**Title:** A new fully automated approach for aligning and comparing shapes

Doug M. Boyer[1,3], Jesus Puente[2], Justin T. Gladman[3,4], Chris Glynn[5], Sayan Muhkerjee[5-7], Gabriel S. Yapuncich[1], Ingrid Daubechies[7]

[1]Department of Evolutionary Anthropology, Duke University
[2]Program In Applied and Computational Mathematics, Princeton University
[3]NYCEP, New York Consortium in Evolutionary Primatology
[4]PhD Program in Anthropology, CUNY Graduate Center
[5]Department of Statistical Science, Duke University
[6]Department of Computer Science, Duke University
[7]Department of Mathematics, Duke University

**Short Title:** Automated 3D Geometric Morphometrics

**Keywords:** *auto3dgm*, minimum spanning tree, R-package, iterative closest points, morphological disparity, transformational homology

**Manuscript breakdown**
**Text pages:** 42 (12 font, Times New Roman, double-spaced)
**References pages:** 11 (12 font, Times New Roman, double-spaced)
**Figures:** 12
**Tables:** 6
**Supplemental Figures:** 3
**Supplemental Tables:** 4

**Abstract**

Three-dimensional geometric morphometric (3DGM) methods for placing landmarks on digitized bones have become increasingly sophisticated in the last 20 years, including greater degrees of automation. One aspect shared by all 3DGM methods is that the researcher must designate initial landmarks. Thus, researcher interpretations of homology and correspondence are required for and influence representations of shape. We present an algorithm allowing fully automatic placement of correspondence points on samples of 3D digital models representing bones of different individuals/species, which can then be input into standard 3DGM software and analyzed with dimension reduction techniques. We test this algorithm against several samples, primarily a dataset of 106 primate calcanei represented by 1,024 correspondence points per bone.

We compared results of our automated analysis of these samples to a published study using a traditional 3DGM approach with 27 landmarks on each bone. Data were analyzed with *morphologika*[2,5] and PAST. Results show strong correlations between principal component scores, similar variance partitioning among components, and similarities between the shape spaces generated by the automatic and traditional methods. While cluster analyses of both automatically generated and traditional datasets produced broadly similar results, there were also differences. Overall these results suggest to us that automatic quantifications can lead to shape spaces that are as meaningful as those based on observer landmarks, thereby presenting potential to save time in data collection, increase completeness of morphological quantification, eliminate observer error, and allow comparisons of shape diversity between different types of bones. We provide an R package for implementing this analysis.

## Introduction

As the theme of this volume is the application of three dimensional (3D) geometric morphometrics (GM) to functional morphology, there is little need to convince most readers about the importance of morphological studies to evolutionary and developmental biological research. However, the utility of detailed morphological information in such research has become increasingly questioned (see Springer et al. [2013] comment on O'Leary et al. [2013a, b]). Therefore, we would like to emphasize that patterns of phenotypic variation (including morphology) among biological structures form the basis for understanding gene function (e.g., Morgan, 1911; Abzhanov et al., 2006), developmental mechanisms (e.g., Harjunmaa et al., 2012), ecological adaptation (e.g., Losos, 1990; Frost et al., 2003), and evolutionary history (e.g., Leakey et al., 1964;

Ostrom, 1975; Gingerich et al., 2001). Given its importance in a diverse set of biological disciplines, we believe that morphological information remains highly relevant to scientific discovery and advancement.

Since the Modern Synthesis of Evolutionary Theory was reached in the 1940s and evolution was appropriately re-defined in its most basic population-genetic context, genomic approaches to studying evolution have exploded. In part, this sea change is a result of increasingly available data and improving computational power. Ever more comprehensive and rapid assessments of genetic variation have been possible as a result (Venter et al., 2003). Since the late 1980s, large-scale automated genomic analyses have flourished and a great deal is now known about genotypic variation (McVean et al., 2005; Houle et al., 2010). Genetic data are even accessible from remains of extinct organisms such as subfossil lemurs (Orlando et al., 2008) and Neandertals (Green et al., 2010).

The utility of morphology is now questioned, in part, because the ability to analyze morphological data has progressed much more slowly than the ability to analyze genomic data. However, there is a call from some evolutionary biologists for the collection and analysis of high-dimensional phenotypic data (Houle et al., 2010) in an analogous high-throughput and automated fashion. This perspective proposes that the utility and information content of genetic data will only reach its fullest extent once data on associated phenotypes can be analyzed at equivalent rates and scales. Ideally, increasing availability of phenomic data would promote comprehension of how the interaction between phenotypic variation and the environment is mediated by the genome and how selective pressures on the phenome are transferred to the genome. Reflecting the perceived importance of such data, the field of phenomics has recently been defined as

that endeavoring to acquire high-dimensional phenotypic data on an organism-wide scale (Houle et al., 2010). Although phenomics is defined in analogy to genomics, the analogy is misleading in one respect. We can come close to characterizing a genome completely but not a phenome, as the information content of phenomes dwarves genomes and is heavily influenced by the mode, tempo, duration, and timing of its observation and quantification (Houle et al., 2010).

By itself, variation in morphological structure (a component of phenomic variation) has higher dimensionality than variation in the genome, which makes it exponentially more difficult to quantify in a meaningful way (e.g., Boyer et al., 2011). This is not to say that significant advances in analysis of morphology are impossible or that the field of morphometrics has stagnated. As emphasized and demonstrated by work in this volume, new and more sophisticated approaches are being developed. More sophisticated statistical contexts (Nunn, 2011) are available thanks to improved computing power and flexible open-source coding languages (Orme et al., 2011; R Coding Team, 2012). Additionally, there is growing automation of shape quantification based on new variations of methods for spreading semi-landmarks over a 3D surface model (Bookstein, 1997; Bookstein et al., 1999; Bookstein et al., 2002; Perez et al., 2006; Harcourt-Smith et al., 2008; Mitteroecker and Gunz, 2009). However, 3D shape analyses are generally tied to at least two-user determined landmarks (Polly and MacLeod, 2008), and 3DGM analyses do not appear to be very meaningful without four or more (Gunz et al., 2005; Wiley et al., 2005). As a result, these approaches continue to have many of the same limitations as morphological studies from 30-40 years ago. Part of the problem is sample size; in most cases the number of measurements, and the sample sizes per study have

changed little (compare Berge and Jouffroy [1986] to Moyà-Solà et al. [ 2012] – though statistical analyses are more sophisticated in the more recent study, there are no substantial differences in measurement complexity or sample sizes in these two studies almost 30 years apart). Other principal limitations to the current traditional approach to morphological studies include: 1) subjectivity/observer-error in interpretation and measurement, 2) time intensiveness for generating large datasets, 3) sparse and potentially incomplete and/or biased representation of specimen morphology and sample variation, and 4) limited accessibility of information encapsulated in morphology due to lack of widespread researcher expertise. All restrictions stem from the necessity that researchers must directly observe, interpret, and actively measure (or mark) every specimen of a study. These limitations may explain why genetic data currently provide a more statistically powerful approach to certain evolutionary questions, and also why questions that can be addressed only by morphology (e.g., what physical traits are functionally beneficial for a certain behavior?) are often less thoroughly examined or appear more controversial despite long histories of analyses.

As discussed by MacLeod et al. (2010), in order to make the study of morphology less of a "cottage industry" and bring it to a new level of objectivity, standardization, efficiency, and accessibility, we should seek more automation in the determination of patterns of morphological similarity and difference. Several researchers (Lohmann, 1983; MacLeod, 1999; Polly and MacLeod, 2008; Sievwright and MacLeod, 2012) have worked to develop techniques that minimize assumptions involved in measuring shape similarity. Initiatives for "automated taxonomy" exist (Weeks et al., 1999; MacLeod, 2007) and have had some degree of success. However, all of these automated approaches

require a “dimension reduction” in the initial analytical stages, which still necessitates that the researcher make a decision, informed by their understanding of important and “equivalent” morphological features, on how to make that reduction. Most automated work has been carried out on 2D outlines or raster-photographs. In such cases, the shape of an outline and the images in a photograph are determined by how the researcher orients the camera with respect to the specimen. Even when attempting the “same” view, two different researchers may have systematic error with respect to one another or different levels of random error in setting up specimens for photography. Furthermore, many techniques described as automated, including those for 2D objects, still require direct interaction with the study materials to determine at least one “corresponding point” common to all the shapes of the study sample (see papers in MacLeod, 2007).

Biomedical and neuroscience research pursued by computer scientists has led to some successful automated quantification procedures in 3D (Styner et al., 2006; Paniagua et al., 2012). However, these methods have been designed with a limited range of variation in mind and applied to monospecific samples. Whether these methods would have meaningful success in a sample with more substantial shape diversity among homologous objects is unknown.

In order to begin testing the limits on the degree to which, and the questions for which, shape analysis can be automated towards a scientifically meaningful end, we present a new fully automated algorithm for aligning digital 3D models of bones and placing landmarks comprehensively on them. We also provide an R package application to promote its testing and use by other researchers. This method builds conceptually on a previously published approach (Boyer et al., 2011) where it was shown that a

151 superficially similar algorithm can 1) reasonably match corresponding points on different
152 instances of the same bone (represented by different individuals and species), 2) estimate
153 shape differences that allow classification of shapes to species with accuracy comparable
154 to, or better than, user selected landmarks on the same specimens, and 3) allow for the
155 entertainment of different "correspondence hypotheses" based on the morphocline (or
156 "path") that is assumed to connect shapes in the dataset. Operationally, the method of
157 Boyer et al. (2011) finds several hundred candidate alignments between conformally-
158 flattened representations of two objects. Each initial alignment is "improved" using a thin
159 plate spline to align automatically identified extremal points (points of high local
160 curvature – i.e., "type II landmarks"). These mappings are then applied to unflattened
161 versions of the two objects and a continuous Procrustes distance is computed (Lipman
162 and Daubechies, 2010). The mapping that results in the minimum continuous Procrustes
163 distance is treated as the best mapping among the many candidate maps. This minimum
164 distance mapping was found to usually represent a biologically meaningful alignment
165 according to criteria 1 and 2 described above.

166 Despite its successes, the method presented by Boyer et al. (2011) has several
167 shortcomings: 1) since correspondences used to determine shape differences are purely
168 pairwise and not transitive, there is an inconsistent template for biological
169 correspondence relating all pairs of shapes in the dataset); 2) the conformal flattening
170 procedure of the analysis limits its application to "disc-type" shapes with an open end
171 (like the tooth crowns or ends of long bones of that dataset); and 3) the MATLAB®
172 application for the analysis is difficult to work with, lacks good visualization tools, and
173 does not yield output that can be widely employed in other analytical procedures.

We overcome these limitations in the new algorithm presented here, which we have developed into an R-package called *auto3dgm*. One of the most exciting prospects of *auto3dgm* is its potential to help quantify morphology more comprehensively and equably (if not exhaustively). It has long been acknowledged that measurements of select characters are less meaningful than more comprehensive approaches:

> "Direct determination of rate of evolution for whole organisms, as opposed to selected characters of organisms, would be of the greatest value for the study of evolution. Matthew wrote, nearly a generation ago (1914), 'to select a few of the great number of structural differences for measurement would be almost certainly misleading; to average them all would entail many thousands of measurements for each genus or species compared.'" (Simpson, 1944: pg.14)

> "Another level of description -of entire surface regions, or of volumetric elements, or of qualitative aspects of structures rather than structures themselves- may in some instances be most meaningful (Roth, 1984, 1991) and bring us closer to identifying the biological processes of interest. Hence the appeal and utility of methods of comparison that interpolate between landmark points, such as D'Arcy Thompson's transformation grids" (Roth, 1993: pg. 53)

Matthew's implied perspective was that increasing the number of measurements would be useful (though impractical) and would approach a representation of the "total taxonomic distance." This taxonomic distance is sometimes referred to as "morphological disparity" and may allow meaningful discussion of the amount, rate and pattern of evolution among a sample of species in certain settings. A greater amount of morphological difference between corresponding and homologous structures is assumed to relate to the amount of evolutionary change that has occurred in the compared taxa since they diverged from their common ancestor. This idea is reflected in the numerical taxonomy movement (Sokal, 1966; Sneath and Sokal, 1973).

A wealth of careful, mathematically-rooted consideration has been aimed at these premises over the years. It has been effectively argued that it is actually impossible to generate a generalized comprehensive view of the total phenetic distance between specimens or taxa (Bookstein, 1980; Bookstein, 1994; MacLeod, 1999). In fact, Bookstein (1991; 1994) argues that morphometrics is purely about documenting covariance among biological forms, stating that morphometric methods are neither suited for "the computation of 'magnitude' of shape change nor for the clustering of individual specimens according to degree of similarity of shape" (Bookstein, 1994, p.205). MacLeod (1999) explains the insufficiency of morphometrics in this regard, saying: "All morphological disparity estimates published thus far represent indices that are inextricably tied to particular methods of morphological representation and particular scales of morphological assessment", that "it seems…unlikely that a generalized estimate of 'morphological disparity,'…can ever be achieved." and finally that it is imperative that

218 “the morphometrician remembers the domain within which he/she operates is strictly
219 limited” (MacLeod, 1999, p.134).
220 We do not suggest the method we present fundamentally resolves any of these issues.
221 It aids in the discussion of morphological disparity because it is more objective and
222 comprehensive in its measurement of shape than previous methods. Though Bookstein
223 (1994) argues that morphometrics must be applied after homology considerations have
224 taken place, we suggest that our method can help identify an “operational homology” or
225 “biological correspondence” (Smith, 1990) more objectively.
226 Of the various types of homology discussed by evolutionary biologists and
227 paleontologists, it is relevant to review at least three different types here: these include
228 transformational, operational, and taxic homology (Patterson, 1982; Smith, 1990). It
229 would seem that transformational homology is of primary importance in an evolutionary
230 sense. It is similar to Darwinian homology (Simpson, 1961), in which features are
231 considered homologous among several taxa if they are equivalent through “descent with
232 modification” from the common ancestor. This also matches Van Valen’s (1982)
233 definition of homology as “continuity of information” through evolution. Of course,
234 comprehension of transformational homology is often fairly elusive, since the
235 morphoclines describing it can be expected to gain accuracy with a more complete fossil
236 record and an accurate phylogeny of life (Van Valen, 1982).
237 Operational homology most generally appears to refer to ontologies defining
238 biological correspondence for the sake of measurement, comparison among taxa, and/or
239 as a working hypothesis of transformational homology. What Macleod (2001, p.3)
240 describes as “geometric (or morphometric) homology (sensu Bookstein 1991)” of

geometric morphometrics can be considered as specific types of operational homologies. In a way, Thompson (1942), as also quoted by Roth (1993), reminds researchers not to forget the distinction between operational homologies and carefully tested hypotheses of transformational homology:

> "The morphologist, when comparing one organism with another, describes the differences between them point by point and "character" by "character" ....and he falls readily into the habit of thinking and talking of evolution as though it had proceeded on the lines of his own descriptions, point by point, and character by character." (Thompson, 1942, p.1036)

Finally, taxic homology is equivalent to "synapomorphy" or "symplesiomorphy" whereby similarity in morphological form (usually referred to as a "character state") of a transformationally homologous feature exhibited by a taxonomic sample of interest is thought to reflect the inheritance of that "state" from a common ancestor. Whether identified taxic homologies help elucidate phylogenetic relationships depends on whether particular character states have evolved numerous times and exhibit homoplasy, as well as whether perceptions of transformational homology are correct. When discussing features on a finer scale than whole bones or organs, hypotheses of transformational homology are usually difficult to test. When the data necessary for such tests are available (e.g., via a dense fossil record [Van Valen, 1982]) the results can be surprising.

The empirical route to homology hypotheses is a recursive one. Van Valen (1982) says that homology is "more than similarity" which means that assessment of shape

similarity *is* involved. Shubin (1994) discusses tests and evaluations of homology hypotheses, saying homology is "only indirectly related to similarity" and that "homologous features may be very dissimilar". But without an a priori phylogeny, how does one postulate homology of dissimilar features? In many cases, operational homology hypotheses are qualitatively rooted in geometric similarities even for matching dissimilar features in two taxa. For skeletal elements, operational homology (= topological correspondence) hypotheses are established by researchers physically or conceptually seriating features of specimens into morphoclines. The correspondence among end-members of the morphocline (the humeri of a whale and a bat – for instance) may be un-interpretable next to each other, but will have more definitive operational homologies if they are compared through the intermediate forms along a taxonomically rich seriated sample. Of course, this task is aided by information beyond the geometry of isolated bones: the position and orientation of the bone in the complete skeleton is also known and used (i.e., cues from "type I" landmarks). Different researchers may see and emphasize different aspects of shape, and samples with different taxa will suggest different morphoclines and possibly different patterns of correspondence among end-members. As Roth (1993, p.53) says "The recognition, and operational definition, of homologous points is a non-trivial problem (Jardine, 1969; Smith, 1990), and one not necessarily with unique solutions." Furthermore, different skeletal element sets from the same taxonomic sample may seriate in morphoclines with different taxonomic orderings. For example, the calcaneus bone of a tarsier has the most extreme form in comparison to any sample of primate species, whereas the astragalus bone of tarsiers can be described as roughly intermediate between that of certain anthropoid and strepsirrhine primates). For a

given taxonomic sample, a consideration of which bones arrange in morphoclines with similar orderings of taxa (and thereby present congruent pictures of operational homology) aids in formulating phylogeny hypotheses. Cladistic parsimony analyses are conceptually related to this practice. Clearly, determination of operational homology is at least partly based on a qualitative consideration of geometric similarity and morphoclines among samples. Our automated procedure, which considers the total surface of bones and the pattern of distances between them, can be implemented toward this end.

Because *auto3dgm* determines feature correspondence objectively (algorithmically) and more comprehensively, it can assess morphological differences in a way that suffers from less measurement sensitivity. This decreased sensitivity makes the shape quantifications of one bone or 'part' more easily generalizable to other parts compared with previous methods (as we will demonstrate with an example). Ultimately, this allows greater insight into patterns in, and the generation of, morphological disparity through the evolutionary process.

## Materials and Methods

*Institutional abbreviations.*— AMNH, American Museum of Natural History, New York, NY; CGM, Egyptian Geological Museum, Cairo, Egypt; DPC, Duke Lemur Center Division of Fossil Primates, Durham, NC; GU, H.N.B Garhwal University, Srinagar, Uttarakhand, India; IGM, Museo Geológico del Instituto Nacional de Investigaciones Geológico-Mineras, Bogotá, Colombia; IRSNB, Institut Royal des Sciences Naturelles del Belgique, Brussels, Belgium; KU, Kyoto University, Kyoto, Japan; MCZ, Museum of Comparative Zoology, Harvard University, Cambridge, MA;

MNHN, Muséum National d'Histoire Naturelle, Paris, France; NMB, Naturhistorisches Museum Basel, Basel, Switzerland; NMNH, Smithsonian Institution National Museum of Natural History, Washington, D.C.; NYCEP, New York Consortium in Evolutionary Primatology, New York, NY; SBU, Stony Brook University, Stony Brook, NY; SDNHM, San Diego Natural History Museum, San Diego, California; SMM, Science Museum of Minnesota, Minneapolis, MN; UCM, University of Colorado Museum of Natural History, Boulder, CO; UCMP, University of California Museum of Paleontology, Berkeley, California; UK, University of Kentucky, Lexington, KY; UM, University of Michigan, Ann Arbor, Michigan; USGS, U.S. Geological Survey, Denver, Colorado.

*Samples.*—We utilize four samples of surface meshes generated from either microCT or laser scans to test *auto3dgm*. Table 1 is a taxonomic list for each dataset with sample sizes per genus (supplemental tables 1-3 give the specimen numbers for each sample). The first sample includes 106 calcaneal bones of 67 genera, and is the exact sample used by Gladman et al. (2013). We test our method by running the same analyses on this sample as Gladman et al. (2013) and compare the results. *auto3dgm* produces landmark datasets that can be analyzed in a manner identical to traditional user-collected landmark datasets. The second sample is comprised of 80 astragali that we analyze and compare to a subset of 80 calcanei from the first sample. The third sample is of 49 distal phalanges representing fossil and extant taxa to demonstrate the method on a bone with a "different quality" of shape variation. Distal phalanges are basically cone-shaped with fewer consistent "feature points" than astragali or calcanei, but exhibit a range of forms from "blade-like" (falcular) to "spatulate" (unguliform) (Fig. 1). Therefore, each bone is less complex, but the range of variation across the sample remains substantial. The fourth

sample also represents astragali and overlaps the second, but includes additional specimens and species (Table 1). This sample is used to demonstrate the semi-supervised alignment procedure of the R-package "Shape_Alignment".

*Sample processing.*—Very little pre-processing is required for *auto3dgm*. Surface files should be in the Open file format (.off) and of sufficient resolution to capture all surface features of interest. It should be noted that the .off format is closely related to more widely known Stanford Polygonal Mesh (.ply) format. The free software MeshLab can be used to convert .ply files to .off files, as well as batch converters (see http://www.stat.duke.edu/~sayan/3DGM/index.shtml). If made from CT scans, the surfaces must be carefully checked and cleaned so they have no internal vertices. Virtually no processing is required for laser-scan generated data aside from smoothing or filling holes in the mesh.

The majority of surface files in our datasets were generated by microCT scanning. Details on both laser- and microCT scanning parameters of the astragalus and calcaneus specimens have been reported on previously in appendices and supplementary tables (Boyer and Seiffert, 2013; Boyer et al., 2013). The distal phalanx dataset is new.

*auto3dgm input and output files.*— The method demonstrated here was developed by Puente (2013) as a major component of a Ph.D. thesis and the mathematical details can be found there. Additional technical papers focusing on mathematics are forthcoming (Puente and Daubechies, in preparation). The input files for the routine are a set of surface mesh files in .off format. The user must also supply a set of "low resolution" versions of the mesh files that will be used by the algorithm to generate summary images. Downsampling of mesh files can be accomplished with visualization programs such as

Meshlab (Cignoni et al., 2012), Avizo (Visualization Sciences Group, 2009), and Geomagic (3D Systems Inc., 2013).

The outputs include 1) an "alignment file", which is a "multi-surface".off file that includes displays of user-supplied low resolution renderings of all specimens shown in the algorithm-determined optimal alignment (Fig. 2); 2) an "MDS file," which is another multi-surface file that embeds the same aligned renderings of specimens in a coordinate space determined by a multi-dimensional scaling (MDS) analysis of the distance matrix of aligned specimens (again for visualization purposes) (Fig. 3); 3) a "scaled".txt file with all of the coordinate data for all specimens scaled to the same centroid size, that can be loaded into, visualized, and analyzed in *morphologika*$^{2.5}$ (O'Higgins and Jones, 2006); 4) an "unscaled".txt file with all of the coordinate data for all specimens at the scale of the original input files which can also be analyzed in *morphologika*$^{2.5}$; and 5) a folder with copies of all the original input files, the coordinates of which have been multiplied by the rotation matrix used in the final alignments.

The purpose of the alignment file is to check for errors generated by the alignment algorithm. If errors are found, we provide functions allowing for a semi-supervised repair, though most likely such errors indicate insufficient degrees of incremental variation in the dataset (i.e., the morphological gaps between a single specimen, or certain groups of specimens, and the rest of the dataset are too large). The purpose of the MDS file is to provide a quick view of the phenetic affinities suggested by the matrix of continuous Procrustes distances between specimens in the analysis. The *morphologika*$^{2.5}$ file allows further analyses of the sample of shapes as aligned by the method. Finally, the aligned versions of the input files provides data for users who wish to standardize

379 alignment before taking measurements that are sensitive to orientation [like relief indices
380 or other topographic variables measured on teeth (Bunn et al., 2011)], or who wish to use
381 the images for figure generation.

382 *Pseudolandmarks and alignment.*— In order to facilitate adoption of this method by
383 3DGM community, this protocol represents and aligns pairs of surfaces with landmark-
384 like feature points. We say these are "landmark-like" because we represent each bone
385 with same number of points (in this study 1,024 points per bone are used, but the
386 algorithm can be set to use more or fewer), and by the final stage of the algorithm each
387 point has a fairly consistent biological identity across all bones of the sample. Each of
388 these points is therefore analogous to an observer-placed landmark. On the other hand,
389 they are not identified based on any of the criteria for determining type I, II, or III
390 landmarks (Zelditch et al., 2004), or even semi-landmarks (Bookstein, 1997;
391 Mitteroecker and Gunz, 2009), and therefore are dubbed "pseudolandmarks" here. Other
392 recent fully automated algorithms (Boyer et al., 2011) do not generate a globally
393 consistent mapping of a set number of points across all specimens of a dataset, and this
394 limits their utility for certain applications.

395 *Major computational steps.*— There are at least four important ingredients to the
396 protocol. The first is re-sampling of surface coordinates to a specified standard number of
397 points (Fig. 4). This is done using approaches that evenly spread points over the surface
398 (Eldar et al., 1997). Once a new sample of bones with a standard number of evenly
399 spread coordinates has been generated, the algorithm attempts to align each pair of bones
400 using an iterative closest points (ICP) procedure (Besl and McKay, 1992). We avoid
401 incorrect local minima known to plague ICP by having our algorithm assume that

principal axes of variation will tend to be homologous in some sense between bones. After computing the principal axes of variation in points for two surfaces, the algorithm attempts alignments where the first principal axes are aligned in one of two possible ways (Fig. 5). There are a total of eight ways to align the first through third principal axes, and these eight possible alignments are our starting points for ICP. They can be run simultaneously, and an approximation of the global minimum Procrustes distance can be found quickly (especially if a low number of pseudolandmarks are used). Of course, a major advantage of the method is the ability to include large numbers of data points on the surface. To resolve the conflict between processing speed and accuracy, our algorithm performs initial alignments with highly down-sampled surfaces using several hundred points (the exact number of pseudolandmarks is a user-defined parameter). Next, more densely sampled surfaces are rigidly transformed to match their down-sampled counterparts, so that only the final "tweaking" of the alignment has to be performed on the full-resolution surface file.

Since the best alignment is found by computing a Procrustes distance, a Procrustes distance matrix is available for computation of a minimum spanning tree (MST) for the sample. The MST connects all cases in the dataset using the shortest edge length possible and is a unique solution, except in datasets where several cases are exactly equidistant from each other. Though not all points will be connected to their nearest neighbors in such a tree, most connections represent a joining of nearest neighbors for one of the cases involved. In datasets with high degrees of shape diversity, it is virtually guaranteed that between certain pairs of bones, the minimum Procrustes alignment will be a biologically meaningless arrangement. However, because the pairs connected by the segments of

425 MST are among the shortest in the distance matrix, they are the most likely to be
426 biologically meaningful and/or precise alignments. Therefore, instead of attempting to
427 directly align pairs of shapes that have a relatively large Procrustes distance separating
428 them, alignments between such pairs are generated by propagating alignments between
429 intermediate shapes, ultimately allowing very different shapes to be aligned indirectly
430 (Fig. 6).

431 *Parameters that must be specified.*—Before the "automated part" of our algorithm can
432 begin, the user must choose values for three parameters. Varying values of these
433 parameters (see below), improves fidelity, detail, and accuracy of alignment in the one
434 direction, and speed of calculation in the other. It may be possible to determine optimal
435 values for these parameters in more or less general conditions by incrementally
436 modifying them, re-running analyses, and checking the results. We have not yet done this
437 systematically. The parameters to be set include 1) the number of points used to represent
438 shapes in the low resolution version of the alignment; 2) the number of points to
439 represent shapes in the high-resolution, or final version of the alignment; and 3) the
440 number of principal alignments (usually this number is set to the eight possible
441 combinations of the alignments along the first three principal axes, but additional random
442 principal alignments can be chosen). In the first three samples we evaluate in this study,
443 we use the following pairs of point numbers: Calcaneus dataset of 106 specimens:
444 initial=150 points, final=1,024 points, 8 principal alignments; paired calcaneus and
445 astragalus datasets: initial=256 points, final=1,024 points, 12 principal alignments; distal
446 phalanx dataset: same as for paired astragalus and calcaneus. In the fourth dataset we use

far fewer points in order to generate problematic alignments: initial = 32, final = 64, 8 principal alignments.

*Fixing errors in the alignment protocol.*—Because it is sometimes the case that at least one specimen is mapped into the MST with an incorrect alignment, it is important to provide options for correcting the problem.

1. Usually such problems stem from insufficient number of initial points (first parameter above). Thus, the first step is to try re-running the initial steps of the algorithm with slightly greater numbers of points per file. However, the problem can also stem from the lack of an adequately similar partner shape in the dataset (from the perspective of its orientation and articulation in the skeleton). This shape represents an “island shape” for which the best geometric alignment (that with the smallest Procrustes distance) to any other shape is a biologically "incorrect" alignment. This property does not guarantee a bad alignment since it may not connect to its nearest neighbor in the minimum spanning tree, but it often allows one. However, it is possible that there are still some shapes in the sample with which the island shape(s) will correctly align. We do not currently have an automated protocol for discovering such shapes, if they exist. We have implemented two different protocols for fixing alignment problems. If there is a **single** misaligned shape: We allow the user to display the results of direct alignments of the island shape to each of the other shapes in the sample using the function branch_pw_distances.r in the R-package. If there are *n* specimens in the sample, this function creates *n-1* multi-surface mesh files. There is one file for every corresponding pair between the island shape and the remaining shapes.

Even if *n* is very large, these can be visually scanned quickly to find a correct alignment. Tiling the multiple files in Meshlab or Aviso is one possible way of quickly arriving at the correct alignment when *n* is large. If the user finds a shape to which the island shape correctly aligns, the MST is re-calculated without the island shape, the global alignment of the remaining shapes is double-checked, and the island shape is connected to the new MST through its successfully aligning partner. The analysis is then completed in the usual way. If there are multiple specimens with which the island shape correctly aligns, the user can choose which to use as the connecting shape, though it seems logical to choose that with the smallest Procrustes distance to the island shape. The pairwise output files from branch_pw_distances.r orders the shape correspondences by their Procrustes distance. The ordering of correspondence will be in the name of the files for clarity.

2. If there are **multiple** island shapes, a more involved protocol is required, because there may be several groups of consistently aligned shapes (Fig. 7). The general problem is that the analysis may return a result in which certain branches are internally consistent, but are misaligned with respect to other such branches. It is therefore necessary to have a protocol allowing the user to chop apart these branches and stick them back together in a way that ensures a globally consistent alignment. The work-flow described below is provided by the example file "alignFix.r" and is available on the first author's website. Documentation that accompanies "alignFix.r" guides the user through a sample problematic dataset

492 (our dataset 4). Users should then be able to edit the code of “alignFix.r” to suit
493 their datasets.

494 a. Observe misaligned regions using alignment and map files (Figs. 7A and
495 7B) together.

496 a.i. If only one misaligned file is observed, follow the procedure
497 described above.

498 a.ii. If **more than one** misaligned file is observed:

499 a.ii.1. Record the alignment numbers of the misaligned
500 files.

501 a.ii.2. View the MDS graph showing the MST
502 connections on points labeled by the alignment number
503 they represent.

504 b. Using the map file and the MST, figure out how many "groups" of
505 misaligned files exist, and how many specimens in each group,
506 and **record this information.**

507 b.i. Specify all "groups greater than 2" (three or more files that are
508 correctly aligned to each other, but not to surrounding shapes) as
509 "groups to analyze separately", since a MST will need to be re-
510 computed within each group.

511 c. For “*b.i.*”, a separate alignment analysis is run on each group of three or
512 more that were internally consistent and all the necessary information is
513 saved (Fig. 7C).

514 d. Now the user must decide how to "re-connect" the separate sub-groups.

d.i. First attempt to analyze all of the shapes in non-connected segments of the minimum spanning tree. For example, with four groups (A, B, C, and D), it is possible that only one will end up connecting to the other three through the MST. If both A, C and D connect to B in the original analysis, and are misaligned with respect to B, it is possible that with B excluded, A, C and D will align correctly. If this is true, skip to "*d.iv.1*" of this description. If not, go to number "*d.ii*."

d.ii. For cases in which the set of non-connecting groups is still an incorrect alignment, the non-connecting groups should be compared in a pairwise fashion. For instance A-C, A-D, and D-C should each be analyzed separately. It is possible that some of these will have correct alignments. If more than two of these are correct, a decision will have to be made on which two to merge, since it has already been demonstrated that all three cannot be. We would suggest merging the two that result in the biggest difference in the number of specimens represented in the final two groups, since this makes the subsequent task of searching for a correct alignment between groups that are not correct via their MST easier. At this stage, the goal should be to merge as many isolated groups together as possible in order to reduce computational demand in the next steps. Ultimately, the user can decide which groups to merge.

d.iii. After managing the isolated but internally consistent segments of the original MST (groups A, C and D above), the user needs to **find a "correct" connection between the isolated groups that were misaligned with respect to each other through the original MST**. Some remnant of the original MST will still be preserved, which can be called the “base tree” (group B in our example). Attempting to reconnect the isolated groups to the base tree using the minimum distance pair will likely generate misalignments, since the MST connections were wrong in the original analysis. However, as MST connections often only represent nearest neighbors for one of the two connected cases, there is still a possibility that one of the cases involved in the incorrectly aligning connection between the base tree and another segment was not connected to its nearest neighbor. This makes it important to look at the minimum distance pairs of the isolated groups and the base tree.

d.iv. Assuming the minimum distance pair is still a misalignment, a protocol for checking alignments between particular shapes in each group must be implemented. This again utilizes the function branch_pw_distances.r.

d.iv.1. The user has the option to check all alignments. The output is *n x m* "summary alignment files" in which *n* is the number of specimens in one group and *m* is the number in

the other group being searched. Each file shows one shape from the group with *n* with one of the *m* specimens of the second group (Fig. 7E). The output files are labeled according to minimum Procrustes distance, so that the first compared specimens are nearest neighbors. The user can then easily identify the correctly aligning pair that also has the minimum Procrustes distance (since there may be more than one correctly aligning pair).

d.iv.2. This process should be repeated for all segments that could not be merged. If there were three remaining segments (e.g., a base tree B, an A-C group and D), there will likely be an option of whether to link each tree to one of two others. We would suggest this linking be done using the option when the Procrustes distance between the linking pair is minimized.

d.iv.3. The user can also opt to only compare specific specimens from one group to specific specimens in the other.

d.v. Finally, all groups are re-aligned using a tree that represents each separate MST connected along user-specified pathways in "*d.iv.2*" This should result in correct alignments for all bones in the sample (Fig. 7G).

583 If the user determines successful alignments between groups of island shapes are
584 impossible, there are two options: 1) remove any island shape groups from the analysis
585 (particularly if their inclusion does not directly address the main questions of the
586 analysis); or 2) add more shapes with the hope of bridging distances between island
587 shapes.

588 *Getting the code for running analyses.*— The R package we developed is called
589 *auto3dgm*. At the time of publication *auto3dgm* has been submitted to CRAN for review,
590 and will ultimately be accessible from their repositories. Until then, *auto3dgm* can be
591 downloaded at www.dougmboyer.com or
592 http://www.stat.duke.edu/~sayan/3DGM/index.shtml. The sample/instructional file for
593 fixing misaligned shapes, alignFix.R, is not part of the R-package itself and will not be
594 available on CRAN. It can however be downloaded from the personal websites
595 mentioned above. Documentation for the packages can be found at these sites as well.

596 *Comparison to results from traditional landmarks.*—In order to maximize our ability
597 to compare and contrast shape information provided by our pseudolandmarks with
598 traditional geometric morphometric datasets, we used the same sample and performed the
599 same analyses on the pseudolandmarked dataset as Gladman et al. (2013) conducted
600 using 27 landmarks and traditional 3DGM techniques.

601 First, the 3D pseudolandmark coordinate-scaled output file from our algorithm was
602 imported into *morphologika*[2.5]. We then ran a General Procrustes Analysis (GPA) with
603 reflections enabled, followed by a Principal Components Analysis (PCA) with "Full
604 Tangent Space Projection" checked for Calculation Options and "Eigenvalues" and "PC
605 Scores" checked for Printing Output Options. The results were saved as a .csv file that

included the PCA output, along with the raw Procrustes distance data in the form of 3D coordinates for each landmarked individual. In *morphologika*$^{2.5}$, the cloud of 1,024 landmarks was visualized and the morphospace of the PC axes was explored. In the traditional 3DGM analysis of this sample, Gladman et al. (2013) added wireframes to the landmarks in order to directly visualize shape changes. Due to the number of pseudolandmarks used by *auto3dgm*, wireframes are currently impractical, but shape changes can easily be observed from transformations of the densely packed pseudolandmarks. All Principal Components (PCs) were examined in *morphologika*$^{2.5}$ by tracking changes in the cloud of 3D landmarks between the extreme morphospace of each axis. The amount and nature of variation represented by these axes in the 1,024 pseudolandmark dataset was then compared to results from the 27 user-determined landmarks of the Gladman et al. (2013) analyses.

Gladman et al. (2013) also used analyses of "generic" means for cluster analyses in their study of the 106 calcanei sample used here. They felt that averaging the few individuals for each genus helped control for any extreme variation that might otherwise dominate the small samples being used to represent extant genera. We replicated their approach with the pseudolandmark coordinates here. Extant genera represented by more than one individual were averaged into a single genus representative (Table 1). As in Gladman et al. (2013), fossil individuals were not averaged together in the analyses. Altogether the dataset was reduced from 106 individuals to 67 generic representatives (Table 1).

In order to generate generic means, the matrix of 3D coordinate Procrustes output data (generated in *morphologika*$^{2.5}$) was imported into PAST statistical software

(Hammer et al., 2001; Hammer et al., 2006). In PAST, all individuals of a single genus were highlighted and averaged using the "Evaluate Expression" function in the "Transform" menu. "Mean (of current column)" was selected in the "Evaluate Expression" menu and then "Compute" in order to change all highlighted rows to the same averaged values. Only one of these newly averaged rows was kept in the dataset to represent a given genus. This technique can be done manually by averaging each X, Y, and Z value separately for each landmark for members of each genus, although with increasingly larger datasets this becomes untenable. Once the averaged dataset was complete, cluster analyses were run within PAST and then compared to the generic mean analyses of Gladman et al. (2013).

*Mixed bone analysis.*—It has been suggested that traditional 3DGM methods could be used to "pool information" from more than one structure (Rohlf, 2002). However, the meaning of results from such an approach is questionable, as the weight of each structure added will depend on the user's choice of landmarks, as well as the number of landmarks used to represent each bone. Furthermore, since there is no basis for collecting landmark data across bone types, it has never been possible to include multiple bone types in the same 3DGM analysis using the same landmark template. Our approach with *auto3dgm*, based on spreading landmarks evenly and selecting alignments based on overall geometric similarity, provides a solution to this problem and allows mutli-bone types of analysis. There are many questions that can be addressed if shape variation can be compared between bone types. For instance, we might wish to ask whether the astragalus has less shape diversity than the calcaneus, due to the former articulating with a greater number of bones and lacking muscular attachments as exhibited by the latter. We might

also be interested in investigating whether the degree of overall shape variation is associated with stronger phylogenetic signal (Nunn, 2011) or stronger functional signals. We performed the first "mixed bone" analysis on a sample of 80 astragali and 80 calcanei representing the same taxa (although sometimes composed of different specimens) and we compare intrinsic levels of overall shape variation.

The basic goal of such an analysis (given the questions above) is to provide a quantitative criterion for comparing size-standardized shape variation between two bones. Since regions on the surface of a calcaneus do not "biologically correspond" in any way to regions on the surface of the astragalus, there is no need to determine a biologically meaningful regional correspondence between them. Therefore, only the most efficient geometric alignment must be established (i.e., the alignment that minimizes the Procrustes distance). However, in a mixed bone analysis, astragali will not only be compared to calcanei, they will also be compared to other astragali. Thus, for some bones in the sample, there *is* a biologically significant alignment that must be discovered before comparisons can be made.

To establish a globally transitive pseudolandmark coordinate dataset for a mixed bone sample, we first ran *auto3dgm* on the calcaneus and astragalus datasets separately to produce two sets of globally consistent pseudolandmark datasets. We then performed searches for the alignment and correspondence between an astragalus and calcaneus that exhibited the minimum Procrustes distance among all such pairs in the combined dataset using the branch_pw_distance.r function. In the second step, we were only concerned with distances since no details about the alignment mattered biologically. Once we found the mixed bone pair with the smallest geometric distance separating them, we used that

675 pair to link the MSTs of the initial analyses, creating a mixed-bone, global-
676 correspondence, 3D pseudolandmark dataset. This dataset was imported into
677 *morphologika*[2.5] and processed with GPA followed by PCA, with results exported as a
678 .csv file, and final analyses performed in PAST like the analyses above.
679 We ran statistics on four samples: 1) pairwise distances separating the calcanei, 2)
680 pairwise distances separating the astragali, 3) the combined dataset of 160 astragali and
681 calcanei, and 4) a combined dataset representing only 40 astragali and 40 calcanei (with
682 taxa matched between the two halves of the sample). We also analyzed the first two PC
683 scores of the astragalus and calcaneus separately looking at their range, variation, and
684 computing their phylogenetic signal. Phylogenetic signal was also calculated on
685 Procrustes distances from the mean for the astragalus dataset and calcaneus dataset.
686 Phylogenetic signal was calculated using *caper* (Orme et al., 2011) in R, and a tree based
687 on v3 of the primate dataset from 10k Trees (Arnold et al., 2010). Testing for
688 phylogenetic signal (Pagel's λ) required using generic means of the sample and reduced
689 the sample size from 80 individuals to 42 genus-averaged individuals.
690
691 **Results**
692 *Alignment success.*— Alignment for the calcaneal dataset of 106 bones was
693 successfully accomplished with a low resolution initial alignment of 150 points, and eight
694 principal alignments (Suppl. Fig. 1). The final high-resolution surface alignment was
695 based on 1,024 points. Successful alignment for the calcaneal dataset of 80 bones was
696 accomplished with a low-resolution initial alignment of 256 points, eight initial positions
697 based on all possible combinations along three principal axes, and a high-resolution final

surface alignment based on 1,024 points. Successful alignment for the astragalar dataset of 80 bones was accomplished with a low-resolution initial alignment of 256 points, 12 initial alignments, and a high-resolution final surface alignment based on 1,024 points (Suppl. Fig. 2).

The distal phalanx dataset was aligned using a low-resolution initial alignment of 256 points, 12 initial alignments, and a high-resolution final surface alignment based on 1,024 points (Suppl. Fig. 3). One specimen, UCMP 217919 (a fossil of unknown taxonomic affinities), had an incorrect alignment to its connecting shape in the MST (a tarsier second digit grooming claw, USNM 196477). We identified a correct alignment with SMM P77.33.517, a claw of *Plesiadapis churchilli*. This is not to say these two bones are very similar. It simply shows that it is usually possible to establish correct alignments for every bone in the sample without manually registering them to each other.

*Comparison to results from traditional landmarks.*— For the PCA of output from *auto3dgm* on individual specimens (n=106, with no genus-level averaging), the first four principal component (PC) axes account for 59.6% of the total variance. This is very close to that explained in the analysis of the same sample using 27 landmarks by Gladman et al. (2013) (Table 2). Generally speaking, major clades were well separated when plotted in morphospace, as in Gladman et al. (2013) (Fig. 8). Examination of the 3D landmark cloud in *morphologika*$^{2.5}$, and the general distribution of specimens in the scatter plots of the PCA morphospace, indicates that PC1 (34.7%) is mostly associated with the overall length and width proportions of the calcanei, with some emphasis on the distal elongation. The distally elongated and narrow-bodied calcanei of omomyiforms and some strepsirrhines dominate one extreme of the PC1 axis, while the distally shorter and

721 wide hominoid calcanei fall on the opposite extreme. This pattern matched well that
722 found by Gladman et al. (2013). Regressing PC1 scores based on manually positioned
723 landmarks against the PC1 scores from analysis of *auto3dgm* output showed high
724 correlations (Table 3). Other axes were more modestly correlated or lacked significant
725 correlations.

726 Variation found in PC2 (13.6%) captured some aspects of the "flexing" of the
727 calcaneus described by Gladman et al. (2013), although the distribution of the taxa within
728 this PC is not identical to the original description. This PC most notably varies in the
729 position of the distal margin of the ectal facet relative to the body of the calcaneus, either
730 raised dorsally off of the body or sunken plantarly. The hominoids are found on one
731 extreme, with ectal facets that sit atop of the calcaneal body, while platyrrhines are the
732 most consistent examples of calcanei with ectal facets depressed into the body. Although
733 more difficult to observe directly from the cloud of pseudolandmarks in *morphologika*$^{2.5}$,
734 there also does seem to be variation in the magnitude, although not the position, of the
735 peroneal tubercle captured in this axis.

736 The variation found in PC3 (6.7%) also resembles some of the "flexing" that has been
737 previously described, although it also includes new variation not recognized in the
738 previous traditional analyses. On the extremes for this PC axis are the hominoids
739 (excluding hylobatids), which have a pronounced proximal plantar heel process and a
740 dorsal bowing of the body of the calcaneus (giving an un-flexed appearance). At the other
741 extreme are most of the colobines (excluding only *Colobus*), which have no proximal
742 plantar heel process and have a more prominent plantar bowing (flexed appearance)
743 caused, in part, by a more prominent angulation of the body at the distal plantar tubercle.

The tradeoff in this axis is between an unflexed calcaneus driven by the presence of a plantar heel and a flexed calcaneus driven by a heightened angle at the distal plantar turbercle.

Finally, similar to PC3 above, PC4 (4.6%) also contributes to variation at the distal plantar tubercle. However, unlike the variation in PC3, the distal plantar tubercle in PC4 only gets larger or smaller in size, and there are no clear changes in the angulation at the tubercle. This PC exhibits variation most notably in the amount of proximal segment elongation and the position of the dorsal heel relative to the ectal facet. While PC1 contained aspects of distal elongation within the larger length and width proportional changes of the calcaneus, PC4 is specifically associated with the elongation of the proximal segment of the calcaneus, measured from the ectal facet to the heel. Additionally, at the extreme of the PC where the proximal segment is shortest, the dorsal heel is near level with the ectal facet, while at the elongated proximal extreme the heel is sub-level to the ectal facet. The fossil euprimates lie at the extremes for this variation, with omomyiforms exhibiting very low amounts of proximal elongation and the adapiforms in this sample with some of the highest levels.

Cluster analyses of the genus-averaged sample provide another way to compare the results of the analyses of *auto3dgm* generated pseudolandmarks to the results of the traditional landmark analyses reported by Gladman et al. (2013). Though there are many differences when comparing the two analyses by their various dendrograms, there are broad similarities as well (Figs. 9-11). Dendrograms for traditional landmark analysis can be viewed in Gladman et al. (2013: their figures 9 & 10, pp. 384-386). We detail comparisons for the Neighbor-Joining (NJ) trees here, and note that similar results are

767 obtained from comparisons between the UPGMA and Wards trees (although these latter
768 two clustering algorithms will not be discussed further).
769 Similarities in the NJ tree (Fig. 9) include the clustering of adapiforms near the taxon
770 chosen as the tree root, *Marcgodinotius indicus*. Extant strepsirrhines and omomyids also
771 cluster together. Within this cluster there are more detailed similarities: *Lepilemur* +
772 *Ourayia* (SDNM 60933) and Omomyid indet. (AMNH 29164) + *Washakius insignis*
773 (AMNH 88824) form two pairs of nearest neighbors, which form a unitary cluster with
774 *Teilhardina* (IRSNB 16786-03) and *Omomys* (UM 98604) in both analyses. *Eulemur*,
775 *Hapalemur*, and *Lemur* form a cluster in both analyses. *Varecia* is external to all
776 members of the strepsirrhine + omomyiform group except *Daubentonia*. All indriids are
777 adjacent to each other. Anthropoids form a unitary cluster separate from non-anthropoids
778 in both analyses, and hominid and pitheciine genera form unitary clusters with respective
779 members of their clades alone (i.e., monophyletic clusters).
780 Major differences include *Daubentonia* falling outside of all clusters and occupying
781 the position closest to the root in the *auto3dgm* based analyses, whereas in Gladman et al.
782 (2013) it clusters with other strepsirrhines. Adapiforms form a unitary cluster with
783 strepsirrhines and omomyiforms in the *auto3dgm* based results, whereas in Gladman et
784 al. (2013), adapiforms formed a unitary cluster basal to all other clusters (in the position
785 of *Daubentonia* in the *auto3dgm* based analysis). In Gladman et al. (2013), the
786 strepsirrhine + omomyiform cluster and the anthropoid cluster group more closely to each
787 other than either does to the adapiform cluster. Though indriids are adjacent in both
788 analyses, they do not form a unitary cluster in the *auto3dgm* based analysis, and
789 *Propithecus* groups with *Avahi*, rather than with *Indri* as in Gladman et al. (2013). In the

*auto3dgm* based analysis, adapiform fossils cluster cleanly by assigned genus with four *Cantius*, two *Smilodectes*, and two *Notharctus* fossils forming three sets of unitary clusters, while in Gladman et al. (2013) these specimens are more mixed. Atelids form a unitary cluster in *auto3dgm* based analysis; in Gladman et al. (2013), they are only adjacent. Hylobatids do not cluster near other hominoids in *auto3dgm* based analysis, whereas hominoids form a unitary cluster in Gladman et al. (2013). *Proteopithecus* (DPC 24776) clusters at the base of a grouping composed primarily of platyrrhines in *auto3dgm* based analysis, whereas it clusters at the base of, and exclusively with, Fayum parapithecid fossils in Gladman et al. (2013). Generally speaking, *auto3dgm* based results were less precise when it comes to interpretable clusters of platyrrhines, cercopithecoids, and hominoids compared to the results of Gladman et al. (2013).

*Mixed bone analysis.*—Because all bones are first scaled to the same unit centroid size (the square root of the sum of squared distances of all landmarks to the centroid of the object), there is a theoretical maximum distance that can accumulate between any pair of bones, and therefore also among all pairs of bones of a given sample size. Nonetheless, the Procrustes distance for any pair of bones and a sample of any size can also approach zero, meaning that shape diversity can be compared by looking at the mean and variance of distances in the distance matrix.

Interestingly, we found that the mean inter-specimen distance and standard deviation were virtually identical for the calcaneal dataset and astragalus dataset treated separately. On the other hand, the mixed samples (both the full 160 specimen sample, and reduced 80 specimen sample - with 40 of each bone type) showed significantly higher mean distance and distance variance (Table 4). That is, results indicate what might be expected

813 intuitively – that there is greater shape diversity in samples containing two kinds of bones
814 than samples containing one kind of bone. Plotting principal component scores reveals
815 obvious taxonomic and phylogenetic clustering (Fig. 12).

816 Comparing phylogenetic signal shows consistently higher estimates of Pagel's
817 lambda in principal component scores of the calcaneus dataset for PCs 1-3 as calculated
818 from both the separate and combined datasets (Table 5). The distance-from-combined-
819 sample-mean dataset ("*mix* MD" in Table 5) for the astragalus had a value of lambda that
820 was higher and more similar to lambda values of the calcaneus datasets. Interestingly,
821 while there was no correlation between PC1 of the astragalus dataset and that of the
822 calcaneus dataset from the separate analyses, those variables from the combined analyses
823 were significantly correlated (Table 6).

824

825 **Discussion**

826 *Comparisons with conventional 3DGM.*— We found the degree of similarity between
827 *auto3dgm* based analyses and those performed on the same sample by Gladman et al.
828 (2013) to be surprising. Compared to our analysis using 1,024 automatically determined
829 points, the carefully selected 27 landmarks used by Gladman et al. (2013) showed similar
830 loadings of shape variance on its Principal Component (PC) axes, similar variance
831 breakdown on the first several PCs, and even a strong correlation between some of the
832 principal component scores (Table 3). The traditional landmark analysis consolidated
833 slightly more variance in its first 4 PCs, though the differences are more pronounced on
834 PCs 3 and 4. Because there are more PCs for the automated analyses than for the manual

one (two orders of magnitude more), it makes sense that the automated method should have a steeper drop-off.

Our automated approach appears more sensitive to errors caused by noise in the surface mesh. This intuitively makes sense and is supported by consideration of some of the clustering "errors" and/or differences between the automated and manual methods. The relatively poor sorting of platyrrhines, hominoids, and cercopithecoids by our automated analysis can be attributed to cases that do not represent mean values, but are the only exemplars of their genus. In particular, the vast majority of catarrhine species in our sample are represented by single specimens, whereas most of our platyrrhines and strepsirrhines are represented by at least two individuals. A single *Colobus* (AMNH 27711) breaks up an otherwise consistent platyrrhine cluster. Though observation of this specimen does not suggest mesh-defects, its lack of any peroneal tubercle projection is anomalous when compared to the prominent peroneal tubercles of all other cercopithecoids in the sample. The lack of a projecting tubercle may give this bone overall length to width proportions that better match the more slender platyrrhines than more robust cercopithecoids. Perhaps the use of a single point in the 27 landmark analysis to represent the peroneal region reduces the effect of this feature's variance on the pattern of morphological affinities (a feature represented by ~100 points in the automated analysis). Similar problems with other specimens likely indicate that having multiple specimen samples is more important generally with our automated approach.

Aside from anomalous individuals, broken specimens and faulty meshes can be expected to "fool" the analysis. A likely example of this is *Leontopithecus* joining a parapithecid (DPC 20576) among a cluster otherwise represented by cercopithecoids.

This fossil is not well preserved in its distal aspect, which likely accentuates the appearance of a strongly sloping lateral border as seen in the callitrichine. It should also be noted however, that Gladman et al. (2013) found that among sampled, extant platyrrhines, *Leontopithecus* has the strongest morphological affinities to cercopithecoids. Both our *auto3dgm* analyses and those of Gladman et al. (2013) suggest morphological affinities uniting Fayum fossil parapithecids with cercopithecoids.

*Comparisons of morphological diversity among parts (mixed bone analysis).*— Our analyses revealed that the astragalus and calcaneus reflect almost identical amounts of shape variation (similar "disparity" as measured with 1,024 evenly distributed points and using either the raw distance matrix, or ordinations based on it). This appears to be a meaningful result since the mixed bone samples (which we believe should express greater shape variation) do, indeed, exhibit significantly greater average distances between shapes.

Interestingly, the phylogenetic signal for a given bone-type was minimally affected (if at all) by running GPA and PCA on a mixed bone sample (Table 5). Despite similar overall variance by almost all measures (Table 4), the calcaneus seems to have developed a stronger phylogenetic signal than the astragalus (Table 5). This suggests that change in calcaneus shape has approximated a Brownian motion model along the branches of the primate phylogenetic tree more so than the astragalus. This difference in mode may be explained functionally by noting that the calcaneus comes into (almost) direct contact with the environment (through the skin, etc.) as the heel, and helps comprise a load arm / lever arm pair that experiences functional demands for leaping and other forms of locomotion (Boyer et al., 2013). In contrast, the astragalus is almost completely isolated

with no part that touches the ground, and no attaching muscles. Therefore, the astragalus may often be insulated from subtle changes in functional demands and be more likely to experience periods of stasis, whereas the calcaneus probably responds more faithfully to small changes in mechanical environment.

The astragalus has long been noted for its high valence in reflecting systematic relationships, while the calcaneus appears less useful. At first pass, this observation seems contradicted by our results. However, if the astragalus has experienced stasis more generally than the calcaneus and developed its comparable morphological variance through more punctuated changes, then the resulting variance may be more clearly associated with more inclusive taxonomic groups (like strepsirrhines, tarsiers, platyrrhines, cercopithecoids, and hominoids) than with species-level differences.

*Biological Significance of Automated Pseudolandmarks.*— The most obvious difference between pseudolandmarks of our method and traditional landmarks is that points associated with a particular feature (e.g., peroneal tubercle) or an articular surface on one bone, may not be located on those features in another bone. This may rub some morphologists the wrong way if they feel that they know that the peroneal tubercle is homologous between two taxa, but the algorithm does not bear this out.

There are several points to be made here. First, as reviewed by MacLeod (2001), Owen's (1846) original definition considered homology as pertaining to "organs" (or we could say "whole bones" here) but did not define mappings of sub-regions therein. In a strict sense, the concept of homology does not apply to features of organs.

Second, the essence of Darwinian homology is that features in different taxa are biologically equivalent if they can be traced to the same feature in a common ancestor

through the process of "descent with modification." This is reflected in a more recent definition stating that homology is a "continuity of information" (Van Valen, 1982). Given that the ultimate arbiter of homology hypotheses is the pattern of transformations that occurred in evolution, it is rare that they can ever be verified.

Third, the critics of the adaptationist programme (Gould and Lewontin, 1979) warn us to beware of "spandrels." One can ask whether the feature of interest exists by genetic design or by developmental context. If the peroneal tubercle "exists" as a genetically specified bump on the side of the calcaneus (in the sense that there are gene products that cause the formation of this bump, and variation in the position or size of the tubercle can be explained by these gene products being expressed at different positions, at different concentrations, and/or for different durations along the shaft of the calcaneus), then it follows that this "bump" should be marked with a landmark of the same identity on any bone regardless of where topologically it occurs. However, it seems equally likely that the form of the bony peroneal tubercle is a mechanical and re-modeling consequence of the paths of the peroneal tendons and where the retinacular ligaments attach. In this alternative scenario, representing the position of this bump by the same "point" regardless of its position on the calcaneus seems misrepresentative. The truth is that the genetic influences and developmental homologies for most features are not known. An informative test of these alternatives (although cruel) would be to remove the tendons at an early stage of development and observe whether and where a peroneal tubercle developed. Even if it were to become known that peroneal tubercle development occurred independent of attaching ligaments and tendons, and the forces they exert, this would only imply evolutionary homology if we assume parsimony in evolution (or Hennig's

927 auxiliary principle) which some researchers are willing to do, but others are not. This also
928 comes down to whether type I or type II landmarks are preferred when the respective
929 criteria suggest different correspondence patterns for a given anatomical region.
930 Finally, in this particular example, there is no widespread agreement on the
931 evolutionary homology of the peroneal tubercle among primates (Decker and Szalay,
932 1974). Variation in features that are plastic and can be modified during life (such as
933 ligament attachment points and articular surface areas and boundary shapes) may be
934 explained by ontogenetic causes. For instance, variation in the development of certain
935 astragalar facets in humans has been explained by different postural tendencies among
936 populations (Barnett, 1954). If we use the distal boundary of the tibial facet as a
937 landmark, this feature point may extend all the way down the astragalar neck in some
938 people, or not approach it at all in others. This would be useful for quantifying variation
939 due to postural differences among humans, but probably not for distinguishing the shape
940 of a human astragalus from a chimpanzee astragalus.
941 Another argument for adding the use of pseudolandmarks to the morphologist's
942 toolkit is the fact that the research community already accepts similar approaches to
943 shape comparison including Fourier analysis (Rohlf and Archie, 1984), eigenshape
944 analysis (MacLeod, 1999), and eigensurface analysis (Polly and MacLeod, 2008). These
945 methods retain no fidelity to specific landmark-like features. The most significant
946 conceptual difference between our approach and eigensurface analysis is that the
947 anatomical axes must be manually set in the latter. A more practical difference is that
948 eigensurface is restricted to "relief-type" or "disc-type" surfaces, whereas *auto3dgm* can
949 be applied to either disc-type or fully 3D surfaces.

The question of whether points or regions on different instances of the same bone are "equivalent" is ultimately a question about transformational homology. Our method provides an "operational homology" (= topological correspondence). The minimum spanning tree used to link forms can be taken as a hypothesis of transformational homology to be tested. The best answer to whether certain "point features" are equivalent must be answered by assessing whether treating them as such results in phenetic patterns that correlate with independent datasets on phylogenetic relationships or functional capacity. This means that if the utility of automated methods is going to increase, then automated correspondence determinations that are more sensitive to feature points (type II landmarks) must also be developed. This requires algorithms based on "non-area preserving maps". The original work of Boyer et al. (2011) presents such a method but lacks applicability to "full 3D" shapes and does not provide a means for inducing transitivity of comparisons. Different patterns of transformational homology will be implied by different phylogenetic hypotheses, which could be evaluated according to different optimization criteria.

*Too many variables, not enough specimens?*– A major challenge in statistical modeling as applied to molecular biology (Golub et al., 1999), genetics (Patterson et al., 2006), image analysis (Roweis and Saul, 2000), and text analysis (Blei et al., 2003) has been the *large P, small N* setting (Poggio and Smale, 2003; West, 2003) where the number of variables is typically much larger than the number of samples. In statistics, the difficulty of modeling data as the number of variables increases and exceeds the number of observations is often called "the curse of dimensionality", a phrase coined by Bellman with respect to optimization problems (Bellman, 1984). However, many of the great

973 advances in the last ten years in statistics, machine learning, and applied mathematics are
974 related to the observation that the relevant dimension of the data is not the number of
975 variables, but the number of independent variables (the intrinsic dimension) (Donoho,
976 2000). For 1,024 landmarks spread on a sample of 80-160 objects, the intrinsic
977 dimensionality will be much lower than the number of landmarks. If the perspective
978 promoted by statisticians dealing with *large P, small N* problems is correct, then the
979 problem of over-determination can be avoided by limiting the number of independent
980 variables generated by data reduction techniques from a landmark dataset with hundreds
981 or thousands of points. The idea that seemingly high-dimensional data have few degrees
982 of freedom, or low intrinsic dimensionality, is central to the methodologies developed in
983 this paper.

984 As a matter of precedent, this philosophy is implicitly acknowledged in papers that
985 use large numbers of evenly (or "optimally") spread semi-landmarks as well as in
986 eigenshape analysis (Polly, 2008; Polly and MacLeod, 2008; Sievwright and MacLeod,
987 2012). Harcourt-Smith et al. (2008) provides a pertinent example, in which a total of nine
988 user-defined landmarks were used to generate 361 semilandmark points on the talo-tibial
989 facets of a sample with 54 specimens representing three species. Another example is
990 Sievright and MacLeod (2012). These authors used 62 points to represent the dorsal
991 surface of the proximal humerus in a sample of 50 falconiform specimens. They
992 projected their coordinates into tangent space and used principal component analysis to
993 generate projection scores. These mutually orthogonal (independent) projection scores
994 were then used to run a Canonical Variates Analysis (=DFA). They limited the number of
995 principal components used in their analysis to 21 (because they argued that this number

represented 95% of the total variation in the dataset and was much less than their n=50). These authors recognize the importance of the number of independent variables, but do not discuss the statistical ramifications of the number of original, yet correlated, variables.

**Summary and Conclusion**

Greater automation and standardization for morphological studies are needed if morphology is to survive as a branch of phenomics with relevance comparable to genomics. The most important level at which such automation must occur is in determining biological/geometric correspondence between shapes. Past attempts to automate such determinations have suffered from the prospect that computations involved were too time intensive (as well as philosophical arguments against the idea of such an approach). Dimension reduction techniques such as working from photographs and outlines have been applied to circumvent this issue, but an observer is needed to orient objects before such application, slightly defeating the purpose of automation. Greater computing power and techniques for simplifying the search for alignment and correspondence mapping between 3D digital models are applied here and an R package for implementing this method has been created.

Our analyses show a surprising and reassuring degree of similarity between quantifications based on user-defined landmarks and our *auto3dgm* approach. Although human interaction must occur at several stages of the analyses to verify that erroneous alignments have not been generated, this approach still represents a step beyond any automation procedures yet applied, because 1) no qualitative decisions about the

geometric equivalence of point features are required and 2) protocols for generating alignments and pseudolandmark datasets lack observer error, since the final procedure for the exact result of the algorithm can be described via the numerical parameter input to the model. Very little familiarity with anatomical terminology or features is required. Only a basic ability to visually compare shapes is necessary in *auto3dgm* in order to verify the absence of misalignments. This method has the potential for adoption by geneticists, molecular biologists, and biomedical engineers who may feel uncomfortable about their ability to take measurements with repeated accuracy or with biological significance to their questions of interest.

One of the most exciting capabilities provided by this algorithm is the ability to compare variance magnitude and patterns for different skeletal elements. Our initial experiments with this approach show that two articulating bones of the skeleton have identical levels of morphological diversity with strong covariance, which makes sense developmentally, but the calcaneus has a consistently stronger phylogenetic signal in its variance patterns than the astragalus.

Future work will explore different types of correspondence algorithms with an emphasis on constructing algorithms that can efficiently determine non-area preserving maps (those that mimic user-defined type II landmarks of 3DGM more closely). Furthermore, we intend to compare variance levels among different regions of the skeleton with the expectation that patterns of covariance and variance magnitudes will differ more between bones that are far apart from each other on the skeleton and are more likely to have different developmental and historical natural-selective contexts. We recognize that these quantities are still dependent on the sample composition, the

parameters of any particular run of *auto3dgm*, and any ordination methods that are used. Nonetheless, we feel that the patterns will be informative for evolutionary questions including those dealing with disparity because the quantification of inter-bone shape distance is objective and more comprehensive *auto3dgm*, and we have articulated a rationale geometric basis for comparing variance between groups of non-homologous elements.

**Acknowledgements**

DMB's involvement in this work was partly funded by support from Brooklyn College, CUNY; Duke University; and NSF BCS 1317525 to DMB and ERS, as well as NSF BCS 1304045 to DMB and EMS. DMB wishes to acknowledge J. Lovoi, J. Butler, and A. Garberg who helped process scans for this study, as well as S. Maiolino who provided some distal phalanx specimens. Work contributed by JTG has been supported by NYCEP IGERT (NSF DGE-0991660), with additional support by a CUNY Graduate Center DSRG and Digital Innovations Grant. NYCEP MORPHOMETRICS GROUP #88. GSY's work has been supported by the Graduate School of Duke University.

**References**

3D Systems Inc. 2013. Geomagic Studio. In. Rock Hill: 3D Systems Inc.

Abzhanov A, Kuo WP, Hartmann C, Grant BR, Grant PR, Tabin CJ. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. Nature 442:563-567.

Arnold C, Matthews LJ, Nunn CL. 2010. The 10kTrees Website: A New Online Resource for Primate Phylogeny. Evolutionary Anthropology 19:114-118.

Barnett CH. 1954. Squatting facets on the European talus. Journal of Anatomy, London 88:509-513.

Bellman RE. 1984. Eye of the Hurricane: an autobiography. Singapore: World Scientific Publishing Company.

Berge C, Jouffroy FK. 1986. Morpho-functional study of *Tarsius's* foot as compared to the galagines': what does an "elongate calcaneuys" mean? In: Taub D, King F, editors. Current Perspectives in Primate Biology. New York: Van Nostrand Reinhold. p 146-158.

Besl PJ, McKay ND. 1992. A Method for Registration of 3-D Shapes. IEEE Trans Pattern Anal Mach Intell 14:239-256.

Blei DM, Ng AY, Jordan MI. 2003. Latent Dirichlet Allocation. Journal of Machine Learning Research 3:993--1022.

Bookstein FL. 1980. When one form is between two others: An application of biorthogonal analysis. American Zoologist 20:627-642.

Bookstein FL. 1991. Morphometric Tools for Landmark Data. Geometry and Biology. Cambridge: Cambridge University Press.

Bookstein FL. 1994. Can biometrical shape be a homologous character? In: Hall BK, editor. Homology: the hierarchical basis of comparative biology. San Diego: Academic Press. p 197-227.

Bookstein FL. 1997. Landmark methods for forms without landmarks: localizing group differences in outline shape. Medical Image Analysis 1:225-243.

Bookstein FL, Schafer K, Prossinger H, Seidler H, Fieder M, Stringer C, Weber GW, Arsuaga J-L, Slice DE, Rohlf FJ, Recheis W, Mariam AJ, Marcus LF. 1999. Comparing frontal ¨cranial profiles in archaic and modern Homo by morphometric analysis. Anatomical Record (New Anatomy) 257:217-224.

Bookstein FL, Streissguth AP, Sampson PD, Connor PD, Barr HM. 2002. Corpus callosum shape and neuropsychological deficits in adult males with heavy fetal alcohol exposure. Neuroimage 15:233-251.

Boyer DM, Lipman Y, St. Clair E, Puente J, Patel BA, Funkhouser TA, Jernvall J, Daubechies I. 2011. Algorithms to automatically quantify the geometric similarity of anatomical surfaces. Proceedings of the National Academy of Science 108:18221-18226.

Boyer DM, Seiffert ER. 2013. Patterns of astragalar fibular facet orientation in extant and fossil primates and their evolutionary implications. American journal of Physical Anthropology 151:420-447.

Boyer DM, Seiffert ER, Gladman JT, Bloch JI. 2013. Evolution and allometry of calcaneal elongation in living and extinct primates. PLoS ONE. 8(7): e67792. doi:10.1371/journal.pone.0067792

Bunn JM, Boyer DM, Jernvall J, Lipman Y, Daubechies I. 2011. Dirichlet normal surface energy of tooth crowns, a new technique of molar shape quantification for dietary inference, compared with previous methods in isolation and in combination. American Journal of Physical Anthropology 145:247-261.

Cignoni P, Ranzuglia G, Callieri M, Corsini M, Ganovelli F, Pietroni N, Tarini M. 2012. MeshLab Visual Computing Lab - ISTI- CNR. In. http://meshlab.sourceforge.net/.

Decker RL, Szalay FS. 1974. Origins and function of the pes in the Eocene Adapidae (Lemuriformes, Primates). In: Jenkins Jr. FA, editor. Primate Locomotion. New York: Academic Press. p 261-291.

Donoho DL. 2000. High-dimensional data analysis: the curses and blessings of dimensionality. In: American Mathematical Society Conference: Math Challenges of the 21st Century. p 1-32.

Eldar Y, Lindenbaum M, Porat M, Zeevi YY. 1997. The farthest point strategy for progressive image sampling. IEEE Transactions: On Image Processing:1315.

Frost SR, Marcus LF, Bookstein F, Reddy DP, Delson E. 2003. Cranial allometry, phylogeography and systematics of large bodied papionins (Primates: Cercopithecinae) inferred from geometric morphometric analysis of landmark data. The Anatomical Record 275A:1048-1072.

Gingerich PD, ul Haq M, Zalmout IS, Khan IH, Malkani MS. 2001. Origin of whales from early artiodactyls: Hands and feet of Eocene Protocetidae from Pakistan. Science 293:2239-2242.

Gladman JT, Boyer DM, Simons EL, Seiffert ER. 2013. A calcaneus attributable to the primitive late Eocene anthropoid *Proteopithecus sylviae*: phenetic affinities and phylogenetic implications. American Journal of Physical Anthropology 151:372-397.

1130 Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh
1131 ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. 1999. Molecular
1132 classification of cancer: class discovery and class prediction by gene expression
1133 monitoring. Science 286:531-537.
1134 Gould SJ, Lewontin RC. 1979. The spandrels of San Marco and the Panglossian pradigm:
1135 a critique of the adaptationist programme. Proceedings of the Royal Society of
1136 London B 205:581-598.
1137 Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H,
1138 Zhai W, Fritz MH-Y, Hansen NF, Durand EY, Malaspinas A-S, Jensen JD,
1139 Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz
1140 R, Aximu-Petri A, Butthof A, Höber B, Höffner B, Siegemund M, Weihmann A,
1141 Nusbaum C, Lander ES, Russ C, Novod N, Affourtit J, Egholm M, Verna C,
1142 Rudan P, Brajkovic D, Kucan Ž, Gušic I, Doronichev VB, Golovanova LV,
1143 Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A, Schmitz RW, Johnson PLF,
1144 Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M, Nielsen R, Kelso J,
1145 Lachmann M, Reich D, Pääbo S. 2010. A Draft Sequence of the Neandertal
1146 Genome. Science 328:710-722.
1147 Gunz P, Mitteroecker P, Bookstein FL. 2005. Semilandmarks in Three Dimensions. In:
1148 Slice DE, editor. Modern Morphometrics in Physical Anthropology. New York:
1149 Kluwer Academic/Plenum Publishers. p 73-98.
1150 Hammer Ø, Harper DAT, Ryan PD. 2001. PAST: Paleontological statistics software
1151 package for education and data analysis. Palaeontologia Electronica 4:9pp.
1152 http://palaeo-electronica.org/2001_2001/past/issue2001_2001.htm.

Hammer Ø, Harper DAT, Ryan PD. 2006. PAST: Paleontological statistics software package for education and data analysis. In, 1.43 ed. Oslo: University of Oslo.

Harcourt-Smith WEH, Tallman M, Frost SR, Wiley DF, Rohlf JF, Delson E. 2008. Analysis of Selected Hominoid Joint SurfacesUsing Laser Scanning and GeometricMorphometrics: A Preliminary Report. In: Dagosto M, Sargis EJ, editors. Mammalian Evolutionary Morphology: A Tribute to Frederick S. Szalay. New York: Springer Science. p 373–383.

Harjunmaa E, Kallonen A, Voutilainen M, Hämäläinen K, Mikkola ML, Jernvall J. 2012. On the difficulty of increasing dental complexity. Nature 483:324-327.

Houle D, Govindaraju DR, Omholt S. 2010. Phenomics: the next challenge. Nature Reviews Genetics 11:855-856.

Leakey LSB, Tobias PV, Napier JR. 1964. A new species of the genus *Homo* from Olduvai Gorge. Nature 202:7-9.

Lipman Y, Daubechies I. 2010. The Continuous Procrustes Distance. In: Technical report: Princeton University.

Lohmann GP. 1983. Eigenshape analysis of microfossils: a general morphometric method for describing changes in shape. Mathematical Geology 15:659-672.

Losos JB. 1990. The evolution of form and function: morphology and locomotor performance in west indian *Anolis* lizards. Evolution 44(5):1189-1203.

MacLeod N. 1999. Generalizing and extending the eigenshape method of shape space visualization and analysis. Paleobiology 25:107-138.

1174 MacLeod N. 2001. Landmarks, Localization and the Use of Morphometrics in
1175 Phylogenetic Analysis. In: Adrain JM, Edgecombe GD, Lieberman BS, editors.
1176 Fossils, Phylogeny, and Form: Springer. p 197-233.
1177 MacLeod N. 2007. Automated taxon identication in systematics: theory, approaches, and
1178 applications. London: CRC Press, Taylor & Francis Group.
1179 MacLeod N, Benfield M, Culverhouse P. 2010. Time to automate identification. Nature
1180 467:154-155.
1181 McVean G, Spencer CCA, Chaix R. 2005. Perspectives on Human Genetic Variation
1182 from the HapMap Project. PLoS Genetics 1:e54.
1183 Mitteroecker P, Gunz P. 2009. Advances in Geometric Morphometrics. Evolutionary
1184 Biology 36:235-247.
1185 Morgan TH. 1911. The Origin of Five Mutations in Eye Color in Drosophila and Their
1186 Modes of Inheritance. . Science 33:534-537.
1187 Moyà-Solà S, Kohler M, Alba DM, Roig I. 2012. Calcaneal proportions in primates and
1188 locomotor inferences in Anchomomys and other Palaeogene Euprimates. Swiss
1189 Journal of Paleontology 131:147-159.
1190 Nunn CL. 2011. The Comparative Method in Evolutionary Anthropology and Biology:
1191 University of Chicago Press.
1192 O'Higgins P, Jones N. 2006. Tools for statistical shape analysis. In.
1193 http://sites.google.com/site/hymsfme/resources.: Hull York Medical School.
1194 O'Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg
1195 SL, Kraatz BP, Luo Z-X, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS,
1196 Rougier GS, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM,

1197 Weksler M, Wible JR, Cirranello AR. 2013a. The Placental Mammal Ancestor
1198 and the Post–K-Pg Radiation of Placentals. Science 339:662-667.
1199 O'Leary MA, Bloch JI, Flynn JJ, Gaudin TJ, Giallombardo A, Giannini NP, Goldberg
1200 SL, Kraatz BP, Luo Z-X, Meng J, Ni X, Novacek MJ, Perini FA, Randall ZS,
1201 Rougier GS, Sargis EJ, Silcox MT, Simmons NB, Spaulding M, Velazco PM,
1202 Weksler M, Wible JR, Cirranello AR. 2013b. Response to Comment on "The
1203 Placental Mammal Ancestor and the Post–K-Pg Radiation of Placentals". Science
1204 341:613.
1205 Orlando L, Calvignac S, Schnebelen C, Douady CJ, Godfrey LR, Hä nni C. 2008. DNA
1206 from extinct giant lemurs links archaeolemurids to extant indriids. BMC
1207 Evolutionary Biology 8:doi:10.1186/1471-2148-1188-1121.
1208 Orme CDL, Freckleton RP, Thomas G, Petzoldt T, Fritz SA, Isaac N, Pearse W. 2011.
1209 caper: Comparative Analysis of Phylogenetics and Evolution in R. In.
1210 Ostrom JH. 1975. On the origin of *Archaeopteryx* and the ancestry of birds. Colloque
1211 International C N R S 218:521-532.
1212 Owen R. 1846. Report on the archetype and homologies of the vertebrate skeleton.
1213 Report of the British Association for the Advancement of Science (Southampton
1214 meeting):169-340.
1215 Paniagua B, Bompard L, Cates J, Whitaker R, Datar M, Vachet C, Styner M. 2012.
1216 Combined SPHARM-PDM and entropy-based particle systems shape analysis
1217 framework. In: SPIE Medical Imaging: International Society for Optics and
1218 Photonics. p 83170L-83170L-83177.

Patterson C. 1982. Morphological characters and homology. In: Joysey KA, Friday AE, editors. Systematics Association Special Volume No. 21, "Problems of Phylogenetic Reconstruction". London and New York: Academic Press. p 21-74.

Patterson N, Price AL, Reich D. 2006. Population Structure and Eigenanalysis. PLoS Genetics 2:e190.

Perez SI, Bernal V, Gonzalez PN. 2006. Differences between sliding semilandmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. Journal of Anatomy 208:769-784.

Poggio T, Smale S. 2003. The Mathematics of Learning: Dealing with Data. Notices of the American Mathematical Society 50:537-544.

Polly PD. 2008. Adaptive Zones and the Pinniped Ankle: A 3D Quantitative Analysis of Carnivoran Tarsal Evolution. In: Sargis EJ, Dagosto M, editors. Mammalian Evolutionary Morphology: A Tribute to Frederick S. Szalay. Dordrecht, The Netherlands: Springer. p 165-194.

Polly PD, MacLeod N. 2008. Locomotion in fossil Carnivora: An application of eigensurface analysis

for morphometric comparison of 3D surfaces. Palaeontologia Electronica 11.

Puente J. 2013. Distances and algorithms to compare sets of shapes for automated biological morphometrics. In: Program in Applied and Computational Mathematics. unpublished: Princeton. p 72.

Puente J, Daubechies I. in preparation. A Generalization of Procrustes Anaylsis.

R Coding Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Rohlf FJ. 2002. Geometric morphometrics and phylogeny. Systematics Association Special Volume 64:175-193.

Rohlf FJ, Archie JW. 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera, Culicidae). Systematic Zoology 3:302-317.

Roth VL. 1993. On three dimensional morphometrics, and on the identification of landmark points. In: Marcus LF, Bello E, Garcia-Valdecasas A, editors. Contributions to Morphometrics. Madrid: Museo Nacional de Ciencias Naturales CSIC. p 41-61.

Roweis S, Saul L. 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 290:2323-2326.

Shubin NH. 1994. History, ontogeny, and evolution of the archetype. In: Grande L, Riepper O, editors. Homology, the Hierarchical Basis of Comparative Biology. New York: The Academic Press, Inc. p 249-271.

Sievwright H, MacLeod N. 2012. Eigensurface analysis, ecology, and modelling of morphological adaptation in the falconiform humerus (Falconiformes: Aves). Zoological Journal of the Linnean Society 165:390-419.

Simpson GG. 1944. Tempo and Mode in Evolution. New York: Columbia University Press.

Simpson GG. 1961. Principles of Animal Taxonomy. New York: Columbia University Press.

Smith GR. 1990. Homology in Morphometrics and Phylogenetics. In: Rohlf FJ, Bookstein FL, editors. Proceedings of the Michigan Morphometrics Workshop: The University of Michigan Museum of Zoology. p 325-338.

Sneath PHA, Sokal FJ. 1973. Principles of Numerical Taxonomy. Second Edition. San Francisco: Freeman.

Sokal RR. 1966. Numerical taxonomy. Scientific American 215:106-116.

Springer MS, Meredith RW, Teeling EC, Murphy WJ. 2013. Technical Comment on "The Placental Mammal Ancestor and the Post–K-Pg Radiation of Placentals". Science 341 613.

Styner M, Oguz I, Xu S, Brechbuler C, Pantazis D, Gerig G. 2006. Framework for the Statistical Shape Analysis of Brain Structures using SPHARM-PDM. Insight Journal 1071:242-250.

Thompson DA. 1942. On Growth and Form. Cambridge: University Press. 1116 pgs

Van Valen LM. 1982. Homology and causes. Journal of Morphology 173:305-312.

Venter JC, Adams MD, Myers EW, Li PW, al. e. 2003. The Sequence of the Human Genome. Science 291:1304-1351.

Visualization Sciences Group. 2009. AVIZO, Version 6.0. In. Burlington: Mercury Computer Systems.

Weeks PJD, O'Neill MA, Gaston KJ, Gauld ID. 1999. Species-identication of wasps using principal component associative memories. Img Vis Comp 17:861-866.

West M. 2003. Bayesian Factor Regression Models in the "Large p, Small n" Paradigm. In: Bernardo JM, Bayarri MJ, Berger OJ, Dawid AP, Heckerman D, Smith AFM, West M, editors. Bayesian Statistics: Oxford University Press. p 733-742.

Wiley DF, Amenta N, Alcantara DA, Ghosh D, Kil YJ, Delson E, Harcourt-Smith W, Rohlf FJ, St. John K, Hamann B. 2005. Evolutionary Morphing. Proceedings of IEEE Visualization 2005:8.

Zelditch ML, Swiderski DL, Sheets DH, Fink WL. 2004. Geometric Morphometrics for Biologists. San Diego: Elsevier Academic Press.

# Tables

## **Table 1.** Taxonomic samples for this study

| Extant | Set 1 | Set 2 | Set 3 | Set 4 |
|---|---|---|---|---|
| **Taxon** | **n Calc.** | **n Ast.** | **n Phal.** | **n Ast.** |
| *Avahi laniger* | 1 | 1 | -- | -- |
| *Microcebus murinus* | -- | -- | 1 | 1 |
| *Cheirogaleus major* | 1 | 1 | 2 | -- |
| *Mirza coquereli* | -- | -- | 1 | -- |
| *Daubentonia madagascariensis* | 1 | 1 | -- | 1 |
| *Eulemur fulvus* | 2 | 2 | 1 | 1 |
| *Hapalemur griseus* | 3 | 3 | 1 | 1 |
| *Indri indri* | 2 | 2 | 1 | -- |
| *Lemur catta* | 3 | 3 | 1 | 1 |
| *Lepilemur mustelinus* | 3 | 3 | -- | 1 |
| *Propithecus verreauxi* | 2 | 2 | 1 | -- |
| *Propithecus diadema* | -- | -- | 1 | -- |
| *Varecia variegata* | 1 | 1 | 1 | -- |
| *Galago senegalensis* | -- | -- | 2 | -- |
| *Otolemur crassicaudatus* | -- | -- | 2 | -- |
| *Loris tardigradus* | -- | -- | -- | 1 |
| *Nycticebus coucang* | -- | -- | -- | 1 |
| *Perodicticus potto* | -- | -- | -- | 1 |
| *Alouatta seniculus,* sp. | 4 | 3 | -- | 1 |
| *Aotus azarae, infulatus,* sp. | 3 | 3 | 2 | 1 |
| *Ateles paniscus, sp.* | 3 | 3 | -- | 1 |
| *Brachyteles arachnoides* | 1 | 1 | -- | -- |
| *Cacajao calvus* | 2 | 2 | -- | 1 |
| *Callicebus donaco., moloch* | 3 | 3 | -- | 1 |
| *Callimico goeldi* | 2 | 2 | -- | -- |
| *Callithrix jacchus* | 2 | 2 | -- | 1 |
| *Cebuella pygmaea* | 2 | 2 | -- | -- |
| *Cebus apella,* sp. | 2 | 2 | -- | 1 |
| *Chiropotes satanus,* sp. | 3 | 3 | -- | -- |
| *Leontopithecus rosalia* | 2 | 2 | -- | -- |
| *Pithecia monachus, pithecia* | 2 | 2 | -- | 1 |
| *Saguinus midas, mystax,* sp. | 4 | 3 | -- | -- |
| *Saimiri boliviensis, sciureus,* sp. | 5 | 3 | -- | -- |
| *Cercopithecus* sp. | 2 | -- | -- | -- |
| *Chlorocebus aethiops, cynosuros* | 2 | 1 | -- | -- |
| *Colobus geureza* | 1 | 0 | -- | -- |
| *Erythrocebus patas* | 1 | 0 | -- | -- |
| *Lophocebus albigena* | 1 | 0 | -- | -- |
| *Macaca nigra, tonkeana* | 2 | 2 | -- | -- |
| *Mandrillus sphinx* | 1 | 0 | -- | -- |
| *Nasalis larvatus* | 1 | 1 | -- | -- |
| *Papio hamadryas* | 1 | -- | -- | -- |
| *Piliocolobus badius* | 2 | -- | -- | -- |
| *Pygathrix nemaeus* | 1 | -- | -- | -- |
| *Theropitheucs gelada* | 1 | -- | -- | -- |
| *Trachypithecus obscurus* | 1 | 1 | -- | -- |
| *Gorilla* sp. | 1 | 1 | -- | -- |
| *Hylobates lar* | 1 | 1 | -- | -- |
| *Pan troglodytes* | 2 | 2 | -- | -- |
| *Pongo pygmaeus* | 1 | 1 | -- | -- |
| *Symphalangus syndactylus* | 1 | 1 | -- | -- |
| *Tarsius pumilus* | -- | -- | 2 | -- |
| *Tarsius bancanus* | -- | -- | 2 | 1 |
| *Tarsius spectrum* | -- | -- | 2 | 1 |
| *Tarsius syrichta* | -- | -- | -- | 1 |
| *Cynocephalus volans* | -- | -- | -- | 2 |
| *Galeopterus variegatus* | -- | -- | -- | 1 |
| *Ptilocercus lowii* | -- | -- | -- | 2 |
| *Tupaia glis* | -- | -- | -- | 2 |
| *Lepus sp.* | -- | -- | -- | 2 |
| *Sylvilagus sp.* | -- | -- | -- | 1 |
| *Ochotona princeps* | -- | -- | -- | 1 |
| *Erethizon sp.* | -- | -- | -- | 1 |
| *Coendou prehensilis* | -- | -- | -- | 1 |
| *Marmota sp.* | -- | -- | -- | 1 |
| *Sciurus sp.* | -- | -- | -- | 1 |
| *Aplodontia rufa* | -- | -- | -- | 1 |
| *Allactaga major* | -- | -- | -- | 1 |
| *Hemiechinus auritus* | -- | -- | 4 | 1 |
| *Erinaceus europaeus* | -- | -- | 3 | 1 |
| *Erinaceus roumanicus* | -- | -- | 4 | -- |
| *Chrysochloris asiatica* | -- | -- | -- | 1 |
| *Crocidura olivieri* | -- | -- | -- | 1 |
| *Desmana moschata* | -- | -- | -- | 1 |
| *Solenodon paradoxus* | -- | -- | -- | 1 |
| *Potos flavus* | -- | -- | -- | 1 |
| *Arctictis binturong* | -- | -- | -- | 1 |
| *Nasua narica* | -- | -- | -- | 1 |
| *Petrodromus tetradactylus* | -- | -- | -- | 1 |
| *Tenrec ecaudatus* | -- | -- | -- | 1 |
| *Setifer setosus* | -- | -- | -- | 1 |
| *Hemicentetes semispinosus* | -- | -- | -- | 1 |
| *Echinops telfairi* | -- | -- | -- | 1 |
| *Potamogale velox* | -- | -- | -- | 1 |
| **TOTAL extant N:** | **82** | **66** | **34** | **52** |

| Fossil | Set 1 | Set 2 | Set 3 |
|---|---|---|---|
| **Taxon** | **Calc. Cat. #** | **Ast. Cat. #** | **Phal. Cat. #** |
| *Cantius abditus* | USGS 6783 | USG 21832 | -- |
| *Cantius* sp. | USGS 6774 | -- | -- |
| *Cantius trigonodus* | AMNH 16852 | -- | -- |
| *Cantius trigonodus* | USGS 21829 | -- | -- |
| *Cebupithecia sarmientoi* | UCMP 38762* | UCMP 38762* | -- |
| *Marcgodinotius indicus* | GU 709 | GU 748 | -- |
| *Mesopithecus pentelici** | MNHN PIK-266 | -- | -- |
| *Neosaimiri fieldsi** | IGM-KU 89202 | IGM-KU | -- |
| *Neosaimiri fieldsi** | IGM-KU 89203 | -- | -- |
| *Notharctus* sp. | AMNH 55061 | AMNH 11474 | -- |
| *Notharctus tenebrosus* | AMNH 11474 | AMNH 129382 | AMNH 143612-3 |
| Omomyid | AMNH 29164 | UM 38321 | -- |
| *Omomys* sp. | UM 98604 | UM 98648 | -- |
| *Oreopithecus bambolii* | NMB 37* | -- | -- |
| *Ourayia uintensis* | SDNM 60933 | -- | -- |
| Parapithecid | DPC 15679 | DPC 5027 | -- |
| Parapithecid | DPC 20576 | DPC 5416A | -- |
| Parapithecid | DPC 2381 | DPC 1001 | -- |
| Parapithecid | DPC 8810 | -- | -- |
| *Proteopithecus sylviae* | DPC 24776 | DPC 22844 | -- |
| *Smilodectes gracilis* | AMNH 131763 | -- | -- |
| *Smilodectes gracilis* | AMNH 131774 | -- | -- |
| *Teihardina belgica* | IRSNB16786-03 | IRSNB16786-01 | -- |
| *Washakius insignis* | AMNH 88824 | UM 99704 | -- |
| *Carpolestes simpsoni* | -- | -- | UM 101963 (x4) |
| *Ignacius clarksforkensis* | -- | -- | UM 82606 |
| *Plesiadapis churchilli* | -- | -- | SMM P77.33.517 |
| *Nannodectes intermedius* | -- | -- | USNM 442229 |
| Incertae sedis | -- | -- | 6 from UCMP |
| **TOTAL fossil N:** | **24** | **14** | **14** |

1292 **Table 2.** Comparison between traditional 3DGM of 106 calcanei sample and FAA of this
1293 study.

| Comparison point | 27 landmark—Manual analysis | 1,024 landmark—Automated |
|---|---|---|
| PC 1 % variance | 35.9 | 34.7 |
| PC 2 % variance | 13.6 | 13.6 |
| PC 3 % variance | 9.5 | 6.7 |
| PC 4 % variance | 6.7 | 4.6 |
| Sum PC 1-4 | 64.9 | 59.6 |
| PC 1 loadings | Overall width/length proportions with emphasis on distal elongation. | Overall width/length proportions with emphasis on distal elongation. |
| PC 2 loadings | Position of lateral peak of the peroneal tubercle relative to both ectal and cuboid facets. | 1) Dorsoplantar elevation of the ectal facet's distal margin relative to the calcaneus body; 2) distinctiveness, but not position, of peroneal tubercle. |
| PC 3 loadings | 1) Proximal segment elongation, shape/orientation of ectal facet, 2) dorsal projection of dorsal heel. | Tradeoff between a prominent proximal plantar heel process and an accentuated angulation at the distal plantar tubercle. |
| PC 4 loadings | Ectal facet position, curvature, and orientation relative to long axis of the calcaneus. | Proximal elongation and dorsal projection of dorsal heel. |

1294

1295 **Table 3.** Correlation (r) and Probability (p) between manual and automated PCs.

| | **Linear correlations (r)** | | | |
|---|---|---|---|---|
| **Manual** | **Automated Pseudolandmarks** | | | |
| **3DGM** | **PC-1** | **PC-2** | **PC-3** | **PC-4** |
| **PC-1** | **-0.96** | -0.16 | 0.09 | 0.07 |
| **PC-2** | 0.11 | **-0.50** | **0.34** | **-0.28** |
| **PC-3** | 0.15 | **-0.64** | 0.03 | 0.18 |
| **PC-4** | -0.01 | 0.06 | **-0.38** | **-0.32** |

| | **Probability of no correlation (P)** | | | |
|---|---|---|---|---|
| **Manual** | **Automated Pseudolandmarks** | | | |
| **3DGM** | **PC-1** | **PC-2** | **PC-3** | **PC-4** |
| **PC-1** | <0.0001 | ns | ns | ns |
| **PC-2** | ns | <0.0001 | 0.0004 | 0.0042 |
| **PC-3** | ns | <0.0001 | ns | ns |
| **PC-4** | ns | ns | <0.0001 | 0.0008 |

1296

**Table 4.** Distance matrices from mixed bone analyses. "Dev. From Mean" represent the distance between each object and the mean object. Thus the number of distances is the same as the sample size. The t-test is done on this sample of deviations from the mean. "Mix" represents the results of analysis of 40 astragali with 40 taxon-matched calcanei.

| **Full Distance Matrix** | | | |
|---|---|---|---|
| **n=3,120** | **Calc.** | **Ast.** | **mix** |
| mean | 0.18 | 0.19 | 0.29 |
| max | 0.40 | 0.37 | 0.54 |
| min | 0.05 | 0.06 | 0.05 |
| sd | 0.06 | 0.05 | 0.11 |
| **Dev. from Mean** | | | |
| **n=80** | **Calc.** | **Ast.** | **mix** |
| Mean dev. | 0.13 | 0.13 | 0.21 |
| max | 0.25 | 0.27 | 0.31 |
| min | 0.07 | 0.07 | 0.16 |
| sd | 0.04 | 0.03 | 0.03 |
| **t-test (on Dev.)** | **df** | **t** | **P** |
| Ast. vs. Calc. | 158 | 0.50 | 0.62 |
| Ast. vs. Mix | 158 | 15.16 | <0.0001 |
| Calc. vs. Mix | 158 | 14.81 | <0.0001 |

**Table 5.** Phylogenetic signal in astragalus and calcaneus shape data based on automated analysis of 1,024 pseudolandmarks. "Mix" preceding the variable name indicates that the data were the result of the sequential GPA and PCA on a "mixed" sample of 160 astragali and calcanei. "MD" stands for mean distance and values represent the continuous Procrustes distance of each specimen from the mean shape. P(0/1) stands for the probability of lambda being zero or one.

| **Phylogenetic Signal** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Astragalus** | | | | **Calcaneus** | | | |
| **Variable** | **lambda(CI)** | **P(0)** | **P(1)** | **Variable** | **lambda(CI)** | **P(0)** | **P(1)** |
| *mix* PC1 | 0.884 (0.578, NA) | <0.0001 | 0.13 | *mix* PC1 | 1.0 (0.924, NA) | <0.0001 | 1 |
| *mix* PC2 | 0.861 (0.623, NA) | <0.0001 | 0.06 | *mix* PC2 | 1.0 (0.919, NA) | <0.0001 | 1 |
| *mix* PC3 | 0.871 (0.638, NA) | <0.0001 | 0.06 | *mix* PC3 | 1.0 (0.954, NA) | <0.0001 | 1 |
| *mix* MD | 1.0 (0.855, NA) | <0.0001 | 1 | *mix* MD | 1.0 (0.949, NA) | <0.0001 | 1 |
| *sep* PC1 | 0.862 (0.641, NA) | <0.0001 | 0.05 | *sep* PC1 | 1.0 (0.945, NA) | <0.0001 | 1 |
| *sep* PC2 | 0.995 (0.856, NA) | <0.0001 | 0.89 | *sep* PC2 | 1.0 (0.942, NA) | <0.0001 | 1 |
| *sep* PC3 | 0.846 (0.339, 0.985) | 0.003 | 0.01 | *sep* PC3 | 1.0 (0.845, NA) | <0.0001 | 1 |
| *sep* MD | 0.990 (0.769, NA) | <0.0001 | 0.91 | *sep* MD | 1.0 (0.929, NA) | <0.0001 | 1 |

**Table 6A.** Correlations between PC scores of astragalus and calcaneus, and correlations between PC scores of mixed and separate bone analyses. Linear correlation (r) values in boxes on the left, (P) values in boxes on the right.

**Between Bone Correlations (comparisons within separate & mixed analyses)**

| **sep.** | **ast.** | | | | | **sep.** | **ast.** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **calc.** | **1** | **2** | **3** | **MD** | | **calc.** | **1** | **2** | **3** | **MD** |
| **1** | 0.86 | -0.17 | -0.13 | -- | | **1** | <0.0001 | ns | ns | -- |
| **2** | -0.08 | 0.86 | 0.05 | -- | | **2** | ns | <0.0001 | ns | -- |
| **3** | -0.16 | -0.02 | 0.02 | -- | | **3** | ns | ns | ns | -- |
| **MD** | -- | -- | -- | 0.57 | | **MD** | -- | -- | -- | <0.0001 |

| **mix.** | **ast.** | | | | | **mix.** | **ast.** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **calc.** | **1** | **2** | **3** | **MD** | | **calc.** | **1** | **2** | **3** | **MD** |
| **1** | 0.68 | 0.86 | 0.57 | -- | | **1** | <0.0001 | <0.0001 | <0.0001 | -- |
| **2** | 0.40 | 0.84 | 0.76 | -- | | **2** | 0.007 | <0.0001 | <0.0001 | -- |
| **3** | -0.25 | -0.76 | -0.80 | -- | | **3** | ns | <0.0001 | <0.0001 | -- |
| **MD** | -- | -- | -- | -0.25 | | **MD** | -- | -- | -- | ns |

**Within Bone Correlations (comparisons between separate & mixed analyses)**

| **calc.** | **mix.** | | | | | **calc.** | **mix.** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **sep.** | **1** | **2** | **3** | **MD** | | **sep.** | **1** | **2** | **3** | **MD** |
| 1 | -0.93 | -0.98 | 0.93 | -- | | 1 | <0.0001 | <0.0001 | <0.0001 | -- |
| 2 | 0.43 | -0.01 | 0.23 | -- | | 2 | 0.004 | ns | ns | -- |
| 3 | -0.08 | -0.01 | -0.05 | -- | | 3 | ns | ns | ns | -- |
| MD | -- | -- | -- | 0.45 | | MD | -- | -- | -- | 0.003 |

| **ast.** | **mix.** | | | | | **ast.** | **mix.** | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **sep.** | **1** | **2** | **3** | **MD** | | **sep.** | **1** | **2** | **3** | **MD** |
| **1** | -0.57 | -0.98 | -0.90 | -- | | **1** | <0.0001 | <0.0001 | <0.0001 | -- |
| **2** | 0.80 | 0.26 | -0.29 | -- | | **2** | <0.0001 | ns | ns | -- |
| **3** | -0.10 | 0.07 | -0.11 | -- | | **3** | ns | ns | ns | -- |
| **MD** | -- | -- | -- | 0.95 | | **MD** | -- | -- | -- | <0.0001 |

1310 **Table 6B.** Phylogenetically informed correlations between astragalus and calcaneus
1311 variables that resulted from sequential GPA followed by PCA on 1,024 pseudolandmarks
1312 per bone. See Table 5A for explanation of variable names.

**PGLS correlations**

| test | lambda(CI) | P(0) | P(1) | slope | r square | P |
|---|---|---|---|---|---|---|
| *sep* PC1 (ast. vs. calc.) | 1.0 (0.946, NA) | <0.0001 | 1 | 0.28 | 0.073 | 0.05 |
| *mix* PC1 (ast. vs. calc.) | 1.0 (0.924, NA) | <0.0001 | 1 | 0.84 | 0.204 | 0.0002 |
| *sep* MD (ast. vs. calc.) | 1.0 (0.925, NA) | <0.0001 | 1 | 0.1 | 0.057 | 0.79 |
| *mix* MD (ast. vs. calc.) | 1.0 (0.952, NA) | <0.0001 | 1 | -0.36 | 0.074 | 0.05 |

1313

**Figures/Captions**

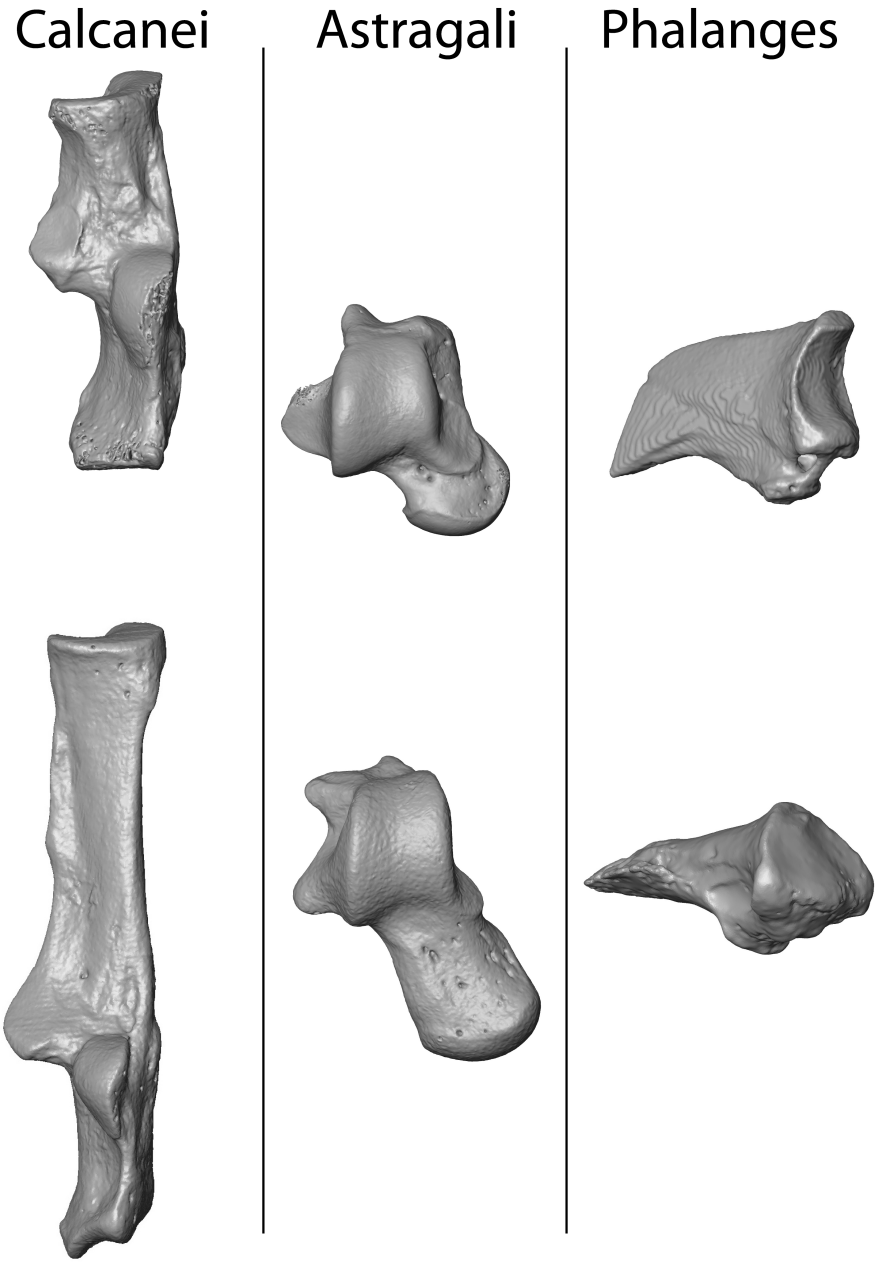



**Figure 1. Bones of the study.** This study utilizes scan datasets of three different types of bones. These datasets are chosen to challenge the automatic alignment algorithm we present with a range of geometric properties. The astragalus and calcaneus datasets are samples that represent geometrically complex bones with seemingly modest sample variance, while the distal phalanges are geometric more simple bones with apparently large sample variance. Analyses include one on a sample of 106 calcanei that is compared to a traditional 3DGM analysis using 27 landmarks by Gladman et al. (2013); one on a sample of 80 calcanei and 80 taxon-matched astragali in a single "mixed-bone" analysis; and one on a sample of 49 distal phalanges (Table 1).

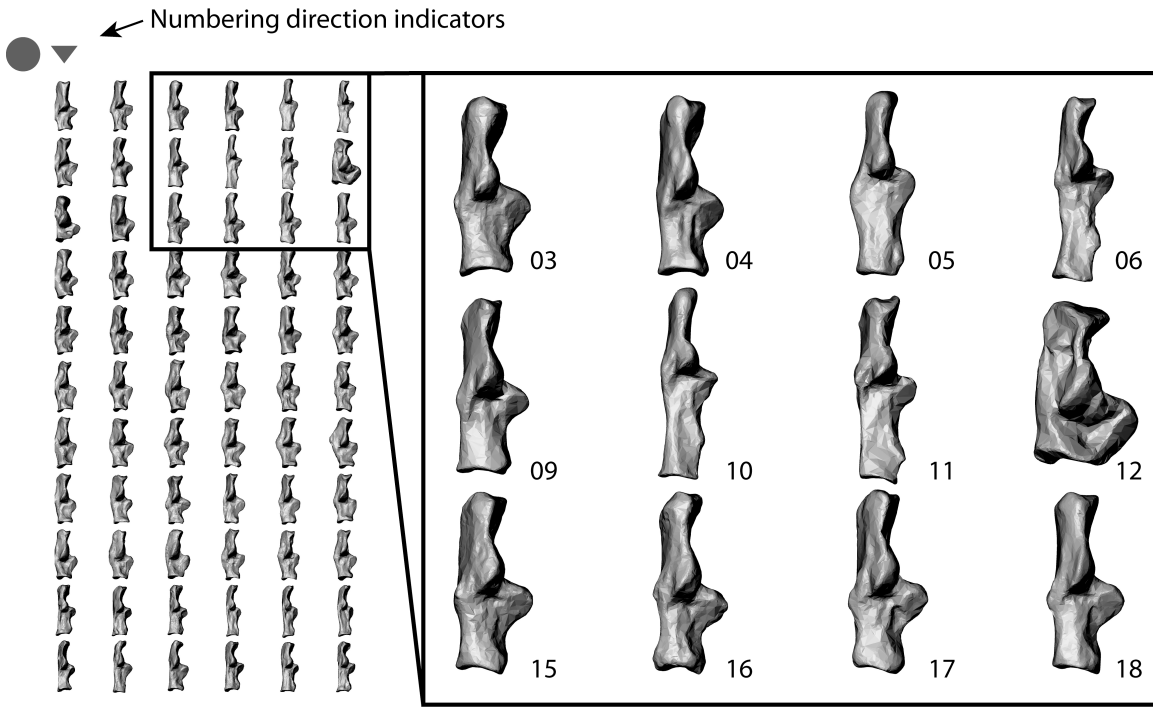



**Figure 2. Example of bones in an alignment file.** One of the outputs of the fully automated alignment algorithm is a 3D mesh file that shows all the specimens of the sample aligned. This allows the researcher to quickly survey the results to determine if he/she should proceed with shape analyses based on the implied correspondence. Sometimes one or more bones may be misaligned. If this results the researcher will catch it at this stage: we present several strategies for correcting such misalignments. The "numbering direction indicators" are mesh objects that show where the #1 bone in the spreadsheet is located. The arrow points down column #1, and numbering proceeds down rows. This allows the researcher to match bones in the alignment file with a spreadsheet containing any metadata on the surface files (like taxonomic information).

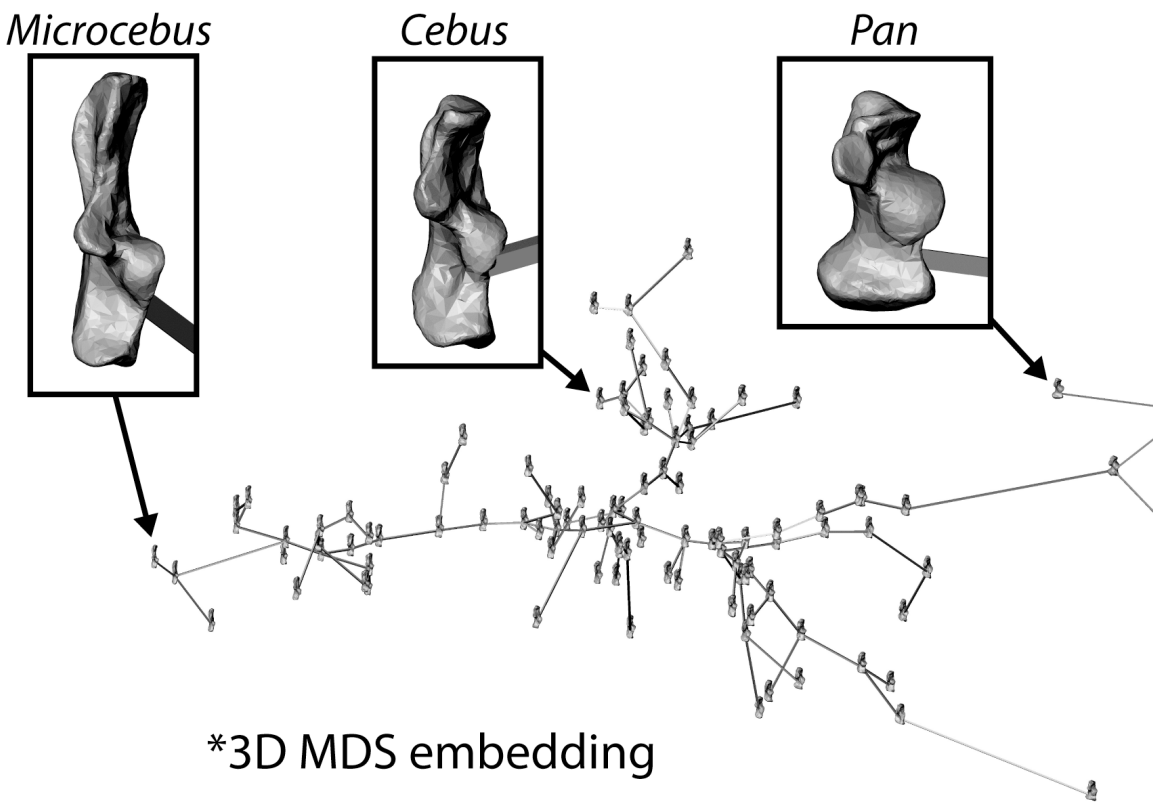



**Figure 3. Multi-Dimensional Scaling (MDS) & Minimum Spanning Tree (MST) embedding file.** This second output is of the same file type as that in Figure 2. It is however, less essential, because it is not useful for visualizing alignments and the data it presents can be re-calculated by the user later. The file simply displays the bones of the sample with their centroids embedded in the coordinate space of an MDS analysis result that is run on the pairwise distance matrix as determined via the MST. The MST is also shown. The point of this file is to give researchers a quick look at the clustering of their specimens.

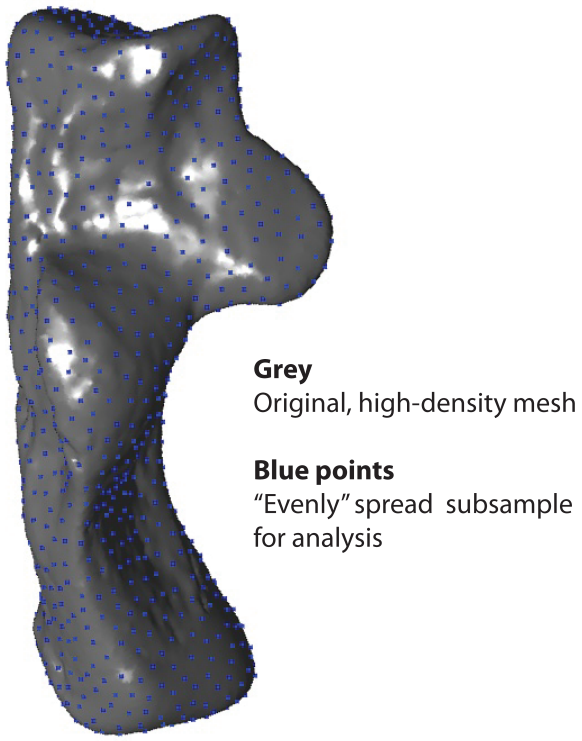



**Figure 4. Down-sampling meshes prior to analysis.** The algorithm is run on point clouds represented by a standard number of points specified by the researcher. These points are chosen by randomly picking a point on the surface, and then picking another point that is farthest from the first point, then by picking a third point whose position on the surface maximizes the sum distance between it and the two existing points, and so on until the specified number of points is achieved.

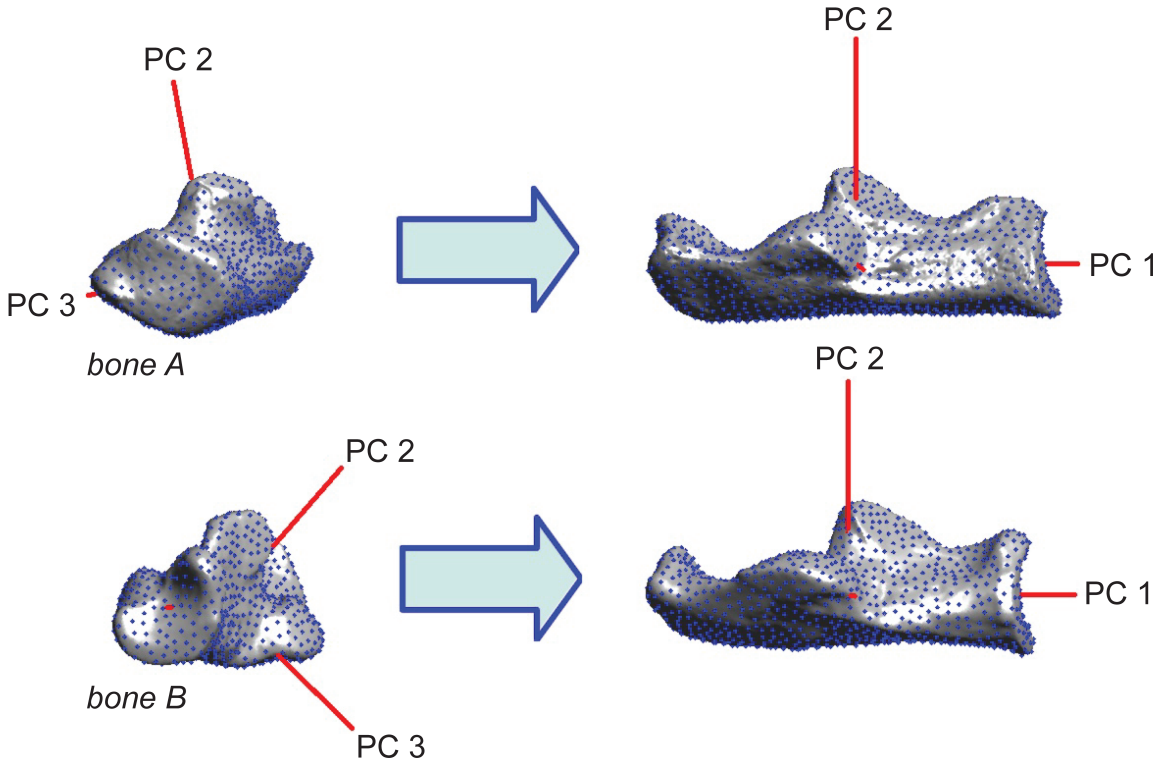



**Figure 5. Principal alignments to improve Iterative Closest Points (ICP) searches.** The best alignment between two bones is almost impossible to find using an ICP approach without any good initial guesses. The problem with supplying an initial guess is that usually this means user intervention is required. Our algorithm supplies at least eight initial guesses withoutuser intervention. It does this by computing the first three principal axes of variance and uses these axes as starting points for ICP. The principal axes along which the smallest continuous Procrustes distance between two shapes is found is almost always correct if the shapes are similar. This is a computationally rapid way of solving a complex problem. The algorithm performs better on samples with many incrementally intermediate shapes (see text and Fig. 4). Red lines on calcaneal surfaces represent principal axes of point variance. Shapes on left have yet to be aligned, while shapes on the right have been aligned so that their principal axes match.

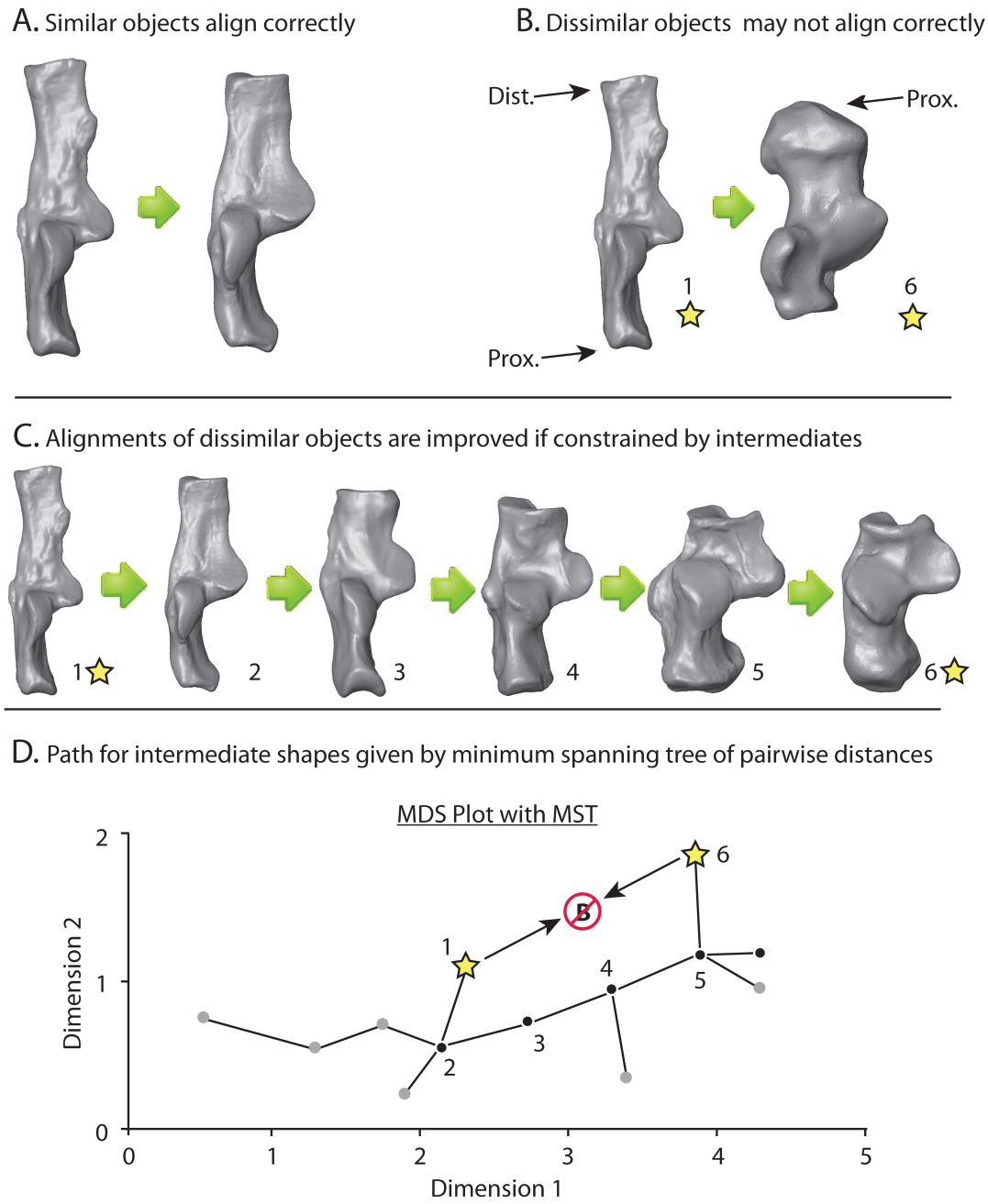
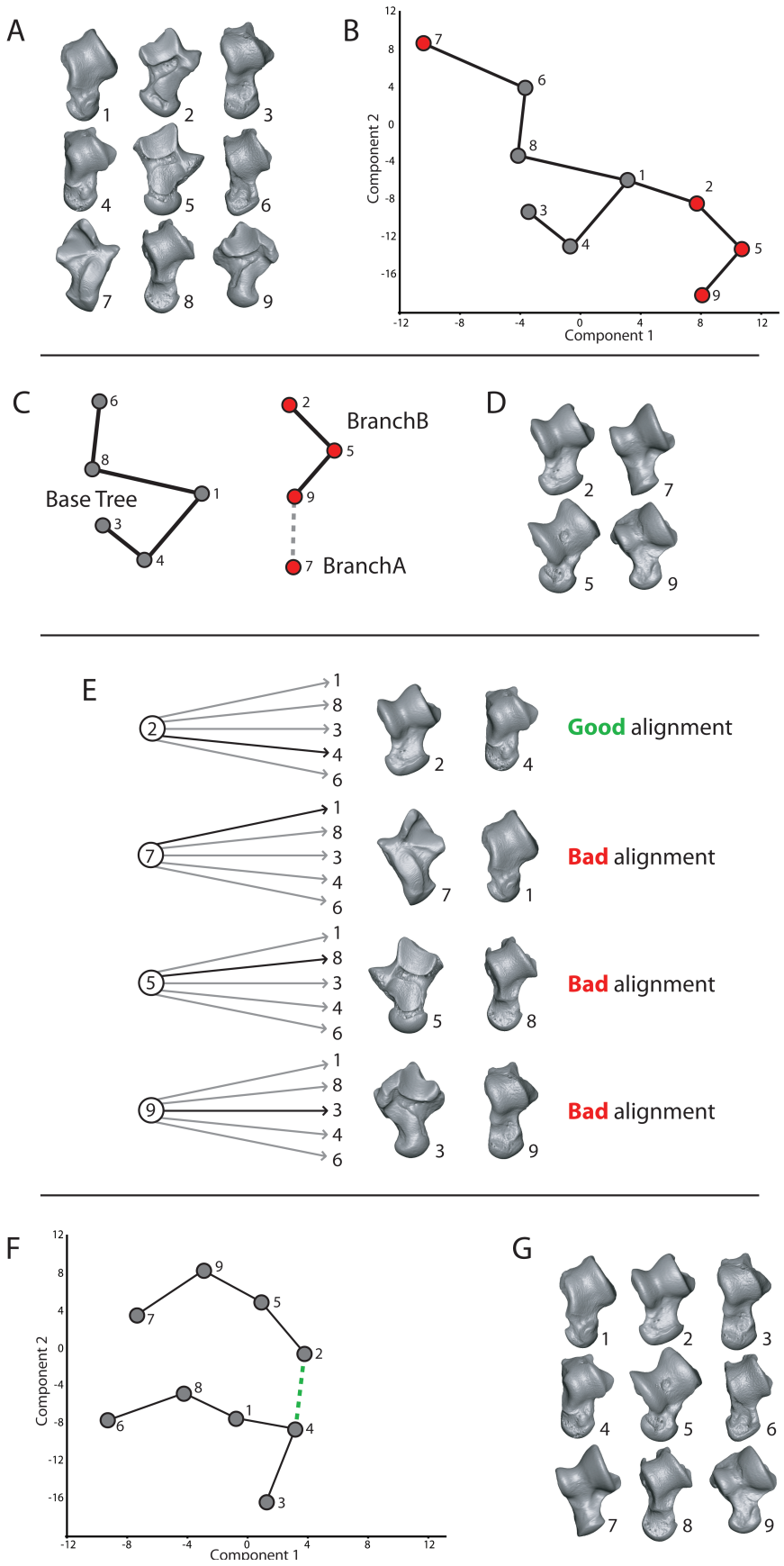



**Figure 6. Method for successfully aligning disparate shapes. A**, the result of applying our version of ICP to two similar shapes. **B**, the incorrect result that emerges when applying our ICP directly to two dis-similar shapes. In the first stage of the analysis, a pairwise distance matrix is calculated using "direct-matches" (even potentially incorrect ones as in B) between all shapes. That distance matrix is used to compute a minimum spanning tree. Because the minimum spanning tree connects only the most similar shapes, these connected pairs almost always represent correct alignments as in "A." **C.** These connections therefore define a path of intermediates that can be used to figure out the correct alignment between different shapes. **D**, The MST route is shown graphically.


A
B
Component 2
Component 1
C
Base Tree
BranchB
BranchA
D
E
Good alignment
Bad alignment
Bad alignment
Bad alignment
F
Component 2
Component 1
G

1379

**Figure 7. Schematic of alignFix protocol.** A) Visual inspection of initial alignment reveals several specimens are misaligned. B) Minimum spanning tree shows misaligned specimens (shown in red) can be found on two branches. C) Minimum spanning tree is broken into three components representing the base tree (in which all alignments are good), and Branches A and B (the misaligned specimens). D) Unsupervised alignment protocol is performed on originally unconnected branches A and B to determine if global alignment exists for those specimens when base tree specimens are excluded from consideration. Here, we show a successful global alignment. If no such alignment exists, then Branches A and B should be treated separately as if they had been a set connected to each other, as each was to the base tree. E) All misaligned specimens are compared to all specimens in the Base Tree to find the appropriate attachment point (i.e., a pair with a correct alignment). Several example alignments from this exhaustive process are shown here. Pairwise comparisons are visually inspected by the user to find an acceptable alignment with the lowest Procrustes distance between the two specimens. F) The designated pair serves as the connection (dotted line) for Branch A+B to the Base Tree. G) Recomputed global alignment using user determined tree in E reveals all specimens to now align correctly.

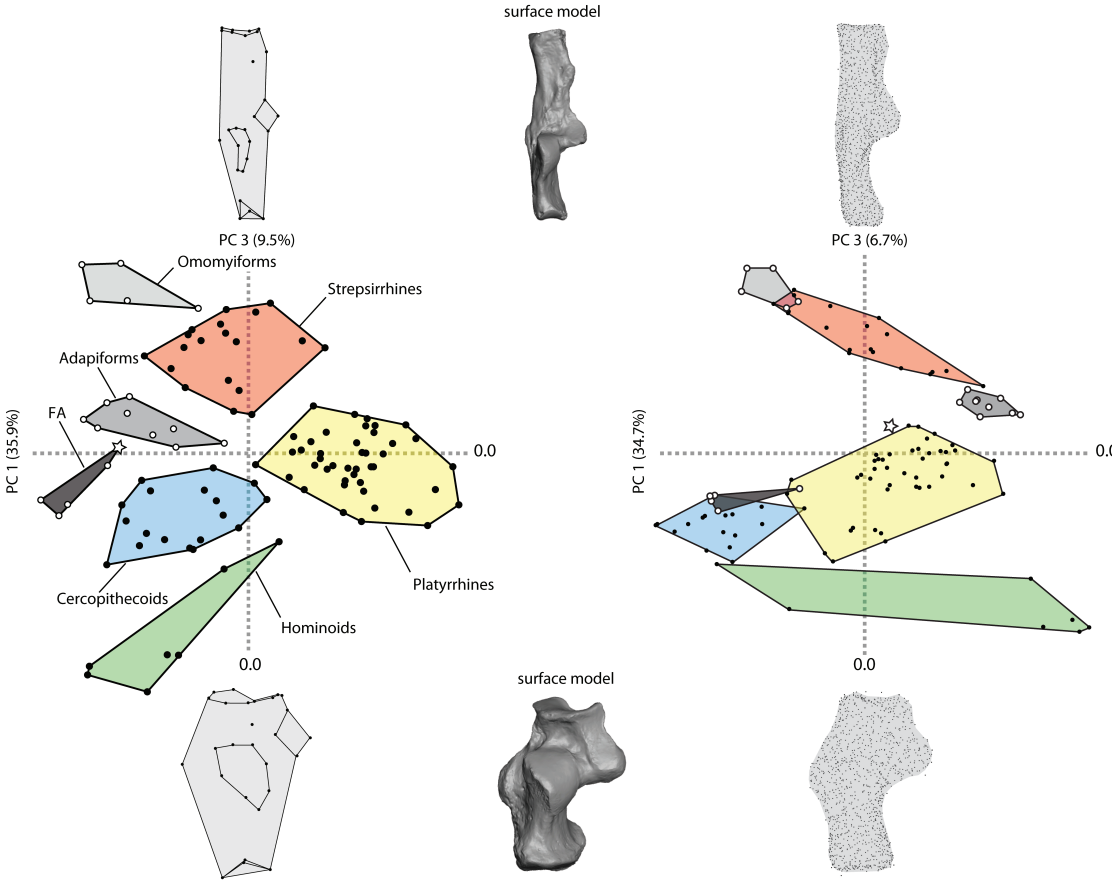



**Figure 8. Shape space of our analysis and comparison to a traditional 3DGM analysis. A**, PCA plot of principal component scores 1 and 3 for data from Gladman et al. (2013) based on 27 landmarks of the calcaneus in a sample of 106 bones. **B**, PCA plot of principal component scores 1 and 3 for the same sample, but as represented by 1,024 pseudolandmark points generated by the algorithm presented here. Both datasets, including our automated output, and that from Gladman et al. (2013) were analyzed with *morphologika*[2.5]. One of the benefits of the output of our algorithm is that it can be analyzed as if it were observer-collected data with traditional statistical software.

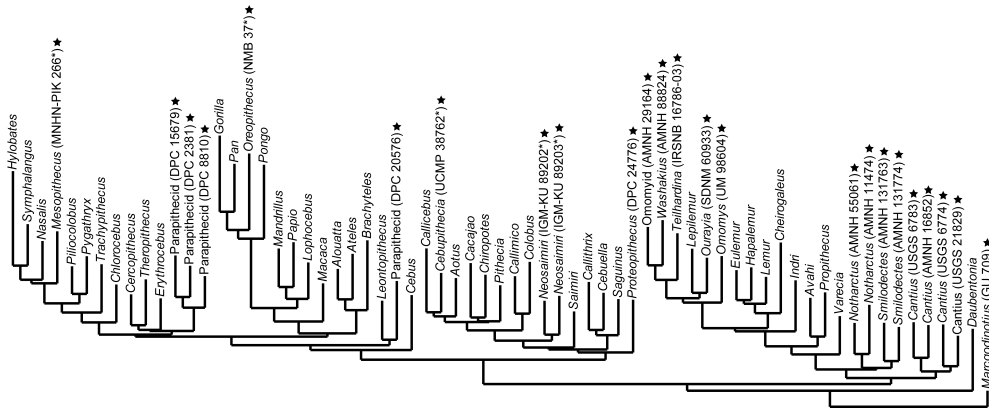



**Figure 9. Neighbor Joining tree.** To explore phenetic affinites implied by pseudolandmarks in the calcaneal dataset we averaged coordinate data from individual specimens into species means as described in the text and then performed three types of clustering algorithms, just as was also done by Gladman et al. (2013) for a 27 landmark traditional dataset. The neighbor-joining tree requires specification of a root to which nearest neighbors are attached. Fossils were not averaged. Therefore stars and specimen numbers represent individual fossils. These analyses were carried out in PAST (Hammer et al. 2001; 2006).

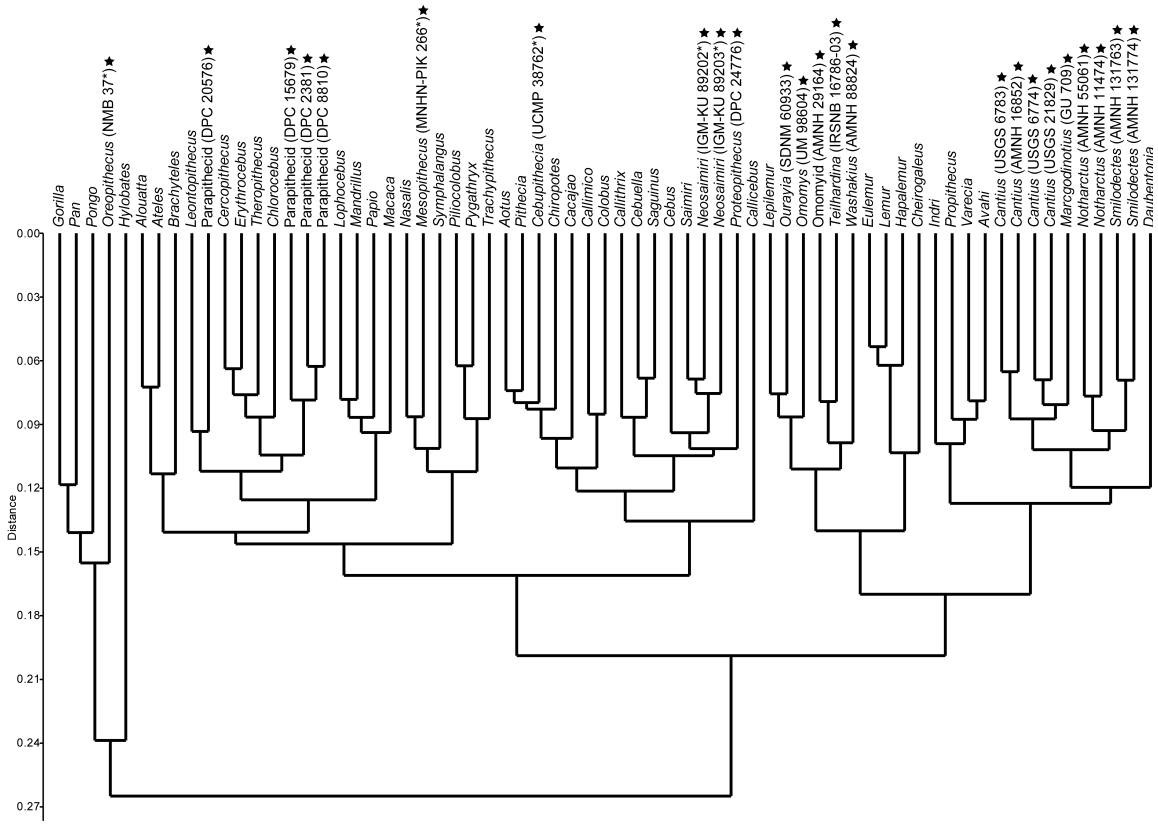



**Figure 10. UPGMA tree.** To explore phenetic affinities implied by pseudolandmarks in the calcaneal dataset we averaged coordinate data from individual specimens into species means as described in the text and then performed three types of clustering algorithms, just as was also done by Gladman et al. (2013) for a 27 landmark traditional dataset. Fossils were not averaged. Therefore stars and specimen numbers represent individual fossils. These analyses were carried out in PAST (Hammer et al. 2001; 2006).

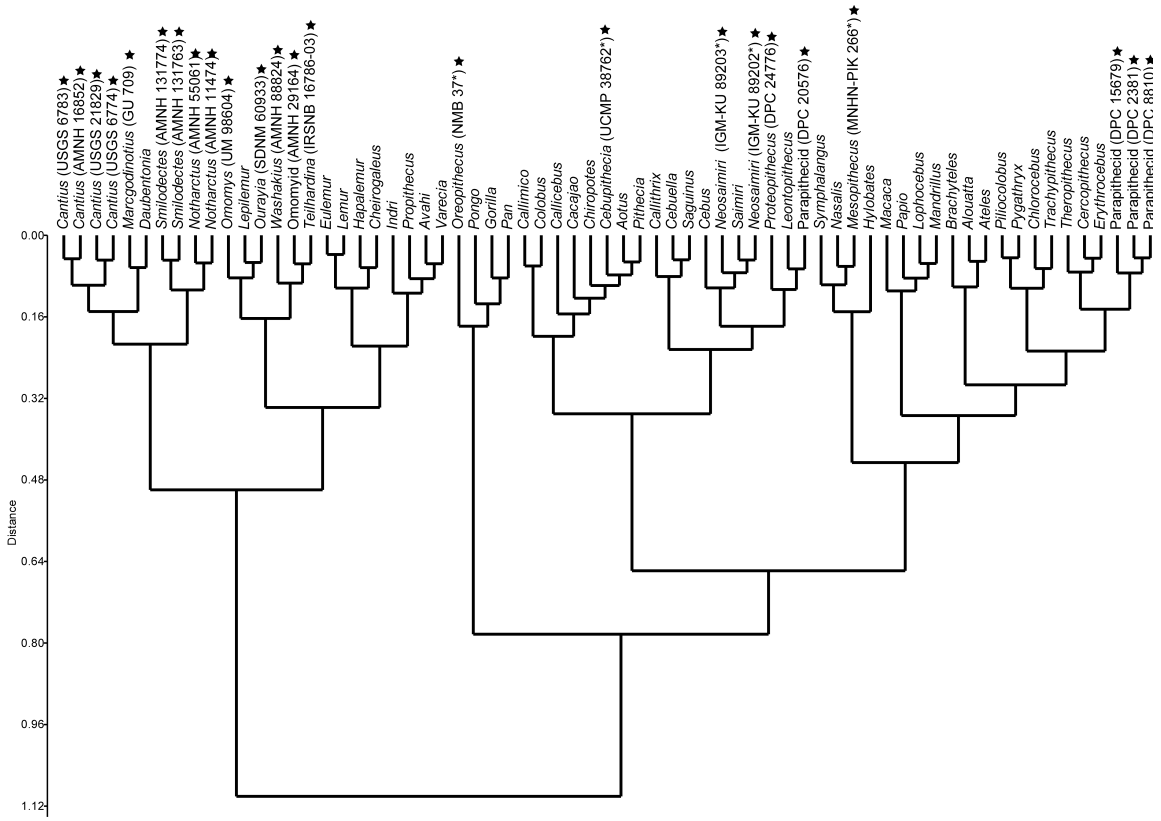
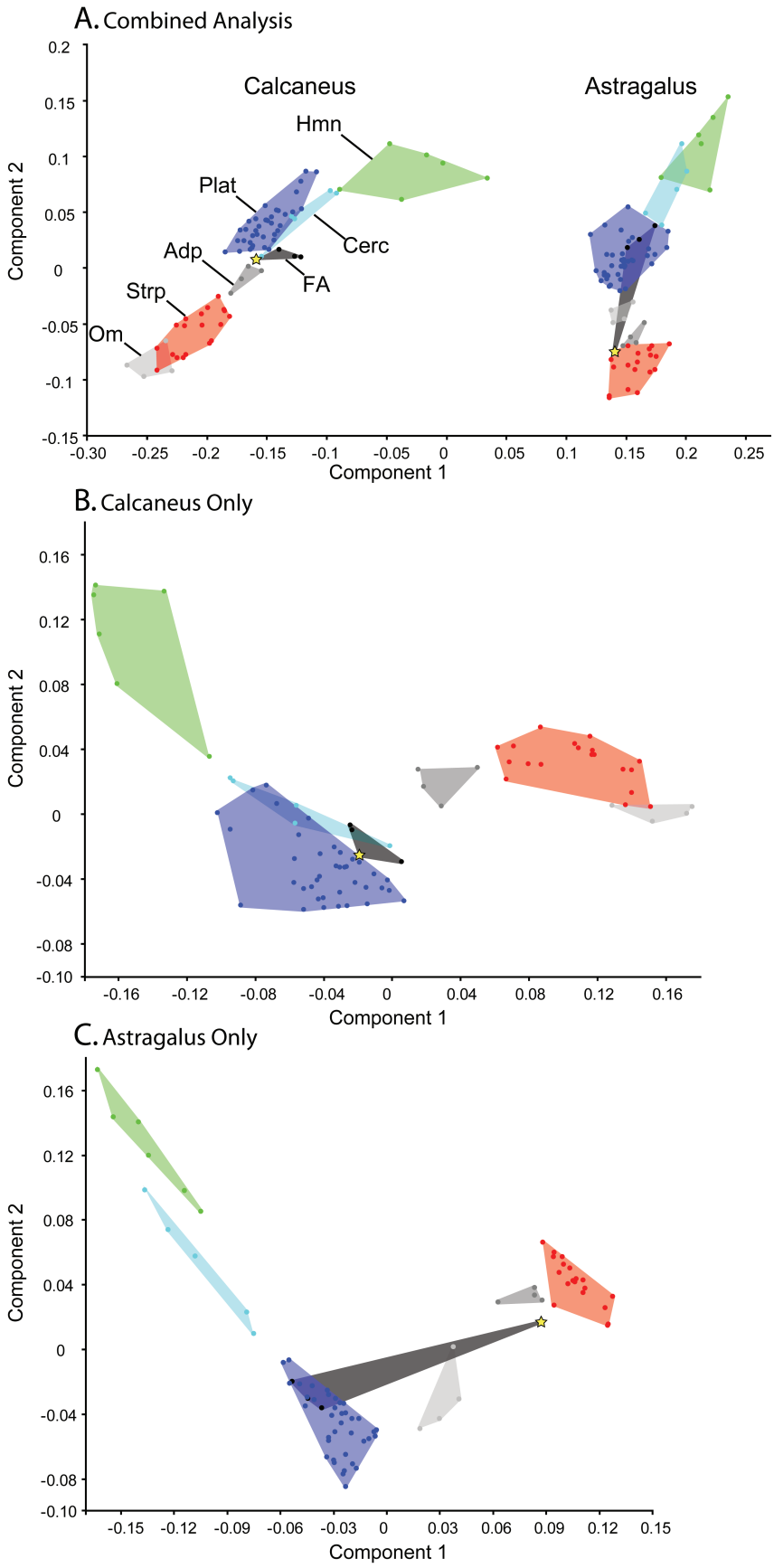



**Figure 11. Wards tree.** To explore phenetic affinities implied by pseudolandmarks in the calcaneal dataset we averaged coordinate data from individual specimens into species means as described in the text and then performed three types of clustering algorithms, just as was also done by Gladman et al. (2013) for a 27 landmark traditional dataset. Fossils were not averaged. Therefore stars and specimen numbers represent individual fossils. These analyses were carried out in PAST (Hammer et al. 2001; 2006).


A. Combined Analysis
Calcaneus
Astragalus
Hmn
Plat
Cerc
Adp
FA
Strp
Om
Component 2
Component 1
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.30
-0.25
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
0.2
0.25
B. Calcaneus Only
Component 2
Component 1
0.16
0.12
0.08
0.04
0
-0.04
-0.08
-0.10
-0.16
-0.12
-0.08
-0.04
0
0.04
0.08
0.12
0.16
C. Astragalus Only
Component 2
Component 1
0.16
0.12
0.08
0.04
0
-0.04
-0.08
-0.10
-0.15
-0.12
-0.09
-0.06
-0.03
0
0.03
0.06
0.09
0.12
0.15

1439

**Figure 12. Mixed bone analyses. A**, PCA plot (PC's 1 and 2) of the mixed bone analysis. MST's were established for each bone type independently using our FAA in the way described above with 1,024 pseudolandmark correspondence points for each set. Then we exhaustively computed the minimum Procrustes distance between every pair of astragalus and calcaneus. We used that pair with smallest distance to connect the calcaneal to the astragalar MST and allow the template to extend between two bones. Then we were able to run GPA and PCA on the mixed bone analysis. **B**, PCA plot (PC's 1 and 2) for the calcaneus when no astragali are included. **C**, PCA plot (PC's 1 and 2) for the astragalar dataset when no calcanei are included. The star represents the Fayum anthropoid *Proteopithecus*. Note that the there is good phylogenetic correlation with and between bones on the same axes whether the analyses are done on mixed or single bone samples. This is demonstrated quantitatively in Tables 6A-B.

**Supplemental information list**

Supplemental Figure 1. Alignment file as 3D pdf of 106 calcanei
Supplemental Figure 2. Alignment file as 3D pdf of 80 astragali
Supplemental Figure 3. Alginment file as 3D of 49 claws
Supplemental Table 1. Specimen numbers for 106 calcanei of first sample.
Supplemental Table 2. Specimen numbers for astragalus & calc pairs of second sample.
Supplemental Table 3. Specimen numbers for claws of third sample.
Supplemental Table 4. Specimen numbers for additional astragalus sample

**Table S1A.** Full calcaneal data set. Bone # column can be used to look up specimens in 3D alignment file as explained in text.

| Taxon | Specimen | Bone # |
|---|---|---|
| *Avahi laniger* | AMNH 170461 | 1 |
| *Cheirogaleus major* | AMNH 100640 | 2 |
| *Daubentonia madagascariensis* | AMNH 185643 | 3 |
| *Eulemur fulvus* | AMNH 17403 | 4 |
| *Eulemur fulvus* | AMNH 31254 | 5 |
| *Hapalemur griseus* | AMNH 170675 | 6 |
| *Hapalemur griseus* | AMNH 170689 | 7 |
| *Hapalemur griseus* | AMNH 61589 | 8 |
| *Indri indri* | AMNH 100504 | 9 |
| *Indri indri* | AMNH 208992 | 10 |
| *Lemur catta* | AMNH 150039 | 11 |
| *Lemur catta* | AMNH 170739 | 12 |
| *Lemur catta* | AMNH 22912 | 13 |
| *Lepilemur mustelinus* | AMNH 170565 | 14 |
| *Lepilemur mustelinus* | AMNH 170568 | 15 |
| *Lepilemur mustelinus* | AMNH 170569 | 16 |
| *Propithecus verreauxi* | AMNH 170463 | 17 |
| *Propithecus verreauxi* | AMNH 170491 | 18 |
| *Varecia variegata* | AMNH 100512 | 19 |
| *Alouatta seniculus* | AMNH 42316 | 20 |
| *Alouatta seniculus* | SBU NAl13 | 21 |
| *Alouatta sp.* | SBU NAl17 | 22 |
| *Alouatta sp.* | SBU NAl18 | 23 |
| *Aotus azarae* | AMNH 211482 | 24 |
| *Aotus infulatus* | AMNH 94992 | 25 |
| *Aotus sp.* | AMNH 201647 | 26 |
| *Ateles paniscus* | SBU NAt10 | 27 |
| *Ateles sp.* | SBU NAt13 | 28 |
| *Ateles sp.* | SBU NAt18 | 29 |
| *Brachyteles arachnoides* | AMNH 260 | 30 |
| *Cacajao calvus* | AMNH 70192 | 31 |
| *Cacajao calvus* | SBU NCj1 | 32 |
| *Callicebus donacophilus* | AMNH 211490 | 33 |
| *Callicebus moloch* | AMNH 244363 | 34 |
| *Callicebus moloch* | AMNH 94977 | 35 |
| *Callimico goeldi* | AMNH 183289 | 36 |
| *Callimico goeldi* | SBU NCa1 | 37 |

| | | |
|---|---|---|
| *Callithrix jacchus* | AMNH 133692 | 38 |
| *Callithrix jacchus* | AMNH 133698 | 39 |
| *Cebuella pygmaea* | AMNH 244101 | 40 |
| *Cebuella pygmaea* | SBU NC1 | 41 |
| *Cebus apella* | SBU NCb4 | 42 |
| *Cebus sp.* | SBU NCb5 | 43 |
| *Chiropotes satanus* | AMNH 95760 | 44 |
| *Chiropotes satanus* | AMNH 96123 | 45 |
| *Chiropotes sp.* | SBU NCh2 | 46 |
| *Leontopithecus rosalia* | AMNH 137270 | 47 |
| *Leontopithecus rosalia* | AMNH 60647 | 48 |
| *Pithecia monachus* | AMNH 187978 | 49 |
| *Pithecia pithecia* | AMNH 149149 | 50 |
| *Saguinus midas* | AMNH 266481 | 51 |
| *Saguinus mystax* | AMNH 188177 | 52 |
| *Saguinus sp.* | SBU NSg12 | 53 |
| *Saguinus sp.* | SBU NSg2 | 54 |
| *Saimiri boliviensis* | AMNH209934 | 55 |
| *Saimiri boliviensis* | AMNH211650 | 56 |
| *Saimiri boliviensis* | AMNH211651 | 57 |
| *Saimiri sciureus* | AMNH188080 | 58 |
| *Saimiri sp.* | SBU NSm2 | 59 |
| *Cercopithecus sp.* | SBU No Number | 60 |
| *Cercopithecus sp.* | SBU No Number | 61 |
| *Chlorocebus aethiops* | SBU OCr7 | 62 |
| *Chlorocebus cynosuros* | AMNH 80787 | 63 |
| *Colobus geureza* | AMNH 27711 | 64 |
| *Erythrocebus patas* | AMNH 34709 | 65 |
| *Lophocebus albigena* | AMNH 52603 | 66 |
| *Macaca nigra* | SBU OCn1 | 67 |
| *Macaca tonkeana* | AMNH 153402 | 68 |
| *Mandrillus sphinx* | AMNH 89367 | 69 |
| *Nasalis larvatus* | AMNH 106272 | 70 |
| *Papio hamadryas* | AMNH 80774 | 71 |
| *Piliocolobus badius* | AMNH 52303 | 72 |
| *Piliocolobus badius* | ED 4651 | 73 |
| *Pygathrix nemaeus* | AMNH 87255 | 74 |
| *Theropitheucs gelada* | AMNH 201008 | 75 |
| *Trachypithecus obscurus* | AMNH 112977 | 76 |
| *Gorilla sp.* | AD 6001 | 77 |
| *Hylobates lar* | AMNH 119601 | 78 |
| *Pan troglodytes* | AMNH 51202 | 79 |
| *Pan troglodytes* | AMNH 51278 | 80 |
| *Pongo pygmaeus* | AMNH 28253 | 81 |
| *Symphalangus syndactylus* | AMNH 106583 | 82 |
| *Cantius abditus* | USGS 6783 | 83 |
| *Cantius sp.* | USGS 6774 | 84 |
| *Cantius trigonodus* | AMNH 16852 | 85 |
| *Cantius trigonodus* | USGS 21829 | 86 |
| *Cebupithecia sarmientoi* | UCMP 38762* | 87 |

| | | |
|---|---|---|
| *Marcgodinotius indicus* | GU 709 | 88 |
| *Mesopithecus pentelici* | MNHN PIK-266* | 89 |
| *Neosaimiri fieldsi* | IGM-KU 89202* | 90 |
| *Neosaimiri fieldsi* | IGM-KU 89203* | 91 |
| *Notharctus sp.* | AMNH 55061 | 92 |
| *Notharctus tenebrosus* | AMNH 11474 | 93 |
| *Omomyid* | AMNH 29164 | 94 |
| *Omomys sp.* | UM 98604 | 95 |
| *Oreopithecus bambolii* | NMB 37* | 96 |
| *Ourayia uintensis* | SDNM 60933 | 97 |
| *Parapithecid* | DPC 15679 | 98 |
| *Parapithecid* | DPC 20576 | 99 |
| *Parapithecid* | DPC 2381 | 100 |
| *Parapithecid* | DPC 8810 | 101 |
| *Proteopithecus sylviae* | DPC 23662A | 102 |
| *Smilodectes gracilis* | AMNH131763 | 103 |
| *Smilodectes gracilis* | AMNH131774 | 104 |
| *Teihardina belgica* | IRSNB 16786-03 | 105 |
| *Washakius insignis* | AMNH 88824 | 106 |

1464

**Table S1B.** Reduced sample of calcaneal specimens for combining with Astragali (Table S2)

| Taxon | Specimen | Bone # |
|---|---|---|
| *Avahi laniger* | AMNH 170461 | 1 |
| *Cheirogaleus major* | AMNH 100640 | 2 |
| *Daubentonia madagascariensis* | AMNH 185643 | 3 |
| *Eulemur fulvus* | AMNH 17403 | 4 |
| *Eulemur fulvus* | AMNH 31254 | 5 |
| *Hapalemur griseus* | AMNH 170675 | 6 |
| *Hapalemur griseus* | AMNH 170689 | 7 |
| *Hapalemur griseus* | AMNH 61589 | 8 |
| *Indri indri* | AMNH 100504 | 9 |
| *Indri indri* | AMNH 208992 | 10 |
| *Lemur catta* | AMNH 150039 | 11 |
| *Lemur catta* | AMNH 170739 | 12 |
| *Lemur catta* | AMNH 22912 | 13 |
| *Lepilemur mustelinus* | AMNH 170565 | 14 |
| *Lepilemur mustelinus* | AMNH 170568 | 15 |
| *Lepilemur mustelinus* | AMNH 170569 | 16 |
| *Propithecus verreauxi* | AMNH 170463 | 17 |
| *Propithecus verreauxi* | AMNH 170491 | 18 |
| *Varecia variegata* | AMNH 100512 | 19 |
| *Alouatta seniculus* | AMNH 42316 | 20 |
| *Alouatta seniculus* | SBU NAl13 | 21 |
| *Alouatta sp.* | SBU NAl17 | 22 |
| *Aotus azarae* | AMNH 211482 | 24 |
| *Aotus infulatus* | AMNH 94992 | 25 |
| *Aotus sp.* | AMNH 201647 | 26 |
| *Ateles paniscus* | SBU NAt10 | 27 |
| *Ateles sp.* | SBU NAt13 | 28 |
| *Ateles sp.* | SBU NAt18 | 29 |
| *Brachyteles arachnoides* | AMNH 260 | 30 |
| *Cacajao calvus* | AMNH 70192 | 31 |
| *Cacajao calvus* | SBU NCj1 | 32 |
| *Callicebus donacophilus* | AMNH 211490 | 33 |
| *Callicebus moloch* | AMNH 244363 | 34 |
| *Callicebus moloch* | AMNH 94977 | 35 |
| *Callimico goeldi* | AMNH 183289 | 36 |
| *Callimico goeldi* | SBU NCa1 | 37 |
| *Callithrix jacchus* | AMNH 133692 | 38 |
| *Callithrix jacchus* | AMNH 133698 | 39 |
| *Cebuella pygmaea* | AMNH 244101 | 40 |
| *Cebuella pygmaea* | SBU NC1 | 41 |
| *Cebus apella* | SBU NCb4 | 42 |
| *Cebus sp.* | SBU NCb5 | 43 |
| *Chiropotes satanus* | AMNH 95760 | 44 |
| *Chiropotes satanus* | AMNH 96123 | 45 |
| *Chiropotes sp.* | SBU NCh2 | 46 |
| *Leontopithecus rosalia* | AMNH 137270 | 47 |
| *Leontopithecus rosalia* | AMNH 60647 | 48 |
| *Pithecia monachus* | AMNH 187978 | 49 |
| *Pithecia pithecia* | AMNH 149149 | 50 |
| *Saguinus midas* | AMNH 266481 | 51 |
| *Saguinus mystax* | AMNH 188177 | 52 |
| *Saguinus sp.* | SBU NSg12 | 53 |

| | | |
|---|---|---|
| *Saimiri boliviensis* | AMNH209934 | 55 |
| *Saimiri boliviensis* | AMNH211650 | 56 |
| *Saimiri boliviensis* | AMNH211651 | 57 |
| *Chlorocebus aethiops* | SBU OCr7 | 62 |
| *Macaca nigra* | SBU OCn1 | 67 |
| *Macaca tonkeana* | AMNH 153402 | 68 |
| *Nasalis larvatus* | AMNH 106272 | 70 |
| *Trachypithecus obscurus* | AMNH 112977 | 76 |
| *Gorilla sp.* | AD 6001 | 77 |
| *Hylobates lar* | AMNH 119601 | 78 |
| *Pan troglodytes* | AMNH 51202 | 79 |
| *Pan troglodytes* | AMNH 51278 | 80 |
| *Pongo pygmaeus* | AMNH 28253 | 81 |
| *Symphalangus syndactylus* | AMNH 106583 | 82 |
| *Cantius abditus* | USGS 6783 | 83 |
| *Cebupithecia sarmientoi* | UCMP 38762* | 87 |
| *Marcgodinotius indicus* | GU 709 | 88 |
| *Neosaimiri fieldsi* | IGM-KU 89203* | 91 |
| *Notharctus sp.* | AMNH 55061 | 92 |
| *Notharctus tenebrosus* | AMNH 11474 | 93 |
| *Omomyid* | AMNH 29164 | 94 |
| *Omomys sp.* | UM 98604 | 95 |
| *Parapithecid* | DPC 15679 | 98 |
| *Parapithecid* | DPC 20576 | 99 |
| *Parapithecid* | DPC 2381 | 100 |
| *Proteopithecus sylviae* | DPC 23662A | 102 |
| *Teihardina belgica* | IRSNB 16786-03 | 105 |
| *Washakius insignis* | AMNH 88824 | 106 |

1466 **Table S2.** Astragalar specimens

| Genus | Specimen | Bone # |
|---|---|---|
| *Cebus* | AMNH 133606 | 1 |
| *Cebus* | AMNH 133608 | 2 |
| *Macaca* | MCZ 34714 | 3 |
| *Cheirogaleus* | DPC 031 | 4 |
| *Chiropotes* | AMNH 95760 | 5 |
| *Chiropotes* | SBU NCh2 | 6 |
| *Chiropotes* | AMNH 95027 | 7 |
| *Daubentonia* | AMNH 185643 | 8 |
| *Eulemur fulvus* | AMNH 170728 | 9 |
| *Eulemur fulvus* | AMNH 31254 | 10 |
| *Hapalemur* | AMNH 61589 | 11 |
| *Hapalemur* | AMNH 170680 | 12 |
| *Hapalemur* | AMNH 170689 | 13 |
| *Hylobates* | MCZ 41456 | 14 |
| *Hylobates* | MCZ 41458 | 15 |
| *Lemur catta* | AMNH 170739 | 16 |
| *Lemur* | AMNH 170740 | 17 |
| *Lemur* | AMNH 170765 | 18 |
| *Lepilemur* | AMNH 170556 | 19 |
| *Lepilemur* | AMNH 170560 | 20 |
| *Lepilemur* | AMNH 170565 | 21 |
| *Nasalis* | MCZ 37327 | 22 |
| *Pan* | AMNH 167343 | 23 |
| *Pan* | NMNH 176229 | 24 |
| *Pithecia* | AMNH 149149 | 25 |
| *Pithecia* | AMNH 187978 | 26 |
| *Alouatta* | AMNH 211585 | 27 |
| *Alouatta* | SBU NAl13 | 28 |
| *Alouatta* | SBU NAl18 | 29 |
| *Pongo* | NMNH 49853 | 30 |
| *Propithecus* | AMNH 170474 | 31 |
| *Propithecus* | AMNH 170463 | 32 |
| *Indri* | AMNH 100504 | 33 |
| *Indri* | AMNH-208992 | 34 |
| *Saguinus* | AMNH 188174 | 35 |
| *Saguinus* | 33B AMNH 97316 | 36 |
| *Saguinus* | AMNH 207726 | 37 |
| *Saimiri* | AMNH 209934 | 38 |
| *Saimiri* | SBU NSm06 | 39 |
| *Saimiri* | SBU Sm2 | 40 |
| *Notharctus* | AMNH 11474 | 41 |
| *Notharctus* | AMNH 129382 | 42 |
| *Omomys* | UM 38321 | 43 |
| *Omomys* | UM 98648 | 44 |
| *Teilhardina* | IRSNB vert-16786-01 | 45 |
| *Aotus* | AMNH 239851 | 46 |
| *Aotus* | AMNH 201647 | 47 |
| *Aotus* | AMNH 94992 | 48 |
| *Cebuella* | AMNH 244101 | 49 |
| *Cebuella* | SBU NC1 | 50 |
| *Marcgodinotius* | GU 748 | 51 |
| *Callimico* | AMNH 183289 | 52 |
| *Callimico* | SBU NCm01 | 53 |

| | | |
|---|---|---|
| *Varecia* | AMNH 100512 | 54 |
| *Apidium* | DPC5027 | 55 |
| *Apidium* | DPC 5416A | 56 |
| *Apidium* | DPC1001 | 57 |
| *Proteopithecus* | DPC22844 | 58 |
| *Ateles* | AMNH 259 | 59 |
| *Ateles* | AMNH 172985 | 60 |
| *Ateles* | SBU NAt10 | 61 |
| *Cacajao* | AMNH-70192 | 62 |
| *Cacajao* | SBU NCj1 | 63 |
| *Callicebus* | AMNH 210393 | 64 |
| *Callicebus* | AMNH 211491 | 65 |
| *Callicebus* | AMNH 211488 | 66 |
| *Callithrix* | AMNH 133698 | 67 |
| *Callithrix* | AMNH 133702 | 68 |
| *Cantius* | USGS 21832 | 69 |
| *Trachypithecus* | AMNH 11297 | 70 |
| *Avahi* | AMNH 170461 | 71 |
| *Leontopithecus* | AMNH 185347 | 72 |
| *Brachyteles* | AMNH 260 | 73 |
| *Washakius* | UM 99074 | 74 |
| *Gorilla* | MCZ 20038 | 75 |
| *Neosaimiri* | Neosiaimiri | 76 |
| *Macaca* | SBU OCN1 | 77 |
| *Chlorocebus* | SBU OCr7 | 78 |
| *Cebupithecia* | UCMP 38762 | 79 |
| *Leontopithecus* | USNM 588177 | 80 |

**Table S3.** Distal Phalanx Specimens

| Taxon | Specimen | Bone | R/L | Bone # |
|---|---|---|---|---|
| *Tarsius bancanus* | AMNH 106754 | P/dp2 | R | 001 |
| *Tarsius bancanus* | AMNH 106754 | P/dp3 | R | 002 |
| *Tarsius spectrum* | AMNH 109367 | P/dp2 | R | 003 |
| *Tarsius spectrum* | AMNH 109367 | P/dp3 | R | 004 |
| *Notharctus tenebrosus* | AMNH 143612-3 | P/dp2 | R | 005 |
| *Hemiechinus auritus* | AMNH 185374A | P/dp4 | L | 006 |
| *Hemiechinus auritus* | AMNH 185374B | P/dp3 | L | 007 |
| *Hemiechinus auritus* | AMNH 185374C | P/dp2 | L | 008 |
| *Hemiechinus auritus* | AMNH 185374D | M/dp4 | L | 009 |
| *Erinaceus europaeus* | AMNH 3770A | P/dp4 | L | 010 |
| *Erinaceus europaeus* | AMNH 3770B | P/dp2 | L | 011 |
| *Erinaceus europaeus* | AMNH 3770C | M/dpX | L | 012 |
| *Erinaceus roumanicus* | AMNH 69553A | P/dp1 | L | 013 |
| *Erinaceus roumanicus* | AMNH 69553B | P/dp2 | L | 014 |
| *Erinaceus roumanicus* | AMNH 69553C | P/dp3 | L | 015 |
| *Erinaceus roumanicus* | AMNH 69553E | P/dp4 | L | 016 |
| *Galago senegalensis* | DPC 003 | P/dp2 | L | 017 |
| *Cheirogaleus medius* | DPC 0130 | P/dp2 | R | 018 |
| *Otolemur crassicaudatus* | DPC 024 | P/dp2 | R | 019 |
| *Microcebus murinus* | DPC 035 | P/dp2 | R | 020 |
| *Mirza coquereli* | DPC 097 | P/dp2 | L | 021 |
| *Galago senegalensis* | DPC 1063F | P/dp2 | R | 022 |
| *Cheirogaleus medius* | DPC 1285 | P/dp2 | L | 023 |
| *Propithecus verreauxi* | DPC 1397 | P/dp2 | L | 024 |
| *Aotus sp.* | DPC nn | P/dp2 | R | 025 |
| *Aotus sp.* | SBU-11 | P/dp2 | R | 026 |
| *Hapalemur griseus* | SBU-12 | P/dp2 | L | 027 |
| *Varecia sp.* | SBU 1383 | P/dp2 | L | 028 |
| *Eulemur fulvus* | SBU-13 | P/dp2 | L | 029 |
| *Indri indri* | SBU 1474 | P/dp2 | R | 030 |
| *Lemur catta* | SBU-14 | P/dp2 | L | 031 |
| *Galago senegalensis* | SBU-15 | P/dp2 | L | 032 |
| *Propithecus diadema* | SBU 1155 | P/dp2 | L | 033 |
| *Otolemur crassicaudatus* | SBU PGa1163 | P/dp2 | R | 034 |
| *Incertae sedis* | UCMP 217999 | P/dp2 | L | 035 |
| *Incertae sedis* | UCMP 218000 | P/dp2 | R | 036 |
| *Tarsius pumilus* | USNM 196477 | P/dp2 | R | 037 |
| *Tarsius pumilus* | USNM 196477 | P/dp3 | R | 038 |
| *Carpolestes simpsoni* | UM 101963A | ?/dpX | ? | 039 |
| *Carpolestes simpsoni* | UM 101963B | ?/dpX | ? | 040 |
| *Carpolestes simpsoni* | UM 101963C | ?/dpX | ? | 041 |
| *Carpolestes simpsoni* | UM 101963D | ?/dpX | ? | 042 |
| *Ignacius clarksforkensis* | UM 82606 | ?/dpX | ? | 043 |
| *Plesiadapis churchilli* | SMM P77.33.517 | ?/dpX | ? | 044 |
| *Nannodectes intermedius* | USNM 442290 | ?/dpX | ? | 045 |
| *Incertae sedis* | UCMP 217919 | ?/dpX | ? | 046 |
| *Incertae sedis* | UCMP 217935 | ?/dpX | ? | 047 |
| *Incertae sedis* | UCMP 218245 | ?/dpX | ? | 048 |
| *Incertae sedis* | UCMP 218246 | ?/dpX | ? | 049 |

1468 **Table S4**. Additional astragali specimens

| Genus | Specimen | Bone # |
|---|---|---|
| *Alouatta* | AMNH 211585 | 001 |
| *Aotus* | AMNH 239851 | 002 |
| *Ateles* | AMNH 259 | 003 |
| *Cacajao* | AMNH 201122 | 004 |
| *Callicebus* | AMNH 210393 | 005 |
| *Callithrix* | AMNH 133688 | 006 |
| *Cebus* | AMNH 133606 | 007 |
| *Cheirogaleus* | DPC 0142 | 008 |
| *Cynocephalus* | AMNH 207001 | 009 |
| *Cynocephalus* | UNSM 15502 | 010 |
| *Galeopterus* | USNM 317118 | 011 |
| *Daubentonia* | USNM 119694 | 012 |
| *Eulemur* | AMNH 170708 | 013 |
| *Hapalemur* | AMNH 61589 | 014 |
| *Lemur* | AMNH 170739 | 015 |
| *Lepilemur* | AMNH 170556 | 016 |
| *Microcebus* | AMNH 174428 | 017 |
| *Nycticebus* | AMNH 90381 | 018 |
| *Perodicticus* | AMNH 269851 | 019 |
| *Pithecia* | AMNH 149149 | 020 |
| *Tarsius* | AMNH 203296 | 021 |
| *Tarsius* | AMNH 106754 | 022 |
| *Tarsius* | AMNH 109369 | 023 |
| *Ptilocercus* | USNM 488055 | 024 |
| *Ptilocercus* | USNM 488058 | 025 |
| *Tupaia* | SBU MIN2 | 026 |
| *Tupaia* | AMNH 215176 | 027 |
| *Loris* | AMNH 150038 | 028 |
| *Lepus* | SBU MLG3 | 029 |
| *Lepus* | SBU MLG4 | 030 |
| *Sylvilagus* | Boyer collection | 031 |
| *Ochotona* | AMNH 124392 | 032 |
| *Erethizon* | Boyer collection | 033 |
| *Coendou* | AMNH 80045 | 034 |
| *Marmota* | Boyer collection | 035 |
| *Sciurus* | SBU MRd10 | 036 |
| *Aplodontia* | AMNH 142747 | 037 |
| *Allactaga* | AMNH 227 | 038 |
| *Tenrec* | AMNH 170513 | 039 |
| *Setifer* | AMNH 170547 | 040 |
| *Hemicentetes* | AMNH 170593 | 041 |
| *Echinops* | AMNH 170607 | 042 |
| *Potamogale* | AMNH 55204 | 043 |
| *Erinaceus* | AMNH 3770 | 044 |
| *Hemiechinus* | AMNH 180318 | 045 |
| *Chrysochloris* | AMNH 205 | 046 |
| *Crocidura* | AMNH 48490 | 047 |
| *Desmana* | AMNH 97807 | 048 |
| *Solenodon* | AMNH 77745 | 049 |
| *Potos* | AMNH 267053 | 050 |
| *Arctictis* | AMNH 119600 | 051 |
| *Nasua* | AMNH 14062 | 052 |
| *Petrodromus* | AMNH 115790 | 053 |

1469