	ae	75	—	—
Anodontoides	aj	2	f	33
Gonidea	af	18	ah	1
Pilsbryoconcha	ag	36	ai	3
Pyganodon	ak	1	ac	28
Simpsonaias	al	1	ad	25
Strophitus	am	1	j	47
Utterbackia	an	1	k	44
Total		**202**		**182**

of 25 cm or with an arbitrary scaling factor dependent on size of the original specimen and the monitor size (Table 4.5). Individual specimens used for this analysis are marked in Appendix E.2. Elliptical Fourier Analysis was run on the TPS files using the program HAngle with the following settings: File extension = TPR, number of header lines = 3, amount of smoothing = 2, reverse outlines = 0, number of Fourier harmonics = 12, normalize to which harmonic = 2. See Appendix A.3 for detailed workflow.

Statistical Tests

A one-way ANOVA was performed on pairwise Euclidean distances to test for equality of WD among genera. Post hoc pairwise comparisons were performed with Tukey's HSD test (Table 4.6). 95% confidence intervals of WD and ΣV were used to examine the variation in these values across genera. A multivariate analysis of variance (MANOVA) of EFA scores was performed to test the equality of generic morphology based on EFA scores. Post hoc comparisons were made with Hotelling's T^2 test. Permutation tests for two multivariate groups were also performed to double-check the MANOVA and Hotelling's T^2 results. PCA and CVA were performed on the

EFA scores to visualize the separation of specimens and genera, respectively (Table 4.6). Results can be found in Chapter 4.8.1.

4.8.6 Within-Group Dispersion and Sum of Variance in Selected Extant Edentulous Freshwater Mussel Species and Comparison with L6516 Unionoids

The size of the morphospace occupied by modern edentulous freshwater mussel species was determined by measuring the WD and ΣV based on the Fourier scores produced by Elliptical Fourier Analysis of individual valves from each species. This resulted in an envelope of morphospace occupation for each species or group of species (in this case being members of those genera that most closely resemble the fossil L6516 specimens in outline) of known size that can be compared with theoretical fossil species-level groups to determine whether they are probable according to modern morphospace occupation.

Data Capture and EFA

Outlines of 361 valves of 24 extant unionoid mussel species were manually digitized at either a standard on-screen length of 25 cm or with an arbitrary scaling factor dependent on size of the original specimen and the monitor size (Table 4.7). Species were chosen from edentulous genera used in prior analyses; only species with more than 3 specimens were used. Individual specimens used for this analysis are marked in Appendix E.2. Elliptical Fourier Analysis was run on the TPS files using the program HAngle with the following settings: File extension = TPR, number of header lines = 3, amount of smoothing = 2, reverse outlines = 0, number of Fourier harmonics = 12, normalize to which harmonic = 2. Detailed workflow in Appendix A.3.

Statistical Tests

A one-way ANOVA was performed on pairwise Euclidean distances to test for equality of WD among species. Post hoc pairwise comparisons were performed with Tukey's HSD test (Table 4.8). 95% confidence intervals of WD and ΣV were used to

Table 4.6: Statistical tests performed to compare the results of Elliptical Fourier Analysis on modern edentulous freshwater mussels as shown by similarity in Fourier scores and size of sum of variance or within-group dispersion of each genus.

Statistical Test	**Assumptions / Met?**	**Comments**
One-way ANOVA with post hoc Tukey's HSD tests.	Normal distribution of pairwise Euclidean distances / No except for *Pilsbryoconcha* (Table D.41 on page 174). Similar variances / No (Levene's means p(same) = 0.0000; medians p(same) = 0.0000).	Welch $F = 1652$, df = 1804, p = 0. Difference still significant with unequal variances (Hammer et al., 2008). Performed in PAST.
WD with 95% confidence intervals .	Normal distribution of pairwise Euclidean distances / No except for *Pilsbryoconcha* (Table D.41 on page 174).	Confidence interval based on the t distribution (Davis, 2002). Performed in R.
Bootstrapped ΣV with 95% confidence intervals.	Unknown.	Performed in R.
MANOVA with post hoc Hotelling's T^2 tests.	Multivariate normal distribution / No (Table D.42 on page 174). Similar covariance matrices / No (Box's $M = 474.06$, p(same) < 0.001).	Box's M may be too sensitive to differences in covariance matrices (Hammer et al., 2008). Compare p(same) values with those of multivariate permutation test. Performed in PAST.
Permutation test for two multivariate groups.	Similar distributions (variances) between paired multivariate groups / Unknown.	Equal to Hotelling's T^2 post-hoc test for MANOVA but with fewer assumptions (Hammer, 2002; Hammer et al., 2008). Performed in PAST.
Principal Component Analysis.	Debated (Hammer et al., 2008).	Performed in PAST.
Canonical Variates Analysis.	Multivariate normal distributions / No (Table D.42 on page 174). Similar covariance matrices / No (Box's $M = 474.06$, p(same) < 0.0001).	Performed in PAST.

Table 4.7: On-screen size of photographed valves during digitization and key to specimen prefixes in files used in analyses.

Genus	**Prefix (25 cm)**	**N (25 cm)**	**Prefix (arbitrary)**	**N (arbitrary)**
Anodonta couperiana	z	7	—	—
Anodonta cygnea	z	5	—	—
Anodonta grandis	z	28	—	—
Anodonta imbecillis	z	9	—	—
Anodonta implicata	z	4	—	—
Anodonta suborbiculata	z	6	—	—
Anodontites elongatus	ae	13	—	—
Anodontites farrarisi	ae	3	—	—
Anodontites irisans	ae	4	—	—
Anodontites moricandi	ae	5	—	—
Anodontites obtusus	ae	3	—	—
Anodontites patagonicus	ae	7	—	—
Anodontites tenebricosus	ae	14	—	—
Anodontites trapesialis	ae	23	—	—
Anodontoides ferussacianus	aj	1	f	30
Gonidea angulata	af	18	ah	1
Pilsbryoconcha exilis	ag	36	ai	3
Pyganodon cataracta	ak	—	ac	5
Pyganodon grandis	ak	—	ac	14
Pyganodon lacustris	ak	—	ac	6
Simpsonaias ambigua	al	1	ad	25
Strophitus subvexus	—	—	j	9
Strophitus undulatus	am	1	j	37
Utterbackia imbecillis	an	1	k	42
Total		**189**		**172**

examine the variation in these values across species.

MANOVA, Hotelling's T^2 , permutation tests for multivariate groups, PCA and CVA were not performed on species-level data due to constraints of the program PAST on the number of multivariate groups that could be included in a single analysis.

Table 4.8: Statistical tests performed to compare the results of Elliptical Fourier Analysis on modern edentulous freshwater mussels as shown by similarity in Fourier scores and size of sum of variance or within-group dispersion of each species.

Statistical Test	**Assumptions / Met?**	**Comments**
One-way ANOVA with post hoc Tukey's HSD tests.	Normal distribution of pairwise Euclidean distances / Yes for 17 out of 24 species (Table D.48 on page 179). Similar variances / No (Levene's means p(same) = 0.0000 medians p(same) = 0.0000).	Welch $F = 191.9$, p(same) = 0.0000. Difference still significant with unequal variances (Hammer et al., 2008). Performed in PAST.
WD with 95% confidence intervals .	Normal distribution of pairwise Euclidean distances / Yes for 17 out of 24 species (Table D.48 on page 179).	Confidence interval based on the t distribution (Davis, 2002). Performed in R.
Bootstrapped ΣV with 95% confidence intervals.	Unknown.	Performed in R.

CHAPTER 5

RESULTS

5.1 Minimizing Digitization Error as Measured by Within-Group Dispersion and Sum of Variance

Repeated-measures ANOVAs for each synthetic specimen showed statistical significance among WD by digitization length ($\alpha = 0.05$, $p < 0.001$) (Table 5.1). WD generally decreases with an increase in digitization length in the interval from 5 cm to 30 cm (Figure 5.1). Post hoc pairwise significance by Tukey's HSD test is summarized in Table 5.2 to show trends. Most synthetic valves possessed statistically significantly different WD values between most digitization lengths (Tables D.5–D.13 on p. 157–159), showing generally that an increase in digitization length results in a decrease in the WD of the five repeated digitizations per specimen (Figure 5.1). The 25 cm and 30 cm digitization lengths showed no statistically significant difference in WD. Using 95% confidence intervals to determine statistical significance yields similar results to those obtained by the ANOVAs (Figure 5.1). Sum of variance (ΣV) trends in specimens closely follow trends in WD (Figure 5.2). 95% confidence intervals for ΣV are generally comparatively larger than those of WD values.

5.1.1 Arbitrary Digitization Lengths versus those Standardized to 25 cm in modern *Anodonta* Specimens

Student's paired-groups t test shows that the WD of specimens digitized at 25 cm (WD = 0.9458) is statistically significantly different from WD of specimens digitized at an arbitrary screen length (WD = 0.9852) ($t = 23.163$, $p = 0.000$, $\alpha = 0.05$). A permutation t test ($N = 10{,}000$) agrees with these results ($p < 0.0001$, $\alpha = 0.05$). *Anodonta* specimens digitized at an arbitrary length have a smaller within-group

Table 5.1: Comparing the statistical significance across amounts of smoothing, power and effect size on the basis of a repeated-measures ANOVA for each genus. F value, power and effect size were calculated with the Lower-bound method, the most conservative output by SPSS. **Bold** values are statistically significant ($\alpha = 0.05$).

Synthetic Valve Outline	Statistical Significance	Power	Effect Size (Partial Eta Squared)	F
S2780	**0.000**	0.999	0.790	33.935
S2800	**0.000**	1.000	0.838	46.711
S2919	**0.000**	1.000	0.769	30.005
S2920	**0.000**	0.999	0.786	33.128
S2921	**0.000**	1.000	0.808	37.997
S2922	**0.000**	0.998	0.775	30.960
S2923	**0.000**	1.000	0.859	54.736
S2924	**0.000**	0.999	0.781	32.062
S2925	**0.000**	1.000	0.807	37.712

dispersion than the specimens digitized at 25 cm. Calculated 95% confidence intervals around WD based on the t distribution (25 cm WD confidence intervals = 0.0973 to 0.0997; arbitrary length WD confidence intervals = 0.0934 to 0.0958) agrees with the permutation and Student's t tests, but not by bootstrapped 95% confidence intervals around ΣV (25 cm ΣV confidence intervals = 0.0033 to 0.0080; arbitrary length ΣV confidence intervals = 0.0029 to 0.0077)).

The permutation test for two multivariate groups (2000 repeats, Mahalanobis distance = 0.06428, p(same) = 0.849, $\alpha = 0.05$), discriminant analysis visualization (Figure 5.3) and Hotellings T^2 test (p(same) = 0.922, $\alpha = 0.05$) each show that there is no statistically significant difference between the two groups of specimens (digitized at 25 cm and at an arbitrary scale) based on the multivariate means of the two groups of Fourier scores.

5.2 Examination of Smoothing Effects During Elliptical Fourier Analysis

Repeated-measures ANOVAs identified statistically significant differences between WD values at different levels of smoothing in all genera ($\alpha = 0.05$, $p < 0.001$) (Table

Table 5.2: Summary of post hoc uncorrected pairwise comparisons of synthetic valve WD at different digitization lengths. Specimens listed at each pairwise comparison are statistically significant. Raw data in Tables D.5–D.12 on p. 157–159.

Digitization Length (cm)	**5**	**10**	**15**	**20**	**25**	**30**
5	—	S2780	S2780	S2780	S2780	S2780
		S2800	S2800	S2800	S2800	S2800
		S2919	S2919	S2919	S2919	S2919
		S2921	S2920	S2920	S2920	S2920
		S2922	S2921	S2921	S2921	S2921
		S2923	S2922	S2922	S2922	S2922
		S2925	S2923	S2923	S2923	S2923
			S2924	S2924	S2924	S2924
			S2925	S2925	S2925	S2925
10		—	S2780	S2780	S2780	S2780
			S2800	S2800	S2800	S2800
			S2920	S2919	S2919	S2919
			S2921	S2920	S2920	S2920
			S2922	S2921	S2921	S2921
			S2923	S2922	S2922	S2922
			S2924	S2923	S2923	S2923
				S2924	S2924	S2924
				S2925	S2925	S2925
15			—	S2929	S2780	S2780
				S2922	S2800	S2800
				S2924	S2919	S2919
				S2925	S2920	S2920
					S2922	S2922
					S2923	S2923
					S2924	S2924
					S2925	S2925
20				—	S2800	S2800
					S2920	S2920
					S2921	S2921
					S2922	S2922
					S2923	S2923
					S2924	S2924
					S2925	S2925
25					—	
30						—

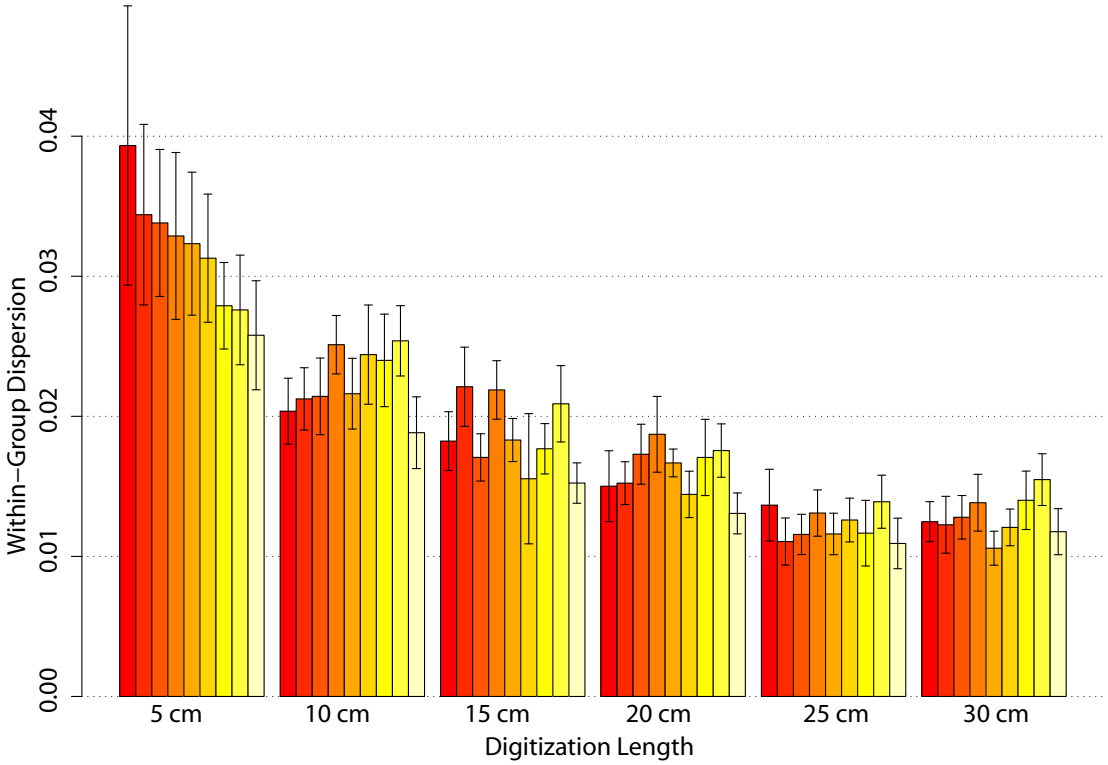

Figure 5.1: Within-group dispersion of five repeated digitizations of nine randomly picked fossil specimens as a function of on-screen length during manual outline digitization. Height of bar represents WD value (= mean pairwise Euclidean distance among specimens digitized at the same length). Error bars represent 95% confidence interval based on the t distribution. Synthetic valves based on the following specimens according to bar number (from left to right): 1–S919, 2–S2925, 3–S2800, 4–S2922, 5–S2923, 6–S2921, 7–S2920, 8–S2924, 9–S2780. Raw data in Table D.2 on p. 154.

5.3). WD decreases as smoothing increases in the interval from smoothing value 2 to smoothing value 20 in all genera examined (Figure 5.4). Almost all post hoc pairwise comparisons (least significant difference tests) between different smoothing values of a single genus are statistically significant ($\alpha = 0.05$, p generally < 0.001). 95% confidence intervals around WD based on the t distribution (and produced with R) for individual genera are shown in Figure 5.4. 95% confidence intervals agree closely with post hoc pairwise comparisons. Genera with a higher within-group dispersion in

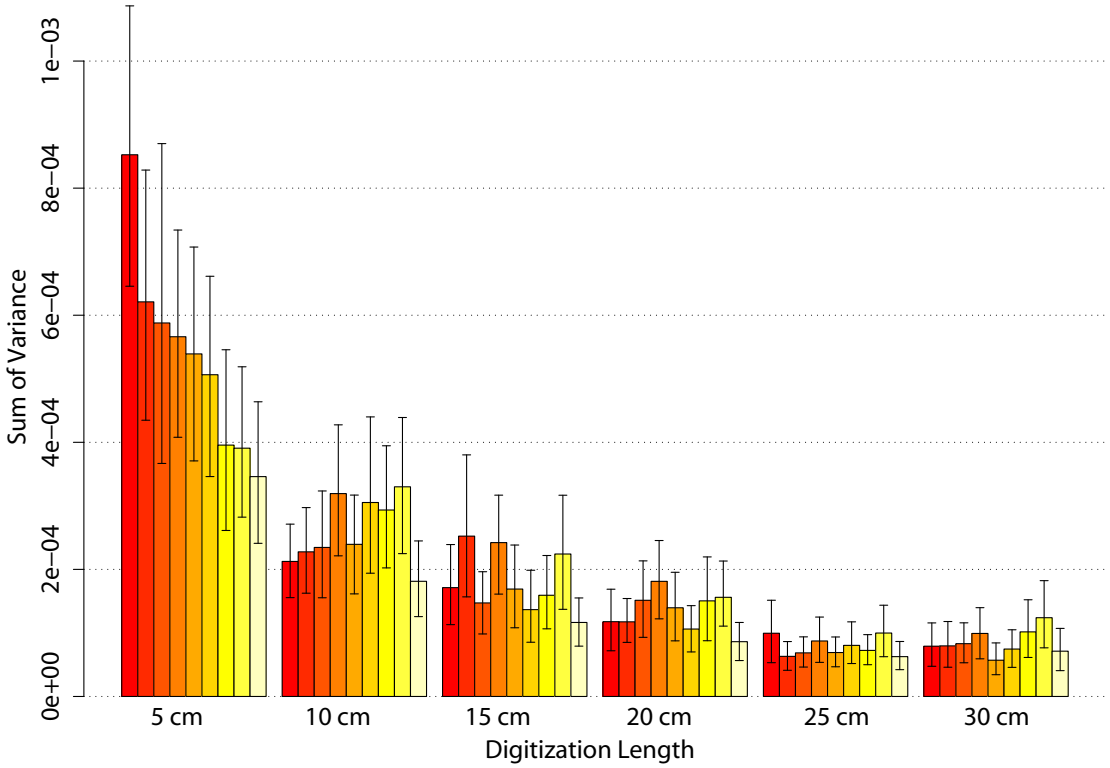

Figure 5.2: Sum of variance of five repeated digitizations of nine randomly picked fossil specimens as a function of on-screen length during manual outline digitization. Height of bar represents ΣV value based on bootstrapped ($N = 1000$) sum of variances, error bars represent bootstrapped ($N = 1000$) 95% confidence intervals. Synthetic valves based on the following specimens according to bar number (from left to right, decreasing ΣV): 1–S2919, 2–S2925, 3–S2800, 4–S2922, 5–S2923, 6–S2921, 7–S2920, 8–S2924, 9–S2780. Raw data in Table D.4 on p. 156.

general show a trend towards more rapid changes in WD than those with lower WD. A similar trend in slope to that of WD in Figure 5.4 can be seen in ΣV in Figure 5.5, although ΣV 95% bootstrapped confidence intervals are comparatively larger,

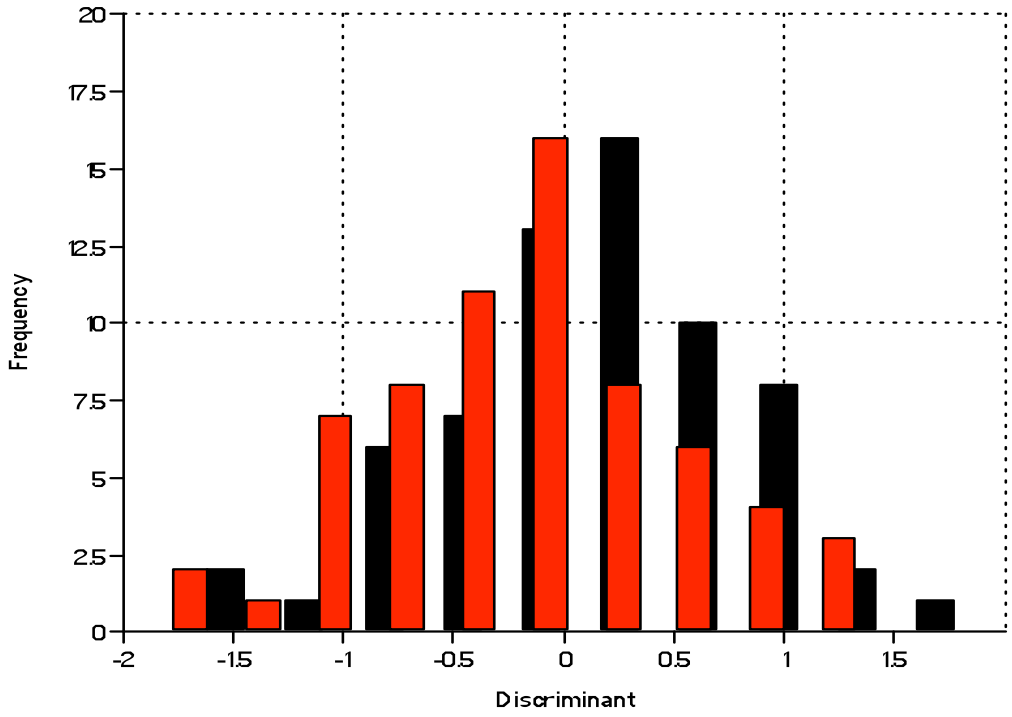

Figure 5.3: Discriminant Analysis visualization comparing multivariate means of specimens digitized at 25 cm (red) and at an arbitrary length (black).

resulting in less statistically significant change in ΣV as smoothing increases.

5.3 Within-Group Dispersion and Sum of Variance of Extant Edentulous Freshwater Mussel Genera and Comparison with Unionoids from Locality L6516

The average within-group dispersion of the mussels used in this analysis is 0.0790 with a range from 0.0427 to 0.1057 (range of 0.0629) (Figure 5.6; Table D.43 on p. 175). The range of 95% confidence intervals is from 0.0421 to 0.1068 (range of 0.0648). A one-way ANOVA found a statistically significant difference among WD values for modern genera ($\alpha = 0.05$, F = 574.4, $p = 0$). Almost all post hoc pairwise

Table 5.3: Comparing the statistical significance across amounts of smoothing, power and effect size, on the basis of a repeated-measures ANOVA for each genus. F value, power and effect size were calculated with the Lower-bound method, the most conservative output by SPSS. **Bold** values are statistically significant ($\alpha = 0.05$).

Genus	Statistical Significance	Power	Effect Size (Partial Eta Squared)	F
Anodonta	**0.000**	1.000	0.758	7115.512
Anodontites	**0.000**	1.000	0.813	12027.761
Anodontoides	**0.000**	1.000	0.761	1890.472
Gonidea	**0.000**	1.000	0.837	879.188
Pilsbryoconcha	**0.000**	1.000	0.259	258.839
Pyganodon	**0.000**	1.000	0.752	1233.066
Simpsonaias	**0.000**	1.000	0.697	745.717
Strophitus	**0.000**	1.000	0.509	1119.645
Utterbackia	**0.000**	1.000	0.523	1083.585

Tukey's HSD tests are statistically significant ($\alpha = 0.05$, $p < 0.01$) (Figure 5.8). 95% confidence intervals based on the t distribution agree with these tests (Figure 5.6).

The average sum of variance of the modern genera used in this analysis is approximately 0.0035 with a range from 0.0009 to 0.0061 (range of 0.0052) (Figure 5.7; Table D.44 on p. 175). The range of 95% confidence intervals is from 0.0008 to 0.01 (range of 0.009). Some pairs of genera were found to possess statistically significant ΣV values based on non-overlapping confidence intervals (Table 5.8). These do not agree with the WD ANOVA and confidence interval results because of the larger confidence intervals around ΣV, however the overall rank order of genera is the same (Figure 5.7).

A MANOVA revealed statistically significant differences among modern genera based on the multivariate means of Fourier scores ($p = 0$, $\alpha = 0.05$; Wilk's $\lambda = 0.0156$, F = 11.23; Pillai trace = 2.89, F = 9.282). All post hoc pairwise Hotelling's T^2 tests show statistically significant differences among genera (Table D.46 on p. 177). These results are supported by the permutation test for two multivariate groups (Table D.47

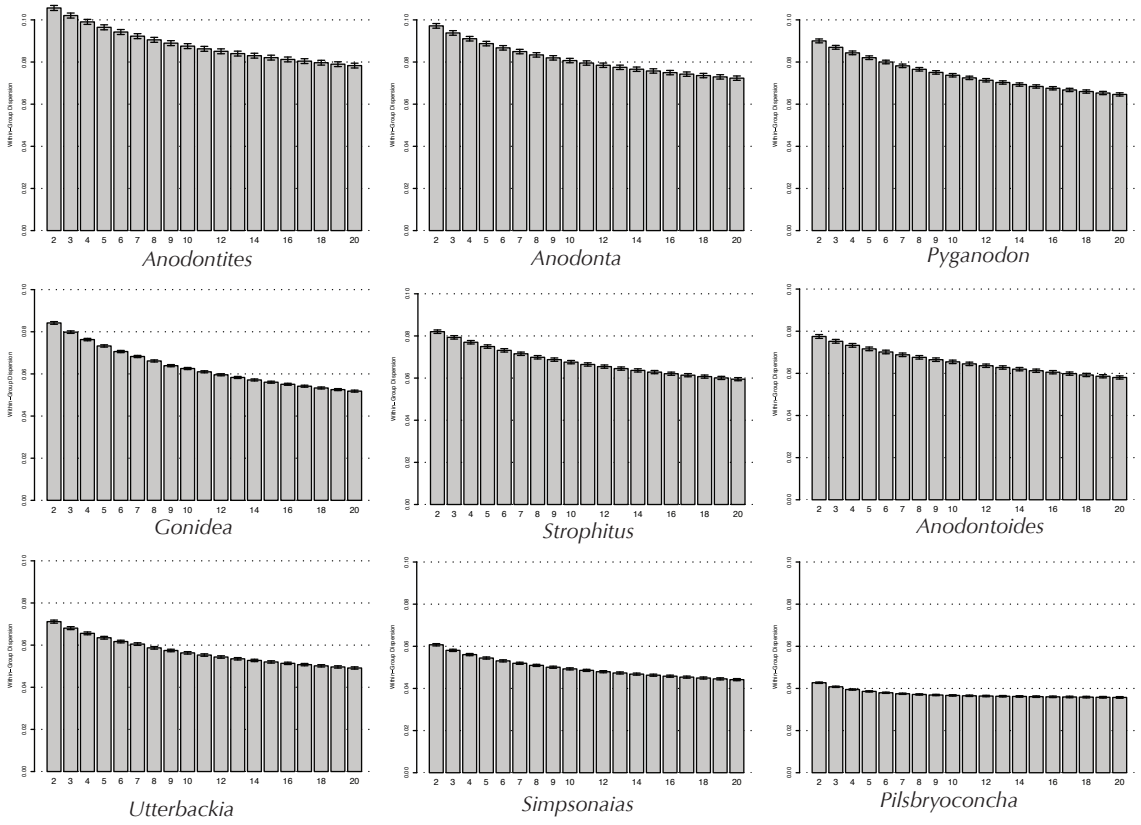

Figure 5.4: WD for individual genera plotted for EFA smoothing values. Error bars show 95% confidence intervals based on the t distribution. Genera are ordered according to maximum WD (at smoothing 2). Graphs produced with the R function `unibar.wd()` from data produced by `wd.upcon()` and `wd.lowcon()` on `euc.group()` output. Raw data in Tables D.23–D.31 on p. 164–168.

on p. 178).

The WD calculated for the selected L6516 specimens was 0.0890 (95% confidence interval from 0.0871 to 0.0910). This falls within the WD range defined by the modern genera examined (Figure 5.6). This value is statistically significantly higher than 66% (6 out of 9) of the selected modern genera based on 95% confidence intervals (Figure 5.6). A one-way ANOVA with post hoc Tukey's HSD tests found a statistically significant difference ($\alpha = 0.05$, F = 520.2, p(same) = 0; Welch F = 1513, p(same) = 0) between the L6516 group and all but two genera (higher than *Gonidea* and lower than *Pyganodon*); the L6516 WD was statistically significantly higher than 56% of

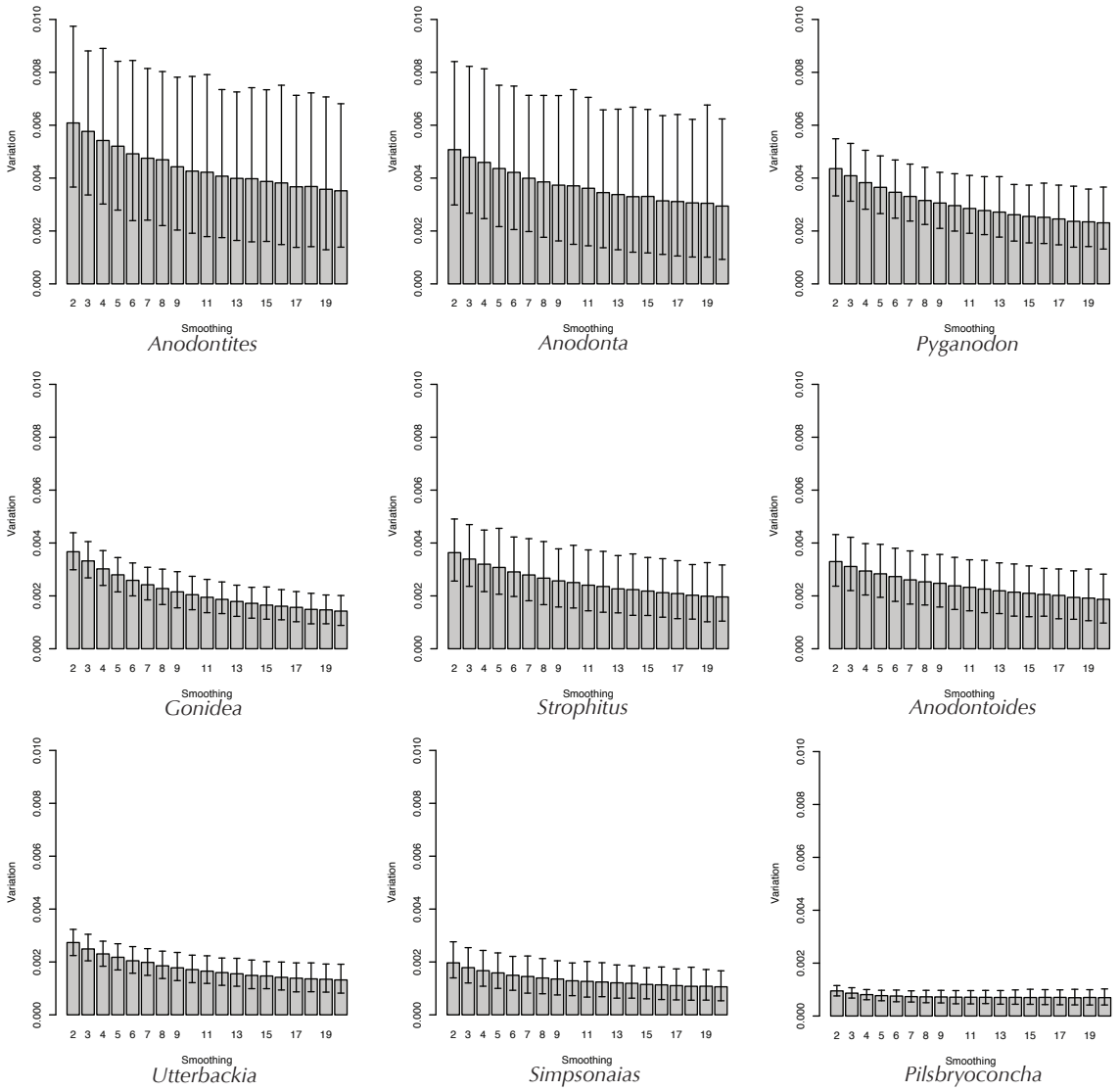

Figure 5.5: ΣV for individual genera plotted for EFA smoothing values. Error bars show bootstrapped 95% confidence intervals ($N = 1000$). Genera are ordered according to maximum ΣV (at smoothing 2). Graphs produced with the R function `justplotbars()` from data produced by `mv.table()`, `mvupcon()`, `mvlowcon()` on `confplotvbar()` output. Raw data in Tables D.32–D.40 on p. 169–173.

selected modern genera (5 out of 9) (Figures 5.6 and 5.8).

ΣV of L6516 specimens was only statistically significantly different than one of the selected genera (*Pilsbryoconcha*), making it statistically significantly greater than 11% of those genera based on 95% bootstrapped confidence intervals (Figures 5.7 and 5.8).

A plot of the first two principal components (variance-covariance matrix, SVD in PAST) of the EFA output for the modern specimens does not show large differentiation between all groups of specimens representing the different genera (Figure 5.9). However, some genera are notably distinct from others when 95% confidence limits are placed around them. The *Gonidea* envelope does not overlap with that of *Anodonta*, *Pilsbryoconcha*, *Strophitus* or *Utterbackia*. *Pilsbryoconcha* is distinct from *Simpsonaias* and *Strophitus* in the same manner. The first three principal components account for 25.8%, 24.9% and 8.6% of the variance, respectively. 95% of the variance is explained in the first 14 principal components (Table D.53). Principal component loadings (Table D.54) show a positive relationship between the first principal component and EFA harmonics B5, B7, and A6 (in descending order), the second principal component and harmonics B2, B5, B7, A3, and B6; a negative relationship of the first principal component with harmonics B2, A3, B4, B8, B6, and A5, and the second principal component and harmonics A3, A5, A7, and B4 (Figures D.1–D.2). A plot of the first two canonical variates designed to maximize between-group differences shows similar results (Figure 5.10).

5.4 Within-Group Dispersion and Sum of Variance of Extant Edentulous Freshwater Mussel Species and Comparison with Unionoids from Locality L6516

The average within-group dispersion of the mussels used in this analysis is 0.0704 with a range from 0.0427 to 0.1042 (range of 0.0615) (D.49, Figure 5.11). The range of 95% confidence intervals is from 0.0421 to 0.1063 (range of 0.0641). A one-way

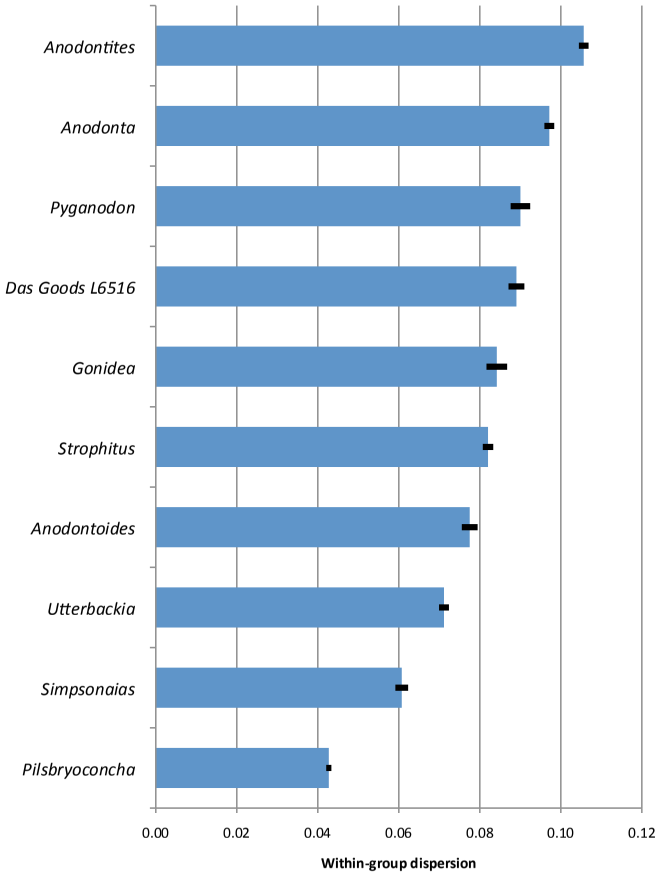

Figure 5.6: Plot comparing within-group dispersion of some edentulous freshwater mussel genera based on outline shape. Height of bar represents WD value (= mean pairwise Euclidean distance among specimens in the same genus). Error bars represent 95% confidence interval based on the t distribution. Raw data in Table D.43 on p. 175.

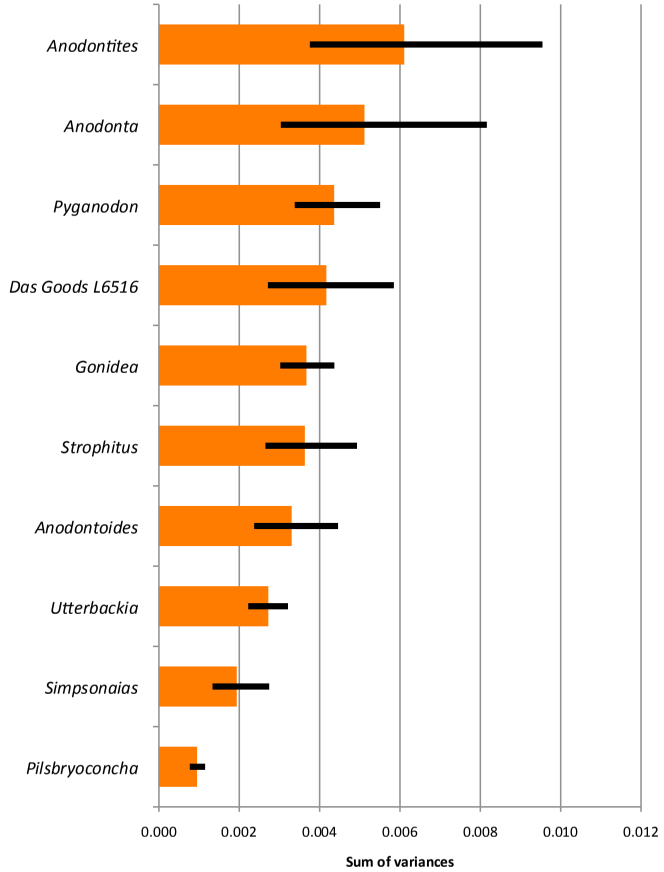

Figure 5.7: Plot comparing sum of variance of some edentulous freshwater mussel genera based on outline shape. Height of bar represents ΣV value based on bootstrapped ($N = 1000$) sum of variances, error bars represent bootstrapped ($N = 1000$) 95% confidence intervals. Raw data in Table D.44 on p. 175.

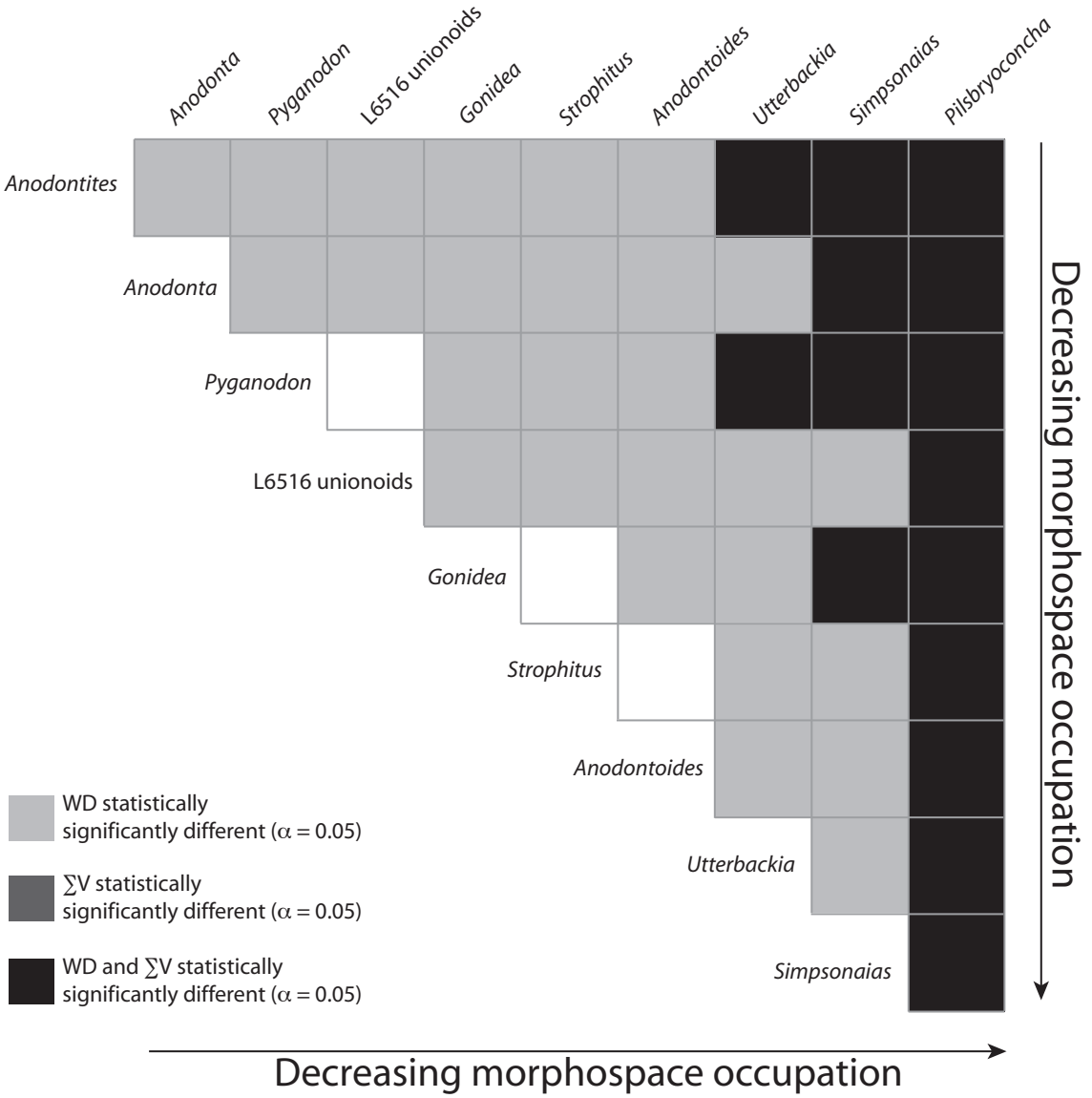

Figure 5.8: Summarized statistically significant differences among WD and ΣV for modern genera and L6516 specimens. Size of morphospace occupation decreases to the bottom right. WD data in Table D.45. ΣV data in D.44 on p. 175.

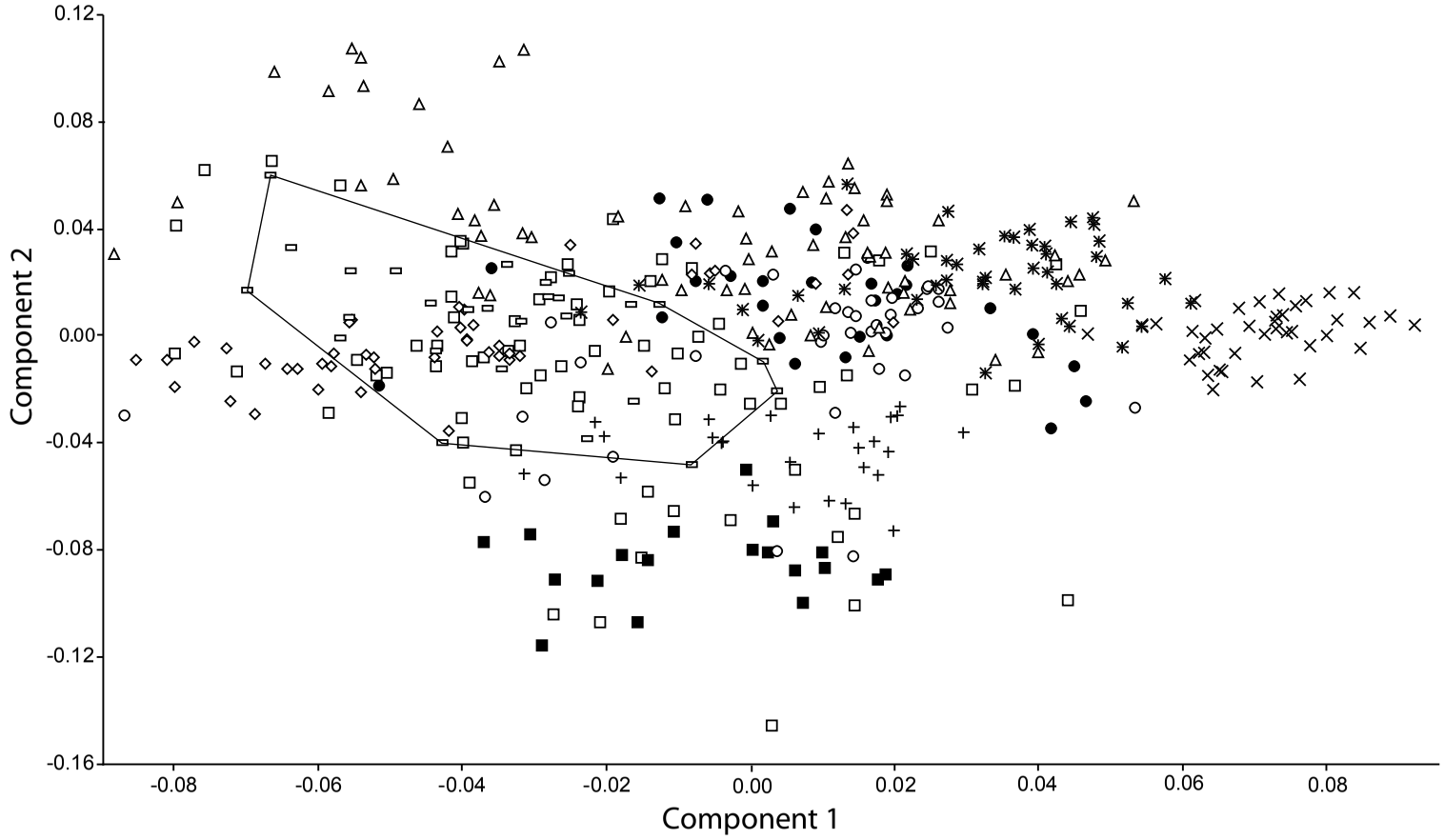

Figure 5.9: Principal component plot of elliptical Fourier coefficients of selected modern edentulous freshwater mussels and fossil unionoids from L6516. L6516 unionoids are outlined. Key to symbols: open triangle–*Anodonta*, open square–*Anodontites*, open circle–*Anodontoides*, filled square–*Gonidea*, x–*Pilsbryoconcha*, filled circle–*Pyganodon*, cross–*Simpsonaias*, open diamond–*Strophitus*, star–*Utterbackia*, open rectangle–L6516 unionoids. Eigenvalues in Table D.53 on p. 184.

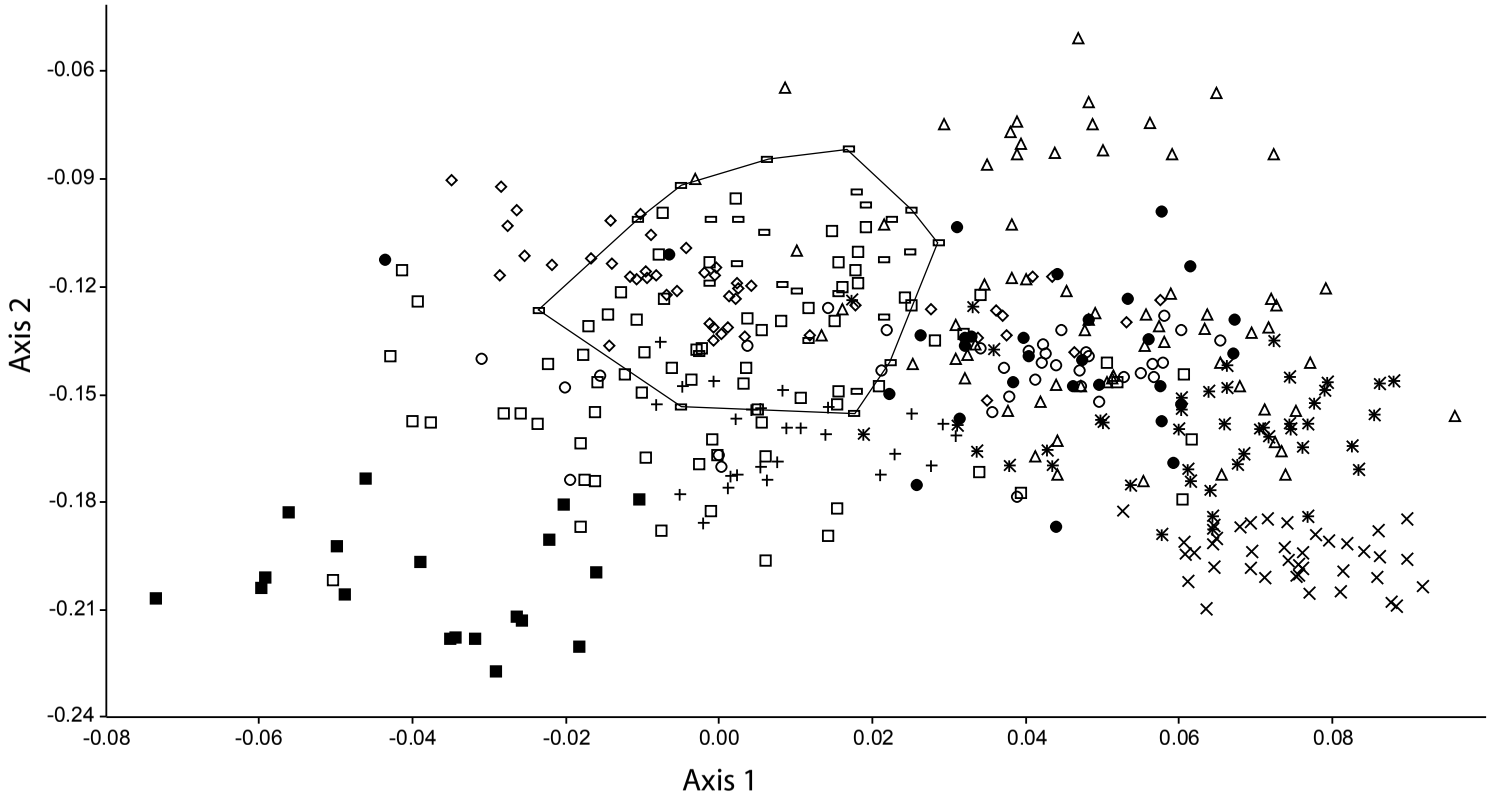

Figure 5.10: Canonical variate plot of elliptical Fourier coefficients of selected modern edentulous freshwater mussels and fossil unionoids from L6516. L6516 unionoids are outlined. Key to symbols: open triangle–*Anodonta*, open square–*Anodontites*, open circle–*Anodontoides*, filled square–*Gonidea*, x–*Pilsbryoconcha*, filled circle–*Pyganodon*, cross–*Simpsonaias*, open diamond–*Strophitus*, star–*Utterbackia*, open rectangle–L6516 unionoids.

ANOVA found a statistically significant difference among WD values for modern species ($\alpha = 0.05$, $F = 103$, $p = 0$; Welch $F = 184.3$, $p(\text{same}) < 0.0001$). Tables D.51 and D.52 on p. 182 and 183 shows post hoc Tukey's HSD test results (Figure 5.13). 95% confidence intervals based on the t distribution agree with these tests (Figure 5.11).

The WD calculated for the selected L6516 specimens was 0.0890 (95% confidence interval from 0.0871 to 0.0910). This falls within the WD range defined by the modern species examined (Figure 5.11). A one-way ANOVA with post hoc Tukeys HSD tests found a statistically significant difference ($\alpha = 0.05$, $F = 113.5$, $p(\text{same}) = 0$; Welch $F = 211.8$, $p(\text{same}) < 0.0001$) between the L6516 group and nine of the modern species; the L6516 WD was statistically significantly higher than 88% of modern species (21 out of 24, see Figure 5.13).

The average sum of variance of the modern species used in this analysis is approximately 0.0028 with a range from 0.0008 to 0.0108 (range of 0.0101) (D.50, Figure 5.12). The range of 95% confidence intervals is from 0.0008 to 0.01 (range of 0.009). Some genera were found to possess statistically significantly different ΣV values based on non-overlapping confidence intervals (Figure 5.13). These do not agree with the WD ANOVA and confidence interval results because of the larger confidence intervals around ΣV, however the overall rank-order relationship between genera appears to be the same (Figure 5.12).

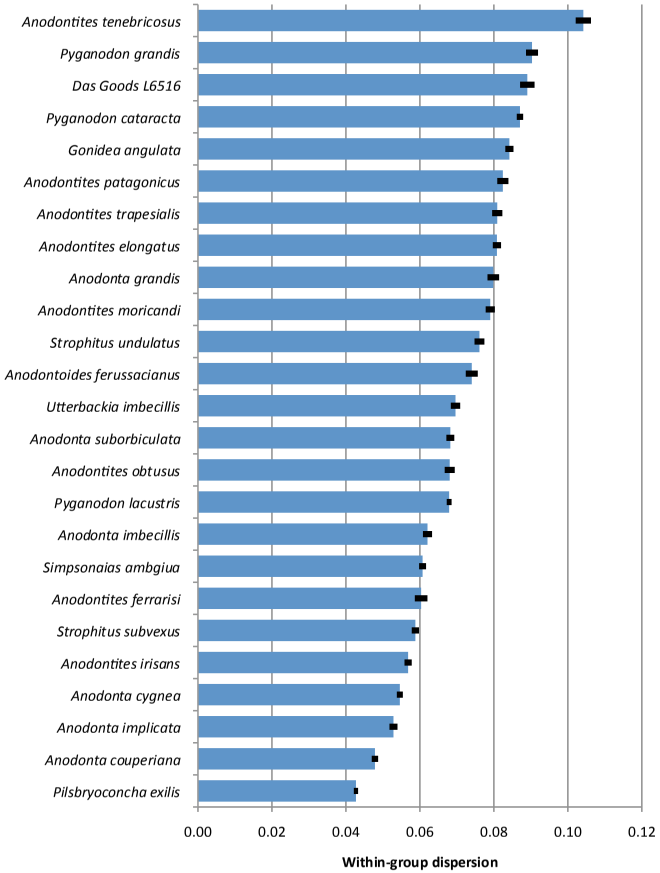

Figure 5.11: Plot comparing within-group dispersion of some edentulous freshwater mussel species based on outline shape. Height of bar represents WD value (= mean pairwise Euclidean distance among specimens in the same species). Error bars represent 95% confidence interval based on the t distribution. Raw data in Table D.49 on p. 180.

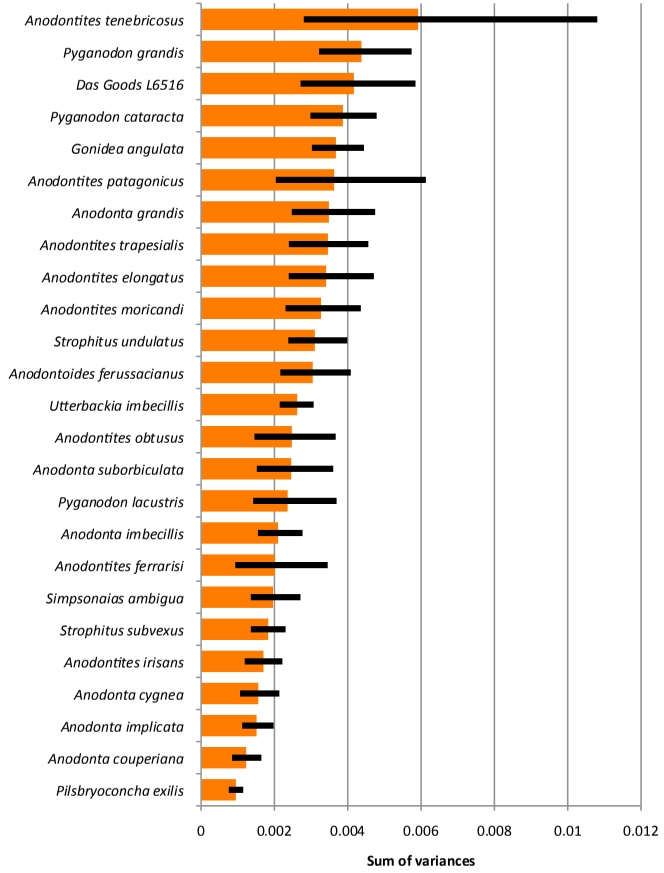

Figure 5.12: Plot comparing morphological variation of some edentulous freshwater mussel species based on outline shape. Height of bar represents MV value based on bootstrapped ($N = 1000$) sum of variances, error bars represent bootstrapped ($N = 1000$) 95% confidence intervals. Raw data in Table D.50 on p. 181.

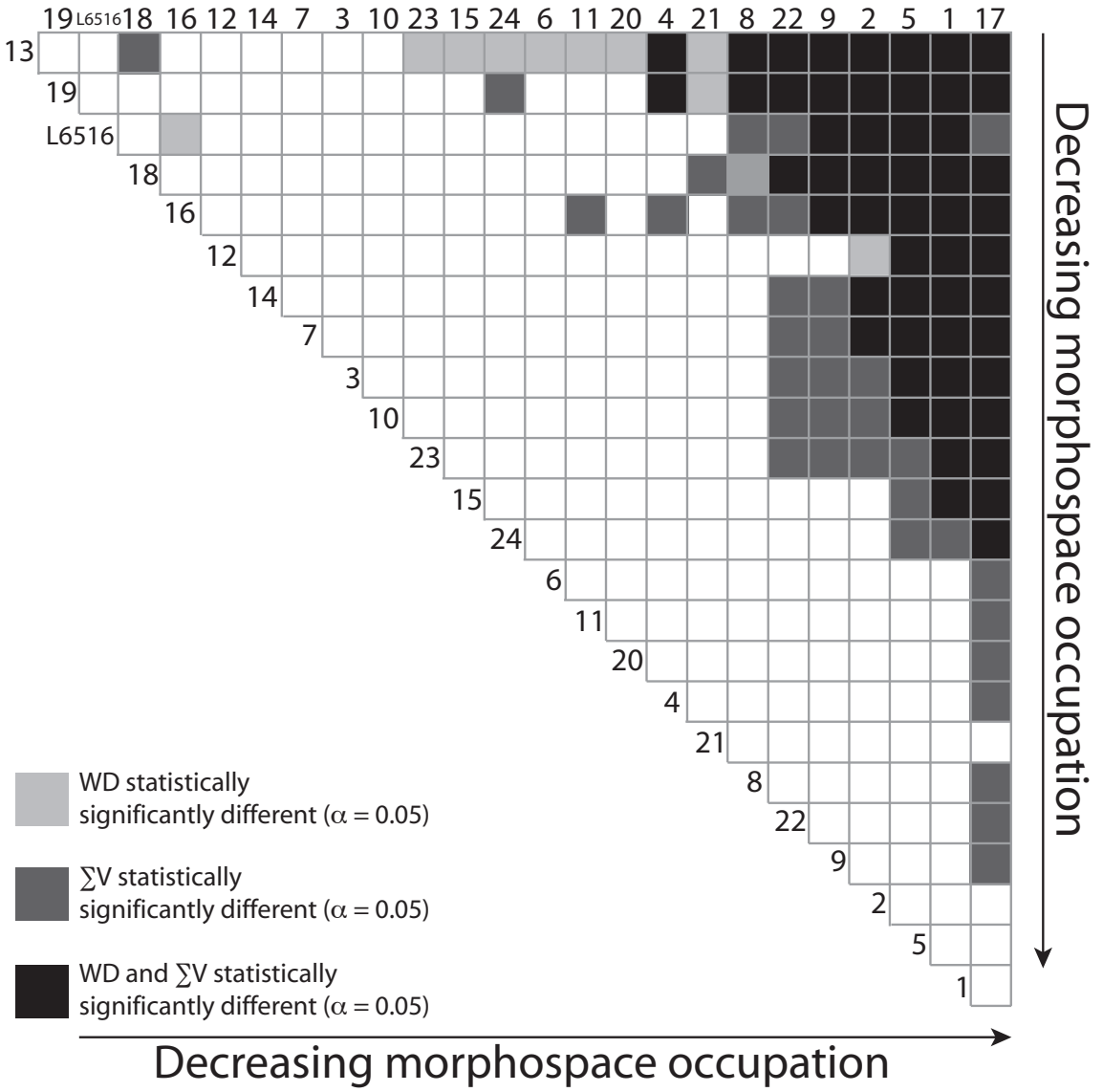

Figure 5.13: Summarized statistically significant differences among WD and ΣV for modern species and L6516 specimens. Size of morphospace occupation decreases to the bottom right. WD data in Tables D.51 and D.52 on p. 182 and 183. ΣV data in Table D.50 on p. 181. Key to species (alphabetical): 1–*Anodonta couperiana*, 2–*Anodonta cygnea*, 3–*Anodonta grandis*, 4–*Anodonta imbecillis*, 5–*Anodonta implicata*, 6–*Anodonta suborbiculata*, 7–*Anodontites elongatus*, 8–*Anodontites ferrarisi*, 9–*Anodontites irisans*, 10–*Anodontites moricandi*, 11-*Anodontites obtusus*, 11–*Anodontites obtusus*, 12–*Anodontites patagonicus*, 13–*Anodontites tenebricosus*, 14–*Anodontites trapesialis*, 15–*Anodontoides ferussacianus*, 16–*Gonidea angulata*, 17–*Pilsbryoconcha exilis*, 18–*Pyganodon cataracta*, 19–*Pyganodon grandis*, 20–*Pyganodon lacustris*, 21–*Simpsonaias ambigua*, 22–*Strophitus subvexus*, 23–*Strophitus undulatus*, 24–*Utterbackia imbecillis*.

CHAPTER 6

DISCUSSION

6.1 Areal Extent of Deposit and Future Sampling Procedures

More localities of this nature undoubtedly exist in this area. A combination of factors contributed to a lack of understanding of the geographical extent and shape of this apparently continuous horizon. Field time was limited, and as this study was primarily designed to examine the within-locality morphology of the unionoids from L6516, excavation at other sites was curtailed after sufficient material was collected to document the existence of other sites. The two buttes containing sites are locally the northern-most extension of rocks containing the Hell Creek/Ludlow contact, limiting further exploration.

The equivalence in stratigraphic position of the Das Goods localities cannot be affirmed without more stratigraphic sections measured in direct proximity to the localities. Generally, however, the sites occur in the interval between the Hell Creek-Ludlow contact coal and the first coal above in the Ludlow Member of the Fort Union Formation, an interval that varies in thickness throughout the area. The presence of similarly shaped unionoid mussels with the same style of preservation in similar sediments strongly suggests that these deposits are penecontemporaneous and part of the same lake, pond, or large meandering river system. More work is needed to determine the extent of this deposit and whether it is truly laterally continuous between sites.

6.2 Preservation

The preservation of the unionoids at L6516 is both good, for the presence of some growth lines, trace fossils, and complete outlines, and bad, for the lack of original shell material and peculiar mold and cast relationships within a single specimen. Luckily, preservation was mainly parallel to the bedding plane, however this led to problems with interpretation in some specimens (particularly S2871a and S2871b), which were truncated along growth lines, making it difficult to tell that the shell had at one time extended farther. The timing of the dissolution of shell material could not be determined except that it had to have taken place after the sediment was dewatered and somewhat hard in order for the molds of the interior and exterior to have been preserved, yet still plastic enough for some of these molds to have produced casts.

6.2.1 Trace Fossils

The overall poor preservation of these specimens makes it difficult to accurately characterize the traces preserved on the valves. There appear to be traces in both concave and convex epirelief on all areas of the valves, meaning that regardless of whether the valve surface is a mold or cast of the interior or exterior, there are two types of traces preserved: one type which represents a structure built atop the shell surface and one type which represents a removal of shell material along the surface. The range in width of the traces was not measured, but suggests either that at least two types of organism were involved (per convex or concave class of traces) or that the producers were of the same type over a growth gradient.

The most striking aspect of these structures is their coverage in some cases of most of the valve, regardless of the exposed surface being internal or external, which may suggest that these are not trace fossils at all but are diagenetic features. While this may be the case, there are a number of organisms that could have produced these or similar traces both before and after the death and disarticulation of the unionoids;

epizoan communities on modern unionid clams can be quite diverse (Curry et al., 1981 and references).

Convex Traces

The traces that originally appeared as convex structures on the surface of the valves could have been produced by a number of organisms. They were initially thought to be root casts, however the seemingly random orientation and non-branching nature of the traces argue against this. Roots would also, if appearing concurrently with the burial of the deceased unionoids, be subject to the same taphonomic processes as the valves themselves, and would not likely have been preserved. The fact that the traces do not cross the edge of the valves into the surrounding matrix supports the idea of an epizoic or endozoic producer. Modern aquatic plants such as *Podestemum* (riverweed) sometimes colonize the surface of living freshwater mussels (Vaughn et al., 2002). The holdfast "root" attachments of freshwater *Podestemum ceratophyllum* are similar to the convex structures (Rutishauser, 1997), however no identifiable aquatic plants were recovered from this horizon.

Bryozoa such as modern *Plumatella* live in freshwater, however the lack of a dendritic arrangement of the structures on the L6516 unionoids does not resemble the modern species. One fossil genus (*Plumatellites*) was described by Fric and Bayer (1901) as encrusting a Cretaceous *Unio*, however this source was not available for review (Bassler, 1953). The narrow linear form and apparent independent growth of these structures does not suggest the remains of freshwater sponges, which do not typically grow in such an organized manner when encrusting unless multiple unrelated colonies are competing for space (Frost, 1991), something of which there is no evidence at L6516. Related encrusting sponge colonies typically have no problems overlapping while growing (Frost, 1991; Lauer et al., 2001).

Obvious arthropod producers of tubes on hard substrates in freshwater are cad-

disfly larvae (Insecta: Trichoptera), which date back to the Triassic and are known locally from the Paleocene (Sentinel Butte Member of the Fort Union Formation) of North Dakota (Erickson, 1983; Grimaldi and Engel, 2005). Both fossil and modern examples of caddisfly cases are proportionally much shorter and wider than the convex structures on the L6516 unionoids. Some types (e.g., members of tribe Tanytarsini) of chironomid larvae (Insecta: Diptera) also produce hard, long, narrow, non-branching tubes on hard substrates, although the majority of modern tube-building taxa do so at the sediment-water interface (Oliver, 1971; Erickson, 1983; Pinder, 1986; Mermillod-Blondin et al., 2003; Ólafsson and Paterson, 2004). In some experimental settings, mucus-sediment tubes up to 40 mm in length have been measured, representing (for example) *Cricotopus*, *Psectrocladius*, and *Tanytarsus* by Chaloner and Wotton (1996) and *Micropsectra* and *Pseudodiamesa* by Pringle (1985), although these tubes were not affixed to a hard substrate and one genus (*Pseudodiamesa*) produced irregular structures rather than single tubes. There are modern examples of commensal or parasitic relationships between chironomid larvae and many freshwater mussel taxa, however these larvae live within the mantle cavity (Roback et al., 1979), in some cases being overgrown by the nacre and appearing as convex surface structures a few millimeters in length (Forsyth and McCallum, 1978).

Only three families (Ampharetidae, Sabellidae, and Serpulidae) of tube-dwelling polychaete annelids are known in freshwater, and of these only the serpulids attach calcareous tubes to hard substrates or aquatic plants (Davies, 1991). These tubes can be many times the length of the animal within (Glasby et al., 2000). Members of the Serpulidae are known from the Palaeozoic Era, many having coiled rather than straight encrusting tubes in marine settings (Howell, 1962). Only a few modern genera of serpulids or sabellids are known in freshwater: *Ficopomatus* (= *Mercierella*; Barnes:1994aa) is known to produce calcareous tubes on hard substrates in fresh-

water (Foster, 1972), however they are built colonially and erect from the substrate (Barnes, 1994); and *Manayunkia* builds tubes of mud or sand and mucus (Smith, 2001). Polychaetes or similar worms tunneling between the surface of the shell and the periostracum might conceivably produce a raised blister on the exterior of the valve, however this has not been reported.

None of the above taxa, unfortunately, have been reported in the literature as having a definitive association with unionoid mussels and producing the type of convex structure seen on the L6517 specimens. Without doing more in-depth actualistic studies on the relationship between freshwater mussels and epizoan traces, the most likely candidates for production of the convex traces as seen at L6516 are chironomid larvae, polychaete annelids, and aquatic plant holdfasts.

Concave Traces

Few possibilities present themselves as producers of the concave traces on the L6516 unionoids. Although viviparid gastropods are also present at L6516, the concave traces on the unionoid valve surfaces do not appear to be radular marks (c.f. *Radulichnus*) due to the smooth surfaces of the traces and the visible continuity along the length of a single structure (Smith, 1988; Seilacher, 2007). The modern freshwater boring polychaete *Caobangia* from southeast Asia is known to infest thiarid, pleurocerid, and viviparid gastropods and unionid mussels; the traces left behind are teardrop-shaped burrows rather than surficial grooves (Jones, 1969, 1974). Organisms feeding on the periostracum itself could produce such traces, however this has not been reported. Polychaetes are postulated to be the most likely producer of these traces, with the recognition that their habit of movement would have been parallel to the surface of the valve rather than perpendicular as is seen in modern forms.

6.3 Digitization Error

The results of the analysis of the within-group dispersion (WD) suggest that the best usable on-screen length of a specimen for during manual digitization is approximately 25 cm in order to minimize error introduced by the person digitizing the outline. Therefore, 25 cm should be chosen as a standard on-screen length for manual digitization of outlines, however some inaccuracies may stem from procedural errors. First, the synthetic outlines used for multiple digitization trials were composed of lines with an absolute width on the screen which changed with changes in magnification. These outlines were digitized by tracing the best approximation of the center of the drawn line, while outlines of modern specimens were digitized by tracing the edge of the valve. (Fossil specimen outlines were drawn prior to digitization as well, which meant the digitization of a line with width rather than an edge with no width.) A better method would have been to digitize the outermost edge of all synthetic valves. Second, synthetic valves were produced from a narrow range of fossil morphologies rather than a wider range of modern genera. The relationship between the amount of digitization error and the digitization length may vary according to the departure of the valves from a perfect ellipse. Third, only one person (the author) digitized all of the valves used in this study. Although this should standardize the digitization error relatively well, the amount of variance in digitization from person to person is unknown, and should be taken into account if trying to compare digitized outlines from different studies. The optimal digitization length for one worker may not be the best for another, although it is hoped that if this is the case more robust methods can be found by empirically testing human causes of variation in outlines. With specimens that are more easily distinguished from a background, automatic edge detection can be used, such as in tpsDig and ImageJ (Abramoff et al., 2004; Rohlf, 2008).

Changes in the outlines of individual valves at different digitization lengths may

be statistically significant. This was not tested, but since WD and MV get smaller as the outline is digitized at a larger size, any study that deals with inter- or intra-group variation should conservatively limit itself to utilizing specimens that were digitized with the same methods throughout—either by using the automated digitization method or by manually digitizing the valves at the same screen size and, as stated above, preferably by the same worker until the human effects have been more closely studied.

This caveat can be applied to the present study, where valves of 47% of modern specimens were digitized at an unrecorded arbitrary length related to screen size, but estimated to average from 20 to 30 cm after the fact. Genera are made up of valves digitized predominantly at one length (either 25 cm or arbitrary), with no more than three valves making up the minor portion. The largest digitization length a specimen could possess on the computer monitor used was approximately 30 cm, although the range is most likely from 10 cm to 40 cm or more (requiring scrolling during digitization). The estimated effect of this mistake is that the WD of some modern genera (*Pyganodon*, *Simpsonaias*, and *Strophitus*) is inflated due to the added digitization error from digitizing some specimens at non-standard sizes. This may have resulted in an inflated sense of the maximum size of morphospace occupation by these modern genera and, as applied to the fossil specimens, an undercounting of the number of possible fossil genera based on the limits set by the within-group dispersion and the morphological variation. This situation is regrettable (however little the effect seems to be) but, due the effort involved in manually redigitizing all modern specimens at a standardized length, currently unavoidable. These three genera do not appear to possess more or less WD or MV than the rest of the genera (i.e., they do not set the upper or lower limit of either statistic).

6.4 Standardized versus Arbitrary Digitization Length

Although both Student's t test and 95% confidence intervals based on the t distribution show a statistically significant difference between WD of *Anodonta* specimens digitized at 25 cm and those at an arbitrary length, these results are suspect due to the non-normality of the differences between the paired Euclidean distances and the non-normality of the Euclidean distances themselves. In contrast, the MV results seem more reliable because both MV values and confidence intervals are based on bootstrapped data, which has been shown to produce more accurate results with non-normal or strongly skewed data (Hesterberg et al., 2005). As shown, digitization length does have an effect on WD and MV. These results suggest that the arbitrary digitization length was in fact averaged around 25 cm to produce these results. If the WD t test and confidence intervals can be accepted, they show a lower WD for the arbitrary digitization length than for the 25 cm length; this is counter to the hypothesized (and intended) effect of standardizing the digitization length. This effect may be due to digitization at a greater length than was tested for (greater than 30 cm), or possibly to an average digitization length that is between 25 cm and 30 cm which was not examined.

Multivariate analyses on the original Fourier scores show that there is no statistically significant difference between the Fourier coefficients of the specimens across digitization lengths. Although these results, coupled with the equality of the MV values based on confidence intervals, might appear to suggest that an equality in morphospace occupation has something to do with equality of shape, this is definitively not the case as many disparate organisms could possess the same amount of WD or MV, while being vastly different shapes.

6.5 Smoothing Effects During Elliptical Fourier Analysis

Performing Elliptical Fourier Analysis on the same generic groups at increasing amounts of smoothing resulted in a marked decrease in within-group dispersion. The decrease shown by sum of variance was not, however, statistically significant due to large 95% confidence intervals. The effect tentatively shown by these two methods is a reduction in the size of morphospace occupation of a group with increased amounts of smoothing, leading to a homogenization of outlines represented in the group. Although such homogenization makes valves more similar within groups, it also has the effect of making specimens more similar between groups, which could lead to more morphospace overlap than is practicable for statistically determining the presence of groups.

The lowest smoothing value was used in further analyses when manually digitizing outlines because smoothing is typically necessary only at high numbers of points per outline (>700) (Crampton and Haines, 1996). A smoothing value of 2 was the lowest value that could be used and not cause HAngle to fail on certain outlines.

Interpolating the varying amounts of outline points per specimen to produce 1024 points before applying smoothing may standardize the valves before smoothing is applied and minimize the overall movement of outline points, essentially unweighting the original points by increasing the overall number of points along the straight-line intervals between them.

6.6 Placement of the L6516 Unionoids

The placement of the L6516 unionoids within the context of the size of morphospace occupation of the modern edentulous genera and species used is not definitive; i.e., the L6516 specimens did not occupy significantly more or significantly less morphospace than all other genera and species tested. Qualitatively, however, the possibility of more than one morphotype at L6516 is clear to the naked eye, and these

could probably be described utilizing current mothodologies of genus and species description to discriminate, for example, between the circular, quadrate, and long and narrow forms.

The underlying questions at stake are 1) whether a quantitative or qualitative method is more useful in determining what defines a genus or species and 2) if quantitative methods are worth the trouble. Identification to the species level of any organism should be based on discrete characters, which can be described qualitatively as well as mathematically, but each method has flaws in the way they can be interpreted: qualitative characters can be argued according to preservation or individual variation or pathology, and quantitative characters can be manipulated with different methods of significance testing or variation in alpha levels.

The failure of the analyses presented here to use size of morphospace occupation to calculate the number of unionoid genera or species present at L6516 is not necessarily based on poor methods, although the methods need to be improved. Perhaps the L6516 unionoids occupy less morphospace than some modern edentulous genera and species and still represent multiple morphotypes that, when the size of morphospace occupation for each is added together, it is still less than those genera with high within-group variation. To solve this problem, subgroups of the L6516 unionoid fauna can be selected manually (based on qualitative assessment of morphotypes) or automatically (based on all possible combinations of specimens) and tested against modern taxa. Future work may involve determining the possible morphotypes at L6516, comparing the shapes of different morphotypes using some of the multivariate methods already discussed, and calculating the size of morphospace occupation which, with an improvement of these methods, should become a more reliable value of comparison.

6.7 Methodological Issues

A number of issues exist with the methods used above that need to be addressed. Although it is not the opinion of the author that the results specified above are inaccurate, criticism can be made of specific aspects of the methodology that can be improved and extraneous variation removed from the calculations of size of morphospace occupation.

6.7.1 Choice of Modern Genera

The primary concern when interpreting these data is whether the extrapolation from the modern forms selected can be applied to the fossil unionoids from Das Goods. This refers specifically to the choice of the modern genera and species used to set the baseline of size of morphospace occupation. The modern genera initially selected were edentulous forms relatively simple to obtain, which is far from a systematic approach. Only nine out of over two dozen genera lacking hinge teeth were analyzed.

This concern is an important one when utilizing modern forms to determine the taxonomic identity of fossil assemblages. Without selecting a specific set of modern genera identified by a quantitative shell character and analyzing a large number of each of those genera, the actual position of the fossil assemblage within the range of size of morphospace occupation can only be a rough estimate. Put another way, had the three genera (*Anodontites*, *Anodonta*, and *Pyganodon*) or two species (*Anodontites tenebricosus* and *Pyganodon grandis*) occupying more morphospace than the L6516 unionoids been left out by accident or design, there would be more support for the possibility that more than one genus-sized group of unionoids occurs at the Das Goods localities.

If all modern edentulous unionoid mussels been included in the analysis, however, the argument would clearly be against the possibility of multiple genera at Das Goods. Unfortunately, by this definition (and working within the sample group) all genera

except for *Anodontites* and all species except for *Anodontites tenebricosus* would be suspect, when clearly they have different valve shapes and soft-part morphology. This is not to discount the work described above, but to recognize that these methods can only be used to find extreme groups at the generic and specific levels.

6.7.2 Ontogeny and Size

Capturing variation in unionoid mussels is difficult because of ontogenetic variation within genera and species. Additionally, because of the environmental plasticity of the unionoids, they are subject to variation in growth rate among habitats even along their ontogenetic trajectories. Optimally, morphospace occupation would be calculated with specimens of the same age from the same site, which would theoretically be the same size due to their common habitat. This would control for ontogeny and size, however an adequate sample size would be difficult to obtain for every genus and species used in this project. It is unclear whether size or ontogeny of freshwater mussels has a more stable relationship to shape; if this were calculated, either age (based on growth lines, and able to be estimated even in the L6516 specimens) or size (based on length if using outlines or centroid size if using landmarks) could be utilized as a measure of standardization, potentially allowing specimens from multiple sites to be used in calculating generic or specific morphospace occupation.

6.7.3 Morphological Plasticity and Convergence

Specimens of the same genera were obtained from multiple museums and multiple publications, and among those collections from a variety of habitats in watersheds throughout the world. Understanding the plasticity of the unionoids according to habitat (leading to convergence through space and time) is key to improving studies, such as the present study, that extrapolate from the present to the past.

Optimally, all modern specimens (of all applicable modern genera and species, discussed above) would be collected from a similar environment as the paleoenviron-

ment represented at Das Goods—the muddy bottom of a long-lived pond or lake. This would help to reduce the amount of calculated morphospace occupation due to specimens from different environments possessing differing morphologies. Additionally, specimens of a single genus or species would be most likely to be similar if collected from the same habitat in the same watershed, although locating hundreds of specimens collected in this manner would be difficult, if not impossible, without a designated collecting expedition. Such a project would create possibilities of comparing the morphospace occupation of taxa from multiple habitats and watersheds with fossil localities, and with each other, to determine the interaction of morphospace occupation with habitat, population dynamics and geography in an attempt to fill in some of the gaps in the fossil record. Investigations of this type have recently been accomplished by (Costa et al., 2008), on marine clams, showing that morphological distances between species can be less than the morphological distance between different populations of the same species.

6.7.4 Taphonomic Deformation

An original goal of this project was to determine the potential amount of deformation undergone by the L6516 fossils due to lithostatic loading and unloading. The gastropod steinkerns in particular have been compressed to some degree parallel to the bedding plane, which suggests that the unionoids have as well. Due to time constraints, physical tests to determine the possible amount of outline deformation due to compaction were not accomplished, and the outlines were used as-is.

6.8 Morphometrics

Outline analysis and geometric morphometrics have a great deal of potential for use in the natural sciences, however there needs to be more cohesiveness within the field regarding standardization and communication. Transformation of data for use between different software packages was extremely arduous. Standardization of data

formats will allow workers to exchange information, freely and without loss, to be used in different programs. The release of different standardized datasets can be used by newcomers to the field to learn how to utilize the methods involved, and by experienced workers to attempt new and better methodologies. Simple, clear communication will be key for the newer morphometric procedures to be used by those who did not create them. Detailed, step-by-step procedural methods need to be recorded and published, not so that newcomers can produce data without understanding morphometric theory, but so mistakes can be avoided, problems identified, and solutions created for difficult tasks. Software (for data capture and statistics) needs to be documented, including reference to the theory behind the point and click interface (for an excellent example, see Hammer et al., 2008).

Many custom scripts and small programs had to be written to streamline the data capture and manipulation for this project; for someone less able or willing to produce such custom software, use of such a large data set may be overwhelming. As newer morphometric procedures are utilized by more workers, more support for the existing software will drive improvement of data manipulation capabilities (for interoperability of different software packages that were not originally designed to work together), while hopefully allowing for detailed control of data when experimenting with new methods.

All of the software used during this project works well and has dedicated support for improvement. Not all software that is designed to work with outline data was tested, and no endorsement (positive or negative) is implied. Software that was *not* used for this project includes SHAPE (Iwata and Ukai, 2002) and Morpheus (Slice, 2007). New and promising statistical and modeling techniques have recently been released, including geodesic distance shape analysis (GDA) (Klassen et al., 2004; Prieto-Marquez et al., 2007), various methods of shape classification (Joshi and Sri-

vastava, 2003; McNeill and Vijayakumar, 2005) soft independent modeling of class analogy (SIMCA), and partial least square discriminant analysis (PLSDA) (Costa et al., 2008).

CHAPTER 7

CONCLUSIONS

This project fell short of its intended goal of determining the number of unionoid taxa present at Das Goods locality L6516, but did produce useful results for comparing fossil and modern taxa at the assemblage level. A close examination of individual specimens yielded a greater understanding of the preservation occurring and the surprising trace fossil assemblage. Methodological problems that were encountered over the course of the project were also addressed with the intent that future studies will produce more taxonomically useful results. Specific conclusions are listed below.

1. Fossil locality L6516 is not a unique occurrence of this type of preservation, which was documented at five additional sites in close stratigraphic proximity within a five hectare area. Similarity between the molluscan fauna, sedimentology, and stratigraphic position of these sites indicates a shared penecontemporaneous environment, such as a lake or pond.

2. Unionoid specimens from L6516 are preserved as a composite of molds of the interior (typically near the umbo) and casts of the exterior (typically near the margin) of the valves, with no visible original shell material present.

3. The freshwater benthic fauna at locality L6516 preserves not only unionoid mussels and viviparid gastropods, but traces of a previously undescribed eipifaunal trace fossil assemblage attributed to polychaete worms, chironomid larvae, or epifloral aquatic plant holdfasts.

4. The size of morphospace occupation of extant edentulous freshwater mussels

can be calculated and ranked according to the within-group dispersion and sum of variance measures, based on elliptical Fourier coefficients of the outlines of the valves.

5. The unionoid mussels preserved at locality L6516 do not possess statistically significantly more or less morphological variation (using the within-group dispersion and sum of variance measures) than the selection of modern genera and species used, based on elliptical Fourier coefficients of the outlines of the valves.

6. Methodological problems, including choice of modern genera and species, ontogeny and size of taxa, morphological plasticity and convergence, and taphonomic deformation of the fossil specimens, contributed to exaggerated size of morphospace occupation.

7. Morphometric techniques, morphometric datasets, and morphometric procedures will need to be standardized before classification based on computer-intensive methods will be practicable.

APPENDICES

APPENDIX A

COMPUTER SOFTWARE AND FILE TYPES

A.1 Computer Software

Computer software packages are listed alphabetically.

CorelDraw®— Image manipulation. Adding vector data to specimen photographs to speed outline digitization for Elliptical Fourier Analysis. Editing PAST chart outputs. Version 12 (Windows XP). Corel (2003).

Excel®— Data manipulation. Reformatting of HAngle and HMatch output to be used as multivariate data in PAST. Simple plotting of results, data recording. Version 9.0 (Windows XP) and 12.1 (Mac OS 10.5). Microsoft (1985-2007).

Illustrator®— Image manipulation. Adding vector data to specimen photographs to speed digitization for Elliptical Fourier Analysis. Editing PAST chart outputs. Version CS2 (Mac OS 10.5). Adobe (1990-2005).

PAST — Statistical testing. Version 1.82b (Windows XP). Hammer (1999-2008).

Photoshop®— Image manipulation. Cropping, rotating, and mirroring specimen photographs. Version 9.0 (Windows XP) and CS2 (Mac OS 10.5). Adobe (1990-2005).

R — Data formatting of HAngle and HMatch output (after Excel) for PAST. Statistical testing using bootstraps and permutation tests. Version 1.24-devel (Mac OS 10.5). R Group (2004-2007).

tpsDig — Data capture of outline data. Used to digitize valve outlines of fossil and modern specimens for use in EFA. Version 2.05 (Windows XP). Rohlf (2006).

A.2 Custom Programs and Scripts

See Appendix B.3.2 for purpose, details for use, and source code.

A.3 File Types

File types are presented in the order they are produced during digitization, Elliptical Fourier Analysis, and data manipulation prior to statistical analyses.

TPS file — (*.tps) Output from tpsDig. First line listing number of landmarks. Second line listing number of curves. Multiple lines listing coordinate pairs of points along each curve (one pair per line). Final line listing file name of source image.

TPR file — (*.tpr) Output from `remove-lastline.bat` or `addfileprefix.bat`. Input for HAngle, HMatch and `del3resapp.bat`. Same data structure as TPS file but without final line.

Fourier Coefficient (FC) — (*.xls or *.txt) Output from "Fourier Format HAngle.xls" or "Fourier Format HMatch.xls". Input for `euc.list()` and `hbootsumvar()` functions in R. First row, Fourier coefficient labels (A2, B2, C3,...). Multiple rows, Fourier coefficients for specimens, one specimen per row. The first column contains specimen identifiers.

HAngle file — (*.unf) Output from HAngle. First row describes EFA settings. Multiple rows, numbered pairs of Fourier coefficients

HMatch file — (*.fit) Output from HMatch. First row describes EFA settings. Multiple rows, numbered pairs of Fourier coefficients

`euc.list()` file — (*.euc) Output from `euc.list()` R function. Input for `euc.group()` R function. Mean pairwise Euclidean distance between all specimens in a group (line 1) and the pairwise Euclidean distances themselves (subsequent lines).

`euc.group()` file — (*.euc) Output from `euc.group()` R function. Input for `meanxbar()` R function. Euc.list files arranged one per column. Missing data is filled with NA.

`meanxbar()` file — (*.euc) Output from `meanxbar()` R function. Mean Euclidean distances per group (from Euc.group files) arranged in columns (one column per

Euc.group input file).

APPENDIX B

DATA MANIPULATION

B.1 Elliptical Fourier Analysis

1. TPR files from chosen specimens were collected in a single directory. Each specimen had a prefix designating the genus or other taxon designation (done with `addfileprefix.bat`). File names were never more than eight characters in order to be read correctly by the software.

2. The batch file makehnames.bat was run to create an "hnames.dat" files containing the names of each TPR file in the directory.

3. HAngle was run, selecting the TPR files, which each have three header lines. Smoothing for most analyses in this thesis was set at 2. As all digitizations were performed with the anterior of the valve to the left, no reversal should be necessary with these specimens. 12 Fourier scores were retained for all analyses, and all valves were normalized to the second Fourier harmonic.

4. The number of TPR files and UNF files in the directory were compared to determine if any specimens were accidentally dropped during the analysis.

5. The batch file `combine-unf.bat` was run to append the UNF files output by HAngle.

6. 'Fourier Format HAngle.xls' was opened to manipulate the data in the output from `combine-unf.bat`.

   (a) "combine-unf-output.txt" was imported to sheet 'HAngle Input' cell A1.

(b) The last row of the input data was checked to make sure it fit in the established workflow of the spreadsheet.

(c) The final line of data in the 'HAngle Output' sheet was typically incomplete. The formulas in the cells from the line above were copied to make sure the missing data is moved from the first sheet to the second.

(d) 'HAngle Output' data was copied and pasted as values into a new workbook and sorted by column Specimen.

(e) Numerical data was copied to another sheet and the first sheet was deleted (this reduced the file size and allowed for use of the file in R).

(f) The file was saved with a standard name of 'FourierCoeffs' followed by the number of the experiment (e.g., 'FourierCoeffs36').

(g) 'Fourier Format HAngle.xls' was deleted to save space.

7. In Excel, column widths were manually increased so that Fourier Scores do not appear in non-decimal notation (e.g., "1E-08"). This prevented extra characters from appearing in the file.

8. FourierCoeffs file was saved as a tab-delimited *.txt file.

9. Data from the FourierCoeffs file were copied into PAST and saved as a PAST *.dat file.

10. In PAST, specimens were arranged by taxon and rows colored to distinguish these genera.

11. The Elliptical Fourier Analysis data were then able to be used in other analyses in both PAST and the R package.

B.2 Minimizing Digitization Error

B.2.1 Based on Sum of Variance

1. Steps 1-11 from Appendix B.1 were performed.

2. R function `split.file()` was run multiple times (once per specimen number prefix denoting digitization length) on the "FourierCoeffs" tab-delimited *.txt file. This separated the Fourier scores according to digitization length (each file contained Fourier scores for all digitization repeats of all specimens at one digitization length).

3. R function `split.file()` was run on the output multiple times (once per specimen number). This separated the Fourier scores according to specimen (each file contained Fourier scores for five repeats of one specimen at one digitization length).

4. R function `confplotvbar()` was run on this `split.file()` output (once per specimen) to produce one bar plot with confidence intervals representing the change in morphological variation of the specimen as a function of the digitization length, as well as one output file summarizing the data making up the plot (data for one specimen at six digitization lengths).

5. R function `justplotbarsmulti()` was used to produce the same plots as `confplotvbar()` (using `confplotvbar()` output files as a source) with a standard y-axis scale across multiple plots.

6. R functions `mv.table()`, `mvlowcon()`, and `mvupcon()` were run (using `confplotvbar()` files as a source) to arrange data for combined bar plots.

7. R function `multibar()` was used to produce combined bar plots.

B.2.2 Based on Within-Group Dispersion

1. Steps 1-3 from Appendix B.1 were performed.

2. R function `euc.list()` was run on the resulting FC files, producing one *.euc file for every FC file input.

3. R function `euc.group()` was run on the euc.list() output, grouping according to digitization length value (e.g., pattern="5cm" etc.).

4. R function `meanxbar()` was run on `euc.group()` output files just produced.

5. Output of `meanxbar()` was copied to PAST. A repeated-measures ANOVA with paired post hoc least significant difference tests was performed to determine if the within-group disparity of each genus at successive amounts of smoothing was significantly different.

6. R functions `wd.upcon()` and `wd.lowcon()` were run (using `euc.group()` output files as input) to arrange data for combined bar plots.

7. R functions `multibar.wd()` and `unibar.wd()` were used to produce combined bar plots.

B.3 Optimizing Amount of Smoothing During EFA

B.3.1 Based on Morphological Variation

1. Steps 1-11 from Appendix B.1 were performed, changing the smoothing value each time. This resulted in 19 directories, containing the EFA output for smoothing values from 2 to 20.

2. R function `split.file()` was run multiple times (once per specimen number prefix denoting genus) on the FourierCoeffs tab-delimited *.txt file. This separated the Fourier scores according to genus. Some genera comprised a combina-

tion of specimens that were digitized at a standard length (25 cm) and specimens that were digitized at an arbitrary length. As the R function `split.file()` separated these groups, they were recombined by hand. The results were nine FC files (one for each genus) at each level of smoothing (2-20), or a total of 171 FC files.

3. R function `confplotvbar()` was run on this `split.file()` output (once per specimen) to produce one bar plot with confidence intervals representing the change in morphological variation of the specimen as a function of the smoothing, as well as one output file summarizing the data making up the plot.

4. R function `justplotbarsmulti()` was used to produce the same plots as `confplotvbar()` (using `confplotvbar()` output files as a source) with a standard y-axis scale across multiple plots.

5. R functions `mv.table()`, `mvlowcon()`, and `mvupcon()` were run (using `confplotvbar()` output files as a source) to arrange data for combined bar plots.

6. Resulting "mvtable.txt" file (representing all of the MV values for each specimen across smoothing) was copied to PAST and transposed. A one-way ANOVA with paired post hoc least significant different tests was performed to determine if the morphological variation of each genus at successive amounts of smoothing was significantly different.

7. R function `multibar()` was used to produce combined bar plots.

B.3.2 Based on Within-Group Dispersion

1. Steps 1-2 from Appendix B.3.1 were performed.

2. R function `euc.list()` was run on the resulting FC files, producing one *.euc file for every FC file input (171 files total).

3. R function `euc.group()` was run on the `euc.list()` output, grouping according to smoothing value (e.g., pattern="smooth02-" etc.).

4. R function `meanxbar()` was run on `euc.group()` output files just produced.

5. Output of `meanxbar()` was copied to PAST. A one-way ANOVA with paird post hoc least significant difference tests was performed to determine if the within-group disparity of each genus at successive amounts of smoothing was significantly different.

6. R functions `wd.upcon()` and `wd.lowcon()` were run (using `euc.group()` output files as input) to arrange data for combined bar plots.

7. R functions `multibar.wd()` and `unibar.wd()` were used to produce combined bar plots.

APPENDIX C

CUSTOM PROGRAMS AND SCRIPTS

C.1 Introduction

A number of small computer programs and scripts were written for this project in order to rapidly modify multiple files and rearrange data for input into various programs. Each program includes a summary (describing use, input and output), the actual program listing (code), a description of user input, requirements (input file formatting and/or necessary programs called from within), and output formatting. Each script includes a summary, the program listing, user input (in the form of variables), requirements, and output formatting. R script listings include the function definition, making them able to be copied and pasted into R. Comments are included where deemed appropriate and where possible.

The batch files and R functions included here are provided to show the methods used to produce the data used for this project. They are not designed for use on other systems without modification and, if used, should be checked against other methods to ensure that they have provided the expected and proper output and format. Even though explanations are provided, do not attempt to use these programs and functions unless you are sure of what they are doing.

C.2 DOS Batch Files

All programs written for this thesis are batch files (*.bat) for use in MS DOS or the command line in Microsoft® Windows® (all versions).

C.2.1 `addfileprefix.bat`

C.2.1.1 Summary

This batch file adds a prefix to all file names with a certain extension (supplied by the user) in the working directory. This program was used to differentiate between multiple files from multiple analyses with the same filename.

C.2.1.2 Requirements

`addfileprefix.bat` must be located in the same directory as the files to be modified.

C.2.1.3 User Input

User is prompted for the three-letter extension of the files that will be modified, and for the prefix (which can be multiple characters) to be applied to the filenames.

C.2.1.4 Output

Prefix will be added to filenames of the type supplied by the user in the working directory.

C.2.1.5 Listing

```
@echo off
set dir=%CD%
set /p type=File type (extension):
set /p prefix=Prefix wanted:
md %prefix%prefix

for /r %%K in (*.%type%) do call :rename "%%K"
goto:eof

:rename
echo Adding prefix to %name%
set name=%~n1
copy %name%.%type% "%dir%\%prefix%prefix\%prefix%%name%.%type%"
goto:eof
```

C.2.2 `combine-fit.bat`

C.2.2.1 Summary

`combine-fit.bat` appends all files with a *.fit extension in the working directory. This program was used to combine output of HMatch.

C.2.2.2 Requirements

`combine-fit.bat` must be located in the same directory as the files to be appended.

C.2.2.3 User Input

None required.

C.2.2.4 Output

A file named "combine-fit-output.txt" with the appended *.fit files within. The source files are not modified.

C.2.2.5 Listing

```
dir *.fit /b > combine-fit-filelist.txt
FOR %%1 in (*.fit) do type %%1 >> combine-fit-output.txt
```

C.2.3 `combine-unf.bat`

C.2.3.1 Summary

`combine-unf.bat` appends all files with a *.unf extension in the working directory. This program was used to combine output of HAngle.

C.2.3.2 Requirements

`combine-unf.bat` must be located in the same directory as the files to be appended.

C.2.3.3 User Input

None required.

C.2.3.4 Output

A file named "combine-unf-output.txt," containing the appended *.unf files, is produced. The source files are not modified.

C.2.3.5 Listing

```
dir *.unf /b > combine-unf-filelist.txt
FOR %%1 in (*.unf) do type %%1 >> combine-unf-output.txt
```

C.2.4 `makehnames.bat`

C.2.4.1 Summary

`makehnames.bat` produces a file named "hnames.dat" that lists all files in the working directory with a *.tpr extension. "hnames.dat" is used by HAngle and HMatch.

C.2.4.2 Requirements

`makehnames.bat` must be located in the same directory as the files to be listed.

C.2.4.3 User Input

None required.

C.2.4.4 Output

A file named "hnames.dat," containing a list of *.tpr files in the working directory, is produced. The source files are not modified.

C.2.4.5 Listing

```
set OLDDIR=%CD%

::Loop through all files with .tpr extension
for %%K in (*.tpr) do call :hnames "%%K"
goto:eof

:hnames
set name=%~n1
echo %name% >> hnames.dat
```

```
goto:eof

:eof
exit
```

C.3 R Scripts

C.3.1 `cbind.all`

C.3.1.1 Summary

Written by Dorai-Raj (2005). Does the same thing as `cbind()`, but will join lists of unequal lengths. Function is used in `euc.group()`.

C.3.1.2 Requirements

None.

C.3.1.3 User Input

A list of column vectors (which can be different lengths) is required as arguments.

C.3.1.4 Output

Output is a data matrix containing the columns entered, arranged side by side.

C.3.1.5 Listing

```
cbind.all <- function(..., fill.with = NA) {
   args <- list(...)
   len <- sapply(args, NROW)
   if(diff(rng <- range(len)) > 0) {
     maxlen <- rng[2]
     pad <- function(x, n) c(x, rep(fill.with, n))
     for(j in seq(along = args)) {
       if(maxlen == len[j]) next
       if(is.data.frame(args[[j]])) {
         args[[j]] <- lapply(args[[j]], pad, maxlen - len[j])
         args[[j]] <- as.data.frame(args[[j]])
       } else if(is.matrix(args[[j]])) {
         args[[j]] <- apply(args[[j]], 2, pad, maxlen - len[j])
       } else if(is.vector(args[[j]])) {
         args[[j]] <- pad(args[[j]], maxlen - len[j])
       } else {
         stop("... must only contain data.frames or arrays.")
```

```
      }
    }
  }
  do.call("cbind", args)
}
```

C.3.2 `confplotv()` / `confplotd()`

C.3.2.1 Summary

Bootstraps the sum of variance of a given group of specimens (for `confplotd()`, a group of group centroids) that have had EFA performed with different amounts of smoothing. Outputs a file with smoothing, sum of variance (= within-group dispersion for `confplotv()`, = morphological disparity for `confplotv()`), upper confidence interval, and lower confidence interval. Plots sum of variance (with confidence intervals) as a function of smoothing. The bootstrap portion of this program was written by H. Scholz (MfN, pers. comm., 2008). The only difference between these two scripts is the use of the label "Variation" vs. "Disparity" and the lack of row labels in the input files for `confplotd()`. The latter can be produced by replacing

```
x <- read.table(i,header=TRUE,row.names=1);
```

with

```
x <- read.table(i,header=TRUE);
```

below. The listing for `confplotv()` is shown.

C.3.2.2 Requirements

`plotrix` library and `split.file()` output files.

C.3.2.3 User Input

path Path to directory holding files of interest (in `split.file()` format).

pattern Pattern to identify files to use in analysis.

vardis Variation or Disparity. Used for y-axis label. Default = "Variation."

smoo Smoothing values used during EFA. Default = c(1:20), meaning smoothing values 1-20 were used. c(1,5,10,15,20) would mean only smoothing values of 1, 5, 10, 15 and 20 were input (in that order).

reps Number of permutations to use for bootstrap.

probs A `c()` containing the boundaries of the calculated confidence interval. The default `c(‘‘0.025’’,‘‘0.975’’)` is the 95% confidence interval from 2.5% to 97.5%.

ext Extension to add to file root to differentiate new file.

C.3.2.4 Output

One file with one row per smoothing amount during EFA. Each row contains smoothing, sum of variance, upper confidence limit and lower confidence limit. One plot to screen and PDF file of sum of variance (with confidence intervals) as a function of smoothing.

C.3.2.5 Listing

```
confplotv <- function(path, pattern, vardis="Variation",
+smoo=c(1:20), reps=1000, ext="-confplotv.txt", probs=c(0.025,0.975))
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for(i in files)
#Loop through files and assign variable names
{
x <- read.table(i,header=TRUE,row.names=1);
assign(i, x);
}
for(j in 1:length(files))
{
SOV <- get(files[j]);
bootSOV <- numeric(reps);      #creates place for bootstrap values
for (k in 1:reps) {bootSOV[k] <- sum(diag(var(sample(SOV,
+replace=TRUE))))};
conf<-quantile(bootSOV,probs=probs);
```

```
toplot<-paste(smoo[j],mean(bootSOV),conf[2],conf[1]);
write.table(toplot,file=paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),quote=FALSE,append=TRUE,
+row.names=FALSE,
+col.names=FALSE)
}
tableback<-read.table(file=paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""));
x<-unlist(tableback[1]);
y<-unlist(tableback[2]);
cui<-unlist(tableback[3]);
cli<-unlist(tableback[4]);
plotCI(x,y,ui=cui,li=cli,xlab=c("Smoothing"),ylab=vardis,
+ylim=c(0,0.025));
pdf(paste(substring(files[1],1,nchar(files[1])-4), ext,
+c(".pdf"), sep=""));
plotCI(x,y,ui=cui,li=cli,xlab=c("Smoothing"),ylab=vardis,
+ylim=c(0,0.025));
dev.off();
}
```

C.3.3 `confplotvbar()`

C.3.3.1 Summary

Bootstraps the sum of variance of a given group of specimens on which EFA has been performed. Outputs a file with smoothing, sum of variance, upper confidence limit, and lower confidence limit. These are absolute values (confidence limits). Plots sum of variance (with confidence intervals) as vertical bars. The bootstrap portion of this program was written by H. Scholz (MfN, pers. comm., 2008).

C.3.3.2 Requirements

`gregmisc package`. `split.file()` output files.

C.3.3.3 User Input

path Path to directory holding files of interest (in `split.file()` format).

pattern Pattern to identify files to use in analysis.

vardis Variation or Disparity. Used for y-axis label. Default = "Variation."

smoo Smoothing values used during EFA. Default = c(1:20), meaning smoothing values 1-20 were used. c(1,5,10,15,20) would mean only smoothing values of 1, 5, 10, 15 and 20 were input (in that order).

xlabel Label for x-axis.

barnames Used to number genera along x-axis. Default numbers bars alphabetically by filename.

reps Number of permutations to use for bootstrap.

probs A `c()` containing the boundaries of the calculated confidence interval. The default `c(''0.025'',''0.975'')` is the 95% confidence interval from 2.5% to 97.5%.

ext Extension to add to file root to differentiate new file.

C.3.3.4 Output

One file with number of input file (1 through 20 maximum), sum of variance, upper confidence limit, lower confidence limit and input file path. One plot to screen and one PDF file (Figure C.1) of sum of variance (with confidence intervals) as vertical bars. Names of output files are based on the first input filename alphabetically, with ext added (default "-confplotvbar.txt" and "-confplotvbar.txt.pdf").

C.3.3.5 Listing

```
confplotvbar<-function(path, pattern, vardis="Variation",
+barnames=c(1:20),xlabel="Genus", reps=1000,
+ext="-confplotvbar.txt", probs=c(0.025,0.975))
{
require(gplots);
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for(i in files)
#Loop through files and assign variable names [This is not necessary]
```

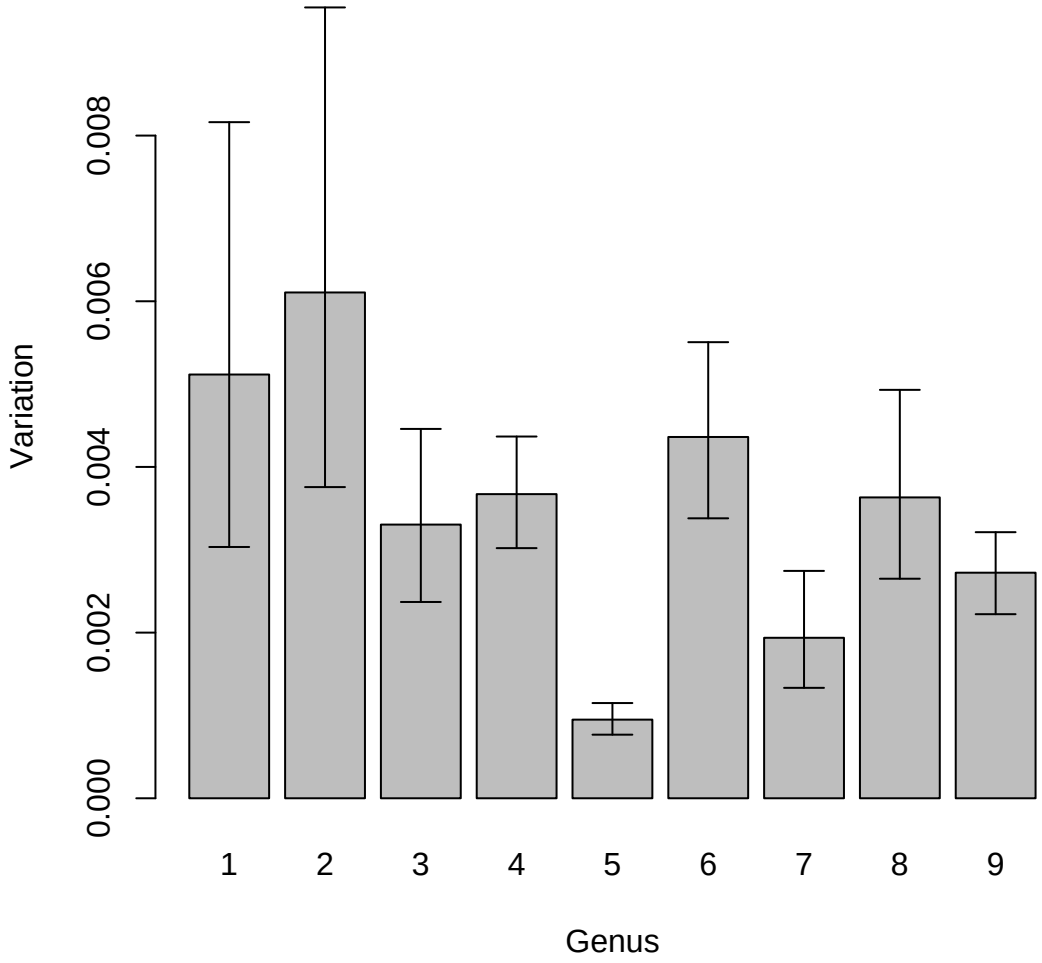

Figure C.1: Example output from `confplotvbar()`.

```
{
x <- read.table(i,header=TRUE);
assign(i, x);
}
for(j in 1:length(files))
{
SOV <- get(files[j]);
bootSOV <- numeric(reps);
 #creates place for bootstrap values
for (k in 1:reps) {bootSOV[k] <- sum(diag(var(sample(SOV,
```

```
+ replace=TRUE))))};
conf<-quantile(bootSOV,probs=probs);
toplot<-paste(barnames[j],mean(bootSOV),
+conf[2],conf[1],files[j]);
write.table(toplot,file=paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),quote=FALSE,append=TRUE,
+row.names=FALSE,col.names=FALSE)
}
tableback<-read.table(file=paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""));
x<-unlist(tableback[1]);
y<-unlist(tableback[2]);
cui<-unlist(tableback[3]);
cli<-unlist(tableback[4]);
barplot2(height=y,names.arg=x,xlab=xlabel,
+ylab=vardis,ci.l=cli,ci.u=cui,plot.ci=TRUE);
pdf(paste(substring(files[1],1,nchar(files[1])-4),
+ext, c(".pdf"), sep=""));
barplot2(height=y,names.arg=x,xlab=xlabel,
+ylab=vardis,ci.l=cli,ci.u=cui,plot.ci=TRUE);
dev.off();
file.rename(paste(substring(files[1],1,nchar(files[1])-4),
+ext, sep=""),paste(path,"/confplotvbar-",pattern,".txt",sep=""));
file.rename(paste(substring(files[1],1,nchar(files[1])-4),
+ext, c(".pdf"), sep=""),paste(path,"/confplotvbar-",
+pattern,".pdf",sep=""));
system("say All done!");
}
```

C.3.4 `euc.group()`

C.3.4.1 Summary

Takes a set of Euclidean distance files (usually output from `euc.list()` and joins them according to name matched to pattern. Results in a summary file arranged with Euclidean distance files by column. First row will contain mean of the column, subsequent rows contain Euclidean distance measurements. To later compare across methodological trials, pattern should match the variable being tested (e.g., digitization size "25 cm" will result in one column per specimen at a 25 cm digitization length). To later compare variables across individual specimens, pattern should

match the specimen identification number (e.g., “S2919” will result in one column per methodological trial).

C.3.4.2 Requirements

Function `cbind.all()`. Single-column input files without column or row names.

C.3.4.3 User Input

path Source path (where to look for files), in quotation marks.

pattern Pattern to match in file name, in quotation marks. “2780”, “.txt”, etc.

len Number of rows in the files in question (number of rows must be equal between files).

ext Extension of output files (will replace current extension, e.g., “S2919.txt” to “S2919.euc”). Can be more than four characters (e.g., “S2919.txt” to “S2919-eucgroup.euc”)

C.3.4.4 Output

One file named for first file in directory that matches pattern, with added **ext**. One column per input file, in alphabetical order of input files. First row is mean of each column, successive rows Euclidean distances.

C.3.4.5 Listing

```
euc.group <- function(path,pattern,ext="-eucgroup.euc",len=11)
 {
 file.list <- list.files(path=path,pattern=pattern,full.names=TRUE);
#List files that match pattern.
 for(i in file.list)
 #Loop through files and assign variable names
  {
                x <- read.table(i);
                assign(i, x);
  }
 output <- cbind.all(get(file.list[1]));  #Initiate output
```

```
for (j in 2:length(file.list))
#Dump everything into that array
 {
 output <- cbind.all(output, get(file.list[j]));
 }
write(t(as.matrix(output)),file=paste(substring(file.list[1],1,
+nchar(file.list[1])-4), ext, sep=""),
+ncolumns=dim(output)[2],sep="\t");
return(output);
}
```

C.3.5 `euc.list()`

C.3.5.1 Summary

Takes a set of Fourier Coefficient files with matching filenames, outputs a new file (one for each input file) listing [line 1] mean Euclidean distance between specimens and [subsequent lines] paired Euclidean distances between all specimens in each file.

C.3.5.2 Requirements

R libraries `Biobase` and `bioDist`; load with `library()`. Input files with column (Fourier Coefficients) and row (Specimen identification numbers) names.

C.3.5.3 User Input

path Path to directory, in quotation marks and no trailing /. "∼/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

folder Name of output destination folder (will be created).

ext Extension of output files (will replace current extension, e.g., "S2919.txt" to "S2919.euc"). Can be more than four characters (e.g., "S2919.txt" to "S2919-eucgroup.euc")

C.3.5.4 Output

Files named with new **ext**, one per input file. First row with mean of Euclidean distances, succeeding rows with Euclidean distances themselves between all pairs of

specimens.

C.3.5.5 Listing

```
euc.list <- function(path,pattern,folder="euc",ext=".euc")
{
files <- list.files(path=path,pattern=pattern,
+full.name=TRUE)
#Get original FC files.
dir.create(paste(path,c("/"),folder,sep=""),
+showWarnings=FALSE);
#Create folder for output files
for(i in 1:length(files))
{
x <- read.table(files[i], header=TRUE, row.names=1);
spec <- numeric();
specimen.euc<-c(euc(as.matrix(x)));
spec[1]<-mean(specimen.euc);
for (j in 1:length(specimen.euc))
{
spec[j+1]<-specimen.euc[j];
#print(specimen.euc[j]);
}
write(spec, paste(substring(files[i],1,
+nchar(files[i])-4),ext, sep=""),ncolumns=1);
dsrc <- paste(substring(files[i],1,nchar(files[i])-4),
+ext, sep="");
ddest <- paste(path,c("/"),folder,c("/"),sep="");
file.copy(dsrc,ddest);
file.remove(dsrc);
}
}
```

C.3.6 `euc.listPLM()`

C.3.6.1 Summary

Takes a set of IMP files with matching filenames, outputs a new file (one for each input file) listing [line 1] mean Euclidean distance between specimens and [subsequent lines] paired Euclidean distances between all specimens in each file. This is different than `euc.list()` because it does not expect a header line in the input file.

C.3.6.2 Requirements

R libraries `Biobase` and `bioDist`; load with `library()`. Input files with column (Fourier Coefficients) and row (Specimen identification numbers) names.

C.3.6.3 User Input

path Path to directory, in quotation marks and no trailing /. "∼/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

folder Name of output destination folder (will be created).

ext Extension of output files (will replace current extension, e.g., "S2919.txt" to "S2919.euc"). Can be more than four characters (e.g., "S2919.txt" to "S2919-eucgroup.euc")

C.3.6.4 Output

Files named with new **ext**, one per input file. First row with mean of Euclidean distances, succeeding rows with Euclidean distances themselves between all pairs of specimens.

C.3.6.5 Listing

```
euc.listPLM <- function(path,pattern,folder="euc",ext=".euc")
{
files <- list.files(path=path,pattern=pattern,full.name=TRUE)
#Get original FC files.
dir.create(paste(path,c("/"),folder,sep=""),showWarnings=FALSE);
#Create folder for output files
for(i in 1:length(files))
{
x <- read.table(files[i], header=FALSE);
spec <- numeric();
specimen.euc<-c(euc(as.matrix(x)));
spec[1]<-mean(specimen.euc);
for (j in 1:length(specimen.euc))
{
```

```
spec[j+1]<-specimen.euc[j];
#print(specimen.euc[j]);
}
write(spec, paste(substring(files[i],1,
+nchar(files[i])-4),ext, sep=""),ncolumns=1);
dsrc <- paste(substring(files[i],1,nchar(files[i])-4),
+ext,sep="");
ddest <- paste(path,c("/"),folder,c("/"),sep="");
file.copy(dsrc,ddest);
file.remove(dsrc);
}
}
```

C.3.7 `hcurveplot()`

C.3.7.1 Summary

Produces a series of PDF files containing specimen outlines as reconstructed by HCurve.

C.3.7.2 Requirements

At least one output file from HCurve, with the extension *.pos, is required.

C.3.7.3 User Input

directory Path to directory, in quotation marks and no trailing /. "~/Desktop/folder"

pattern File extension to match, in quotation marks. Default is "POS".

header_lines Number of header lines in input files before coordinate pairs begin. Default is 1.

type Type of input file. Default is "pos", other options are possible.

C.3.7.4 Output

Output is a series of PDF files, each containing an outline from a single input file, with equal horizontal and vertical scales. No standardization is supplied by this function.

C.3.7.5 Listing

```
hcurveplot<-function(directory,pattern=".POS",header_lines=1,
+type="pos")
{
files<-list.files(path=directory,pattern=pattern,full.names=TRUE);
for (i in 1:length(files))
     {
      points<-read.table(files[i],sep="",skip=header_lines);
     if (type=="pos")
        {
        points<-cbind(points[2],points[1]);
        }
      pdf(paste(substring(files[i],1,nchar(files[i])-4),
+c(".pdf"), sep=""));
   plot(points,type="l",asp=1,main=files[i],xlab="",ylab="",
+xaxt="n",yaxt="n");
   dev.off();
   }
 }
```

C.3.8 `hcurveplotover()`

C.3.8.1 Summary

This function is similar to `hcurveplot()`, but instead of outputting separate files for each input file, all outlines are printed on top of one another to visually show the amount of variation in the group.

C.3.8.2 Requirements

At least one output file from HCurve, with the extension *.pos, is required.

C.3.8.3 User Input

directory Path to directory, in quotation marks and no trailing /. "∼/Desktop/folder"

pattern File extension to match, in quotation marks. Default is "POS".

header_lines Number of header lines in input files before coordinate pairs begin. Default is 1.

type Type of input file. Default is "pos", other options are possible.

C.3.8.4 Output

Output is a single PDF file containing overlain outlines of all input files, with equal horizontal and vertical scales. No standardization is supplied by this function.

C.3.8.5 Listing

```
hcurveplotover<-function(directory,pattern=".POS",
+header_lines=1,type="pos")
{
files<-list.files(path=directory,pattern=pattern,
+full.names=TRUE);
for (i in 1:length(files))
      {
      points<-read.table(files[i],sep="",skip=header_lines);
     if (type=="pos")
        {
        points<-cbind(points[2],points[1]);
        }
par(new=T);   plot(points,type="l",asp=1,main=files[i],xlab="",
+ylab="",xaxt="n",yaxt="n");
   }
 }
```

C.3.9 `justplotbars()`

C.3.9.1 Summary

This function produces a bar plot with confidence intervals (exported as a PDF file), based on data from `confplotvbar()`.

C.3.9.2 Requirements

The R package `gplots` is required. Load with `library()`.

C.3.9.3 User Input

path Path to directory, in quotation marks and no trailing /. "∼/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

vardis Label for y-axis. Default is "Variation."

barnames A `c()` containing names for individual bars. Default is the values 1–20.

ylimit Limit of the y-axis, to be used for standardizing multiple plots.

xlabel Label for x-axis. Default is "Digitization Length."

reps

ext Extension of output files (will replace current extension, e.g., "S2919.txt" to "S2919.euc"). Can be more than four characters (e.g., "S2919.txt" to "S2919-eucgroup.euc")

C.3.9.4 Output

Output is a bar plot in PDF format with confidence intervals.

C.3.9.5 Listing

```
justplotbars<-function(path, pattern, vardis="Variation",
+barnames=c(1:20),ylimit,xlabel="Digitization Length", ext="-justplotbars.txt")
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
tableback<-read.table(file=files[1]);
x<-unlist(tableback[1]);
y<-unlist(tableback[2]);
cui<-unlist(tableback[3]);
cli<-unlist(tableback[4]);
barplot2(height=y,names.arg=x,xlab=xlabel,
+ylab=vardis,ylim=ylimit,ci.l=cli,ci.u=cui,plot.ci=TRUE);
pdf(paste(substring(files[1],1,nchar(files[1])-4),
+ext, c(".pdf"), sep=""));
barplot2(height=y,names.arg=x,xlab=xlabel,
+ylab=vardis,ylim=ylimit,ci.l=cli,ci.u=cui,plot.ci=TRUE);
dev.off();
}
```

C.3.10 `justplotbarsmulti()`

C.3.10.1 Summary

This function works like `justplotbars()`, but produces a clustered bar plot based on multiple input files.

C.3.10.2 Requirements

The R package `gplots` is required. Load with `library()`.

C.3.10.3 User Input

path Path to directory, in quotation marks and no trailing /. "~/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

vardis Label for y-axis. Default is "Variation."

barnames A `c()` containing names for individual bars. Default is the values 1–20.

ylimit Limit of the y-axis, to be used for standardizing multiple plots.

xlabel Label for x-axis. Default is "Digitization Length."

reps

ext Extension of output files (will replace current extension, e.g., "S2919.txt" to "S2919.euc"). Can be more than four characters (e.g., "S2919.txt" to "S2919-eucgroup.euc")

C.3.10.4 Output

Output is a bar plot in PDF format with confidence intervals.

C.3.10.5 Listing

```
justplotbarsmulti<-function(path, pattern, vardis="Variation",
+barnames=c(1:20),ylimit,xlabel="Digitization Length", ext="-justplotbars.txt")
{
```

```
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for (i in 1:length(files)) {
tableback<-read.table(file=files[i]);
x<-unlist(tableback[1]);
y<-unlist(tableback[2]);
cui<-unlist(tableback[3]);
cli<-unlist(tableback[4]);
barplot2(height=y,names.arg=x,xlab=xlabel,
+ylab=vardis,ylim=ylimit,ci.l=cli,ci.u=cui,plot.ci=TRUE);
pdf(paste(substring(files[i],1,nchar(files[1])-4),
+ext, c(".pdf"), sep=""));
barplot2(height=y,names.arg=x,xlab=xlabel,
+ylab=vardis,ylim=ylimit,ci.l=cli,ci.u=cui,plot.ci=TRUE);
dev.off();
}
}
```

C.3.11 `list.centroids()`

C.3.11.1 Summary

Takes output of `split.file()` and produces a list of group centroids based on the Euclidean average of each group/`split.file()` output. Will output header row with Fourier Coefficient labels, followed by one row per group with the Fourier Coefficients of the groups centroid. Output can later be used to calculate morphological disparity between groups with `hbootsumvar()`.

C.3.11.2 Requirements

One or more tab-delimited files with Fourier Coefficients for one group per file and header (FC labels) and row (specimens) names (e.g., the output of `split.file()`).

C.3.11.3 User Input

path Path to directory, in quotation marks and no trailing /. "~/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

ext Extension of output files (will replace current extension, e.g., "S2919.txt" to "S2919.euc"). Can be more than four characters (e.g., "S2919.txt" to "S2919-

eucgroup.euc")

C.3.11.4 Output

One file of Fourier Coefficients arranged with the Euclidean centroid of one group per row. Header row with FC labels.

C.3.11.5 Listing

```
list.centroids <- function(path, pattern, ext=-centroid.txt)
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for(i in files)
 #Loop through files and assign variable names
{
x <- read.table(i,header=TRUE,row.names=1);
assign(i, x);
}
output <- list();
for (j in 1:length(files))      #Dump everything into that array
{
place <- mean(get(files[j]));
output <- rbind(output, place);
}
write.table(output, file=paste(substring(files[1],
+1,nchar(files[1])-4), ext, sep=""), sep="\t",
+row.names=FALSE); return(output);
}
```

C.3.12 `meanxbar()`

C.3.12.1 Summary

Takes a set of "-eucgroup.euc" files with names that match entered pattern and outputs the values in the first line (usually the mean values of the columns) of each file as a separate column. Used for comparing average Euclidean distances (of trace repeats of the same specimen) across methods. Will output one column per methodological trial with one row per specimen.

C.3.12.2 Requirements

At least one output file from `euc.group()` is required.

C.3.12.3 User Input

path Path to directory, in quotation marks and no trailing /. "~/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

ext Extension of output files (will replace current extension, e.g., "S2919.txt" to "S2919.euc"). Can be more than four characters (e.g., "S2919.txt" to "S2919-eucgroup.euc")

C.3.12.4 Output

One file named for first file in directory that matches pattern, with added **ext**. One column per input file, in alphabetical order of input files. Ignore "NA" in first column. Each column represents transposed row of means of columns in one input file.

C.3.12.5 Listing

```
meanxbar <- function(path,pattern,ext="-meanxbar.euc")
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);

for(i in files)
#Loop through files and assign variable names
{
x <- read.table(i);
assign(i, x);
}
output <- array();
for (j in 1:length(files))
#Dump everything into that array
{
place <- get(files[j]);
output <- rbind(output, place[1,]);
}
write(as.matrix(output), file=paste(substring(files[1],
+1,nchar(files[1])-4), ext, sep=""),
+ncolumns=dim(output)[1],sep=\t);
```

```
return(t(output));
}
```

C.3.13 multibar()

C.3.13.1 Summary

Used to make plots of the sum of variance at different digitization lengths. Basis for this function was written by Schwartz (2006).

C.3.13.2 Requirements

The R package `gplots` is required. Load with `library()`. Input is output of `mv.table()`, `mvupcon()` and `mvlowcon()`.

C.3.13.3 User Input

No input. The following lines in the source need to be replaced by the locations of mvtable.txt, mvlowcon.txt, mvupcon.txt, and the output file and name to be created :

“ /Desktop/EFA28-29 MV/Same Scale/mvtable.txt”

“ /Desktop/EFA28-29 MV/Same Scale/mvlowcon.txt”

“ /Desktop/EFA28-29 MV/Same Scale/mvupcon.txt”

“ /Desktop/EFA28-29 MV/Same Scale/+mvplot.pdf”

C.3.13.4 Output

Output is a single PDF file containing a bar plot of sum of variance with error bars across digitization length.

C.3.13.5 Listing

```
multibar<-function() {
hh <- t(read.table("~/Desktop/EFA28-29 MV/Same Scale/mvtable.txt"));
mybarcol <- "gray20";
ci.l <- t(as.matrix(read.table("~/Desktop/EFA28-29
+MV/Same Scale/mvlowcon.txt")));
ci.u <- t(as.matrix(read.table("~/Desktop/EFA28-29
+MV/Same Scale/mvupcon.txt")));
```

```
mp <- barplot2(hh, beside = TRUE,
        ylim = c(0,1.2e-3),
        ylab = "Morphological Variation",
        main = "", font.main = 4,
        sub = "", col.sub = mybarcol,
        cex.names = 1.5, plot.ci = TRUE, ci.l = ci.l, ci.u = ci.u,
        plot.grid = TRUE) ;
mtext(side = 1, at = colMeans(mp), line = 2,
      text = paste("Mean", formatC(colMeans(hh))), col = "red");
box();
pdf("~/Desktop/EFA28-29 MV/Same Scale/+mvplot.pdf",width=11,
+height=8.5,pointsize=12);
mp <- barplot2(hh, beside = TRUE,
        ylim = c(0,1.2e-3),
        ylab = "Morphological Variation",
        main = "", font.main = 4,
        sub = "", col.sub = mybarcol,
        cex.names = 1.5, plot.ci = TRUE, ci.l = ci.l, ci.u = ci.u,
        plot.grid = TRUE) ;
mtext(side = 1, at = colMeans(mp), line = 2,
      text = paste("Mean", formatC(colMeans(hh))), col = "red");
dev.off();
}
```

C.3.14 `multibar.wd()`

C.3.14.1 Summary

This script is used to produce the WD plot of synthetic valves at different digitization lengths. Input is output of `meanxbar()`, output of `wd.upcon()` and output of `wd.lowcon()`.

C.3.14.2 Requirements

The R package `gplots` is required. Load with `library()`. Input is output of `meanxbar()`, output of `wd.upcon()` and output of `wd.lowcon()`.

C.3.14.3 User Input

The source below needs to be edited to reflect the proper filenames for meanxbar.euc, wd_lowcon.txt, and wd_upcon.txt. Depending on the number of digitization lengths tested, the variable "section.names" needs to be adjusted to have the correct number

of names (for this project, 6).

C.3.14.4 Output

Output is a single PDF file containing a bar plot of within-group dispersion with error bars across digitization length.

C.3.14.5 Listing

```
multibar.wd<-function() {
hh <- as.matrix(read.table("~/Desktop/39EFA/euc/FourierCoeffs39-
+smooth02-Anodonta-bysmoothing-eucgroup-meanxbar.euc",
+na.strings="NA")); #This needs to be a variable.
hh <- hh[,-1];
mybarcol <- "gray20";
ci.l <- as.matrix(read.table("~/Desktop/39EFA/euc/wd_lowcon.txt"));
ci.u <- as.matrix(read.table("~/Desktop/39EFA/euc/wd_upcon.txt"));
section.names=c(1:6);
mp <- barplot2(hh, beside = TRUE, names.arg=section.names,
        #ylim = c(0,.05),
        ylab = "Within-Group Dispersion",
        main = "", font.main = 4,
        sub = "", col.sub = mybarcol,
        cex.names = 1.5, plot.ci = TRUE, ci.l = ci.l, ci.u = ci.u,
        plot.grid = TRUE) ;
mtext(side = 1, at = colMeans(mp), line = 2,
      text = paste("Mean", formatC(colMeans(hh))), col = "red");
box();
pdf("~/Desktop/wdplot.pdf",width=11,height=8.5,pointsize=12);
mp <- barplot2(hh, beside = TRUE, names.arg=section.names,
        #ylim = c(0,.05),
        ylab = "Within-Group Dispersion",
        main = "", font.main = 4,
        sub = "", col.sub = mybarcol,
        cex.names = 1.5, plot.ci = TRUE, ci.l = ci.l, ci.u = ci.u,
        plot.grid = TRUE) ;
mtext(side = 1, at = colMeans(mp), line = 2,
      text = paste("Mean", formatC(colMeans(hh))), col = "red");
dev.off();
}
```

C.3.15 `mvlowcon()`

C.3.15.1 Summary

This script is used to make a table of lower confidence limit values that is congruent to the upper confidence limit table and the ΣV value table for digitization length.

C.3.15.2 Requirements

Input is a series of output files from `confplotvbar()`.

C.3.15.3 User Input

path Path to source directory, in quotation marks, with no trailing slash.

pattern Pattern in filename to be used to select files.

ext Extension to be added to output file. Default is "-mvlowcon.txt."

C.3.15.4 Output

Output is one file per input file containing a table of lower confidence limit values that us congruent to the input files.

C.3.15.5 Listing

```
mvlowcon<-function(path, pattern, ext="-mvlowcon.txt")
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for (i in 1:length(files)) {
tableback<-read.table(file=files[i]);
if(file.exists(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""))){
oldfile<-read.table(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""));
write.table(cbind(oldfile,unlist(tableback[4])),
+paste(substring(files[1],1,nchar(files[1])-4), ext, sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE)
}
else {write.table(tableback[4],paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),col.names=FALSE,
+row.names=FALSE,quote=FALSE)}
}
```

```
file.rename(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),paste(path,"/mvlowcon.txt",sep=""))
}
mv.table<-function(path, pattern,ext="-mvtable.txt")
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for (i in 1:length(files)) {
tableback<-read.table(file=files[i]);
if(file.exists(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""))){
oldfile<-read.table(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""));
write.table(cbind(oldfile,unlist(tableback[2])),
+paste(substring(files[1],1,nchar(files[1])-4), ext, sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE)
}
else {write.table(tableback[2],paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),
+col.names=FALSE,row.names=FALSE,
+quote=FALSE)}
}
file.rename(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),paste(path,"/mvtable.txt",sep=""))
}
```

C.3.16 `mvupcon()`

C.3.16.1 Summary

Makes a table of upper confidence limit values that is congruent to the lower confidence limit table and the ΣV value table for digitization length.

C.3.16.2 Requirements

Input is a series of output files from `confplotvbar()`.

C.3.16.3 User Input

path Path to source directory, in quotation marks, with no trailing slash.

pattern Pattern in filename to be used to select files.

ext Extension to be added to output file. Default is "-mvupcon.txt."

C.3.16.4 Output

Output is one file per input file containing a table of upper confidence limit values that us congruent to the input files.

C.3.16.5 Listing

```
mvupcon<-function(path, pattern,ext="-mvupcon.txt")
{
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for (i in 1:length(files)) {
tableback<-read.table(file=files[i]);
if(file.exists(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""))){
oldfile<-read.table(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""));
write.table(cbind(oldfile,unlist(tableback[3])),
+paste(substring(files[1],1,nchar(files[1])-4), ext,sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE)
}
else {write.table(tableback[3],paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),col.names=FALSE,
+row.names=FALSE,quote=FALSE)}
}
file.rename(paste(substring(files[1],1,
+nchar(files[1])-4), ext, sep=""),paste(path,"/mvupcon.txt",sep=""))
}
```

C.3.17 `shapiro.wilk()`

C.3.17.1 Summary

Performs a Shapiro-Wilk test on all columns of a `euc.group()` file.

C.3.17.2 Requirements

Output file from `euc.group()`.

C.3.17.3 User Input

file Path to file, in quotation marks.

C.3.17.4 Output

Output is a file containing a list of Shapiro-Wilk test results, ordered according to column order in input file.

C.3.17.5 Listing

```
shapiro.wilk <- function(file)
{
a<-read.table(file,header=FALSE);
for (i in 1:length(a)) {b<-shapiro.test(a[,i]);
print(paste(i+1,": ",b$statistic,b$p.value,sep=" "),quote=FALSE)}
}
```

C.3.18 `split.file()`

C.3.18.1 Summary

Splits tab-delimited Excel file into multiple files based on the row (specimen) name. Make sure all values are visible in Excel before exporting to tab-delimited (i.e., "-.00000001" rather than "-1E-9"), otherwise extra slashes (/) and equal signs (=) will be added before "E" and a negative number ("-1E-9"), respectively.

C.3.18.2 Requirements

One or more tab-delimited files with Fourier Coefficients for one group per file and header (FC labels) and row (specimens) names.

C.3.18.3 User Input

path Path to directory, in quotation marks and no trailing /. "~/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

C.3.18.4 Output

Fourier Coefficients as in original file, one output file per group of specimens matching pattern (header and row names retained).

C.3.18.5 Listing

```
split.file <- function(path,pattern)
{
for (i in 1:length(pattern))
{
input.file <- read.table(path,header=TRUE,row.names=1);
nm <- rownames(input.file);
 #Gets the list of row names.
matching.rows <- nm %in% grep(pattern[i], nm, value=TRUE);
 #greps those row names that match pattern
output <- subset(input.file,matching.rows);
 #Get the subset of  the rows that match
write.table(output, file=paste(substring(path,1,nchar(path)-4), -,
 +pattern[i],.txt, sep=""),sep=\t);
}
}
```

C.3.19 `unibar.wd()`

C.3.19.1 Summary

This function is similar to multibar, but plots one at a time.

C.3.19.2 Requirements

The R package `gplots` is required. Load with `library()`. Input is output of `mv.table()`, `mvupcon()` and `mvlowcon()`.

C.3.19.3 User Input

meanxbar.in Location of `meanxbar()` output.

wd_lowcon.in Location of `wd_lowcon()` output.

wd_upcon.in Location of `wd_upcon()` output.

column Which column of input files to use.

file Location and name of output file.

C.3.19.4 Output

A bar plot with confidence intervals is produced based on one column of the input files.

C.3.19.5 Listing

```
unibar.wd<-function(meanxbar.in,wd_lowcon.in,wd_upcon.in,
+column=1,filename="~/Desktop/wdplot.pdf") {
require(gplots);
hh <- as.matrix(read.table(meanxbar.in,na.strings="NA"));
hh <- hh[,-1];
mybarcol <- "gray20";
ci.l <- as.matrix(read.table(wd_lowcon.in));
ci.u <- as.matrix(read.table(wd_upcon.in));
section.names=c(2:20);
mp <- barplot2(hh[column,], beside = TRUE, names.arg=section.names,
        #ylim = c(0,ylimit),
        ylab = "Within-Group Dispersion",
        main = "", font.main = 4,
        sub = "", col.sub = mybarcol,
        cex.names = 1.5, plot.ci = TRUE,
+ci.l = ci.l[column,], ci.u = ci.u[column,],
        plot.grid = TRUE) ;
#mtext(side = 1, at = colMeans(mp), line = 2, text =
+paste("Mean", formatC(colMeans(hh))), col = "red");
box();
pdf(filename,width=11,height=8.5,pointsize=12);
mp <- barplot2(hh[column,], beside = TRUE, names.arg=section.names,
        #ylim = c(0,ylimit),
        ylab = "Within-Group Dispersion",
        main = "", font.main = 4,
        sub = "", col.sub = mybarcol,
        cex.names = 1.5, plot.ci = TRUE,
+ci.l = ci.l[column,], ci.u = ci.u[column,],
        plot.grid = TRUE) ;
#mtext(side = 1, at = colMeans(mp), line = 2, text =
+paste("Mean", formatC(colMeans(hh))), col = "red");
dev.off();
}
```

C.3.20 `wd.lowcon()`

C.3.20.1 Summary

Takes input of `euc.group()` files and calculates the lower limit of the confidence interval based on the t distribution. N is automatically calculated.

C.3.20.2 Requirements

Requires a `euc.group()` output file.

C.3.20.3 User Input

path Path to directory, in quotation marks and no trailing /. "∼/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

C.3.20.4 Output

Output is "wd_lowcon.txt" with lower confidence limits.

C.3.20.5 Listing

```
wd.lowcon<-function(path,pattern,n=9) {
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for (i in 1:length(files)) {
temp<-read.table(files[i],skip=1);
    for (j in 1:length(temp))
    {
    n<-length(temp[j][,1]);
    xbar<-mean(temp[j],na.rm=TRUE);
    s<-sd(temp[j],na.rm=TRUE);
    error<-qt(0.975,df=n-1)*s/sqrt(n);

    write.table(xbar-error,paste(path,"/wd_lowcon-temp.txt",sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE,append=TRUE);

}
    tableback<-read.table(paste(path,"/wd_lowcon-temp.txt",sep=""));
    file.remove(paste(path,"/wd_lowcon-temp.txt",sep=""));
    if(file.exists(paste(path,"/wd_lowcon.txt",sep="")))
        {
        oldfile<-read.table(paste(path,"/wd_lowcon.txt",sep=""));
        write.table(cbind(oldfile,tableback),
```

```
+paste(path,"/wd_lowcon.txt",sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE);
        }

    else write.table(tableback,paste(path,"/wd_lowcon.txt",sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE);
    }
}
```

C.3.21 wd.upcon()

C.3.21.1 Summary

Takes input of `euc.group()` files and calculates the upper limit of the confidence interval based on the t distribution. N is automatically calculated.

C.3.21.2 Requirements

Requires a `euc.group()` output file.

C.3.21.3 User Input

path Path to directory, in quotation marks and no trailing /. "∼/Desktop/folder"

pattern Pattern to match in file name, in quotation marks. "2780", ".txt", etc.

C.3.21.4 Output

Output is "wd_upcon.txt" with upper confidence limits.

C.3.21.5 Listing

```
wd.upcon<-function(path,pattern) {
files<-list.files(path=path,pattern=pattern,full.names=TRUE);
for (i in 1:length(files)) {
temp<-read.table(files[i],skip=1);
    for (j in 1:length(temp))
    {
    n<-length(temp[j][,1]);
    xbar<-mean(temp[j],na.rm=TRUE);
    s<-sd(temp[j],na.rm=TRUE);
    error<-qt(0.975,df=n-1)*s/sqrt(n);

    write.table(xbar+error,paste(path,"/wd_upcon-temp.txt",sep=""),
```

```
+col.names=FALSE,row.names=FALSE,quote=FALSE,append=TRUE);

}
    tableback<-read.table(paste(path,"/wd_upcon-temp.txt",sep=""));
    file.remove(paste(path,"/wd_upcon-temp.txt",sep=""));
    if(file.exists(paste(path,"/wd_upcon.txt",sep="")))
        {
        oldfile<-read.table(paste(path,"/wd_upcon.txt",sep=""));
        write.table(cbind(oldfile,tableback),
+paste(path,"/wd_upcon.txt",sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE);
        }

    else write.table(tableback,paste(path,"/wd_upcon.txt",sep=""),
+col.names=FALSE,row.names=FALSE,quote=FALSE);
    }
}
```

APPENDIX D

ADDITIONAL STATISTICAL DATA

Table D.1: Shapiro-Wilk test for normality results of synthetic specimens digitized at diffeent lengths. **Bold** value indicates statistically significant departure from a normal distribution. See Table 4.1 on p. 59.

Specimen	**Digitization Length (cm)**	**Shapiro-Wilk W**	**Shapiro-Wilk p(normal)**
S2780	5	0.9354	0.5030
	10	0.9451	0.6113
	15	0.9406	0.5592
	20	0.9366	0.5161
	25	0.9295	0.4433
	30	0.8607	0.0777
S2800	5	0.8780	0.1237
	10	0.9066	0.2584
	15	0.9333	0.4814
	20	0.9459	0.6206
	25	0.9200	0.3568
	30	0.9317	0.4651
S2919	5	0.9037	0.2406
	10	0.9556	0.7354
	15	0.9556	0.7348
	20	0.8942	0.1888
	25	0.8922	0.1796
	30	0.9470	0.6330
S2920	5	0.9869	0.9915
	10	0.9825	0.9771
	15	0.9581	0.7637
	20	0.9811	0.9708
	25	0.9610	0.7973
	30	0.8754	0.1156
S2921	5	0.9352	0.5012
	10	0.9133	0.3042
	15	0.7369	**0.0025**
	20	0.8720	0.1054
	25	0.8912	0.1751
	30	0.9242	0.3933
S2922	5	0.9254	0.4042
	10	0.9725	0.9130
	15	0.9514	0.6847
	20	0.9703	0.8933
	25	0.9346	0.4945
	30	0.9519	0.6905
S2923	5	0.9609	0.7965
	10	0.9635	0.8244
	15	0.9601	0.7871
	20	0.9687	0.8788
	25	0.8495	0.0573
	30	0.9419	0.5742
S2924	5	0.9111	0.2889
	10	0.9485	0.6503
	15	0.9332	0.4796
	20	0.9305	0.4524
	25	0.9198	0.3556
	30	0.9344	0.4927
S2925	5	0.9426	0.5817
	10	0.9368	0.5177
	15	0.9660	0.8512
	20	0.9278	0.4270
	25	0.9405	0.5583
	30	0.9078	0.2661

Table D.2: Within-group dispersion of synthetic valves after repeat manual digitizations. Data are shown in Figure 5.1 on p. 76. Specimen numbers indicate specimen on which each synthetic valve was based.

Specimen	Digitization Length (cm)	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
S2780	5	0.0258	0.0219	0.0297
	10	0.0188	0.0163	0.0214
	15	0.0152	0.0138	0.0167
	20	0.0131	0.0116	0.0145
	25	0.0109	0.0091	0.0127
	30	0.0118	0.0101	0.0134
S2800	5	0.0338	0.0286	0.0391
	10	0.0214	0.0187	0.0242
	15	0.0171	0.0154	0.0188
	20	0.0173	0.0152	0.0194
	25	0.0116	0.0101	0.0130
	30	0.0128	0.0112	0.0143
S2919	5	0.0393	0.0294	0.0493
	10	0.0204	0.0180	0.0227
	15	0.0182	0.0161	0.0203
	20	0.0150	0.0125	0.0175
	25	0.0137	0.0111	0.0162
	30	0.0125	0.0111	0.0139
S2920	5	0.0279	0.0248	0.0310
	10	0.0240	0.0207	0.0273
	15	0.0177	0.0159	0.0195
	20	0.0171	0.0143	0.0198
	25	0.0117	0.0093	0.0140
	30	0.0140	0.0119	0.0161
S2921	5	0.0313	0.0267	0.0359
	10	0.0244	0.0209	0.0279
	15	0.0155	0.0109	0.0202
	20	0.0144	0.0128	0.0161
	25	0.0126	0.0110	0.0142
	30	0.0121	0.0108	0.0134
S2922	5	0.0329	0.0269	0.0388
	10	0.0251	0.0230	0.0272
	15	0.0219	0.0198	0.0240
	20	0.0187	0.0160	0.0214
	25	0.0131	0.0114	0.0147
	30	0.0138	0.0118	0.0159
S2923	5	0.0323	0.0272	0.0374
	10	0.0216	0.0191	0.0241
	15	0.0183	0.0168	0.0198
	20	0.0167	0.0157	0.0177
	25	0.0116	0.0101	0.0131
	30	0.0106	0.0094	0.0118
S2924	5	0.0276	0.0237	0.0315
	10	0.0254	0.0229	0.0279
	15	0.0209	0.0182	0.0236
	20	0.0176	0.0157	0.0195
	25	0.0139	0.0120	0.0158
	30	0.0155	0.0136	0.0173
S2925	5	0.0344	0.0280	0.0408
	10	0.0212	0.0190	0.0235
	15	0.0221	0.0193	0.0249
	20	0.0152	0.0137	0.0168
	25	0.0111	0.0094	0.0127
	30	0.0123	0.0102	0.0143

Table D.3: Results of tests for multivariate normality of Fourier scores for Discriminant Analysis on standard and nonstandard digitization lengths. Data are normal for kurtosis but non-normal for skewness. See also Table 4.2 on p. 61. **Bold** indicates statistical significance.

Statistic	**Value**
Mardia skewness coefficent	155.7
Mardia skewness statistic	3425
Mardia skewness df	2024
Mardia skewness p(normal)	**0.0000**
Mardia kurtosis coefficient	530.8
Mardia kurtosis statistic	0.5013
Mardia kurtosis p(normal)	0.6162
Doornik and Hansen omnibus Ep	69.43
Doornik and Hansen omnibus p(normal)	**0.0086**

Table D.4: Data shown in Figure 5.2 on p. 77. Specimen numbers indicate specimen on which each synthetic valve was based.

Specimen	Digitization Length (cm)	MV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
S2780	5	0.00035	0.00024	0.00046
	10	0.00018	0.00013	0.00024
	15	0.00012	0.00008	0.00016
	20	0.00009	0.00006	0.00012
	25	0.00006	0.00004	0.00009
	30	0.00007	0.00004	0.00011
S2800	5	0.00059	0.00037	0.00087
	10	0.00023	0.00016	0.00032
	15	0.00015	0.00010	0.00020
	20	0.00015	0.00009	0.00021
	25	0.00007	0.00005	0.00009
	30	0.00008	0.00005	0.00012
S2919	5	0.00085	0.00065	0.00109
	10	0.00021	0.00016	0.00027
	15	0.00017	0.00011	0.00024
	20	0.00012	0.00007	0.00017
	25	0.00010	0.00005	0.00015
	30	0.00008	0.00005	0.00012
S2920	5	0.00040	0.00026	0.00055
	10	0.00029	0.00020	0.00039
	15	0.00016	0.00011	0.00022
	20	0.00015	0.00009	0.00022
	25	0.00007	0.00005	0.00010
	30	0.00010	0.00006	0.00015
S2921	5	0.00051	0.00035	0.00066
	10	0.00031	0.00019	0.00044
	15	0.00014	0.00009	0.00020
	20	0.00011	0.00007	0.00014
	25	0.00008	0.00005	0.00012
	30	0.00007	0.00005	0.00010
S2922	5	0.00057	0.00041	0.00073
	10	0.00032	0.00022	0.00043
	15	0.00024	0.00016	0.00032
	20	0.00018	0.00012	0.00025
	25	0.00009	0.00005	0.00012
	30	0.00010	0.00006	0.00014
S2923	5	0.00054	0.00037	0.00071
	10	0.00024	0.00016	0.00032
	15	0.00017	0.00011	0.00024
	20	0.00014	0.00009	0.00020
	25	0.00007	0.00005	0.00009
	30	0.00006	0.00003	0.00008
S2924	5	0.00039	0.00028	0.00052
	10	0.00033	0.00022	0.00044
	15	0.00022	0.00014	0.00032
	20	0.00016	0.00011	0.00021
	25	0.00010	0.00006	0.00014
	30	0.00012	0.00008	0.00018
S2925	5	0.00062	0.00043	0.00083
	10	0.00023	0.00016	0.00030
	15	0.00025	0.00016	0.00038
	20	0.00012	0.00009	0.00015
	25	0.00006	0.00004	0.00009
	30	0.00008	0.00005	0.00012

Table D.5: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2780 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.008**	**0.000**	**0.000**	**0.000**	**0.000**
5 cm		—	**0.039**	**0.000**	**0.001**	**0.000**
5 cm			—	0.077	**0.001**	**0.012**
5 cm				—	0.072	0.125
5 cm					—	0.329
5 cm						—

Table D.6: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2800 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**
10 cm		—	**0.003**	**0.014**	**0.000**	**0.000**
15 cm			—	0.841	**0.000**	**0.001**
20 cm				—	**0.000**	**0.000**
25 cm					—	0.251
30 cm						—

Table D.7: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2919 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.004**	**0.000**	**0.000**	**0.000**	**0.000**
10 cm		—	0.099	**0.005**	**0.002**	**0.000**
15 cm			—	**0.028**	**0.002**	**0.000**
20 cm				—	0.070	0.077
25 cm					—	0.307
30 cm						—

Table D.8: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2920 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	0.058	**0.000**	**0.000**	**0.000**	**0.000**
10 cm		—	**0.001**	**0.011**	**0.000**	**0.000**
15 cm			—	0.687	**0.001**	**0.006**
20 cm				—	**0.012**	**0.041**
25 cm					—	0.071
30 cm						—

Table D.9: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2921 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.009**	**0.000**	**0.000**	**0.000**	**0.000**
10 cm		—	**0.001**	**0.000**	**0.000**	**0.000**
15 cm			—	0.562	0.191	0.086
20 cm				—	**0.001**	0.029
25 cm					—	0.598
30 cm						—

Table D.10: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2922 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.030**	**0.005**	**0.001**	**0.000**	**0.000**
10 cm		—	**0.011**	**0.000**	**0.000**	**0.000**
15 cm			—	**0.029**	**0.000**	**0.000**
20 cm				—	**0.001**	**0.018**
25 cm					—	0.443
30 cm						—

Table D.11: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2923 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.001**	**0.000**	**0.000**	**0.000**	**0.000**
10 cm		—	**0.032**	**0.004**	**0.000**	**0.000**
15 cm			—	0.058	**0.000**	**0.000**
20 cm				—	**0.000**	**0.000**
25 cm					—	0.235
30 cm						—

Table D.12: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2924 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	0.118	**0.007**	**0.000**	**0.000**	**0.000**
10 cm		—	**0.013**	**0.004**	**0.000**	**0.000**
15 cm			—	**0.010**	**0.000**	**0.002**
20 cm				—	**0.006**	**0.012**
25 cm					—	0.174
30 cm						—

Table D.13: Post hoc uncorrected least significant difference pairwise comparisons of specimen S2925 WD at different digitization lengths after ANOVA, summarized in Table 5.2 on p. 75. **Bold** indicates statistical significance.

	5 cm	**10 cm**	**15 cm**	**20 cm**	**25 cm**	**30 cm**
5 cm	—	**0.001**	**0.006**	**0.000**	**0.000**	**0.000**
10 cm		—	0.394	**0.000**	**0.000**	**0.000**
15 cm			—	**0.000**	**0.000**	**0.000**
20 cm				—	**0.002**	**0.002**
25 cm					—	0.344
30 cm						—

Table D.14: Shapiro-Wilk test for normality results of modern *Anodonta* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Anodonta	2	0.9858	**0.0000**
	3	0.9845	**0.0000**
	4	0.9831	**0.0000**
	5	0.9817	**0.0000**
	6	0.9802	**0.0000**
	7	0.9786	**0.0000**
	8	0.9770	**0.0000**
	9	0.9753	**0.0000**
	10	0.9737	**0.0000**
	11	0.9722	**0.0000**
	12	0.9707	**0.0000**
	13	0.9693	**0.0000**
	14	0.9679	**0.0000**
	15	0.9666	**0.0000**
	16	0.9654	**0.0000**
	17	0.9642	**0.0000**
	18	0.9631	**0.0000**
	19	0.9621	**0.0000**
	20	0.9611	**0.0000**

Table D.15: Shapiro-Wilk test for normality results of modern *Anodontites* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Anodontites	2	0.9727	**0.0000**
	3	0.9692	**0.0000**
	4	0.9663	**0.0000**
	5	0.9637	**0.0000**
	6	0.9616	**0.0000**
	7	0.9598	**0.0000**
	8	0.9582	**0.0000**
	9	0.9569	**0.0000**
	10	0.9557	**0.0000**
	11	0.9547	**0.0000**
	12	0.9538	**0.0000**
	13	0.9530	**0.0000**
	14	0.9522	**0.0000**
	15	0.9516	**0.0000**
	16	0.9510	**0.0000**
	17	0.9505	**0.0000**
	18	0.9501	**0.0000**
	19	0.9496	**0.0000**
	20	0.9492	**0.0000**

Table D.16: Shapiro-Wilk test for normality results of modern *Anodontoides* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Anodontoides	2	0.9605	**0.0000**
	3	0.9592	**0.0000**
	4	0.9581	**0.0000**
	5	0.9570	**0.0000**
	6	0.9560	**0.0000**
	7	0.9552	**0.0000**
	8	0.9545	**0.0000**
	9	0.9540	**0.0000**
	10	0.9535	**0.0000**
	11	0.9530	**0.0000**
	12	0.9526	**0.0000**
	13	0.9521	**0.0000**
	14	0.9517	**0.0000**
	15	0.9512	**0.0000**
	16	0.9508	**0.0000**
	17	0.9504	**0.0000**
	18	0.9500	**0.0000**
	19	0.9497	**0.0000**
	20	0.9493	**0.0000**

Table D.17: Shapiro-Wilk test for normality results of modern *Gonidea* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Gonidea	2	0.9834	0.0384
	3	0.9809	**0.0182**
	4	0.9786	**0.0094**
	5	0.9772	**0.0063**
	6	0.9784	**0.0087**
	7	0.9808	**0.0174**
	8	0.9827	**0.0306**
	9	0.9846	0.0542
	10	0.9862	0.0891
	11	0.9872	0.1206
	12	0.9886	0.1813
	13	0.9900	0.2718
	14	0.9914	0.3931
	15	0.9927	0.5397
	16	0.9936	0.6564
	17	0.9945	0.7729
	18	0.9952	0.8602
	19	0.9959	0.9220
	20	0.9965	0.9622

Table D.18: Shapiro-Wilk test for normality results of modern *Pilsbryoconcha* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
	2	0.9965	0.1013
	3	0.9976	0.3558
	4	0.9982	0.6366
	5	0.9982	0.6503
	6	0.9977	0.4165
	7	0.9970	0.1955
	8	0.9968	0.1490
	9	0.9967	0.1234
	10	0.9962	0.0691
Pilsbryoconcha	11	0.9953	**0.0217**
	12	0.9942	**0.0057**
	13	0.9930	**0.0014**
	14	0.9917	**0.0004**
	15	0.9904	**0.0001**
	16	0.9892	**0.0000**
	17	0.9879	**0.0000**
	18	0.9867	**0.0000**
	19	0.9856	**0.0000**
	20	0.9846	**0.0000**

Table D.19: Shapiro-Wilk test for normality results of modern *Pyganodon* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
	2	0.9697	**0.0000**
	3	0.9681	**0.0000**
	4	0.9670	**0.0000**
	5	0.9665	**0.0000**
	6	0.9663	**0.0000**
	7	0.9664	**0.0000**
	8	0.9666	**0.0000**
	9	0.9668	**0.0000**
	10	0.9672	**0.0000**
Pyganodon	11	0.9676	**0.0000**
	12	0.9680	**0.0000**
	13	0.9685	**0.0000**
	14	0.9691	**0.0000**
	15	0.9696	**0.0000**
	16	0.9702	**0.0000**
	17	0.9708	**0.0000**
	18	0.9713	**0.0000**
	19	0.9718	**0.0000**
	20	0.9723	**0.0000**

Table D.20: Shapiro-Wilk test for normality results of modern *Simpsonaias* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Simpsonaias	2	0.9681	**0.0000**
	3	0.9671	**0.0000**
	4	0.9656	**0.0000**
	5	0.9643	**0.0000**
	6	0.9635	**0.0000**
	7	0.9628	**0.0000**
	8	0.9623	**0.0000**
	9	0.9620	**0.0000**
	10	0.9618	**0.0000**
	11	0.9616	**0.0000**
	12	0.9613	**0.0000**
	13	0.9610	**0.0000**
	14	0.9608	**0.0000**
	15	0.9604	**0.0000**
	16	0.9601	**0.0000**
	17	0.9597	**0.0000**
	18	0.9591	**0.0000**
	19	0.9585	**0.0000**
	20	0.9581	**0.0000**

Table D.21: Shapiro-Wilk test for normality results of modern *Strophitus* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	**Smoothing**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Strophitus	2	0.9870	**0.0000**
	3	0.9852	**0.0000**
	4	0.9832	**0.0000**
	5	0.9813	**0.0000**
	6	0.9796	**0.0000**
	7	0.9779	**0.0000**
	8	0.9744	**0.0000**
	9	0.9754	**0.0000**
	10	0.9745	**0.0000**
	11	0.9737	**0.0000**
	12	0.9731	**0.0000**
	13	0.9727	**0.0000**
	14	0.9723	**0.0000**
	15	0.9721	**0.0000**
	16	0.9719	**0.0000**
	17	0.9718	**0.0000**
	18	0.9717	**0.0000**
	19	0.9717	**0.0000**
	20	0.9717	**0.0000**

Table D.22: Shapiro-Wilk test for normality results of modern *Utterbackia* paired Euclidean distances. See Table 4.3 on p. 65. **Bold** value indicates statistically significant departure from a normal distribution.

Genus	Smoothing	Shapiro-Wilk W	Shapiro-Wilk p(normal)
Utterbackia	2	0.9784	**0.0000**
	3	0.9815	**0.0000**
	4	0.9837	**0.0000**
	5	0.9856	**0.0000**
	6	0.9870	**0.0000**
	7	0.9897	**0.0000**
	8	0.9886	**0.0000**
	9	0.9889	**0.0000**
	10	0.9890	**0.0000**
	11	0.9889	**0.0000**
	12	0.9886	**0.0000**
	13	0.9882	**0.0000**
	14	0.9877	**0.0000**
	15	0.9870	**0.0000**
	16	0.9864	**0.0000**
	17	0.9857	**0.0000**
	18	0.9850	**0.0000**
	19	0.9842	**0.0000**
	20	0.9835	**0.0000**

Table D.23: Within-group dispersion of modern *Anodonta* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Anodonta	2	0.0972	0.0961	0.0983
	3	0.0938	0.0927	0.0949
	4	0.0911	0.0900	0.0921
	5	0.0887	0.0877	0.0898
	6	0.0867	0.0857	0.0878
	7	0.0850	0.0839	0.0860
	8	0.0834	0.0823	0.0844
	9	0.0820	0.0809	0.0830
	10	0.0807	0.0797	0.0817
	11	0.0795	0.0785	0.0806
	12	0.0785	0.0774	0.0795
	13	0.0775	0.0765	0.0785
	14	0.0766	0.0756	0.0776
	15	0.0758	0.0747	0.0768
	16	0.0750	0.0739	0.0760
	17	0.0743	0.0732	0.0753
	18	0.0736	0.0725	0.0746
	19	0.0729	0.0719	0.0740
	20	0.0723	0.0713	0.0734

Table D.24: Within-group dispersion of modern *Anodontites* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Anodontites	2	0.1057	0.1045	0.1068
	3	0.1021	0.1009	0.1032
	4	0.0990	0.0979	0.1002
	5	0.0965	0.0953	0.0977
	6	0.0942	0.0931	0.0954
	7	0.0923	0.0911	0.0935
	8	0.0905	0.0894	0.0917
	9	0.0890	0.0878	0.0901
	10	0.0875	0.0864	0.0887
	11	0.0863	0.0851	0.0874
	12	0.0851	0.0839	0.0862
	13	0.0840	0.0829	0.0852
	14	0.0830	0.0819	0.0842
	15	0.0821	0.0809	0.0832
	16	0.0812	0.0801	0.0824
	17	0.0804	0.0793	0.0816
	18	0.0797	0.0785	0.0808
	19	0.0790	0.0778	0.0801
	20	0.0783	0.0772	0.0794

Table D.25: Within-group dispersion of modern *Anodontitoides* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Anodontoides	2	0.0775	0.0766	0.0784
	3	0.0752	0.0743	0.0761
	4	0.0733	0.0724	0.0742
	5	0.0716	0.0708	0.0725
	6	0.0701	0.0693	0.0710
	7	0.0688	0.0680	0.0697
	8	0.0676	0.0668	0.0684
	9	0.0665	0.0657	0.0673
	10	0.0655	0.0646	0.0663
	11	0.0645	0.0637	0.0653
	12	0.0636	0.0628	0.0644
	13	0.0628	0.0620	0.0636
	14	0.0620	0.0612	0.0628
	15	0.0612	0.0605	0.0620
	16	0.0605	0.0598	0.0613
	17	0.0599	0.0591	0.0606
	18	0.0592	0.0585	0.0600
	19	0.0586	0.0579	0.0594
	20	0.0580	0.0573	0.0588

Table D.26: Within-group dispersion of modern *Gonidea* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Gonidea	2	0.0842	0.0836	0.0848
	3	0.0798	0.0792	0.0804
	4	0.0763	0.0757	0.0769
	5	0.0733	0.0727	0.0738
	6	0.0706	0.0701	0.0711
	7	0.0683	0.0678	0.0688
	8	0.0662	0.0657	0.0667
	9	0.0639	0.0634	0.0644
	10	0.0626	0.0621	0.0630
	11	0.0610	0.0605	0.0615
	12	0.0596	0.0591	0.0601
	13	0.0583	0.0578	0.0588
	14	0.0571	0.0567	0.0576
	15	0.0561	0.0556	0.0565
	16	0.0551	0.0546	0.0556
	17	0.0542	0.0537	0.0546
	18	0.0533	0.0529	0.0538
	19	0.0526	0.0521	0.0530
	20	0.0518	0.0514	0.0523

Table D.27: Within-group dispersion of modern *Pilsbryoconcha* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Pilsbryoconcha	2	0.0427	0.0424	0.0430
	3	0.0408	0.0405	0.0411
	4	0.0395	0.0392	0.0398
	5	0.0386	0.0383	0.0389
	6	0.0380	0.0376	0.0383
	7	0.0375	0.0372	0.0378
	8	0.0371	0.0368	0.0375
	9	0.0369	0.0366	0.0372
	10	0.0367	0.0364	0.0370
	11	0.0365	0.0362	0.0369
	12	0.0364	0.0361	0.0367
	13	0.0363	0.0359	0.0366
	14	0.0362	0.0358	0.0366
	15	0.0361	0.0357	0.0365
	16	0.0360	0.0357	0.0364
	17	0.0360	0.0356	0.0363
	18	0.0359	0.0355	0.0363
	19	0.0358	0.0354	0.0362
	20	0.0357	0.0353	0.0361

Table D.28: Within-group dispersion of modern *Pyganodon* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Pyganodon	2	0.0900	0.0891	0.0909
	3	0.0870	0.0861	0.0879
	4	0.0844	0.0835	0.0852
	5	0.0821	0.0812	0.0829
	6	0.0800	0.0792	0.0809
	7	0.0782	0.0774	0.0790
	8	0.0766	0.0757	0.0774
	9	0.0751	0.0743	0.0759
	10	0.0737	0.0729	0.0745
	11	0.0725	0.0717	0.0733
	12	0.0713	0.0705	0.0721
	13	0.0703	0.0695	0.0711
	14	0.0693	0.0685	0.0701
	15	0.0684	0.0676	0.0692
	16	0.0676	0.0668	0.0683
	17	0.0668	0.0660	0.0675
	18	0.0660	0.0652	0.0668
	19	0.0653	0.0645	0.0661
	20	0.0646	0.0639	0.0654

Table D.29: Within-group dispersion of modern *Simpsonaias* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Simpsonaias	2	0.0607	0.0602	0.0613
	3	0.0581	0.0576	0.0586
	4	0.0561	0.0555	0.0566
	5	0.0544	0.0539	0.0549
	6	0.0531	0.0526	0.0536
	7	0.0519	0.0514	0.0524
	8	0.0509	0.0504	0.0514
	9	0.0501	0.0496	0.0506
	10	0.0493	0.0488	0.0498
	11	0.0486	0.0481	0.0491
	12	0.0479	0.0474	0.0484
	13	0.0473	0.0468	0.0478
	14	0.0468	0.0463	0.0473
	15	0.0463	0.0458	0.0468
	16	0.0458	0.0453	0.0463
	17	0.0454	0.0449	0.0459
	18	0.0450	0.0445	0.0454
	19	0.0446	0.0441	0.0451
	20	0.0442	0.0437	0.0447

Table D.30: Within-group dispersion of modern *Strophitus* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Strophitus	2	0.0820	0.0812	0.0828
	3	0.0793	0.0785	0.0801
	4	0.0770	0.0762	0.0778
	5	0.0750	0.0742	0.0758
	6	0.0732	0.0724	0.0740
	7	0.0715	0.0707	0.0723
	8	0.0699	0.0691	0.0707
	9	0.0688	0.0680	0.0696
	10	0.0676	0.0668	0.0683
	11	0.0665	0.0657	0.0672
	12	0.0654	0.0647	0.0662
	13	0.0645	0.0637	0.0653
	14	0.0636	0.0629	0.0644
	15	0.0628	0.0621	0.0636
	16	0.0621	0.0613	0.0628
	17	0.0614	0.0606	0.0621
	18	0.0607	0.0600	0.0614
	19	0.0601	0.0593	0.0608
	20	0.0595	0.0587	0.0602

Table D.31: Within-group dispersion of modern *Utterbackia* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.4 on p. 80.

Genus	Smoothing	WD	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Utterbackia	2	0.0712	0.0704	0.0719
	3	0.0681	0.0674	0.0688
	4	0.0656	0.0649	0.0663
	5	0.0635	0.0629	0.0641
	6	0.0617	0.0611	0.0623
	7	0.0605	0.0599	0.0611
	8	0.0587	0.0581	0.0593
	9	0.0575	0.0569	0.0580
	10	0.0563	0.0557	0.0569
	11	0.0553	0.0547	0.0559
	12	0.0544	0.0538	0.0549
	13	0.0535	0.0529	0.0541
	14	0.0527	0.0522	0.0533
	15	0.0520	0.0515	0.0526
	16	0.0514	0.0508	0.0519
	17	0.0507	0.0502	0.0513
	18	0.0502	0.0496	0.0507
	19	0.0496	0.0491	0.0502
	20	0.0492	0.0486	0.0497

Table D.32: Sum of variance of modern *Anodonta* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Anodonta	2	0.0051	0.0030	0.0084
	3	0.0048	0.0027	0.0082
	4	0.0046	0.0025	0.0081
	5	0.0044	0.0022	0.0075
	6	0.0042	0.0021	0.0075
	7	0.0040	0.0020	0.0071
	8	0.0039	0.0018	0.0071
	9	0.0037	0.0016	0.0071
	10	0.0037	0.0015	0.0073
	11	0.0036	0.0014	0.0071
	12	0.0035	0.0014	0.0066
	13	0.0034	0.0013	0.0066
	14	0.0033	0.0012	0.0067
	15	0.0033	0.0012	0.0066
	16	0.0031	0.0011	0.0064
	17	0.0031	0.0011	0.0064
	18	0.0031	0.0010	0.0062
	19	0.0030	0.0010	0.0068
	20	0.0029	0.0009	0.0062

Table D.33: Sum of variance of modern *Anodontites* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Anodontites	2	0.0061	0.0037	0.0097
	3	0.0058	0.0034	0.0088
	4	0.0054	0.0030	0.0089
	5	0.0052	0.0028	0.0084
	6	0.0049	0.0024	0.0084
	7	0.0048	0.0024	0.0081
	8	0.0047	0.0022	0.0080
	9	0.0044	0.0020	0.0078
	10	0.0043	0.0019	0.0078
	11	0.0042	0.0018	0.0079
	12	0.0041	0.0017	0.0073
	13	0.0040	0.0016	0.0073
	14	0.0040	0.0016	0.0074
	15	0.0039	0.0016	0.0073
	16	0.0038	0.0015	0.0075
	17	0.0037	0.0014	0.0071
	18	0.0037	0.0014	0.0072
	19	0.0036	0.0013	0.0071
	20	0.0035	0.0014	0.0068

Table D.34: Sum of variance of modern *Anodontoides* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Anodontoides	2	0.0033	0.0024	0.0043
	3	0.0031	0.0022	0.0042
	4	0.0029	0.0020	0.0040
	5	0.0028	0.0019	0.0039
	6	0.0027	0.0018	0.0038
	7	0.0026	0.0017	0.0037
	8	0.0025	0.0017	0.0036
	9	0.0025	0.0016	0.0036
	10	0.0024	0.0015	0.0035
	11	0.0023	0.0014	0.0034
	12	0.0023	0.0014	0.0033
	13	0.0022	0.0013	0.0032
	14	0.0021	0.0012	0.0032
	15	0.0021	0.0012	0.0031
	16	0.0021	0.0012	0.0030
	17	0.0020	0.0011	0.0030
	18	0.0019	0.0011	0.0029
	19	0.0019	0.0011	0.0030
	20	0.0019	0.0010	0.0028

Table D.35: Sum of variance of modern *Gonidea* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Gonidea	2	0.0037	0.0030	0.0044
	3	0.0033	0.0027	0.0041
	4	0.0030	0.0024	0.0037
	5	0.0028	0.0022	0.0034
	6	0.0026	0.0020	0.0032
	7	0.0024	0.0018	0.0031
	8	0.0023	0.0017	0.0030
	9	0.0022	0.0015	0.0029
	10	0.0020	0.0015	0.0027
	11	0.0019	0.0014	0.0026
	12	0.0019	0.0013	0.0025
	13	0.0018	0.0012	0.0024
	14	0.0017	0.0012	0.0023
	15	0.0017	0.0011	0.0023
	16	0.0016	0.0011	0.0022
	17	0.0016	0.0010	0.0022
	18	0.0015	0.0009	0.0021
	19	0.0015	0.0009	0.0020
	20	0.0014	0.0009	0.0020

Table D.36: Sum of variance of modern *Pilsbryoconcha* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Pilsbryoconcha	2	0.0010	0.0008	0.0012
	3	0.0009	0.0007	0.0011
	4	0.0008	0.0006	0.0010
	5	0.0008	0.0006	0.0010
	6	0.0008	0.0005	0.0010
	7	0.0007	0.0005	0.0010
	8	0.0007	0.0005	0.0010
	9	0.0007	0.0005	0.0010
	10	0.0007	0.0005	0.0010
	11	0.0007	0.0005	0.0010
	12	0.0007	0.0005	0.0010
	13	0.0007	0.0004	0.0010
	14	0.0007	0.0004	0.0010
	15	0.0007	0.0004	0.0010
	16	0.0007	0.0004	0.0010
	17	0.0007	0.0004	0.0010
	18	0.0007	0.0004	0.0010
	19	0.0007	0.0004	0.0010
	20	0.0007	0.0004	0.0010

Table D.37: Sum of variance of modern *Pyganodon* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Pyganodon	2	0.0044	0.0033	0.0055
	3	0.0041	0.0031	0.0053
	4	0.0038	0.0028	0.0051
	5	0.0037	0.0027	0.0048
	6	0.0035	0.0025	0.0047
	7	0.0033	0.0024	0.0045
	8	0.0032	0.0022	0.0044
	9	0.0031	0.0021	0.0042
	10	0.0030	0.0020	0.0042
	11	0.0029	0.0019	0.0041
	12	0.0028	0.0019	0.0041
	13	0.0027	0.0018	0.0041
	14	0.0026	0.0016	0.0038
	15	0.0026	0.0016	0.0037
	16	0.0025	0.0015	0.0038
	17	0.0025	0.0015	0.0037
	18	0.0024	0.0014	0.0037
	19	0.0023	0.0014	0.0036
	20	0.0023	0.0013	0.0037

Table D.38: Sum of variance of modern *Simpsonaias* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Simpsonaias	2	0.0020	0.0014	0.0028
	3	0.0018	0.0012	0.0025
	4	0.0017	0.0011	0.0024
	5	0.0016	0.0010	0.0023
	6	0.0015	0.0009	0.0022
	7	0.0014	0.0008	0.0022
	8	0.0014	0.0008	0.0021
	9	0.0014	0.0008	0.0020
	10	0.0013	0.0007	0.0020
	11	0.0013	0.0007	0.0020
	12	0.0012	0.0007	0.0020
	13	0.0012	0.0006	0.0019
	14	0.0012	0.0006	0.0019
	15	0.0011	0.0006	0.0018
	16	0.0011	0.0006	0.0018
	17	0.0011	0.0006	0.0017
	18	0.0011	0.0006	0.0018
	19	0.0011	0.0006	0.0017
	20	0.0011	0.0005	0.0017

Table D.39: Sum of variance of modern *Strophitus* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Strophitus	2	0.0036	0.0026	0.0049
	3	0.0034	0.0024	0.0047
	4	0.0032	0.0022	0.0045
	5	0.0031	0.0021	0.0046
	6	0.0029	0.0020	0.0042
	7	0.0028	0.0018	0.0042
	8	0.0027	0.0017	0.0041
	9	0.0026	0.0016	0.0038
	10	0.0025	0.0015	0.0039
	11	0.0024	0.0014	0.0037
	12	0.0024	0.0014	0.0037
	13	0.0023	0.0014	0.0035
	14	0.0022	0.0013	0.0036
	15	0.0022	0.0013	0.0035
	16	0.0021	0.0012	0.0034
	17	0.0021	0.0011	0.0033
	18	0.0020	0.0011	0.0032
	19	0.0020	0.0010	0.0033
	20	0.0020	0.0010	0.0032

Table D.40: Sum of variance of modern *Utterbackia* based on EFA with varying amounts of smoothing. Data are graphed in Figure 5.5 on p. 81.

Genus	Smoothing	ΣV	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Utterbackia	2	0.0027	0.0022	0.0032
	3	0.0025	0.0020	0.0030
	4	0.0023	0.0018	0.0028
	5	0.0022	0.0017	0.0027
	6	0.0020	0.0016	0.0026
	7	0.0020	0.0015	0.0025
	8	0.0019	0.0014	0.0024
	9	0.0018	0.0013	0.0024
	10	0.0017	0.0012	0.0023
	11	0.0017	0.0012	0.0022
	12	0.0016	0.0011	0.0021
	13	0.0016	0.0011	0.0021
	14	0.0015	0.0010	0.0021
	15	0.0015	0.0010	0.0020
	16	0.0014	0.0009	0.0020
	17	0.0014	0.0009	0.0020
	18	0.0014	0.0009	0.0020
	19	0.0013	0.0009	0.0019
	20	0.0013	0.0008	0.0019

Table D.41: Results of tests for normality of Euclidean distances used to calculate WD. See Table 4.4 on p. 67. Performed in PAST.

Genus	**Probability Plot PPCC**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Anodonta	0.993	0.9858	**0.0000**
Anodontites	0.9862	0.9727	**0.0000**
Anodontoides	0.981	0.9603	**0.0000**
Gonidea	0.9901	0.9838	**0.0448**
Pilsbryoconcha	0.9984	0.9965	**0.0989**
Pyganodon	0.9845	0.9697	**0.0000**
Simpsonaias	0.9834	0.9684	**0.0000**
Strophitus	0.9937	0.9869	**0.0000**
Utterbackia	0.9892	0.9784	**0.0000**

Table D.42: Results of tests for multivariate normality of Fourier scores for MANOVA. Data are non-normal for kurtosis and skewness. See Table 4.6 on p. 70. **Bold** indicates statistical significance.

Statistic	**Value**
Mardia skewness coefficent	71.64
Mardia skewness statistic	4585
Mardia skewness df	2024
Mardia skewness p(normal)	**0.000**
Mardia kurtosis coefficient	615.8
Mardia kurtosis statistic	26.46
Mardia kurtosis p(normal)	**0.000**
Doornik and Hansen omnibus Ep	223.9
Doornik and Hansen omnibus p(normal)	**0.000**

Table D.43: Within-group dispersion values of some edentulous freshwater mussel genera based on outline shape. Values for L6516 unionoids included for comparison. Ranked from highest to lowest size of morphospace occupation. Plotted in Figure 5.6 on p. 83.

Genus	**WD**	**Lower 95% Confidence Limit**	**Upper 95% Confidence Limit**
Anodontites	0.1057	0.0959	0.0984
Anodonta	0.0972	0.1045	0.1068
Pyganodon	0.0900	0.0756	0.0795
L6516 unionoids	0.0890	0.0870	0.0910
Gonidea	0.0842	0.0816	0.0867
Strophitus	0.0820	0.0421	0.0434
Anodontoides	0.0775	0.0876	0.0924
Utterbackia	0.0712	0.0592	0.0623
Simpsonaias	0.0607	0.0807	0.0833
Pilsbryoconcha	0.0427	0.0699	0.0724

Table D.44: Sum of variance values of some edentulous freshwater mussel genera based on outline shape. Values for L6516 unionoids included for comparison. Ranked from highest to lowest size of morphospace occupation. Plotted in Figure 5.7 on p. 83.

Genus	**ΣV**	**Lower 95% Confidence Limit**	**Upper 95% Confidence Limit**
Anodontites	0.0061	0.0038	0.0095
Anodonta	0.0051	0.0030	0.0082
Pyganodon	0.0044	0.0034	0.0055
L6516 unionoids	0.0042	0.0027	0.0058
Gonidea	0.0037	0.0030	0.0044
Strophitus	0.0036	0.0027	0.0049
Anodontoides	0.0033	0.0024	0.0045
Utterbackia	0.0027	0.0022	0.0032
Simpsonaias	0.0019	0.0013	0.0027
Pilsbryoconcha	0.0009	0.0008	0.0011

Table D.45: Results of post hoc Tukey's HSD test to test equality of mean within-group dispersion between genera. Tukey's Q is found below the diagonal and p(same) above. **Bold** values are statistically significant ($\alpha = 0.05$). Data are summarized in Figure 5.8 on p. 84.

Genus	*Anodonta*	*Anodontites*	*Anodontoides*	*Gonidea*	*Pilsbryoconcha*	*Pyganodon*	*Simpsonaias*	*Strophitus*	*Utterbackia*
Anodonta	—	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**0.0004**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Anodontites	7.3750	—	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Anodontoides	17.06	24.44	—	**0.0013**	**< 0.0001**	**< 0.0001**	**< 0.0001**	0.1226	**0.0031**
Gonidea	11.25	18.63	5.809	—	**< 0.0001**	**0.0104**	**< 0.0001**	0.9218	**< 0.0001**
Pilsbryoconcha	47.27	54.64	30.20	36.01	—	**< 0.0001**	**< 0.0001**	**< 0.0001**	**< 0.0001**
Pyganodon	6.193	13.57	10.87	5.061	41.07	—	**< 0.0001**	**< 0.0001**	**< 0.0001**
Simpsonaias	31.63	39.00	14.57	20.37	15.64	25.44	—	**< 0.0001**	**< 0.0001**
Strophitus	13.14	20.51	3.925	1.884	34.13	6.945	18.49	—	**< 0.0001**
Utterbackia	22.58	29.95	5.512	11.32	24.69	16.38	9.053	9.437	—

Table D.46: Post hoc Hotelling's T^2 pairwise tests for equivalence of multivariate means among modern edentulous freshwater mussel genera based on Fourier scores of outlines. See Chapter 4.8.1 on p. 78.

Genus	*Anodonta*	*Anodontites*	*Anodontoides*	*Gonidea*	*Pilsbryoconcha*	*Pyganodon*	*Simpsonaias*	*Strophitus*	*Utterbackia*
Anodonta	—	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0004**	**0.0000**	**0.0000**	**0.0004**
Anodontites	**0.0000**	—	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Anodontoides	**0.0000**	**0.0000**	—	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Gonidea	**0.0000**	**0.0000**	**0.0000**	—	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Pilsbryoconcha	**0.0000**	**0.0000**	**0.0000**	**0.0000**	—	**0.0000**	**0.0000**	**0.0000**	**0.0000**
Pyganodon	**0.0133**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	—	**0.0000**	**0.0000**	**0.0000**
Simpsonaias	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	—	**0.0000**	**0.0000**
Strophitus	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	—	**0.0000**
Utterbackia	**0.0159**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.0005	**0.0000**	**0.0000**	–

Table D.47: Permutation test for two multivariate groups to test for equivalence of multivariate means among modern edentulous freshwater mussel genera based on Fourier scores of outlines. p(same) values above diagonal, Mahalanobis distances below. See Chapter 4.8.1 on p. 78. **Bold** indicates statistical significance.

Genus	*Anodonta*	*Anodontites*	*Anodontoides*	*Gonidea*	*Pilsbryoconcha*	*Pyganodon*	*Simpsonaias*	*Strophitus*	*Utterbackia*
Anodonta	—	**<0.0005**	**<0.0005**	**<0.005**	**<0.0005**	**< 0.0005**	**<0.0005**	**<0.0005**	**<0.0005**
Anodontites	0.2644	—	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**
Anodontoides	0.2055	0.2055	–	**<0.0005**	**<0.0005**	**0.001**	**<0.0005**	**<0.0005**	**<0.0005**
Gonidea	1.015	0.4662	1.07	—	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**
Pilsbryoconcha	0.4227	0.5497	1.036	1.906	—	**<0.0005**	**<0.0005**	**<0.0005**	**<0.0005**
Pyganodon	0.2103	0.2855	0.341	1.333	0.8323	—	**<0.0005**	**<0.0005**	**<0.0005**
Simpsonaias	0.5249	0.3319	0.6234	1.25	1.973	0.8516	—	**<0.0005**	**<0.0005**
Strophitus	0.3604	0.3083	0.4614	1.364	1.04	0.4233	0.892	—	**<0.005**
Utterbackia	0.1955	0.3753	0.5131	1.255	0.5497	0.4272	1.084	0.6843	—

Table D.48: Results of tests for normality of Euclidean distances used to calculate WD. See Table 4.8 on p. 72. **Bold** indicates statistically significant deviation from normality. Performed in PAST.

Species	**Probability Plot PPCC**	**Shapiro-Wilk** W	**Shapiro-Wilk p(normal)**
Anodonta couperiana	0.9778	0.9462	0.2648
Anodonta cygnea	0.9713	0.9323	0.4344
Anodonta grandis	0.9902	0.9803	**0.0000**
Anodonta imbecillis	0.9849	0.9620	0.2342
Anodonta implicata	0.9779	0.9268	0.5474
Anodonta suborbiculata	0.9793	0.9658	0.7672
Anodontites elongatus	0.9727	0.9463	**0.0023**
Anodontites ferrarisi	0.9159	0.8348	0.1807
Anodontites irisans	0.9775	0.9617	0.8335
Anodontites moricandi	0.9459	0.9167	0.2921
Anodontites obtusus	0.9712	0.9615	0.7881
Anodontites patagonicus	0.9867	0.9726	0.7704
Anodontites tenebricosus	0.9946	0.9855	0.4044
Anodontites trapesialis	0.9917	0.9836	**0.0051**
Anodontoides ferussacianus	0.9670	0.9328	**0.0000**
Gonidea angulata	0.9899	0.9834	**0.0384**
Pilsbryoconcha exilis	0.9984	0.9965	0.1013
Pyganodon cataracta	0.8387	0.7377	0.0014
Pyganodon grandis	0.9906	0.9847	0.3609
Pyganodon lacustris	0.9869	0.9593	0.6489
Simpsonaias ambigua	0.9832	0.9681	**0.0000**
Strophitus subvexus	0.9867	0.9609	0.2154
Strophitus undulatus	0.9857	0.9715	**0.0000**
Utterbackia imbecillis	0.9843	0.9687	**0.0000**

Table D.49: Within-group dispersion values of some edentulous freshwater mussel species based on outline shape. Values for L6516 unionoids included for comparison. Ranked from highest to lowest size of morphospace occupation. Plotted in Figure 5.11 on p. 88.

Species	**Mean**	**Lower 95% Confidence Limit**	**Upper 95% Confidence Limit**
Anodontites tenebricosus	0.1042	0.1021	0.1063
Pyganodon grandis	0.0903	0.0887	0.0919
L6516 unionoids	0.0890	0.0870	0.0910
Pyganodon cataracta	0.0871	0.0862	0.0879
Gonidea angulata	0.0842	0.0831	0.0853
Anodontites patagonicus	0.0824	0.0809	0.0839
Anodontites trapesialis	0.0809	0.0795	0.0823
Anodontites elongatus	0.0808	0.0797	0.0819
Anodonta grandis	0.0798	0.0783	0.0814
Anodontites moricandi	0.0790	0.0778	0.0803
Strophitus undulatus	0.0761	0.0748	0.0774
Anodontoides ferussacianus	0.0740	0.0724	0.0756
Utterbackia imbecillis	0.0696	0.0684	0.0709
Anodonta suborbiculata	0.0682	0.0672	0.0693
Anodontites obtusus	0.0681	0.0667	0.0694
Pyganodon lacustris	0.0679	0.0673	0.0686
Anodonta imbecillis	0.0621	0.0608	0.0633
Simpsonaias ambgiua	0.0607	0.0598	0.0617
Anodontites ferrarisi	0.0603	0.0586	0.0620
Strophitus subvexus	0.0588	0.0578	0.0598
Anodontites irisans	0.0568	0.0558	0.0578
Anodonta cygnea	0.0546	0.0538	0.0554
Anodonta implicata	0.0529	0.0518	0.0539
Anodonta couperiana	0.0478	0.0470	0.0487
Pilsbryoconcha exilis	0.0427	0.0421	0.0433

Table D.50: Morphological variation values of some edentulous freshwater mussel species based on outline shape. Values for L6516 unionoids included for comparison. Ranked from highest to lowest size of morphospace occupation. Plotted in Figure 5.12 on p. 88.

Species	**Mean**	**Lower 95% Confidence Limit**	**Upper 95% Confidence Limit**
Anodontites tenebricosus	0.0059	0.0028	0.0108
Pyganodon grandis	0.0044	0.0032	0.0057
L6516 unionoids	0.0042	0.0027	0.0058
Pyganodon cataracta	0.0039	0.0030	0.0048
Gonidea angulata	0.0037	0.0030	0.0044
Anodontites patagonicus	0.0036	0.0020	0.0061
Anodonta grandis	0.0035	0.0025	0.0047
Anodontites trapesialis	0.0035	0.0024	0.0046
Anodontites elongatus	0.0034	0.0024	0.0047
Anodontites moricandi	0.0033	0.0023	0.0044
Strophitus undulatus	0.0031	0.0024	0.0040
Anodontoides ferussacianus	0.0030	0.0022	0.0041
Utterbackia imbecillis	0.0026	0.0021	0.0031
Anodontites obtusus	0.0025	0.0015	0.0037
Anodonta suborbiculata	0.0025	0.0015	0.0036
Pyganodon lacustris	0.0024	0.0014	0.0037
Anodonta imbecillis	0.0021	0.0015	0.0028
Anodontites ferrarisi	0.0020	0.0009	0.0035
Simpsonaias ambigua	0.0020	0.0014	0.0027
Strophitus subvexus	0.0018	0.0014	0.0023
Anodontites irisans	0.0017	0.0012	0.0022
Anodonta cygnea	0.0016	0.0011	0.0021
Anodonta implicata	0.0015	0.0011	0.0020
Anodonta couperiana	0.0012	0.0008	0.0016
Pilsbryoconcha exilis	0.0009	0.0008	0.0011

Table D.51: Results of post hoc Tukey's HSD test to test equality of mean within-group dispersion between modern species. Tukey's Q is found below the diagonal, p(same) above. **Bold** values are statistically significant ($\alpha = 0.05$). Summarized in Figure 5.13 on p. 89. Continued on next page.

	Anodonta couperiana	*Anodonta cygnea*	*Anodonta grandis*	*Anodonta imbecillis*	*Anodonta implicata*	*Anodonta suborbiculata*	*Anodontites elongatus*	*Anodontites ferrarisi*	*Anodontites irisans*	*Anodontites moricandi*	*Anodontites obtusus*	*Anodontites patagonicus*	*Anodontites tenebricosus*
Anodonta couperiana	–	1.0000	**0.0014**	0.9488	1.0000	0.3784	**0.0008**	0.9888	0.9999	**0.0024**	0.3962	**0.0003**	**0.0000**
Anodonta cygnea	1.3550	–	0.0611	1.0000	1.0000	0.9669	**0.0383**	1.0000	1.0000	0.0885	0.9712	**0.0167**	**0.0000**
Anodonta grandis	6.4420	5.0870	–	0.6711	**0.0263**	0.9958	1.0000	0.4737	0.1574	1.0000	0.9948	1.0000	0.0926
Anodonta imbecillis	2.8600	1.5050	3.5820	–	0.9999	1.0000	0.5595	1.0000	1.0000	0.7584	1.0000	0.3799	**0.0000**
Anodonta implicata	1.0090	0.3465	5.4340	1.8520	–	0.8896	**0.0157**	1.0000	1.0000	**0.0399**	0.9001	**0.0063**	**0.0000**
Anodonta suborbiculata	4.1040	2.7490	2.3390	1.2430	3.0950	–	0.9875	1.0000	0.9967	0.9986	1.0000	0.9493	**0.0001**
Anodontites elongatus	6.6390	5.2840	0.1968	3.7790	5.6310	2.5350	–	0.3671	0.1063	1.0000	0.9852	1.0000	0.1386
Anodontites ferrarisi	2.5140	1.1590	3.9280	0.3467	1.5050	1.5900	4.1250	–	1.0000	0.5692	1.0000	0.2219	**0.0000**
Anodontites irisans	1.8050	0.4504	4.6370	1.0550	0.7969	2.2980	4.8340	0.7083	–	0.2134	0.9973	0.0521	**0.0000**
Anodontites moricandi	6.2750	4.9200	0.1667	3.4150	5.2670	2.1720	0.3636	3.7620	4.4700	–	0.9982	1.0000	0.0641
Anodontites obtusus	4.0700	2.7150	2.3730	1.2090	3.0610	0.0339	2.5690	1.5560	2.2640	2.2060	–	0.9427	**0.0001**
Anodontites patagonicus	6.9610	5.6060	0.5189	4.1010	5.9530	2.8580	0.3221	4.4470	5.1560	0.6856	2.8910	–	0.2485
Anodontites tenebricosus	11.3400	9.9870	4.8990	8.4810	10.3300	7.2380	4.7030	8.8280	9.5360	5.0660	7.2720	4.3810	–
Anodontites trapesialis	6.6570	5.3020	0.2145	3.7960	5.6480	2.5530	0.0176	4.1430	4.8510	0.3812	2.5870	0.3044	4.6850
Anodontoides ferussacianus	5.2700	3.9150	1.1720	2.4090	4.2610	1.1660	1.3690	2.7560	3.4640	1.0060	1.2000	1.6910	6.0720
L6516 unionoids	7.3170	5.9620	0.8751	4.4570	6.3090	3.2140	0.6783	4.8040	5.5120	1.0420	3.2480	0.3562	4.0240
Gonidea angulata	1.0330	2.3880	7.4750	3.8930	2.0410	5.1360	7.6720	3.5470	2.8380	7.3080	5.1020	7.9940	12.3700
Pilsbryoconcha exilis	7.8930	6.5380	1.4510	5.0320	6.8840	3.7890	1.2540	5.3790	6.0870	1.6170	3.8230	0.9317	3.4490
Pyganodon cataracta	8.5490	7.1940	2.1070	5.6880	7.5400	4.4450	1.9100	6.0350	6.7430	2.2730	4.4790	1.5880	2.7930
Pyganodon grandis	4.0390	2.6840	2.4030	1.1780	3.0300	0.0646	2.6000	1.5250	2.2340	2.2370	0.0307	2.9220	7.3030
Pyganodon lacustris	2.5930	1.2380	3.8490	0.2675	1.5840	1.5110	4.0460	0.0793	0.7876	3.6820	1.4770	4.3680	8.7490
Simpsonaias ambigua	2.2000	0.8454	4.2420	0.6600	1.1920	1.9030	4.4390	0.3133	0.3950	4.0750	1.8690	4.7610	9.1410
Strophitus subvexus	5.6880	4.3330	0.7540	2.8280	4.6800	1.5850	0.9509	3.1740	3.8830	0.5873	1.6190	1.2730	5.6530
Strophitus undulatus	4.3850	3.0300	2.0580	1.5240	3.3760	0.2811	2.2540	1.8710	2.5790	1.8910	0.3150	2.5760	6.9570
Utterbackia imbecillis	8.2930	6.9380	1.8510	5.4320	7.2840	4.1890	1.6540	5.7790	6.4870	2.0170	4.2230	1.3320	3.0490

Table D.52: (Continued from previous page). Results of post hoc Tukey's HSD test to test equality of mean within-group dispersion between modern species. Tukey's Q is found below the diagonal, p(same) above. **Bold** values are statistically significant ($\alpha = 0.05$). Summarized in Figure 5.13 on p. 89.

	Anodontites trapesialis	*Anodontoides ferussacianus*	*L6516 unionoids*	*Gonidea angulata*	*Pilsbryoconcha exilis*	*Pyganodon cataracta*	*Pyganodon grandis*	*Pyganodon lacustris*	*Simpsonaias ambigua*	*Strophitus subvexus*	*Strophitus undulatus*	*Utterbackia imbecillis*
Anodonta couperiana	**0.0007**	**0.0396**	**0.0001**	1.0000	**0.0000**	**0.0000**	0.4126	0.9834	0.9983	**0.0134**	0.2468	**0.0000**
Anodonta cygnea	**0.0367**	0.4815	**0.0061**	0.9943	**0.0010**	**0.0001**	0.9748	1.0000	1.0000	0.2685	0.9092	**0.0003**
Anodonta grandis	1.0000	1.0000	1.0000	**0.0001**	1.0000	0.9991	0.9938	0.5188	0.3098	1.0000	0.9994	0.9999
Anodonta imbecillis	0.5493	0.9936	0.2183	0.4936	0.0692	**0.0134**	1.0000	1.0000	1.0000	0.9547	1.0000	**0.0264**
Anodonta implicata	**0.0149**	0.3007	**0.0022**	0.9995	**0.0003**	**0.0000**	0.9090	1.0000	1.0000	0.1449	0.7772	**0.0001**
Anodonta suborbiculata	0.9863	1.0000	0.8475	0.0545	0.5533	0.2228	1.0000	1.0000	0.9998	1.0000	1.0000	0.3350
Anodontites elongatus	1.0000	1.0000	1.0000	**0.0000**	1.0000	0.9998	0.9828	0.4088	0.2253	1.0000	0.9975	1.0000
Anodontites ferrarisi	0.3581	0.9658	0.1131	0.6903	**0.0302**	**0.0050**	1.0000	1.0000	1.0000	0.8624	0.9999	**0.0104**
Anodontites irisans	0.1025	0.7336	**0.0215**	0.9529	**0.0042**	**0.0005**	0.9978	1.0000	1.0000	0.4996	0.9844	**0.0012**
Anodontites moricandi	1.0000	1.0000	1.0000	**0.0001**	1.0000	0.9972	0.9978	0.6146	0.3934	1.0000	0.9998	0.9996
Anodontites obtusus	0.9838	1.0000	0.8340	0.0590	0.5338	0.2100	1.0000	1.0000	0.9999	1.0000	1.0000	0.3186
Anodontites patagonicus	1.0000	1.0000	1.0000	**0.0000**	1.0000	1.0000	0.9362	0.2536	0.1235	1.0000	0.9846	1.0000
Anodontites tenebricosus	0.1435	**0.0044**	0.4206	**0.0000**	0.7415	0.9604	**0.0001**	**0.0000**	**0.0000**	**0.0147**	**0.0003**	0.9037
Anodontites trapesialis	0.0000	1.0000	1.0000	**0.0000**	1.0000	0.9998	0.9813	0.3994	0.2186	1.0000	0.9972	1.0000
Anodontoides ferussacianus	1.3870	0.0000	0.9994	**0.0022**	0.9809	0.8210	1.0000	0.9756	0.8976	1.0000	1.0000	0.9111
L6516 unionoids	0.6607	2.0480	0.0000	**0.0000**	1.0000	1.0000	0.8212	0.1328	0.0570	1.0000	0.9338	1.0000
Gonidea angulata	7.6900	6.3030	8.3500	0.0000	**0.0000**	**0.0000**	0.0633	0.6466	0.8398	**0.0006**	**0.0274**	**0.0000**
Pilsbryoconcha exilis	1.2360	2.6230	0.5754	8.9260	0.0000	1.0000	0.5162	**0.0368**	**0.0132**	0.9982	0.7108	1.0000
Pyganodon cataracta	1.8920	3.2790	1.2310	9.5820	0.6559	0.0000	0.1989	**0.0063**	**0.0019**	0.9488	0.3475	1.0000
Pyganodon grandis	2.6180	1.2310	3.2780	5.0720	3.8540	4.5100	0.0000	1.0000	0.9999	1.0000	1.0000	0.3041
Pyganodon lacustris	4.0640	2.6770	4.7240	3.6260	5.3000	5.9560	1.4460	0.0000	1.0000	0.8895	0.9999	**0.0130**
Simpsonaias ambigua	4.4560	3.0690	5.1170	3.2330	5.6920	6.3480	1.8380	0.3925	0.0000	0.7215	0.9984	**0.0042**
Strophitus subvexus	0.9685	0.4184	1.6290	6.7210	2.2050	2.8610	1.6490	3.0950	3.4880	0.0000	1.0000	0.9824
Strophitus undulatus	2.2720	0.8852	2.9330	5.4170	3.5080	4.1640	0.3457	1.7920	2.1840	1.3040	0.0000	0.4852
Utterbackia imbecillis	1.6360	3.0230	0.9755	9.3260	0.4001	0.2558	4.2540	5.7000	6.0920	2.6050	3.9080	0.0000

Table D.53: Results of principal component analysis of modern genera and L6516 assemblage based on EFA scores. PCA plot is in Figure 5.9 on p. 85.

Principal Component	**Eigenvalue**	**% variance**	**Cumulative variance**
1	0.0015381	25.415	25.415
2	0.00146646	24.231	49.646
3	0.000503011	8.3116	57.9576
4	0.000450725	7.4477	65.4053
5	0.000315823	5.2186	70.6239
6	0.000303143	5.0091	75.633
7	0.000227065	3.752	79.385
8	0.000203792	3.3674	82.7524
9	0.000181258	2.9951	85.7475
10	0.000158851	2.6248	88.3723
11	0.000109744	1.8134	90.1857
12	0.000108035	1.7851	91.9708
13	0.000102186	1.6885	93.6593
14	8.42×10^{-05}	1.3917	95.051
15	7.10×10^{-05}	1.1734	96.2244
16	5.89×10^{-05}	0.97259	97.19699
17	5.28×10^{-05}	0.87215	98.06914
18	3.85×10^{-05}	0.63608	98.70522
19	3.35×10^{-05}	0.55411	99.25933
20	2.46×10^{-05}	0.40636	99.66569
21	2.02×10^{-05}	0.33382	99.99951
22	1.93×10^{-16}	3.19×10^{-12}	99.99951

Table D.54: Loadings of principal components from PCA of modern genera and L6516 assemblage based on EFA scores. PCA plot is in Figure 5.9 on p. 85.

EFA harmonic	**Axis 1**	**Axis 2**
A2	-9.47×10^{-10}	6.91×10^{-10}
B2	-0.7105	0.6399
A3	-0.3673	-0.4378
B3	0.04192	0.1972
A4	-0.03013	0.08938
B4	-0.2854	-0.1397
A5	-0.1465	-0.2469
B5	0.3849	0.3872
A6	0.1268	0.1574
B6	-0.1461	-0.05584
A7	0.01806	-0.1406
B7	0.1578	0.2019
A8	0.04992	0.09912
B8	-0.1647	0.006745
A9	-0.02054	0.0949
B9	0.02966	0.05249
A10	-0.07529	0.006403
B10	-0.006515	-0.08702
A11	0.04749	0.03201
B11	0.07123	0.006789
A12	-0.001703	-0.03949
B12	-0.004156	-0.03524

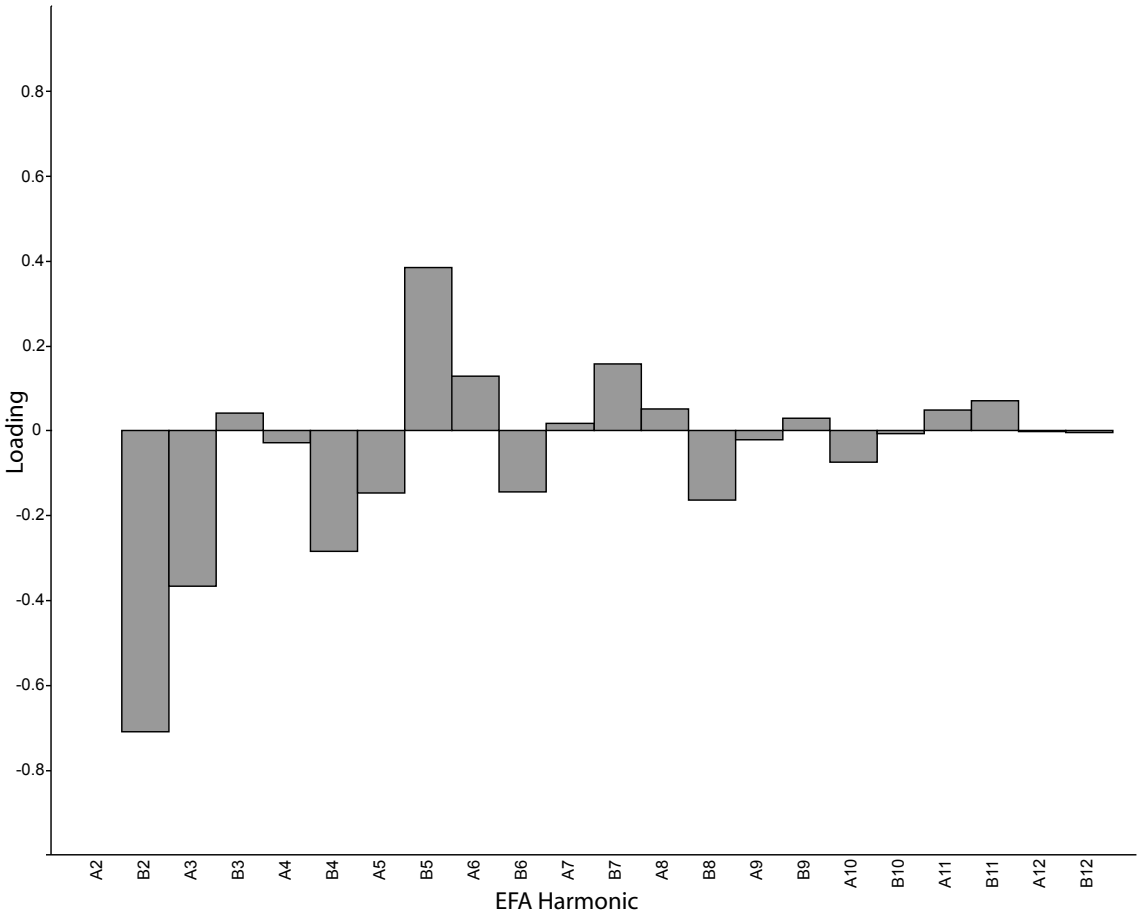

Figure D.1: Plot of principal component loadings on the first principal component from Figure 5.9 on p. 85. Raw data in Table D.54.

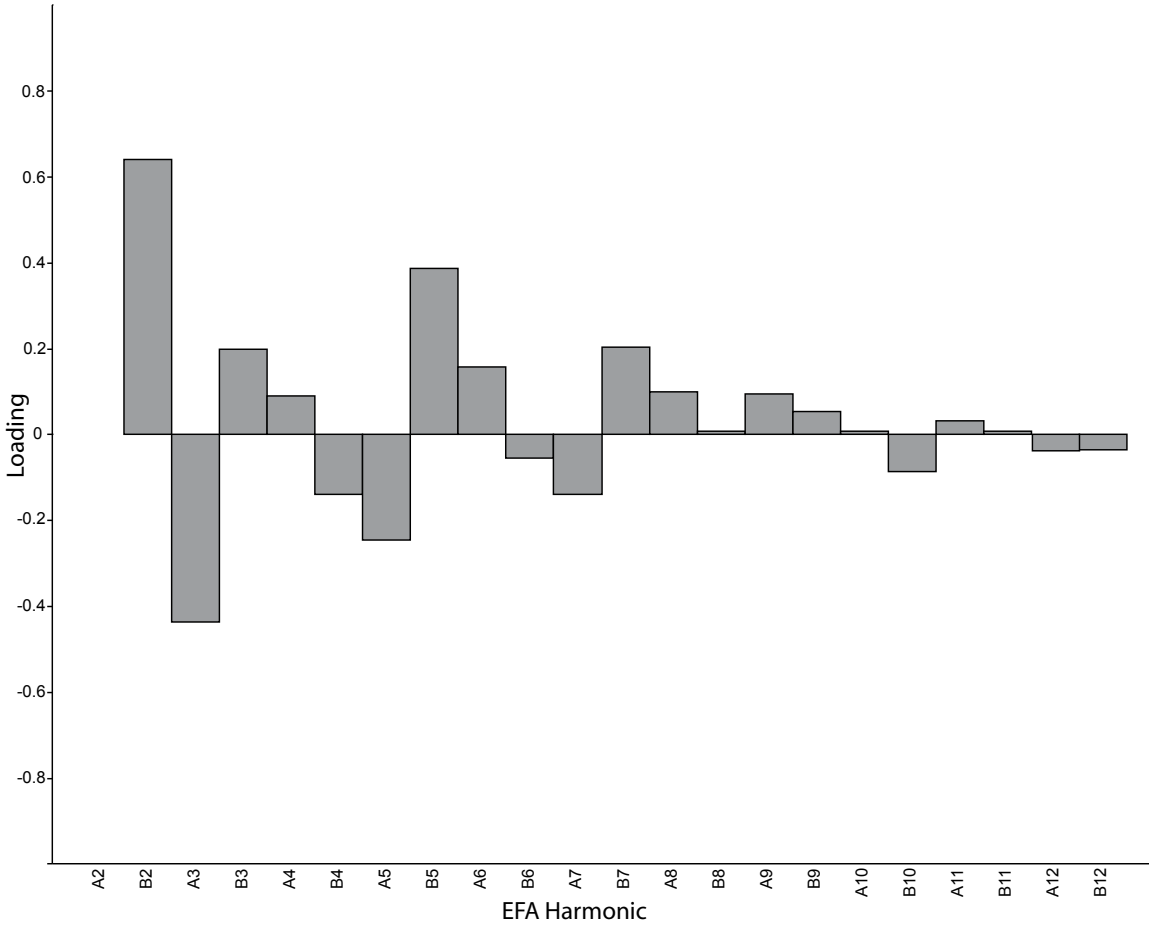

Figure D.2: Plot of principal component loadings on the second principal component from Figure 5.9 on p. 85. Raw data in Table D.54.

APPENDIX E

SPECIMEN INFORMATION

E.1 Filename Prefixes

a Specimen digitized at a length of 5 cm.

b Specimen digitized at a length of 10 cm.

c Specimen digitized at a length of 15 cm.

d *Alasmidonta* specimen digitized at an arbitrary length.

e *Anodonta* specimen digitized at an arbitrary length.

f *Anodontoides* specimen digitized at an arbitrary length.

flp Flipped horizontally from original (on image files).

g *Arcidens* specimen digitized at an arbitrary length.

h *Cristaria* specimen digitized at an arbitrary length.

i *Lasmigona* specimen digitized at an arbitrary length.

j *Strophitus* specimen digitized at an arbitrary length.

k *Utterbackia* specimen digitized at an arbitrary length.

l Unused as a prefix.

m Specimen digitized at a length of 20 cm.

n Specimen digitized at a length of 25 cm.

o Specimen digitized at a length of 30 cm.

p *Glebula* specimen with a smoothing of 1 during EFA.

q *Glebula* specimen with a smoothing of 2 during EFA.

r *Glebula* specimen with a smoothing of 3 during EFA.

S Specimen.

s *Glebula* specimen with a smoothing of 4 during EFA.

T Type or temporary specimen.

t *Glebula* specimen with a smoothing of 5 during EFA.

t1,t2,t3... Traced, number indicates digitization replicate.

tps For digitization (on image files).

u *Glebula* specimen with a smoothing of 6 during EFA.

v *Glebula* specimen with a smoothing of 7 during EFA.

w *Glebula* specimen with a smoothing of 8 during EFA.

x *Glebula* specimen with a smoothing of 9 during EFA.

y *Glebula* specimen with a smoothing of 10 during EFA.

z *Anodonta* specimen digitized at a length of 25 cm.

aa *Anodonta* specimen digitized at an arbitrary length.

ab *Glebula* specimen digitized at a length of 25 cm.

ac *Pyganodon* specimen digitized at an arbitrary length.

ad *Simpsonaias* specimen digitized at an arbitrary length.

ae *Anodontites* specimen digitized at a length of 25 cm.

af *Gonidea* specimen digitized at a length of 25 cm.

ag *Pilsbryoconcha* specimen digitized at a length of 25 cm.

ah *Gonidea* specimen digitized at an arbitrary length.

ai *Pilsbryoconcha* specimen digitized at an arbitrary length.

aj *Anodontoides* specimen digitized at a length of 25 cm.

ak *Pyganodon* specimen digitized at a length of 25 cm.

al *Simpsonaias* specimen digitized at a length of 25 cm.

am *Strophitus* specimen digitized at a length of 25 cm.

an *Utterbackia* specimen digitized at a length of 25 cm.

E.2 Specimens of Extant Freshwater Mussels

Table E.1: Specimens used for this project. Species identifications were identified by the source. T numbers were used for this project only. Valves are identified as left (l) or right (r), and whether the interior or exterior was photographed. The last five columns mark whether each specimen (and which valve) was used in that analysis: EFA genus–elliptical Fourier analysis to compare extant genera with L6516 unionoids, EFA species–elliptical Fourier analysis to compare extant species with L6516 unionoids, Dig. length–determining the effect of digitization length on variation, Smoothing–determining the effect of smoothing values on variation.

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Anodonta anatina	T0009	r ext	Menker (2005)	R		R	R
Anodonta beringiana	T0010	r ext	Menker (2005)	R		R	R
Anodonta californiensis	T0011	r ext	Menker (2005)	R		R	R
Anodonta coarctata	T0012	r ext	Menker (2005)	R		R	R
Anodonta couperiana	T0013	r ext	Menker (2005)	R	R	R	R
Anodonta couperiana	T0127	ext	A. Bogan (NCSM)	LR	LR	LR	LR
Anodonta couperiana	T0128	ext	A. Bogan (NCSM)	LR	LR	LR	LR
Anodonta couperiana	T0129	ext	A. Bogan (NCSM)	LR	LR	LR	LR
Anodonta cygnea	T0014	l ext	Menker (2005)	L	L	L	L
Anodonta cygnea	T0130	ext	A. Bogan (NCSM)	LR	LR	LR	LR
Anodonta cygnea	T0131	ext	A. Bogan (NCSM)	LR	LR	LR	LR
Anodonta globusa	T0015	r ext	Menker (2005)	R		R	R
Anodonta grandis	T0085	l int, r ext	Howells et al. (1996)	LR	LR	LR	LR
Anodonta grandis	T0086	l ext, r int	Howells et al. (1996)	LR	LR	LR	LR
Anodonta grandis	T0087	r ext	Howells et al. (1996)	R	R	R	R
Anodonta grandis	T0088	l ext	Howells et al. (1996)	L	L	L	L
Anodonta grandis	T0089	r ext	Howells et al. (1996)	R	R	R	R
Anodonta grandis	T0090	l ext	Howells et al. (1996)	L	L	L	L
Anodonta grandis	T0132	l int, r ext	Howells et al. (1996)	LR	LR	LR	LR
Anodonta grandis	T0133	l ext, r int	Howells et al. (1996)	LR	LR	LR	LR

continued on next page

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Anodonta grandis	T0134	int	UND-PC	LR	LR	LR	LR
Anodonta grandis	T0135	l int	UND-PC	L	L	L	L
Anodonta grandis	T0136	int	UND-PC	LR	LR	LR	LR
Anodonta grandis	T0137	r int	UND-PC	R	R	R	R
Anodonta grandis	T0138	int	UND-PC	LR	LR	LR	LR
Anodonta grandis	T0139	r int	UND-PC	R	R	R	R
Anodonta grandis	T0140	r int	UND-PC	R	R	R	R
Anodonta grandis	T0141	int	UND-PC	LR	LR	LR	LR
Anodonta grandis	T0142	int	UND-PC	LR	LR	LR	LR
Anodonta grandis	T0143	int	UND-PC	LR	LR	LR	LR
Anodonta imbecillis	T0091	l ext, r int	Howells et al. (1996)	LR	LR	LR	LR
Anodonta imbecillis	T0092	l int, r ext	Howells et al. (1996)	LR	LR	LR	LR
Anodonta imbecillis	T0093	l ext, r int	Howells et al. (1996)	LR	LR	LR	LR
Anodonta imbecillis	T0144	l ext	Howells et al. (1996)	L	L	L	L
Anodonta imbecillis	T0145	r ext	Howells et al. (1996)	R	R	R	R
Anodonta imbecillis	T0146	r ext	Howells et al. (1996)	R	R	R	R
Anodonta implicata	T0016	r ext	Menker (2005)	R	R	R	R
Anodonta implicata	T0094	l int, r ext	Strayer and Jirka (1997)	LR	LR	LR	LR
Anodonta implicata	T0095	l ext	Strayer and Jirka (1997)	L	L	L	L
Anodonta kennerlyi	T0017	r ext	Menker (2005)	R		R	R
Anodonta nuttalliana	T0018	r ext	Menker (2005)	R			R
Anodonta sp.	T0096	l int, r ext	Howells et al. (1996)	LR	LR	LR	LR
Anodonta suborbiculata	T0019	r ext	Menker (2005)	R	R	L	R
Anodonta suborbiculata	T0076	l ext	Parmalee and Bogan (1999)	L	L	L	L
Anodonta suborbiculata	T0097	l ext	Howells et al. (1996)	L	L	L	L
Anodonta suborbiculata	T0147	r ext	Howells et al. (1996)	R	R	R	R
Anodonta suborbiculata	T0148	l ext	Howells et al. (1996)	L	L	L	L
Anodonta suborbiculata	T0386	r ext		R	R		R

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Anodontites crispatus	T0414	r ext	Simone (2006)	R			R
Anodontites elongatus	T0415		Simone (2006)	L	L		L
Anodontites elongatus	T0416		Simone (2006)	L	L		L
Anodontites elongatus	T0417		Simone (2006)	LR	LR		LR
Anodontites elongatus	T0418		Simone (2006)	LR	LR		LR
Anodontites elongatus	T0419		Simone (2006)	LR	LR		LR
Anodontites elongatus	T0420		Simone (2006)	LR	LR		LR
Anodontites elongatus	T0421		Simone (2006)	LR	LR		LR
Anodontites elongatus	T0422		Simone (2006)	R	R		R
Anodontites ferrarisi	T0423		Simone (2006)	LR	LR		LR
Anodontites ferrarisi	T0424		Simone (2006)	R	R		R
Anodontites irisans	T0425		Simone (2006)	LR	LR		LR
Anodontites irisans	T0426		Simone (2006)	LR	LR		LR
Anodontites moricandi	T0427		Simone (2006)	LR	LR		LR
Anodontites moricandi	T0428		Simone (2006)	L	L		L
Anodontites moricandi	T0429		Simone (2006)	L	L		L
Anodontites moricandi	T0430		Simone (2006)	R	R		R
Anodontites obtusus	T0431		Simone (2006)	R	R		R
Anodontites obtusus	T0432		Simone (2006)	LR	LR		LR
Anodontites patagonicus	T0433		Simone (2006)	LR	LR		LR
Anodontites patagonicus	T0434		Simone (2006)	LR	LR		LR
Anodontites patagonicus	T0435		Simone (2006)	LR	LR		LR
Anodontites patagonicus	T0436		Simone (2006)	R	R		R
Anodontites tenebricosus	T0437		Simone (2006)	LR	LR		LR

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Anodontites tenebricosus	T0443		Simone (2006)	L	L		L
Anodontites tenebricosus	T0444		Simone (2006)	R	R		R
Anodontites tenebricosus	T0445		Simone (2006)	R	R		R
Anodontites tenebricosus	T0446		Simone (2006)	R	R		R
Anodontites tenebricosus	T0447		Simone (2006)	R	R		R
Anodontites tortilis	T0448		Simone (2006)	LR			LR
Anodontites trapesialis	T0449		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0450		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0451		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0452		Simone (2006)	R	R		R
Anodontites trapesialis	T0453		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0454		Simone (2006)	R	R		R
Anodontites trapesialis	T0455		Simone (2006)	L	L		L
Anodontites trapesialis	T0456		Simone (2006)	L	L		L
Anodontites trapesialis	T0457		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0458		Simone (2006)	L	L		L
Anodontites trapesialis	T0459		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0460		Simone (2006)	LR	LR		LR
Anodontites trapesialis	T0461		Simone (2006)	L	L		L
Anodontites trapesialis	T0462		Simone (2006)	R	R		R
Anodontites trapesialis	T0463		Simone (2006)	L	L		L
Anodontites trapesialis	T0464		Simone (2006)	L	L		L
Anodontoides connasaugaensis	T0030	r ext	Menker (2005)	R			R
Anodontoides denigrata	T0031	r ext	Menker (2005)	R	R		R

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Anodontoides ferussacianus	T0098	l int, r ext	Strayer and Jirka (1997)	LR	LR		LR
Anodontoides ferussacianus	T0099	r ext	Strayer and Jirka (1997)	R	R		R
Anodontoides ferussacianus	T0100	r ext	Strayer and Jirka (1997)	R	R		R
Anodontoides ferussacianus	T0149	ext	A. Bogan (NCSM)	LR	LR		LR
Anodontoides ferussacianus	T0150	ext	A. Bogan (NCSM)	LR	LR		LR
Anodontoides ferussacianus	T0151	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0152	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0153	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0154	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0192	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0193	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0194	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0195	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0196	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0197	int	UND-PC	LR	LR		LR
Anodontoides ferussacianus	T0388	r ext	Cicerello and Schuster (2003)	R	R		R
Anodontoides radiatus	T0033	r ext	Menker (2005)	R			R
Gonidea angulata	T0041	r ext	Menker (2005)	R	R		R
Gonidea angulata	T0347	r int	K. Cummings (INHS)	R	R		R
Gonidea angulata	T0348	int	K. Cummings (INHS)	LR	LR		LR
Gonidea angulata	T0349	int	K. Cummings (INHS)	LR	LR		LR
Gonidea angulata	T0350	int	K. Cummings (INHS)	LR	LR		LR
Gonidea angulata	T0351	int	K. Cummings (INHS)	LR	LR		LR
Gonidea angulata	T0352	l int	K. Cummings (INHS)	L	L		L
Gonidea angulata	T0353	int	K. Cummings (INHS)	LR	LR		LR
Gonidea angulata	T0354	l int	K. Cummings (INHS)	L	L		L
Gonidea angulata	T0355	r int	K. Cummings (INHS)	R	R		R
Gonidea angulata	T0356	r int	K. Cummings (INHS)	R	R		R

continued on next page

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Gonidea angulata	T0357	l int	K. Cummings (INHS)	L	L		L
Gonidea angulata	T0358	r int	K. Cummings (INHS)	R	R		R
Gonidea angulata	T0359	l int	K. Cummings (INHS)	L	L		L
Pilsbryoconcha exilis	T0051	l ext	Menker (2005)	L	L		L
Pilsbryoconcha exilis	T0175	ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0469	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0470	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0471	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0472	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0473	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0474	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0475	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0476	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0477	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0478	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0479	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0480	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0481	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0482	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0483	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0484	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0485	l int, r ext	A. Bogan (NCSM)	LR	LR		LR
Pilsbryoconcha exilis compressa	T0486	int	A. Bogan (NCSM)	LR	LR		LR
Pyganodon cataracta	T0106	l ext	Strayer and Jirka (1997)	L	L		L
Pyganodon cataracta	T0107	l ext	Strayer and Jirka (1997)	L	L		L
Pyganodon cataracta	T0108	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon cataracta cataracta	T0054	r ext	Menker (2005)	R	R		R
Pyganodon cataracta marginata	T0055	r ext	Menker (2005)	R	R		R
Pyganodon doliaris	T0056	l ext	Menker (2005)	L			L

continued on next page

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Pyganodon gibbosa	T0057	l ext	Menker (2005)	L			L
Pyganodon grandis	T0078	l ext	Parmalee and Bogan (1999)	L	L		L
Pyganodon grandis	T0109	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon grandis	T0110	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon grandis	T0111	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon grandis	T0112	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon grandis	T0113	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon grandis	T0177	ext	A. Bogan (NCSM)	LR	LR		LR
Pyganodon grandis	T0178	r ext	A. Bogan (NCSM)	R	R		R
Pyganodon grandis	T0396	r ext	Cicerello and Schuster (2003)	R	R		R
Pyganodon grandis corpulenta	T0058	r ext	Menker (2005)	R	R		R
Pyganodon grandis grandis	T0059	r ext	Menker (2005)	R	R		R
Pyganodon grandis simpsoniana	T0060	r ext	Menker (2005)	R	R		R
Pyganodon grandis stewartiana	T0061	r ext	Menker (2005)	R	R		R
Pyganodon hallenbecki	T0062	r ext	Menker (2005)	R			R
Pyganodon lacustris	T0114	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon lacustris	T0115	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon lacustris	T0116	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon lacustris	T0117	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon lacustris	T0118	l ext	Strayer and Jirka (1997)	L	L		L
Pyganodon lacustris	T0119	r ext	Strayer and Jirka (1997)	R	R		R
Pyganodon teres	T0063	r ext	Menker (2005)	R			R
Simpsonaias ambigua	T0064	r ext	Menker (2005)	R	R		R
Simpsonaias ambigua	T0079	l ext	Parmalee and Bogan	L	L		L

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Simpsonaias ambigua	T0273	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0274	l ext	A. Bogan (NCSM)	L	L		L
Simpsonaias ambigua	T0275	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0276	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0277	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0278	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0279	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0280	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0281	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0282	ext	A. Bogan (NCSM)	LR	LR		LR
Simpsonaias ambigua	T0397	r ext	Cicerello and Schuster (2003)	R	R		R
Strophitus subvexus	T0070	r ext	Menker (2005)	R	R		R
Strophitus subvexus	T0283	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus subvexus	T0284	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus subvexus	T0285	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus subvexus	T0286	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0081	l ext	Parmalee and Bogan (1999)	L	L		L
Strophitus undulatus	T0120	l int, r ext	Strayer and Jirka (1997)	LR	LR		LR
Strophitus undulatus	T0121	l int, r ext	Strayer and Jirka (1997)	LR	LR		LR
Strophitus undulatus	T0122	l ext, r int	Howells et al. (1996)	LR	LR		LR
Strophitus undulatus	T0189	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0287	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0288	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0289	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0290	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0291	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0292	ext	A. Bogan (NCSM)	LR	LR		LR

continued on next page

continued from previous page

Species	No.	Valves	Source	EFA genus	EFA species	Dig. length	Smoothing
Strophitus undulatus	T0293	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0294	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0295	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0296	l ext	A. Bogan (NCSM)	L	L		L
Strophitus undulatus	T0297	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0298	ext	A. Bogan (NCSM)	LR	LR		LR
Strophitus undulatus	T0299	r ext	A. Bogan (NCSM)	R	R		R
Strophitus undulatus	T0300	r ext	A. Bogan (NCSM)	R	R		R
Strophitus undulatus	T0398	r ext	Cicerello and Schuster (2003)	R	R		R
Strophitus undulatus pavonia	T0123	r ext	Strayer and Jirka (1997)	R	R		R
Strophitus undulatus tennessensis	T0071	r ext	Menker (2005)	R	R		R
Strophitus undulatus undulatus	T0072	r ext	Menker (2005)	R	R		R
Strophitus connasaugaensis	T0080	l ext	Parmalee and Bogan (1999)	L			L
Utterbackia imbecillis	T0073	r ext	Menker (2005)	R	R		R
Utterbackia imbecillis	T0082	l ext	Parmalee and Bogan (1999)	L	L		L
Utterbackia imbecillis	T0124	r ext	Strayer and Jirka (1997)	R	R		R
Utterbackia imbecillis	T0190	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0191	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0301	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0302	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0303	l ext	A. Bogan (NCSM)	L	L		L
Utterbackia imbecillis	T0304	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0305	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0306	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0307	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0308	ext	A. Bogan (NCSM)	LR	LR		LR

continued on next page

continued from previous page

Species	**No.**	**Valves**	**Source**	**EFA genus**	**EFA species**	**Dig. length**	**Smoothing**
Utterbackia imbecillis	T0309	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0310	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0311	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0312	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0313	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0314	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0315	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0316	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0317	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0318	ext	A. Bogan (NCSM)	LR	LR		LR
Utterbackia imbecillis	T0400	r ext	Cicerello and Schuster (2003)	R	R		R
Utterbackia peggyae	T0074	r ext	Menker (2005)	R			R
Utterbackia peninsularis	T0075	r ext	Menker (2005)	R			R

REFERENCES

Abramoff, M. D., Magelhaes, P. J., and Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International*, 11(7):36–42. 95

Aldridge, D. C. (1999). The morphology, growth and reproduction of Unionidae (Bivalvia) in a fenland waterway. *Journal of Molluscan Studies*, 65:47–60. 47

Athreya, S. (2006). Patterning of geographic variation in Middle Pleistocene *Homo* frontal bone morphology. *Journal of Human Evolution*, 50:627–643. 62

Balla, S. A. and Walker, K. F. (1991). Shape variation in the Australian freshwater mussel *Alathyria jacksoni* Iredale (Bivalvia, Hyriidae). *Hydrobiologia*, 220:89–98. 48

Barnes, R. S. K. (1994). *The Brackish-Water Fauna of Northwestern Europe*. Cambridge University Press. 94

Bassler, R. S. (1953). Part G Bryozoa. In Moore, R. C., editor, *Treatise on Invertebrate Paleontology*. University of Kansas and the Geological Society of America. 92

Bauer, G. (2001). Introduction. In Bauer, G. and Wächtler, K., editors, *Ecology and Evolution of the Freshwater Mussels Unionoidea*. Springer-Verlag. 1

Bogan, A. E. (2008). Global diversity of freshwater mussels (Mollusca, Bivalvia) in freshwater. *Hydrobiologia*, 595:139–147. 1

Bond, J. E. and Beamer, D. A. (2006). A morphometric analysis of mygalomorph spider carapace shape and its efficacy as a phylogenetic character (Areneae). *Invertebrate Systematics*, 20:1–7. 62

Bookstein, F. L. (1991). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press. 3, 47

Burton-Kelly, M. E. and Hartman, J. H. (2007). Discrimination of end-Cretaceous anodontine Unionoidea from North Dakota: How many taxa make sense? In Jordaens, K., van Houtte, N., van Goethem, J., and Backeljau, T., editors, *World Congress of Malacology 15-20 July 2007*, pages 28–29. 53

Cardini, A. and Slice, D. E. (2004). Mandibular shape in the genus *Marmota* (Rodentia, Scuridae): a preliminary analysis using outlines. *Italian Journal of Zoology*, 71:17–25. 62

Chaloner, D. T. and Wotton, R. S. (1996). Tube building by larvae of 3 species of midge (Diptera: Chironomidae). *Journal of the North American Benthological Society*, 15(3):300–307. 93

Cicerello, R. R. and Schuster, G. A. (2003). *A Guide to the Freshwater Mussels of Kentucky*. Kentucky State Nature Preserves Commission. 193, 195, 196, 198, 199, 200, 201

Clemens, W. A. (2002). Evolution of the mammalian fauna across the Cretaceous-Tertiary boundary in northeastern Montana and other areas of the Western Interior. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous*, chapter 11, pages 217–246. Geological Society of America. 6

Costa, C., Aguzzi, J.and Menesatti, P., Antonucci, F., Rimatori, V., and Mattoccia, M. (2008). Shape analysis of different populations of clams in relation to their geographical structure. *Journal of Zoology*, 276:71–80. 102, 104

Crampton, J. S. and Haines, A. J. (1996). User's manual for programs HANGLE, HMATCH, and HCURVE for the Fourier shape analysis of two-dimensional outlines. Technical Report 96/37, Institute of Geological & Nuclear Sciences Limited. 47, 48, 62, 98

Curry, M. G., Everitt, B., and Vidrine, M. F. (1981). Haptobenthos on shells of living freshwater clams in Louisiana. *The Wasmann Journal of Biology*, 39(1-2):56–62. 92

Cvancara, A. M. (1983). *Aquatic mollusks of North Dakota, North Dakota Geological Survey Report of Investigation No. 78.* North Dakota Geological Survey. 1

Daegling, D. J. and Jungers, W. L. (2000). Elliptical Fourier analysis of symphyseal shape in great ape mandibles. *Journal of Human Evolution*, 39:107–122. 62

Davies, R. W. (1991). Annelida: Leeches, polychaetes, and acanthobdellids. In Thorp, J. H. and Covich, A. P., editors, *Ecology and classification of North American freshwater invertebrates*, pages 437–480. Academic Press. 93

Davis, J. C. (2002). *Statistics and Data Analysis in Geology.* John Wiley & Sons, 3rd edition. 59, 61, 65, 67, 70, 72

Dorai-Raj, S. (2005). [R] R: cbind from Sundar Dorai-Raj on 2005-08-08 (2005-August.txt). http://finzi.psych.upenn.edu/R/Rhelp02a/archive/59302.html. Accessed 10 November 2008. 121

Dunca, E., Schöne, B. R., and Mutvei, H. (2005). Freshwater bivalves tell of past climates: but how clearly do shells from polluted rivers speak? *Palaeogeography, Palaeoclimatology, Palaeoecology*, 228:43–57. 1

Eager, R. M. C. (1948). Variation in shape of shell with respect to ecological station. A review dealing with Recent Unionidae and certain species of the Anthrocosiidae in Upper Carboniferous times. *Proceedings of the Royal Society of Edinburgh, Section B*, 63:130–148. 47, 48

Eager, R. M. C. (1974). Shape of shell of *Carbonicola* in relation to burrowing. *Lethaia*, 7:219–238. 47, 48

Eager, R. M. C. (1977). Shape of shell in relation to weight of margaritifera margaritifera. *Journal of Conchology*, 29:207–218. 47, 48

Eager, R. M. C. (1978). Shape and function of the shell: a comparison of some living and fossil bivalve mussels. *Biological Reviews of the Cambridge Philosophical Society*, 53:169–210. 47

Erickson, J. M. (1983). *Trichopterodomus leonardi*, a new genus and species of psychomyid caddisfly (Insecta: Trichoptera) represented by retreats from the Paleocene of North Dakota. *Journal of Paleontology*, 57(3):560–567. 93

Ferson, S., Rohlf, F. J., and Koehn, R. K. (1985). Measuring shape variation of two-dimensional outlines. *Systematic Zoology*, 34(1):59–68. 3, 47, 48, 62

Foote, M. (1989). Perimeter-based Fourier analysis: a new morphometric method applied to the trilobite cranidium. *Journal of Paleontology*, 63(6):880–885. 47, 49, 55

Forsyth, D. J. and McCallum, I. D. (1978). Xenochironomus canterhuryensis (Diptera: Chironomidae) an insectan inquiline commensal of Hydridella menziesi (Mollusca: Lamellibranchia). *Journal of Zoology*, 186(33):1–334. 93

Foster, N. (1972). *Freshwater polychaetes (Annelida) of North America*. Environmental Protection Agency, Washington, D.C. 94

Fric, A. and Bayer, E. (1901). Studien im gebiete der bohmischen kreideformation. vii. perucer schichten. *Arch. anturw. Ladesf. Bohmen*, 11(2):184. 92

Frost, T. M. (1991). Porifera. In Thorp, J. H. and Covich, A. P., editors, *Ecology and classification of North American freshwater invertebrates*, pages 95–124. Academic Press. 92

Gauldie, R. W. and Crampton, J. S. (2002). An eco-morphological explanation of individual variability in the shape of the fish otolith: comparison of the otolith of *Hoplostethus atlanticus* with other species by depth. *Journal of Fish Biology*, 60:1204–1221. 62

Glasby, C. J., Hutchings, P. A., Fauchald, K., Paxton, H., Rouse, G. W., Watson Russel, C., and Watson Russel, R. S. (2000). Class polychaeta. In Beesley, P. L., Ross, G. B., and Glasby, C. J., editors, *Polychaetes & Allies: The Southern Synthesis. Fauna of Asutralia.*, volume 4A Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula. CSIRO Publishing. 93

Graf, D. L. and Cummings, K. S. (2006). Palaeoheterodont diversity (Mollusca: Trigonioida + Unionoida): what we know and what we wish we knew about freshwater mussel evolution. *Zoological Journal of the Linnean Society*, 148:343–394. 1

Grimaldi, D. A. and Engel, M. S. A. (2005). *Evolution of the Insects*. Cambridge University Press. 93

Haines, A. J. and Crampton, J. S. (2000). Improvements to the method of fourier shape analysis as applied in morphometric studies. *Palaeontology*, 43:765–783. 3, 47, 48

Hammer, O. (2002). *Morphometrics – brief notes*. Paläontologisches Institut und Museum, Zurich. 61, 70

Hammer, O., Harper, D. A. T., and Ryan, P. D. (2008). *PAST - PAlaeontological STatistics, ver. 1.81*. 61, 67, 70, 72, 103

Hartman, J. H. (1992). Biochronology of uppermost Cretaceous and lower Tertiary nonmarine Mollusca of the northern Great Plains. In Lidgard, S. and Crane, P. R., editors, *Fifth North American Paleontological Convention (Chicago), Abstracts with Programs*, page 123. The Paleontological Society Special Publication Number 6. 1

Hartman, J. H. (1996a). Decimation of the freshwater molluscan fauna near the end of the Cretaceous - a North American perspective. In Bardet, N. and Buffetaut, E., editors, *The Cretaceous-Tertiary Boundary - Biological and Geological Aspects: Séance spécialisé de la Société géologique, Abstracts*, page 28. 1

Hartman, J. H. (1996b). Extinction of sculptured nonmarine bivalves about the Cretaceous-Tertiary boundary. In Wolberg, D. L. and Stump, E., editors, *DinoFest II International Symposium, Programs and Abstracts (April 18-21)*, page 58. 1, 2

Hartman, J. H. (1998). The biostratigraphy and paleontology of the Latest Cretaceous and Paleocene freshwater bivalves from the western Williston Basin, Montana, U.S.A. In Johnston, P. A. and Haggart, J. W., editors, *Bivalves: an eon*

of evolution; paleontological studies honoring Norman D. Newell, pages 317–245. University of Calgary Press. 1, 5

Hartman, J. H., Burton-Kelly, M. E., and Sweet, A. R. (2007). Interpreting the last molluscan Unionoidea from the Cretaceous of North Dakota. In Jordaens, K., van Houtte, N., van Goethem, J., and Backeljau, T., editors, *World Congress of Malacology 15-20 July 2007*, page 92. Unitas Malacologica. x, 10, 11

Hartman, J. H. and Butler, R. D. (1995). Extinction and recovery of nonmarine molluscan assemblages in the late Cretaceous and early Tertiary. *Geological Society of America Abstracts with Programs*, 27(4):13. 1

Hartman, J. H., Johnson, K. R., and Nichols, D. J. (2001). The last freshwater molluscan assemblage of the Cretaceous? A new locality from the Ludlow Formation of North Dakota. *Proceedings of the North Dakota Academy of Science*, 55:63. 2, 10

Hesterberg, T., Moore, D. S., Monaghan, S., Clipson, A., and Epstein, R. (2005). Bootstrap methods and permutation tests. In Moore, D. S. and McCabe, G. P., editors, *Introduction to the practice of statistics*. W. H. Freeman. 56, 57, 97

Hicks, J. F., Johnson, K. R., Obradovich, J. D., Tauxe, L., and Clark, D. (2002). Magnetostratigraphy and geochronology of the Hell Creek and basal Fort Union Formations of southwestern North Dakota and a recalibration of the age of the Cretaceous-Tertiary boundary. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous*, chapter 3, pages 35–56. Geological Society of America. 6, 7

Hinch, S. G. and Bailey, R. C. (1988). Within- and among-lake variation in shell morphology of the freshwater clam *Elliptio complanata* (Bivalvia: Unionidae) from south-central Ontario lakes. *Hydrobiologia*, 157:27–32. 48

Howell, B. F. (1962). Worms. In Moore, R. C., editor, *Treatise on Invertebrate Paleontology, Part W Miscellanea*. University of Kansas and the Geological Society of America. 93

Howells, R. G., Neck, R. W., and Murray, H. D. (1996). *Freshwater Mussels of Texas*. Texas Parks and Wildlife Press. 192, 193, 199

Hunter, J. P. and Archibald, J. D. (2002). Mammals from the end of the age of the dinosaurs in North Dakota and southeastern Montana, with a reappraisal of geographic differentiation among Lancian mammals. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous.*, pages 191–216. Geological Society of America. 6

Hunter, J. P., Hartman, J. H., and Krause, D. W. (1997). Mammals and mollusks across the Cretaceous-Tertiary boundary from Makoshika State Park and vicinity (Williston Basin), Montana. *Rocky Mountain Geology*, 32(1):61–114. 6

Innes, D. J. and Bates, J. A. (1999). Morphological variation of mytilus edulis and mytilus trossulus in eastern newfoundland. *Marine Biology*, 133:691–699. 47, 62

Iwata, H. and Ukai, Y. (2002). SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *The Journal of Heredity*, 93(5):384–385. 103

Johnson, K. R. (2002). The megaflora of the Hell Creek and lower Fort Union Formations in the western Dakotas: vegetational response to climate change, the

Cretaceous-Tertiary boundary event, and rapid marine transgression. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous.*, pages 329–392. Geological Society of America. 7, 10, 48, 49, 50

Jones, M. (1969). Boring of shell by *Caobangia* in freshwater snails of southeast Asia. *American Zoologist*, 9:829–835. 94

Jones, M. L. (1974). On the Caobangiidae, a new family of the Polychaeta, with a redescription of *Caobangia billeti* Giard. *Smithsonian Contributions to Zoology*, 175:1–55. 94

Joshi, S. H. and Srivastava, A. (2003). An algorithm for clustering objects according to their shapes. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.7375. 103

Justham, T. P. (2008). Geochemical analysis of ironstone preserved molluscan fossils of the Hell Creek Formation (Cretaceous) and Ludlow Member of the Fort Union Formation (Paleogene) of southwestern North Dakota. Master's thesis, University of North Dakota. 2

Klassen, E., Srivastava, A., Mio, W., and Joshi, S. H. (2004). Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):372–383. 103

Kroeger, T. J. (2002). Palynology of the Hell Creek Formation (Upper Cretaceous, Maastrichtian) in northwestern South Dakota: effects of paleoenvironment on the composition of palynomorph assemblages. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary*

boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous, chapter 17, pages 457–472. Geological Society of America. 6

Kuhl, F. and Giardina, C. R. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18:236–258. 3, 47

Kuiper, K. F., Deino, A., Hilgen, F. J., Krijgsman, W., Renne, P. R., and Wijbrans, J. R. (2008). Synchronizing rock clocks of earth history. *Science*, 320:500–504. 1

Lauer, T. E., Spacie, A., and Barnes, D. K. (2001). The distribution and habitat preference of freshwater sponges (Porifera) in four souther Lake Michigan harbors. *America Midland Naturalist*, 146:243–253. 92

Lestrel, P. E., Takahashi, O., and Kanazawa, E. (2004). A quantitative approach for measuring crowding in the dental arch: Fourier descriptors. *American Journal of Orthodontics and Dentofacial Orthopedics*, 125(6):716–725. 3, 62

Marcus, L. F. (1990). Traditional morphometrics. In Rohlf, F. J. and Bookstein, F. L., editors, *Proceedings of the Michigan Morphometrics Workshop*, pages 77–122. University of Michigan Museum of Zoology Special Publication No. 2. 3

McLellan, T. and Endler, J. A. (1998). The relative success of some methods for measuring and describing the shape of complex objects. *Systematic Biology*, 47(2):264–281. 3

McNeill, G. and Vijayakumar, S. (2005). 2d shape classification and retrieval. In Kaelbling, L. P. and Saffiotti, A., editors, *IJCAI-05, Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence, Edinburgh, Scotland, UK, July 30-August 5, 2005*. Professional Book Center. 104

Mehlhop, P. and Cifelli, R. (1997). A comparison of morphometric techniques to distinguish sympatric mussel species (family unionidae) with similar shell morphology. In Yates, T. L., Gannon, W. I., and Wilson, D. E., editors, *Life Among the Muses: Papers in Honor of James S. Findley*. The Museum of Southwestern Biology, The University of New Mexico, Albuquerque. 62

Menker, T. (2005). The Freshwater Mussels (Unionacea) of the Mollusc Division, The Museum of Biological Diversity, The Ohio State University. http://www.biosci.ohio-state.edu/ molluscs/gallery/index.htm. 192, 193, 195, 196, 197, 198, 199, 200, 201

Mermillod-Blondin, F., Gaudet, J.-P., Gérino, M., Desrosiers, G., and Creuzé des Châtelliers, M. (2003). Influence of macroinvertebrates on physico-chemical and microbial processes in hyporheic sediments. *Hydrological Processes*, 17:779–794. 93

Murphy, E. C., Hoganson, J. W., and Johnson, K. R. (2002). Lithostratigraphy of the Hell Creek Formation in North Dakota. In *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous.*, pages 9–34. Geological Society of America. 6, 10

Nichols, D. J. (2002). Palynology and palynostratigraphy of the Hell Creek Formation in North Dakota: a microfossil record of plants at the end of Cretaceous time. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous*, chapter 16, pages 393–456. Geological Society of America. 7

Nichols, D. J. (2007). Selected plant microfossil records of the terminal Cretaceous event in terrestrial rocks, western North America. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 255:22–34. 6

Nichols, D. J. and Johnson, K. R. (2002). Palynology and microstratigraphy of Cretaceous-Tertiary boundary sections in southwestern North Dakota. In Hartman, J. H., Johnson, K. R., and Nichols, D. J., editors, *The Hell Creek Formation and the Cretaceous-Tertiary boundary in the northern Great Plains: an integrated continental record of the end of the Cretaceous*, chapter 6, pages 95–144. Geological Society of America. 7

Ólafsson, J. S. and Paterson, D. M. (2004). Alteration of biogenic structyre and physical properties by tube-building chironomid larvae in cohesive sediments. *Aquatic Ecology*, 38:219–229. 93

Oliver, D. R. (1971). Life history of the chironomidae. *Annual Review of Entomology*, 16:211–230. 93

Osgood Jr., R. G. (1987). Trace fossils. In Boardman, R. S., Cheetham, A. H., and Rowell, A. J., editors, *Fossil Invertabrates*, pages 663–674. Blackwell Scientific Publications. 18

Parmalee, P. W. and Bogan, A. E. (1999). *The Freshwater Mussels of Tennessee*. The University of Tennessee Press. 193, 195, 198, 199, 200

Pearson, D. A., Schaefer, T., Johnson, K. R., and Nichols, D. J. (2001). Palynologically calibrated vertebrate record from North Dakota consistent with abrupt dinosaur extinction at the Cretaceous-Tertiary boundary. *Geology*, 29(1):39–42. 6

Pinder, L. C. V. (1986). Biology of the freshwater chironomidae. *Annual Review of Entomology*, 31:1–23. 93

Pothin, K., Gonzalez-Salas, C., Chabanet, P., and Lecomte-Finiger, R. (2006). Distinction between *Mulloidicthys flavolineatus* juveniles from Reunion Island and Mauritius Island (south-west Indian Ocean) based on otolith morphometrics. *Journal of Fish Biology*, 69:38–53. 62

Prieto-Marquez, A., Gignac, P. M., and Joshi, S. (2007). Neontological evaluation of pelvic skeletal attributes purported to reflect sex in extinct non-avian archosaurs. *Journal of Vertebrate Paleontology*, 27(3):603–609. 103

Pringle, C. M. (1985). Effects of chironomid (Insecta: Diptera) tube-building activities on stream diatom communities. *Journal of Phycology*, 21:185–194. 93

Raveloson, H., Le Minor, J.-M., Rumpler, Y., and Schmittbuhl, M. (2005). Shape of the lateral mandibular outline in Lemuridae: a quantitative analysis of variability using elliptical Fourier analysis. *Folia Primatology*, 76:245–261. 62

Roback, S. S., Bereza, D. J., and Vidrine, M. F. (1979). Description of an *Ablubesmyia* (Diptera: Chironomidae: Tanypodinae) symbiont of unionid fresh-water mussels (Mollusca: Bivalvia: Unionacea), with notes on its biology and zoogeography. *Trans. Am. Entomol. Sot.*, 105:577–620. 93

Rodgers, J. L. (1999). The bootstrap, the jacknife, and the randomization text: a sampling taxonomy. *Multivariate Behavioral Research*, 34(4):441–456. 56

Rohlf, F. J. (2008). *TPSDIG: A program to digitize images*. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, New York. 53, 55, 95

Rohlf, F. J. and Archie, J. W. (1984). A comparison of Fourier methods for the description of wing shapes in mosquitos (Diptera: Culicidae). *Systematic Zoology*, 33:302–317. 47, 62

Rutishauser, R. (1997). Structural and developmental diversity in Podostemaceae (river-weeds). *Aquatic Botany*, 57:29–70. 92

Schmittbuhl, M., Allenbach, B., Le Minor, J.-M., and Schaaf, A. (2003). Elliptical descriptors: some simplified morphometric parameters for the quantification of complex outlines. *Mathematical Geology*, 35(7):853–871. 47, 62

Schmittbuhl, M., Le Minor, J. M., Taroni, F., and Mangin, P. (2001). Sexual dimorphism of the human mandible: demonstration by elliptical fourier analalysis. *Int J Legal Med*, 115:100–101. 62

Scholz, H. (2003). Taxonomy, ecology, ecomorphology, and morphodynamics of the Unionoida (Bivalvia) of Lake Malawi (East-Africa). *Beringeria*, 33:1–86. 1, 48, 62

Scholz, H. and Hartman, J. H. (2007a). Fourier analysis and the extinction of unionoid bivalves near the cretaceousäìtertiary boundary of the western interior, usa: Pattern, causes, and ecological significance. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 255:48–63. 1, 62

Scholz, H. and Hartman, J. H. (2007b). Palaeoenvironmental reconstruction of the Upper Cretaceous Hell Creek Formation of the Williston Basin, Montana, USA: implications from the quantitative analysis of unionoid bivalve taxonomic diversity and morphologic disparity. *Palaios*, 22:24–24. 1, 3, 62

Scholz, H. and Scholz, A. (2007). Comparison of traditional morphometrics, elliptical Fourier analysis, and sliding semi-landmark method on unionoid bivalves from the Pliocene–Pleistocene Koobi Fora Formation of the Turkana Basin, Kenya. *Beringeria*, 37:161–174. 3, 47

Schwartz, M. (2006). R Graph Gallery (54) barplot2 : Enhanced Bar Plots. http://addictedtor.free.fr/graphiques/RGraphGallery.php?graph=54. Accessed 10 November 2008. 140

Seilacher, A. (2007). *Trace Fossil Analysis*. Springer-Verlag. 94

Sengupta, D. P., Sengupta, D., and Ghosh, P. (2005). Bilaterally symmetrical Fourier approximations of the skull outlines of temnospondyl amphibians and their bearing on shape comparison. *Journal of Biosciences*, 30(3):377–390. 62

Simone, L. R. L. (2006). *Land and Freshwater Molluscs of Brazil*. Museu de Zoologia Universidade de São Paulo, São Paulo, Brazil. 194, 195

Slice, D. E. (2007). Morpheus et al. http://www.morphometrics.org/morpheus.html. Accessed 18 November 2008. 103

Smith, D. A. (1988). Radular kinetics during grazing in *Helisoma trivolvis* (Gastropoda: Pulmonata). *Journal of Experimental Biology*, 136:89–102. 94

Smith, D. G. (2001). *Pennak's Freshwater Invertebrates of the United States*. Wiley, 4th edition. 94

Stransky, C. (2002). Otolith shape analysis of irminger sea redfish (sebastes mentella): preliminary results. *Journal of Marine Science*. 62

Strayer, D. L. and Jirka, K. J. (1997). *The Pearly Mussels of New York State*. The New York State Education Department. 193, 196, 197, 198, 199, 200

Sweet, A. R. (2006). Palynological report on 14 outcrop samples from the Das Goods Cretaceous/Tertiary contact section; Slope County, North Dakota (NTS071-I-08). Technical report, Geological Survey of Canada. x, 6, 10, 11

Tanaka, H. (1999). Numerical analysis of the proximal humeral outline: bilateral shape differences. *American Journal of Human Biology*, 11:343–357. 62

Tangchaitrong, K., Messer, L. B., Thomas, C. D. L., and Townsend, G. C. (2000). Fourier analysis of facial profiles of young twins. *American Journal of Physical Anthropology*, 113:369–379. 62

Tort, A. (2003). Elliptical Fourier functions as a morphological descriptor of the genus *Stenosarina* (Brachiopoda, Terebratulida, New Caledonia). *Mathematical Geology*, 35(7):873–885. 62

Tort, A. and Finizola, A. (2005). The buried caldera of Misti volcano, Peru, revealed by combining a self-potential survey with elliptic Fourier function analysis of topography. *Journal of Volcanology and Geothermal Research*, 141:283–297. 62

Tracey, S. R., Lyle, J. M., and Duhamel, G. (2006). Application of elliptical Fourier analysis of otolith form as a tool for stock identification. *Fisheries Research*, 77:138–137. 62

Vaughn, C. C., Spooner, D. E., and Hoagland, B. W. (2002). River weed growing epizoically on freshwater mussels. *The Southwestern Naturalist*, 47(4):604–605. 92

Wächtler, K., Dreher-Mansur, M. C., and Richter, T. (2001). Larval types and early postlarval biology in Naiads (Unionoida). In Bauer, G. and Wächtler, K., editors, *Ecology and Evolution of the Freshwater Mussels Unionoidea*, pages 93–125. Springer. 1

Watters, G. T. (1993). Some aspects of the functional morphology of the shell of infaunal bivalves (Mollusca). *Malacologia*, 35(2):315–342. 48

Winston, J. E. (1999). *Describing Species*. Columbia Unversity Press. 3

Zelditch, M. L., Swiderski, D. L., Sheets, H. D., and Fink, W. L. (2004). *Geometric Morphometrics for Biologists: A Primer*. Elsevier Academic Press, London. 3, 4, 47, 56