A Resampling Approach and Implications for Estimating the Phalangeal Index From Unassociated Hand Bones in Fossil Primates

Vivek V. Venkataraman,[1] Campbell Rolian,[2] Adam D. Gordon,[3] and Biren A. Patel[4]*

[1]Department of Biological Sciences, Class of 1978 Life Sciences Center, Dartmouth College, Hanover, NH 03755
[2]Department of Comparative Biology and Experimental Medicine, Faculty of Veterinary Medicine, University of Calgary, Calgary, Alberta, Canada T2N4N1
[3]Department of Anthropology, University at Albany–SUNY, Albany, NY 12222
[4]Department of Cell and Neurobiology, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033

ABSTRACT   Primate fossil assemblages often have metacarpals and phalanges from which functional/behavioral interpretations may be inferred. For example, intrinsic hand proportions can indicate hand function and substrate use. But, estimates of intrinsic hand proportions from unassociated hand elements can be imperfect due to digit misattribution. Although isolated metacarpals can be identified to a specific digit, phalanges are difficult to assign to a specific ray. We used a resampling approach to evaluate how estimates of intrinsic hand proportions are affected by such uncertainty. First, the phalangeal index—intermediate phalanx length plus proximal phalanx length divided by metacarpal length—for the third digit was calculated for associated specimens of terrestrial, semiterrestrial, and arboreal taxa. We then used resampling procedures to generate distributions of "composite digits" based on resampled ratios in which phalanges from the second, fourth, and fifth rays, and from different individuals, were chosen randomly. Results confirm that the phalangeal index for associated third digits significantly discriminates groups. We also found that resampled ratios had significantly lower means, indicating that using composite digits is prone to systematic underestimation. Resampled ratios also generated distributions with greater variance around the means that obscured distinctions between groups, although significant differences between the most arboreal and terrestrial taxa are maintained. We conclude that using unassociated phalanges to calculate a phalangeal index is prone to sampling bias. Nevertheless, a resampling approach has the potential to inform estimates of hand proportions for fossil taxa, provided that the comparative sample is constrained to mimic the fossil composition. Am J Phys Anthropol 151:280–289, 2013.   © 2013 Wiley Periodicals, Inc.

The skeletal intrinsic hand proportions of mammals, including primates, have close functional associations with locomotor behaviors and substrate preferences. For example, arboreal primates have long intermediate and proximal phalanges relative to metacarpal length, which facilitates grasping around circular substrates during above-branch quadrupedal or below-branch suspensory locomotion (Washburn, 1951; Napier, 1993; Lemelin, 1999; Hamrick, 2001; Kirk et al., 2008). In contrast, terrestrial primates have short digits relative to palm length to help reduce costly bending moments acting on fragile hand bones and joints while traveling on flatter, less compliant surfaces that result in high substrate reaction forces (Midlo, 1934; Etter, 1973; Susman, 1979; Nieschalk and Demes, 1993; Richmond, 1998; Patel et al., 2009; Patel and Wunderlich, 2010; Patel et al., 2012).

Associations between hand morphology and naturalistic hand function among extant primates have the potential to aid in functional interpretations of fossil primates and to aid in the reconstruction of their substrate preferences. Indeed, manual phalanges are among the most commonly recovered postcranial elements in primate fossil assemblages, and on rare occasions, they can be found as part of an articulated skeleton. Such examples include articulated partial hands of *Darwinius* (Franzen et al., 2009), *Epipliopithecus* (Zapfe, 1960), *Proconsul* (Napier and

Davis, 1959); *Pierolapithecus (*Almécija et al., 2009), *Hispanopithecus* (Almécija et al., 2007), *Equatorius* (Sherwood et al., 2002), *Simiolus* (Rossie et al., 2012), *Theropithecus* (Jablonski, 2002), *Ardipithecus ramidus* (Lovejoy et al., 2009a), *Australopithecus afarensis* (Alemseged et al., 2006), *Australopithecus cf. africanus* (Clarke 1999, 2008), and *Australopithecus sediba* (Kivell et al., 2011). But, in most cases, the majority of fossil hand bones are unassociated and fragmentary (e.g., Godinot and Beard, 1991), and in some cases, abundant unassociated

elements are available (e.g., Bush et al., 1982, for *Australopithecus afarensis*; Begun et al., 1994, for *Proconsul*; Nakatsukasa et al., 2003, for *Nacholapithecus*; Ersoy et al., 2008, for *Griphopithecus*; Jungers et al., 2005, for *Archaeolemur*). Although some morphological features can be reconstructed based on fragmentary and damaged elements (e.g., phalangeal curvature; Deane and Begun, 2008), measures of intrinsic hand proportions are more problematic because they are typically based on two or more elements from a single individual. For example, the phalangeal index (PI) as defined by Napier and Napier (1967) requires one metacarpal and a proximal and intermediate phalanx from the same digit, typically Digit 3 (e.g., Kirk et al., 2008). Further, given the general similarities in shape and size among the respective proximal and intermediate phalanges of Digits 2 through 5 in some anthropoids (e.g., Susman, 1979), attribution of phalanges to a specific digit can be problematic to nonspecialists, particularly when informative features are not preserved. In addition, when several individuals are present in a fossil assemblage, it is virtually impossible to assign elements to specific digits or individuals. The risk of attributing a phalangeal element to the incorrect individual or digit thus poses severe methodological problems for estimating measures of intrinsic hand proportions, such as the PI, for fossil primates. Accordingly, the lack of such estimates for fossil taxa has limited our understanding of substrate and locomotor preferences and their evolution in key taxa that preserve unassociated phalanges and metacarpals.

The primary aim of this article is to determine whether locomotor groups and substrate preferences can be reliably inferred based on intrinsic hand proportions estimated using unassociated phalangeal and metacarpal elements. Previous studies have used *ad hoc* justifications to argue that intrinsic hand proportions computed from "composite" hands (a reconstruction based on elements from more than one individual) are accurate despite the possibility of misattribution to digit or individual (Jungers et al., 2005; Kirk et al., 2008; but see Patel et al., 2009, and Rolian et al., 2011). While estimates of intrinsic hand proportions are subject to bias when particular unassociated elements are combined (for example, elements from small and large individuals or males and females), it is unclear the extent to which such bias obscures more general signals of substrate preference and locomotion. In this study, we used a resampling approach to test the hypothesis that calculations of intrinsic hand proportions from unassociated digits are robust to varying degrees of uncertainty regarding phalangeal misattribution. Resampling approaches, which are useful for coping with small sample size and the taphonomic uncertainty surrounding unassociated fossil elements, have been widely used in paleoanthropology to compare ratio-based metrics among taxa, including sexual dimorphism (Lockwood et al., 1996; Gordon et al., 2008), limb proportions (Green et al., 2007), metacarpal proportions (Green and Gordon, 2008), and overall hand proportions (Alba et al., 2003).

MATERIALS AND METHODS
Sample

Individuals (n = 460) from extant catarrhine taxa representing a wide range of locomotor categories were included in the study. Samples are adults (mostly wild-shot) of known sex with no pathology in the appendicular skeleton. Specimens were considered adult when their long bone epiphyses were fused or in the process of fusing (Rolian, 2009). Following Patel et al. (2009), taxa were grouped into one of five categories: (1) AQC: arboreal quadrupedal cercopithecoids-*Trachypithecus cristatus* and *Cercopithecus* spp.; (2) SQC: semiterrestrial quadrupedal cercopithecoids-*Macaca mulatta*; (3) TQC: terrestrial quadrupedal cercopithecoids-*Papio hamadryas* and *Erythrocebus patas*; (4) STQAH: suspensory/terrestrial quadrupedal African hominoids-*Pan troglodytes* and *Gorilla gorilla*; or (5) SAH: suspensory Asian hominoids-*Hylobates lar* and *Pongo pygmaeus*. Details of the comparative sample are provided in Table 1.

Data

Skeletal element lengths were derived from a single hand in each individual. Individuals were included only when the bones of all five digits were present in the skeletal preparation. Distal phalanges were not included as they are rarely preserved in skeletal preparations. In most cases, at least one side was available in disarticulated form. In some individuals where a phalanx was missing or damaged, the element was replaced with its antimere from the other hand. The identity of the phalanges to specific ray can be easily determined by their size and morphological differences when all bones are present (Susman, 1979; Christensen, 2009) or when available by comparison with the articulated side of the same skeleton. Measurements were obtained for each hand bone from digital images of specimens placed ventrally in anatomical position on a flatbed scanner (Microtek i320 ScanMaker, Carson, CA) and scanned in TIFF format at 300 pixels per inch (ppi). The scanner method has been validated elsewhere (Hallgrímsson et al., 2002). ImageJ (Rasband, 2010) and *TPSDig2* (Rohlf, 2005) software packages were used to measure element length to the nearest 0.1 mm. Length measurements were collected from the third metacarpal, and the proximal and intermediate phalanges of Digits 2 through 5. Metacarpal length was defined as the length between the dorsal-most aspect of the proximal articular surface and the distal-most aspect of the metacarpal head (i.e., the distal articular surface). For the proximal and intermediate phalanges, length was defined as the distance between the center of the proximal articular surface and the center of the distal articular surface.

Analyses

Recent studies have used resampling procedures involving logarithms to describe and compare relative measures of size between taxa from ratios (e.g., Green et al., 2007; Green and Gordon, 2008). Following previous studies of intrinsic hand proportions (Napier and Napier, 1967; Etter, 1973; Lemelin, 1999; Jungers et al., 2005; Kirk et al., 2008; Patel and Wunderlich, 2010), we used the PI as a measure of digital proportions and relative finger length (Fig. 1). This index has been shown to reliably discriminate between grades of arboreality and terrestriality in primates and other mammals (Kirk et al., 2008; Patel and Wunderlich, 2010). Whether this index is influenced by size remains unclear, but the discrimination between functional groups appears robust. For example, although terrestrial primates tend to be larger than arboreal species, there are still many large-bodied arboreal primates with long fingers (e.g., *Nasalis*,

TABLE 1. Sample composition.

Taxon	n (Female, Male, Unknown)	Substrate/ Locomotor Group	Provenance
Hylobates lar (*Hyl*)	21 (9, 9, 3)	SAH	MCZ
Pongo pygmaeus (*Pon*)	43 (29, 14, 0)	SAH	NMNH, BSM, MCZ
Trachypithecus cristatus (*Tra*)	20 (14, 6, 0)	AQC	BSM, MCZ
Cercopithecus spp. (*Cer*)	52 (24, 26, 2)	AQC	MCZ
Pan troglodytes (*Pan*)	88 (56, 32, 0)	STQAH	PCM, AMNH, NMNH, MCZ
Gorilla gorilla (*Gor*)	92 (53, 39, 0)	STQAH	PCM, AMNH, NMNH, MCZ
Macaca mulatta (*Mac*)	45 (20, 25, 0)	SQC	CPRC
Erythrocebus patas (*Ery*)	54 (32, 21, 1)	TQC	CPRC
Papio hamadryas (*Pap*)	45 (19, 22, 4)	TQC	UT
TOTAL	460		

The *Cercopithecus* sample comprised *C. mitis* and *C. ascanius*.
AQC, arboreal quadrupedal cercopithecoid; SAH, suspensory Asian hominoid; SQC, semi-terrestrial quadrupedal cercopithecoid; STQAH, suspensory/terrestrial quadrupedal African hominoid; TQC, terrestrial quadrupedal cercopithecoid; AMNH, American Museum of Natural History, New York; BSM, Bayerische Staatssammlung, Munich; CMNH, Cleveland Museum of Natural History, Cleveland; MCZ, Harvard Museum of Comparative Zoology, Cambridge; NMNH, National Museum of Natural History, Washington; PCM, Powell-Cotton Museum, Birchington, UK; CPRC, Laboratory for Primate Morphology and Genetics, Caribbean Primate Research Center, San Juan; UT, Department of Anthropology, University of Texas-Austin.

Pongo). The PI is calculated as the sum of the lengths of intermediate phalanx and proximal phalanx of the third digit (IP3 and PP3, respectively), divided by the length of the third metacarpal (MC3). We first calculated the PI for third digits from associated museum specimens (Fig. 1A). This provides the baseline against which other resampled distributions are compared.[1] Next, we simulated two additional scenarios reflecting different states of knowledge about the allocation of digital elements to both ray and individual. The three scenarios progressively increase uncertainty with respect to digit allocation.

All resampling procedures were performed using custom routines in Matlab R2009b (Natick, MA). For Scenario 1 (S1), an exact resampling approach was used in which an individual was selected once from each taxon, and the PI for this individual was computed as the ratio of the summed IP3 and PP3 lengths to the MC3 length (Fig. 1A,B). The natural logarithm (ln) of PI was then calculated (hereafter, "log" refers to the natural logarithm unless otherwise noted). The procedure was repeated for each individual to generate a distribution and mean of log PI values for each taxon. For Scenario 2 (S2), a Monte Carlo resampling approach was used. At each iteration, three individuals were selected at random, and the log PI was obtained by randomly selecting the MC3 from one individual, the PP3 from another individual, and the IP3 from the last individual (Fig. 1C). The procedure was repeated 10,000 times with replacement to generate a resampled distribution and mean of log PI values for each taxon. The resampling

procedure for Scenario 3 (S3) was similar to S2, except that the proximal and intermediate phalanges were selected at random from Digits 2, 3, 4, or 5, as opposed to only from Digit 3 (Fig. 1D). This procedure was also repeated 10,000 times with replacement to generate a resampled distribution and mean of log PI values for each taxon.

Identifying the effect of multiple individuals and phalangeal uncertainty

This study investigates three distinct questions, each of which is based on the same set of resampled log ratios generated using the procedure described above. The first question is whether generating ratios from multiple individuals (S2) or with multiple individuals and phalanges of uncertain attribution to ray (S3) affects the mean and/or variance of resampled distributions within a taxon. The second question is whether these scenarios affect the ability of analyses to find significant difference in sample means between taxa. Finally, because fossil samples are often limited in size, the third question is whether these scenarios affect the probability that a single resampled ratio from one sample (e.g., a fossil sample) will fall within the range of resampled ratios from another sample (e.g., a comparative taxon). We describe the analyses used to address each of these questions below.

Comparing means and variability between scenarios. Within each taxon, means were compared between S1 and S2, S1 and S3, and S2 and S3. For comparisons involving S1 (in which the log PI is calculated exactly once for each individual in the sample), an iterative resampling procedure is used in which each iteration samples with replacement a number of log ratios equal to the sample size for that taxon (n_{TAXON}) from the population of 10,000 resampled values for the other scenario (either S2 or S3). For example, in comparing the S1 and S2 means for *P. troglodytes*, in each iteration a sample of 88 log ratios is drawn from the set of 10,000 resampled ratios for S2. The mean of those log ratios is then calculated and subtracted from the mean of log ratios in S1.

[1] We chose not calculate baseline PI values for the other digits (i.e., MC3 with PP2 and IP2; MC3 with PP4 and IP4: and MC3 with PP5 and IP5) because this should produce the same discrimination patterns as PIs with digit 3 elements only, but with absolutely lower index values since the phalanges of the other digits are shorter than those in digit 3. A possible exception to this would be for *Pongo* which can have longer ray 4 hand bones (Susman, 1979). Also, our approach represents the ideal situation (because elements from the third digit are the easiest to identify in museum collections, even by nonspecialists) to the actual situation paleontologists find with unassociated fossils (i.e., most researchers cannot be absolutely certain which rays (2–5) are in their fossil sample if they do not have them all).
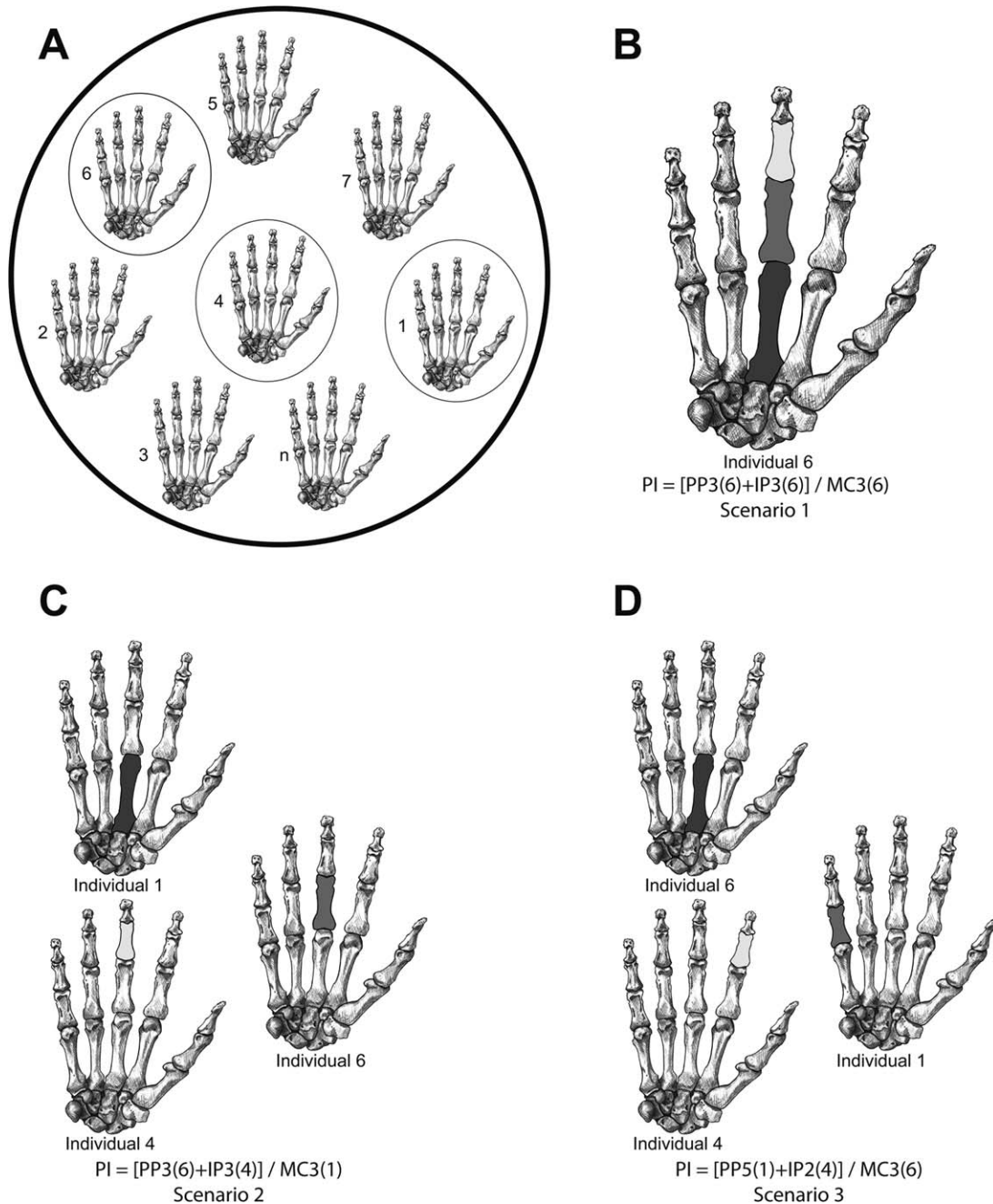
Fig. 1. Three different scenarios for calculating the phalangeal index (PI) used in the resampling protocols. (A) Sample of complete hands for a single taxon. (B) Scenario 1 in which PI is calculated from the bones of the third digit of a single individual. In this example, PI is calculated from Individual 6. (C) Scenario 2 in which PI is calculated from the bones of the third digit of three different individuals. In this example, PI is calculated from the MC3 of Individual 1, the PP3 of Individual 6, and the IP3 of Individual 4. (D) Scenario 3 in which PI is calculated from the MC3 of one individual and a proximal and intermediate phalanx (randomly chosen from Digits 2–5) from two other individuals. In this example, PI is calculated from the MC3 of Individual 6, the PP5 of Individual 1, and the IP2 of Individual 4.

A positive difference indicates that the scenario which is subtracted from S1 has lower ratios, indicating shorter fingers relative to palms. A negative difference indicates the opposite. This procedure is repeated 10,000 times, and two-tailed P values are calculated from the resulting distribution of differences using established methods based on the proportions of negative and positive differences thus generated (Manly, 1997; Gordon et al., 2008;

Green and Gordon, 2008). In the case of comparing means between S2 and S3, the procedure is identical except that samples equal in size to n_{TAXON} are drawn from both S2 and S3 in each iteration.

To test for differences in variability of distributions between scenarios, a resampled Levene test is used. In a Levene test, variation in a distribution is measured as the mean of the absolute value of deviations or differences

of each observation from the median of all observations, with higher mean deviations indicating more variable observations (Van Valen, 2005). Here, we use the same procedure as described above for comparing log ratio means with the following modification. In each iteration, the mean of the absolute value of deviations from the median is calculated for the resampled n_{TAXON} values for each scenario, and the difference in means is calculated. For scenario A minus scenario B, a positive difference indicates that scenario A is more variable than scenario B, while a negative difference indicates the opposite. P values are calculated in the same manner as for differences in means. Significant differences from zero in this case indicate that the distributions generated by the two scenarios under consideration differ significantly in variability.

Comparing taxon means within scenarios. Within each scenario, mean log ratios were compared for all pairs of taxa. For comparisons within S1, comparisons were calculated using a randomized t test (Manly, 2007). Comparisons within S2 and S3 followed a procedure similar to that outlined above for comparisons of means between scenarios. An iterative resampling procedure was used in which each iteration involved sampling with replacement a number of log ratios from the first taxon equal to that taxon's sample size and a number of log ratios from the second taxon equal to that taxon's sample size. For example, in a comparison of taxon means in S2 between *Gorilla* and *Pan*, in each iteration a sample of 92 values would be drawn with replacement from the set of 10,000 S2 values for *Gorilla*, and a sample of 88 values would be drawn with replacement from the set of 10,000 S2 values for *Pan*. Means were then calculated for each taxon's sample, and difference in means was calculated. This procedure was repeated 10,000 times, and P-values were calculated based on the resulting distributions as described above.

Calculating the probability that individual ratios will fall within the range of each taxon. Within each scenario, all pairs of taxa A and B were compared to determine the probability that a single log ratio from taxon A would fall within the range of log ratios in taxon B, and the probability that a single log ratio from taxon B would fall within the range of log ratios in taxon A. (These two probabilities are not necessarily equal; e.g., consider two distributions with the same mean but different variances such that all values of taxon A are contained in taxon B, but the reverse is not true.) Within S1, an exact resampling approach was used in which each individual log ratio from taxon A was compared with the minimum and maximum of log ratios from taxon B, and the proportion of log ratios from taxon A, which fell inside that range, was calculated. The procedure was then repeated with A and B swapped. For S2 and S3, a Monte Carlo resampling procedure was used with 10,000 iterations. In each iteration, one log ratio was selected at random from the 10,000 resampled values for taxon A and compared with the minimum and maximum for a sample of resampled values for taxon B, where the sample size is equal to the number of individuals in that taxon's sample. For example, in a comparison of *Gorilla* and *Pan* for S2, one S2 resampled *Gorilla* log ratio would be compared to the maximum and minimum of a sample of 88 *Pan* log ratios sampled with

replacement from the set of 10,000 *Pan* values for S2. After repeating this procedure 10,000 times, the proportion of times in which the single log ratio from taxon A fell within the range of the sample from taxon B was calculated. A proportion (P value) of less than 0.05 indicates that the log ratio from taxon A is outside the range of taxon B log ratios more than 95% of the time and, thus, is not likely to be a value that could plausibly belong to taxon B. This procedure was then repeated with A and B swapped.

RESULTS

Distributions were generated of actual log PI values for articulated third digits (S1) and resampled log PI values under scenarios of misattribution to incorrect individual (S2) and incorrect individual and digit (S3) (Fig. 2). The first column of Table 2 provides mean logged ratios (MLR) for each taxon under each of the three scenarios. The results for associated digits (S1) show the following order for the PI among studied taxa: $Hyl > Tra > Cer > Pon > Pan > Gor > Mac > Ery > Pap$. Thus, suspensory Asian hominoids (*Hylobates* and *Pongo*) generally have the longest digits relative to metacarpal length, followed by arboreal quadrupedal cercopithecoids (*Trachypithecus* and *Cercopithecus*), African great apes (*Pan* and *Gorilla*), semiterrestrial (*Macaca*) and terrestrial cercopithecoids (*Erythrocebus* and *Papio*). An important deviation from this pattern is that *Cercopithecus* and *Trachypithecus* have higher phalangeal indices than *Pongo*. Only *Erythrocebus* and *Papio* had negative ratio values, indicating that the denominator of the ratio (palm length) is greater than the numerator (finger length without the distal phalanx).

With regard to the effect of the different sampling scenarios on the log ratios produced within a given taxon, calculating PI based on multiple individuals (S2) as opposed to a single individual (S1) does not have an appreciable effect on the mean of the distribution of log ratios. For all within-taxon comparisons between S1 and S2, differences in means are not significant, and P values for those differences are all above 0.8 (Table 3; compare blue to red distributions in Fig. 2). However, distributions of log ratios using multiple individuals (S2) have significantly greater variability than distributions based on articulated digits (S1) in all taxa (Table 4). When distributions of log PI that are based on phalanges of unknown ray from multiple individuals (S3) are compared to those of known ray from multiple individuals (S2), variability does not differ significantly within taxa in most cases (Table 4; compare red to gray distributions in Fig. 2). The only exception is *Pan*, with significantly higher variability for S3. However, means are significantly lower for S3 than for S2 in all taxa except *Papio*, where the difference approaches significance ($P = 0.076$, Table 3). For all taxa, S3 has a significantly lower mean than S1 and is significantly more variable (Tables 3 and 4; compare blue to gray distributions in Fig. 2).

The impact of these differences in variability and mean on the ability to make accurate inferences based on log PI values can be evaluated by determining how patterns of log ratio relationships between taxa change across these three scenarios. For example, as mentioned above, the order of actual log PI values (S1) among taxa in this study is as follows: $Hyl > Tra > Cer > Pon > Pan > Gor > Mac > Ery > Pap$. This order remains the
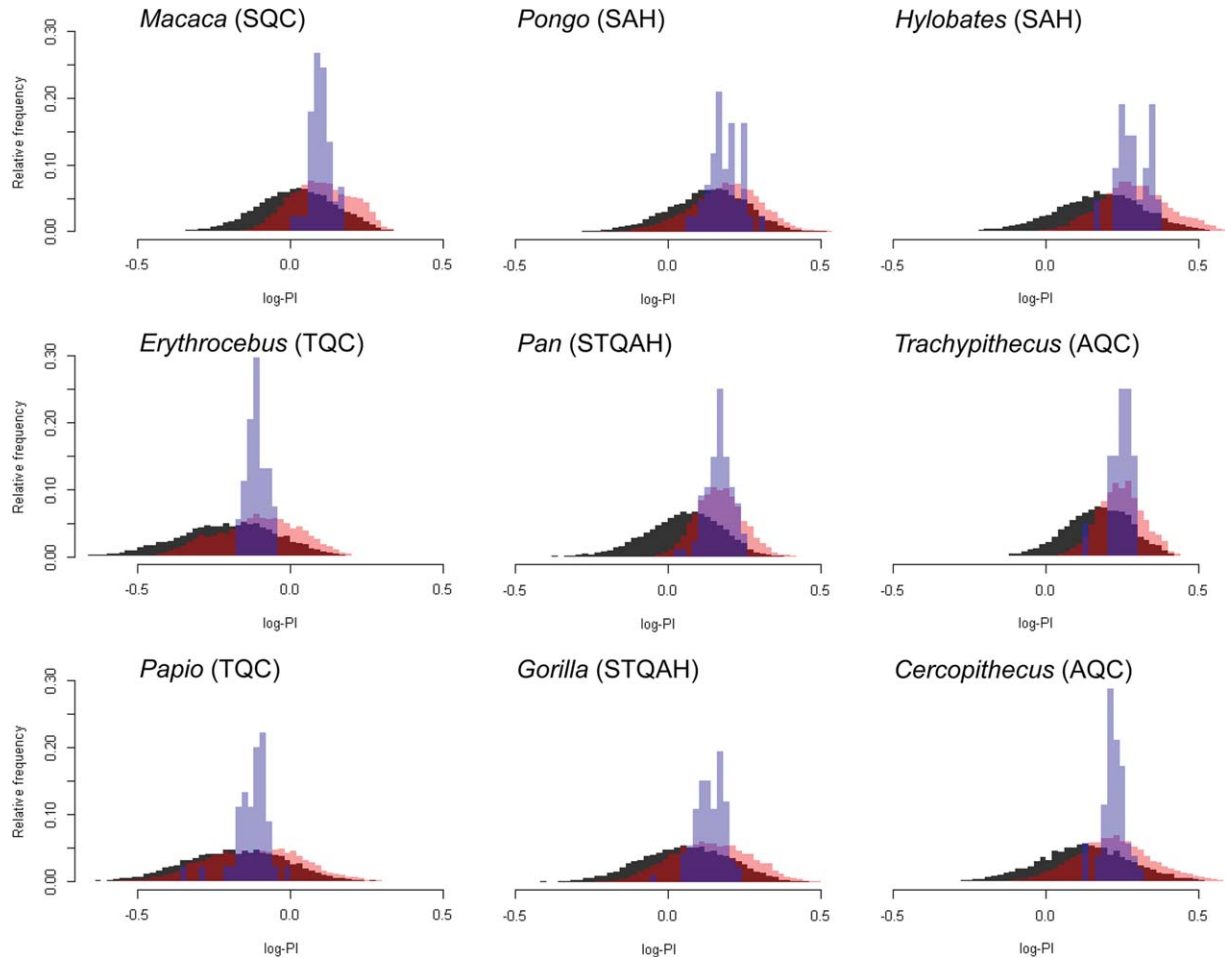
Fig. 2. Histograms for distributions of actual (S1) and resampled (S2 and S3) log PI values for each taxon (substrate/locomotor group abbreviations in parentheses after taxon name). All histograms are plotted on the same scale. Within each taxon, distributions based on S1 are plotted in blue, S2 in red, and S3 in gray. Note that, in any given taxon, S1 and S2 distributions tend to have very similar means but different variances, whereas S2 and S3 distributions tend to have similar variances but different means. S1 and S3 distributions tend to differ in both mean and variance.

same in S2. However, in S3 the PI order changes to *Hyl* > *Tra* > **Pon** > *Cer* > *Gor* > *Pan* > *Mac* > **Pap** > **Ery** (bold face indicates taxa that are in a different position relative to the order of S1). This result indicates that phalangeal misattribution can lead to incorrect estimates regarding comparative PI values among taxa.

Distributions of PIs drawn from associated third digits (S1) generally preserve differences between locomotor groups, although considerable overlap is apparent between taxa with similar degrees of arboreality in their locomotor repertoire (blue distributions in Fig. 2). Despite this overlap, distributions are relatively narrow and thus means differ significantly between all taxa except between *Erythrocebus* and *Papio*, and that difference approaches significance with $P = 0.099$ (Table 2, Scenario 1). Increasing uncertainty by incorporating multiple individuals (S2) slightly decreases the probability of finding a significant difference in means between taxa, although means remain significantly different for all pairwise comparisons except between pairs that are most similar in mean values (Table 2, Scenario 2). In some cases, this results in no significant difference in means between two taxa belonging to different substrate/locomotor groups; e.g., *Hylobates* and *Trachypithecus* (SAH and AQC,

respectively), *Pongo* and *Cercopithecus* (SAH and AQC, respectively), *Pongo* and *Pan* (SAH and STQAH, respectively), and *Gorilla* and *Macaca* (STQAH and SQC, respectively). Increasing uncertainty even more by incorporating phalangeal elements from Rays 2 to 5 (S3) further decreases the probability of finding a significant difference in means between taxa and results in nonsignificant mean differences for more pairs of taxa in different substrate/locomotor categories (Table 2, Scenario 3).

Also under investigation is the effect of increasing uncertainty on the ability to make inferences based on a single log ratio such as one might generate for a fossil sample. Due to the high overlap of taxon distributions of log PI even when considering articulated third rays (S1), there is generally a high probability that any single log ratio could be drawn from most of the taxa considered here. Although single log PI values from *Erythrocebus* and *Papio* fall significantly outside the range of all non-TQC taxa and single log PI values from non-TQC taxa fall significantly outside the range of *Erythrocebus* and *Papio*, the majority of non-TQC log PI values could be drawn from any of the non-TQC taxa (Table 5, S1). The only exceptions to this generalization are *Macaca* log ratios, which fall significantly outside the range of

TABLE 2. *Taxon means (MLR) and* P *values for pairwise comparisons of difference in means between taxa within each of the three scenarios.*[a]

Scenario 1	MLR$_1$	Hyl	Tra	Cer	Pon	Pan	Gor	Mac	Ery
Hylobates	0.288	–							
Trachypithecus	0.246	**0.005**	–						
Cercopithecus	0.221	**<0.001**	**0.010**	–					
Pongo	0.188	**<0.001**	**<0.001**	**<0.001**	–				
Pan	0.167	**<0.001**	**<0.001**	**<0.001**	**0.013**	–			
Gorilla	0.135	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	–		
Macaca	0.102	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	–	
Erythrocebus	−0.110	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	–
Papio	−0.125	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.099
Scenario 2	MLR$_2$	Hyl	Tra	Cer	Pon	Pan	Gor	Mac	Ery
Hylobates	0.291	–							
Trachypithecus	0.244	0.108	–						
Cercopithecus	0.222	**0.020**	0.366	–					
Pongo	0.189	**<0.001**	**0.016**	0.178	–				
Pan	0.169	**<0.001**	**<0.001**	**0.007**	0.293	–			
Gorilla	0.137	**<0.001**	**<0.001**	**<0.001**	**0.021**	**0.038**	–		
Macaca	0.104	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.093	–	
Erythrocebus	−0.106	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	–
Papio	−0.119	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.654
Scenario 3	MLR$_3$	Hyl	Tra	Cer	Pon	Pan	Gor	Mac	Ery
Hylobates	0.180	–							
Trachypithecus	0.172	0.828	–						
Cercopithecus	0.128	0.165	0.165	–					
Pongo	0.132	0.188	0.191	0.870	–				
Pan	0.052	**0.001**	**<0.001**	**0.003**	**<0.001**	–			
Gorilla	0.057	**0.001**	**<0.001**	**0.005**	**0.002**	0.807	–		
Macaca	0.023	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.197	0.146	–	
Erythrocebus	−0.205	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	–
Papio	−0.179	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.420

Bold values indicate that means differ significantly between taxa at $\alpha = 0.05$ level.
[a] Scenarios indicated by subscript. Abbreviations as in Table 1.

TABLE 3. *Differences in mean logged ratios (DMLR) between scenarios within each taxon.*[a]

Taxon	DMLR$_{1-2}$	P	DMLR$_{2-3}$	P	DMLR$_{1-3}$	P
Hylobates	−0.004	0.876	0.111	**0.006**	0.108	**0.001**
Trachypithecus	0.001	0.928	0.072	**0.009**	0.074	**0.001**
Cercopithecus	−0.001	0.938	0.094	**0.001**	0.093	**<0.001**
Pongo	−0.001	0.976	0.057	**0.031**	0.055	**0.004**
Pan	−0.001	0.881	0.117	**<0.001**	0.115	**<0.001**
Gorilla	−0.002	0.867	0.080	**<0.001**	0.078	**<0.001**
Macaca	−0.002	0.889	0.081	**0.001**	0.079	**<0.001**
Erythrocebus	−0.004	0.809	0.099	**<0.001**	0.095	**<0.001**
Papio	−0.006	0.803	0.060	0.076	0.054	**0.026**

Bold values indicate statistically significant difference in distribution means at $\alpha = 0.05$ level.
[a] Scenarios indicated by subscript.

Hylobates values; *Hylobates* log ratios, which fall significantly outside the ranges of *Macaca* and *Gorilla* values; and *Trachypithecus* log ratios, which fall significantly outside the range of *Macaca* values (Table 5, S1). When uncertainty is increased by calculating log PI values based on multiple individuals (S2), the increase in variability of resampled distributions dramatically drives up the probability of sampling any single log ratio value from any taxon in this analysis (Table 5, S2). In fact, no taxon has log ratios that fall significantly outside of the range of any other taxon under this scenario, although single values from *Erythrocebus* and *Papio* approach significant difference from distributions for the two taxa with the longest relative finger length, *Hylobates* and *Trachypithecus*

(P values between 0.060 and 0.099, Table 5, S2). Even this caveat is removed when considering a scenario where phalangeal ray is uncertain (S3); in that case, the lowest P value for any comparison is 0.149 (Table 5, S3).

DISCUSSION

Skeletal intrinsic hand proportions are useful indicators of locomotor preference in mammals, including primates, and can reliably distinguish between arboreal, semi-arboreal, and terrestrial substrate preferences when sample means are compared. While this is true for extant comparative material, measures of intrinsic hand proportions

TABLE 4. *Differences in mean absolute deviations from median (DMADM) between scenarios within each taxon.[a]*

Taxon	$DMADM_{1-2}$	P	$DMADM_{2-3}$	P	$DMADM_{1-3}$	P
Hylobates	−0.048	**<0.001**	−0.023	0.329	−0.071	**<0.001**
Trachypithecus	−0.032	**<0.001**	−0.025	0.127	−0.057	**<0.001**
Cercopithecus	−0.074	**<0.001**	−0.019	0.281	−0.093	**<0.001**
Pongo	−0.051	**<0.001**	−0.010	0.545	−0.061	**<0.001**
Pan	−0.029	**<0.001**	−0.036	**<0.001**	−0.064	**<0.001**
Gorilla	−0.065	**<0.001**	−0.013	0.270	−0.078	**<0.001**
Macaca	−0.057	**<0.001**	−0.017	0.201	−0.074	**<0.001**
Erythrocebus	−0.083	**<0.001**	−0.020	0.226	−0.102	**<0.001**
Papio	−0.089	**<0.001**	−0.005	0.791	−0.093	**<0.001**

Bold values indicate statistically significant difference in distribution variability at $\alpha = 0.05$ level.
[a] Scenarios indicated by subscript.

TABLE 5. *Probability that a randomly sampled log ratio from row taxon will fall within the range of a sample of log ratios from the column taxon (see text for specifics of analysis).*

	Hyl	Tra	Cer	Pon	Pan	Gor	Mac	Ery	Pap
					Scenario 1				
Hylobates	_	0.619	0.619	0.619	0.238	**0.048**	**0.048**	**<0.001**	**<0.001**
Trachypithecus	0.950	_	1.000	1.000	0.450	0.250	**0.050**	**<0.001**	**<0.001**
Cercopithecus	0.904	0.923	_	0.884	0.981	0.769	0.538	0.096	**<0.001**	**<0.001**
Pongo	0.605	0.814	0.884	_	0.860	0.767	0.419	**<0.001**	**<0.001**
Pan	0.409	0.773	0.784	0.977	_	0.920	0.602	**<0.001**	**<0.001**
Gorilla	0.228	0.478	0.533	0.891	0.989	_	0.761	**0.011**	**0.011**
Macaca	**0.022**	0.111	0.133	0.756	0.978	1.000	_	**<0.001**	**<0.001**
Erythrocebus	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.037	**<0.001**	_	1.000
Papio	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	0.044	**<0.001**	0.889	_
					Scenario 2				
Hylobates	_	0.708	0.965	0.856	0.724	0.863	0.526	0.116	0.254
Trachypithecus	0.965	_	0.999	0.983	0.925	0.988	0.742	0.124	0.333
Cercopithecus	0.811	0.685	_	0.907	0.818	0.925	0.711	0.303	0.461
Pongo	0.789	0.710	0.966	_	0.996	0.873	0.978	0.808	0.349	0.541
Pan	0.806	0.754	0.996	0.996	_	0.998	0.938	0.436	0.661
Gorilla	0.631	0.562	0.920	0.925	0.814	_	0.842	0.541	0.682
Macaca	0.546	0.486	0.929	0.949	0.845	0.996	_	0.661	0.795
Erythrocebus	0.086	0.060	0.390	0.441	0.242	0.644	0.490	_	0.986
Papio	0.099	0.083	0.369	0.419	0.248	0.589	0.456	0.914	_
					Scenario 3				
Hylobates	_	0.790	0.965	0.907	0.814	0.910	0.704	0.301	0.436
Trachypithecus	0.974	_	0.996	0.976	0.905	0.974	0.790	0.280	0.461
Cercopithecus	0.882	0.756	_	0.908	0.874	0.938	0.803	0.450	0.585
Pongo	0.927	0.832	0.984	_	0.919	0.973	0.837	0.422	0.571
Pan	0.859	0.734	0.970	0.938	_	0.991	0.944	0.658	0.781
Gorilla	0.823	0.684	0.954	0.899	0.947	_	0.897	0.634	0.740
Macaca	0.796	0.649	0.961	0.913	0.981	0.992	_	0.741	0.834
Erythrocebus	0.289	0.149	0.544	0.422	0.662	0.710	0.604	_	0.962
Papio	0.339	0.203	0.593	0.475	0.696	0.744	0.642	0.951	_

Bold values indicate statistical significance at $\alpha = 0.05$ level.
Abbreviations as in Table 1.

such as the PI are rarely estimated for fossil taxa because most phalangeal elements in the primate fossil record are unassociated and are thus difficult to reliably attribute to the correct individual or digit. In this article, we used resampling methods to determine if estimates of PI derived from unassociated elements are reliable. Our results show two important findings. First, we found that, within a taxon, computing PIs from composite digits can produce estimates that are significantly lower than those generated by associated digits. Resampling scenarios are prone to underestimating MLR and inflating the variance, when compared to the baseline of the more traditional method of using only elements of the third digit (e.g., Kirk et al., 2008; Patel and Wunderlich, 2010), because at

each iteration, there is a 75% chance of drawing a phalangeal element (belonging to Digits 2, 4, or 5), which is shorter than the corresponding element of Digit 3.[2] An implication of this finding is that two composite samples drawn from a single population of primates could have statistically different intrinsic hand proportions depending on which digital rays were present in each sample. In terms of the fossil record, this issue, a case of inflated frequency of Type I error (false positive), would be further compounded by, for example, third digit elements being

[2] Note that these specific findings only apply to anthropoid primate hands since strepsirrhines tend to have relatively longer fourth digits (e.g., Jouffroy et al., 1991).

sampled from one fossil specimen and second, fourth, or fifth digit elements from another.

In terms of the increased variance, we had expected that, in sexually size dimorphic taxa, there would be substantial widening of the tails of the resampled distributions. Such an effect would stem from iterations of the resampling protocol that combine large (male) with small (female) metacarpals and phalanges, and vice versa. Although, in all cases, variance significantly increased in resampled distributions, this is likely due to other factors in addition to sexual dimorphism. For example, the most sexually dimorphic hominoids (*Pongo* and *Gorilla*) had higher differences in absolute deviation from the median between S1 and S2 but only marginally higher than that of *Hylobates*, which exhibits low sexual dimorphism. The largest differences for increased variability between S1 and S2 were for two other highly dimorphic taxa, *Erythrocebus* and *Papio* (-0.083 and -0.089 log units, respectively).

We also tested whether estimates of PI in primates based on composite hands constructed from unassociated metacarpal and phalangeal elements reliably discriminate between different locomotor modes. We found that composite hands from arboreal taxa display phalangeal indices that overlap extensively, particularly in resampled distributions due to the increased variance. Thus, comparisons between taxa to test for differences in locomotor category may tend to have inflated frequencies of Type II error (false negative) due to increased variance of resampled distributions. While overlap is extensive in PI values among locomotor categories under our resampling scenarios (red and gray distributions in Fig. 2), means for most taxa in the highly arboreal and terrestrial categories do remain distinct at $\alpha = 0.05$ and $\alpha = 0.10$ significance levels in S2 and S3 (Table 2). Our results indicate that, although all taxa evince systematic decreases in MLR and increases in variance as the misattribution of phalangeal elements increases, terrestrial taxa (TQC) and highly suspensory and arboreal taxa (SAH and AQC) appear to be the locomotor groups least susceptible to misidentification with each other as a result of the magnitude of dissimilarity in taxon means between these groups.

However, comparisons across all locomotor groups are very sensitive to the increase in variance when considering whether a particular log PI value could have come from any particular locomotor category. As the results in Table 5 demonstrate, terrestrial digitigrade quadrupeds can be distinguished from all other locomotor categories based on a single ratio, and vice versa, but only when that ratio comes from an associated individual third digit. The ability to rule out membership in any locomotor category drops dramatically when multiple individuals are included in a ratio, with only terrestrial quadrupeds being distinct from the taxa with the longest relative phalangeal length (*Hylobates* and *Trachypithecus*) at $\alpha = 0.10$. Even this distinction disappears when identification of phalangeal ray is uncertain. This is because terrestrial primates tend to exhibit more uniform digit length across Rays 2 through 5 than arboreal primates (Etter, 1973), and there is considerable variation in digit row formulae within and between hominoid species (Susman, 1979). This leads to the expectation that replacing third digit phalangeal elements with those from the second, fourth, or fifth digit would alter the mean little between S1 and the resampling scenarios within TQC taxa (e.g., *Papio* in Fig. 2 and Table 3), but

means for arboreal taxa will decrease substantially under S3, making them more similar to means for terrestrial taxa and thus more difficult to distinguish from them.

Taken together, these considerations imply that estimates of intrinsic hand proportions based on composite and associated hands are not directly comparable. To enable statistically robust comparisons, the sampling regime must be constrained to mimic the fossil sample as closely as possible, as done in previous studies (Green et al., 2007; Green and Gordon, 2008). For example, if only three proximal phalanges and two intermediate phalanges are in the sample attributed to any given fossil taxon, the resampling procedure should also have the same number and types of elements in the comparative samples. This procedure could be applied to the numerous fossil assemblages containing unassociated hand elements for taxa with potentially disputed locomotor affinities, including but not limited to *Equatorius africanus*, *Ardipithecus ramidus*, *Australopithecus afarensis*, *Australopithecus africanus*, and *Homo habilis*. These species evince postcranial morphologies consistent with predominately terrestrial locomotion and controversy largely focuses on the extent to which they were arboreal (see references in Wood and Richmond, 2000; Ward, 2002; Green et al., 2007; Patel et al., 2009; Lovejoy et al., 2009a,b). Using resampling protocols (similar to outlined here and elsewhere) and sufficient sample size (e.g., all possible known elements), accurate estimates for intrinsic hand (and foot) proportions for these species could be achieved. Finally, in the few cases where articulated digits are available from the fossil record, we encourage researchers to include in their manuscripts additional analyses that compare only ratios based on completely associated rays due to their much lower variance and associated greater statistical power.

LITERATURE CITED

Alba DM, Moyà-Solà S, Köhler M. 2003. Morphological affinities of the *Australopithecus afarensis* hand on the basis of manual proportions and relative thumb length. J Hum Evol 44:225–254.

Alemseged Z, Spoor F, Kimbel WH, Bobe R, Geraads D, Reed D, Wynn JG. 2006. A juvenile early hominin skeleton from Dikika, Ethiopia. Nature 443:296–301.

Almécija S, Alba DM, Moyà-Solà S. 2009. *Pierolapithecus* and the functional morphology of Miocene ape hand phalanges:

paleobiological and evolutionary implications. J Hum Evol 57:284–297.

Begun DR, Teaford MF, Walker A. 1994. Comparative and functional anatomy of *Proconsul* phalanges from the Kaswanga primate site, Rusinga Island, Kenya. J Hum Evol 26:89–165.

Bush ME, Lovejoy CO, Johanson DC, Coppens Y. 1982. Hominid carpal, metacarpal, and phalangeal bones recovered from the Hadar formation: 1974–1977 collections. Am J Phys Anthropol 57:651–677.

Christensen AM. 2009. Techniques for siding manual phalanges. Forensic Sci Int 193:84–87.

Clarke RJ. 1999. Discovery of complete arm and hand of the 3.3-million-year-old *Australopithecus* skeleton from Sterkfontein. S Afr J Sci 96:477–480.

Clarke RJ. 2008. Latest information on Sterkfontein's *Australopithecus* skeleton and a new look at *Australopithecus*. S Afr J Sci 104:443–449.

Deane AS, Begun DR. 2008. Broken fingers: retesting locomotor hypotheses for fossil hominoids using fragmentary proximal phalanges and high-resolution polynomial curve fitting (HR-PCF). J Hum Evol 55:691–701.

Ersoy A, Kelley J, Andrews P, Alpagut B. 2008. Hominoid phalanges from the middle Miocene site of Paşalar, Turkey. J Hum Evol 54:518–529.

Etter HF. 1973. Terrestrial adaptations in the hands of Cercopithecinae. Folia Primatol 20:331–350.

Green DJ, Gordon AD. 2008. Metacarpal proportions in *Australopithecus africanus*. J Hum Evol 54:705–719.

Green DJ, Gordon AD, Richmond BG. 2007. Limb-size proportions in *Australopithecus afarensis* and *Australopithecus africanus*. J Hum Evol 52:187–200.

Godinot M, Beard KC. 1991. Fossil primate hands: a review and an evolutionary inquiry emphasizing early forms. Hum Evol 6:307–354.

Gordon AD, Green DJ, Richmond BG. 2008. Strong postcranial size dimorphism in *Australopithecus afarensis*: results from two new multivariate resampling methods for multivariate data sets with missing data. Am J Phys Anthropol 135:311–328.

Franzen JL, Gingerich PD, Habersetzer J, Hurum JH, von Koenigswald W, Smith BH. 2009. Complete primate skeleton from the Middle Eocene of Messel Germany: morphology and paleobiology. PLoS One 4:e5723.

Hallgrímsson B, Willmore K, Hall BK. 2002. Canalization, developmental stability, and morphological integration in primate limbs. Am J Phys Anthropol 119:131–158.

Hamrick MW. 2001. Primate origins: evolutionary change in digital ray patterning and segmentation. J Hum Evol 40:339–351.

Jablonski NG. 2002. A new skeleton of *Theropithecus brumpti* (Primates: Cercopithecidae) from Lomekwi, West Turkana, Kenya. J Hum Evol 43:887–923.

Jouffroy FK, Godinot M, Nakano Y. 1991. Biometrical characteristics of primate hands. Hum Evol 6:269–306.

Jungers WL, Lemelin P, Godfrey LR, Wunderlich RE, Burney DA, Simons EL, Chatrath PS, James HF, Randria GFN. 2005. The hands and feet of *Archaeolemur*: metrical affinities and their functional significance. J Hum Evol 49:36–55.

Kirk EC, Lemelin P, Hamrick MW, Boyer DM, Bloch JI. 2008. Intrinsic hand proportions of euarchontans and other mammals: implications for the locomotor behavior of plesiadapiforms. J Hum Evol 55:278–299.

Kivell TL, Kibii JM, Churchill SE, Schmid P, Berger LR. 2011. *Australopithecus sediba* hand demonstrates mosaic evolution of locomotor and manipulative abilities. Science 333:1411–1417.

Lemelin P. 1999. Morphological correlates of substrate use in didelphid marsupials: implications for primate origins. J Zool 247:165–175.

Lockwood CA, Richmond BG, Jungers WL, Kimbel WH. 1996. Randomization procedures and sexual dimorphism in *Australopithecus afarensis*. J Hum Evol 31:537–548.

Lovejoy CO, Simpson SW, White TD, Asfaw B, Suwa G. 2009a. Careful climbing in the Miocene: the forelimbs of *Ardipithecus ramidus* and humans are primitive. Science 326:70e1–70e8.

Lovejoy CO, Latimer B, Suwa G, Asfaw B, White TD. 2009b. Combining prehension and propulsion: the foot of *Ardipithecus ramidus*. Science 326:72e1–72e8.

Manly BJF. 1997. Randomization, bootstrap and Monte Carlo methods in biology, 2nd ed. London: Chapman and Hall.

Midlo C. 1934. Form of the hand and foot in primates. Am J Phys Anthropol 19:337–389.

Nakatsukasa M, Kunimatsu Y, Nakano Y, Takano T, Ishida H. 2003. Comparative and functional anatomy of phalanges in *Nacholapithecus kerioi*, a Middle Miocene hominoid from northern Kenya. Primates 44:371–412.

Napier JR, Napier PH. 1967. A handbook of living primates. New York: Academic Press.

Napier JR. 1993. Hands (revised edition). Princeton: Princeton University Press.

Napier JR, Davis PR. 1959. The fore-limb skeleton and associated remains of *Proconsul africanus*. Fossil Mamm Afr 16:1–69.

Nelson E, Rolian C, Cashmore L, Shultz S. 2011. Digit ratios predict polygyny in early apes, *Ardipithecus*, Neanderthals and early humans but not in *Australopithecus*. Proc R Soc Lond B 278:1556–1563.

Nieschalk U, Demes B. 1993. Biomechanical determinants of reduction of the second ray in Lorisinae. In: Preuschoft H, Chivers DJ, editors. Hands of primates. New York: Springer-Verlag. p 225–234.

Patel BA, Susman RL, Rossie JB, Hill A. 2009. Terrestrial adaptations in the hands of *Equatorius africanus* revisited. J Hum Evol 57:763–772.

Patel BA, Larson SG, Stern JT. 2012. Electromyography of wrist and finger flexor muscles in olive baboons (*Papio anubis*). J Exp Biol 215:115–123.

Patel BA, Wunderlich RE. 2010. Dynamic pressure patterns in the hands of olive baboons (*Papio anubis*) during terrestrial locomotion: implications for cercopithecoid primate hand morphology. Anat Rec 293:710–718.

Rasband WS. 2010. ImageJ. Available from: http://imagej.nih.gov/ij/.

Richmond BG. 1998. Ontogeny and biomechanics of phalangeal form in primates. Ph.D. Dissertation. Stony Brook, NY: State University of New York.

Rohlf FJ. 2005. TPSDig2. Stony Brook, NY: State University of New York.

Rolian C, Gordon AD, Hallgrímsson B. 2011. Assessing manual proportions in *Australopithecus afarensis* using Monte Carlo resampling. Am J Phys Anthropol S144:256.

Rolian C. 2009. Integration and evolvability in primate hands and feet. Evol Biol 36:100–117.

Rossie JB, Gutierrez M, Goble E. 2012. Fossil forelimbs of *Simiolus* from Moruorot, Kenya. Am J Phys Anthropol 147(S54):252.

Sherwood RJ, Ward S, Hill A, Duren DL, Brown B, Downs W. 2002. Preliminary description of the *Equatorius africanus* partial skeleton (KNM-TH 28860) from Kipsaramon, Tugen Hills, Baringo District, Kenya. J Hum Evol 42:63–73.

Susman RL. 1979. Comparative and functional morphology of hominoid fingers. Am J Phys Anthropol 50:215–236.

Van Valen L. 2005. The statistics of variation. In: Hallgrímsson B, Hall BK, editors. Variation: a central concept in biology. London: Elsevier. p 29–47.

Ward CV. 2002. Interpreting the posture and locomotion of *Australopithecus afarensis*: where do we stand? Yearb Phys Anthropol 45:185–215.

Washburn SL. 1951. The analysis of primate evolution with particular reference to the origin of man. Cold Spring Harbor Symp Quant Biol 15:67–77.

Wood B, Richmond BG. 2000. Human evolution: taxonomy and paleobiology. J Anat 196:19–60.

Zapfe H. 1960. Die Primatenfunde aus der Miozanen Spaltenfullung von Neudorf an der March (Devinzka Nova Ves), Tchechoslowakei. Schweizerische Paleontologische Abhandlungen 78:4–293.