Three-Dimensional Geometry of Phalanges as a Proxy for Pair-Matching: Mesh Comparison Using an ICP Algorithm

Konstantina Tsiminikaki, Mara A. Karell,
Despoina Nathena, Demetrios Halazonetis,
Konstantinos Spanakis and Elena F. Kranioti

Abstract

Forensic anthropologists are frequently faced with the challenge of individualizing and sorting commingled remains in a variety of scenarios. A number of protocols have been proposed to standardize the methodological approach to individuating commingled remains, some of which are focused on pair-matching. A recent study by Karell et al. (2016) proposed a virtual method for pair-matching humeri using a semi-automatic procedure that gave encouraging results. With regards to the phalanges, there are only a handful of studies focusing on identifying and siding phalanges, as well as exploring their directional and functional asymmetry. Yet, they are still as important as every other bone when sorting commingled human remains in various situations, such as archaeological common burials and mass graves, commingled decomposed remains resulting from atrocities, accidents or natural disasters. This study investigates a new method for pair-matching, a common individualization technique, using digital three-dimensional models of bone: mesh-to-mesh value comparison (MVC) as proposed by Karell et al. (2016). The MVC method digitally compares the entire three-dimensional geometry of two bones using an iterative closest point (ICP) algorithm to produce a single value as a proxy for their similarity. The method is automated with the use of Viewbox software 4.1 *beta* for a simultaneous comparison of all possible pairs. For this study, 515 phalanges from 24 individuals of mixed ancestry were digitized using CT scans and the 3D modeling program AMIRA 5.3.3. The models were also hollowed (internal information of compact and trabecular bone removed) to test the method with simulated surface scan models. The subsequent data—over 73,000 comparisons—were assessed using sensitivity and specificity rates via ROC analysis to indicate how well the automated version of MVC pair-matched phalanges. The best bone in terms of pair-matching was the proximal phalanx of Digit 3 with 87.5% sensitivity and 92.4% specificity rates at a threshold value of 0.488 for the unhollowed bones. The specificity drops slightly (91.1%) when the hollowed

K. Tsiminikaki · M. A. Karell · E. F. Kranioti (✉)
Edinburgh Unit for Forensic Anthropology,
University of Edinburgh, Edinburgh, UK
e-mail: ekranioti@outlook.com

D. Nathena · E. F. Kranioti
Forensic Medicine Unit, Department of Forensic
Sciences, University of Crete, Heraklion, Greece

D. Halazonetis
Department of Orthodontics, School of Dentistry,
National and Kapodistrian University of Athens,
Athens, Greece

K. Spanakis
Department of Medical Imaging, University of Crete,
School of Medicine, Heraklion University Hospital,
Heraklion, Greece

models are compared. To compare the performance of the method in all phalanges, the specificity was set to 95%—allowing for a 5% acceptable error—and the adjusted sensitivity was compared. The highest sensitivity, namely 68.8%, was noted for Digit 2 proximal phalanx for both unhollowed and hollowed models. Thus far, our preliminary results indicate that the MVC method performs well when pair-matching phalanges, though it is less accurate than pair-matching other types of bones. The introduction of 95% specificity threshold allows for rejecting pairs in great confidence, which could, for instance, significantly reduce the number of DNA comparisons required for the remaining possible matches. In addition, the similar results obtained from hollowed and unhollowed models indicate that the internal information included in the unhollowed models adds little to the identification of true pairs. This means that if a CT scan is not available, the method could be applied to surface models produced by light and laser scanners as well. While additional work needs to be done to verify these preliminary results, this research has the potential to expand the repertoire of individualization methods.

Keywords

Forensic anthropology · Pair-matching · Hand phalanges · 3D models · Mesh-to-mesh value comparison · 3D pattern recognition · ICP

1 Introduction

The adult human skeleton consists of 206 bones, and each one of them has proven to be of extreme importance for the exploration of evolutionary, bioarchaeological and forensic questions. Some skeletal elements, like the skull and the pelvis, have received special attention, mainly due to their ability to elucidate issues such as herit¢ability and locomotion in modern and past populations (Martínez-Abadías et al. 2009; Gruss et al. 2015), while others, like the small hand and feet bones, have been comparatively neglected. This can be attributed to the poor preservation of such small bones in the archaeological and fossil record and the potential problems of identification and siding. Nevertheless, they are still as important as every other skeletal element in forensic investigations, for extracting biological information, that can lead to identification (e.g. unique conditions, handedness) (Danforth and Thompson 2008a; Varas and Thompson 2011) and/or provide evidence of violent events (e.g. defence cut marks, healed fractures) (Saukko and Knight 2016). The current study focuses on the morphological variation of phalanges in an effort to develop a pair-matching method applicable in commingled contexts.

Pair-matching techniques have been developed to improve the sorting of commingled human remains in various situations, such as, archaeological common burials, and mass graves, commingled decomposed remains produced by human atrocities, accidents or natural disasters (Garrido Varas and Intriago Leiva 2012; Karell et al. 2016). Such techniques can be based on visual assessment, osteometric sorting (Adams and Byrd 2006; Thomas et al. 2013) and/or more complicated methods of pattern comparisons such as geometric-morphometrics (Garrido-Varas et al. 2015) and point cloud comparison (Karell et al. 2016). For the purpose of developing pair-matching approaches, both virtual models and physical measurements have been used to develop methods for the scapula, the calcaneus and the metatarsals (Thomas et al. 2013; Garrido-Varas et al. 2015; Karell et al. 2016; Lynch 2017). Some individualisation techniques have successfully explored matching articulations based on metrics and regression analysis, but these are limited to large articulations of the lower limbs (Anastopoulou et al. 2018a, b). There is lack of methodological approaches for sorting smaller bones, such as the wrist bones, tarsal bones and phalanges.

Phalanges, in particular, pose a challenge for pair matching due to the fact that phalanges from

different digits of the hand can be easily mixed and confused when sorted with the naked eye, which complicates identification. There is no study to date that presents a clear methodology on how to pair-match phalanges belonging to the same individual accurately and reliably. Thus, the aim of this study was to develop a methodology of pair-matching phalanges based on 3D models from reconstructions of Computed Tomography scans following a similar methodology as Karell et al. (2016) developed for the humerus.

part of a different project (Virtopsy.GR). The Virtopsy.GR is a research project that explores the validity of postmortem computed tomography (PMCT) as an additional technique to the autopsy findings in forensic investigations of death on the island of Crete (Kranioti et al. 2017). Each individual had undergone PMCT just a few hours after death. The CT scan data were anonymized and each case was given an identification number with basic demographic data (age and sex). The project was approved by the Ethics Committee of the University Hospital of Heraklion in Crete in June 2016.

2 Materials and Methods

2.1 Materials

The current study employed a sample of 515 phalanges from 41 hands (18 right and 23 left) belonging to 24 individuals. Proximal (PP), middle (MP) and distal (DP) phalanges from every available digit were used. A detail list of the available samples for each digit can be found in Table 1.

The material of the study derived from individuals that were submitted to postmortem CT as

2.2 Methods

2.2.1 Scanning Protocol

The CT scans were acquired by the General University Hospital of Heraklion in Crete, using a Revolution GSI system (General Electric Medical Systems, USA). This system provides up to 128 slices per tube rotation and offers more than 15 applications for routine use in cardiac, oncology, neurology, spine, urology, musculoskeletal and more. In the GSI mode the system switches the tube potential from 80 to 140 kVp at

Table 1 Number of bones modelled from 24 individuals

Phalanges	Right	Left	Pairs
Digit 1 PP	17	23	16
Digit 1 DP	17	23	16
Digit 2 PP	16	22	15
Digit 2 MP	16	22	15
Digit 2 DP	17	22	16
Digit 3 PP	15	22	14
Digit 3 MP	17	22	16
Digit 3 DP	17	22	16
Digit 4 PP	15	22	14
Digit 4 MP	17	22	16
Digit 4 DP	16	22	15
Digit 5 PP	17	21	15
Digit 5 MP	16	22	15
Digit 5 DP	15	22	14
Total	228	287	213

a fast rate of up to 4.8 kHz. Thus, it allows the reconstruction of spectral images in the range from 40 to 140 keV. GE's Smart Spectral tools, such as GSI Assist and GSI Viewer 3D, enable one-click workflow on the console. GSI ASiR delivers dose neutral Spectral CT protocols. The scanning protocol for the Virtopsy.GR project uses a tube current of 50 mA, tube voltage of 120 kV, slice thickness of 0.625 mm and slice increment of 0.5 mm. Scans with a field of view of 250×250 mm (matrix 512×512) were made in the coronal (transverse) plane. Voxel size was $0.5 \times 0.5 \times 0.5$ mm. Data were saved as a Digital Imaging and Communications in Medicine (DICOM) format and then converted to a High Dynamic Range file. So, only the CT scan data (the HDR files) were used for the purpose of this study.

2.2.2 Scanning Method and Segmentation

Following the completion of these scans, each scan was "cropped" using Amira 6.0 software, so as to include only the hands. Each hand was loaded on Amira 6.0 software and each phalanx was manually segmented using the "brush" tool. This procedure resulted in 515 three-dimensional phalange models which were extracted as stereolithography (.stl) and wavefront [.obj] format and were then randomized prior to analysis.

2.2.3 Model Manipulation

For the purposes of this paper two datasets were analyzed. D1 included the original 3D models created from the segmentation of the CT scans and D2 the models, in which the internal material (compact and trabecular bone) was removed, maintaining only the surface. D2 models were created using a function of Viewbox 4.1 *beta* software.

2.2.4 Mesh-to-Mesh Value Comparison (MVC) Method

All right phalange models were mirror-imaged using the free software NetFabb basic and were named as mirror-imaged right models (MIR). Two folders containing all left (L) and MIR models were compared automatically using

Viewbox 4*beta* software following the guidelines set by Karell and colleagues (Karell et al. 2016). The software uses a trimmed ICP algorithm (Besl and McKay 1992; Chetverikov et al. 2002) to compare all homologous points between two models (meshes) and computes a single value which expresses the similarity between the shapes of the two models. The single value is called mesh-to mesh value (MTMV) and is expressed in mm. The software runs simultaneously all possible comparisons between the two folders using the following settings: The estimated overlap for the scan is set to 100%, whereas the number of initial positions for rough alignment was set at 20. This alignment used the nearest neighbor search "Approximate fast", with a point sampling of 1%, so that it matched point to point with one hundred iterations. On the other hand, the fine alignment used the nearest neighbor search "Exact with normal compatibility", with a point sampling of 100%, which matched point to plane with one hundred iterations. Finally, completing the mesh-to-mesh comparison, the program automatically creates an Excel spreadsheet of all the MTMV for analysis. The lowest MTMV are hypothesized to belong to true pairs, meaning the left and mirror-right phalanges belong to the same individual.

2.2.5 ROC Analysis

MedCalc software was used to conduct a Receiver Operating Characteristics (ROC) Analysis on the MTMV values. ROC Analysis is currently used to evaluate medical tests (Bewick et al. 2004), such as whether a person is positive or negative for a medical condition, as for instance, the presence of a virus. Here, ROC curves were employed in the evaluation of MTMV between potential pairs of bones as effective predictors of true pairs (belonging to the same individual). The hypothesis tested was whether a pair (L-MIR phalanx) is a correct match (positive) or not (negative). If both diagnosis (true match) and test (predicted match) are positive, the result is called true positive (TP), whereas if diagnosis is positive and the test is negative, the result is called false positive (FP). Similarly, a negative diagnosis with a negative

test is called true negative (TN) and a negative diagnosis with a positive test is called false positive (FP). The quality of the test can be measured with sensitivity and specificity (Kranioti and Tzanakis 2015). Sensitivity (true positive probability) is the proportion of true matches that are correctly identified by the test, while specificity (true negative probability) is the percentage of non-pairs that are correctly identified by the test. Predictive value of a positive test is defined as: $PVP = TP/(TP + FP)$ and predictive value of a negative test is defined as: $PVN = TN/(TN + FN)$ (Bewick et al. 2004; Kranioti and Tzanakis 2015). To help decide whether a pair is a match or not, a cut-off point of the MTMVs is chosen. ROC curve is widely accepted as a method for selecting an optimal cut-off point for a test and to make comparisons between tests (Akobeng 2007). The ROC curve is created by calculating sensitivity and specificity and creating a plot with $y = sensitivity$ and $x = 1 - specificity$ for the entire range of cut-off points (Kranioti and Tzanakis 2015). A large area under the curve (AUC) reflects good performance of the test (Bewick et al. 2004; Akobeng 2007).

3 Results

A total of 73,000 comparisons were automatically conducted, MTMVs were calculated and then analyzed using ROC analysis. Sensitivity, specificity, area under the curve (AUC) and cut-off points were calculated for datasets D1 and D2 and compared. For example, for D1 for Digit 2 proximal phalanx sensitivity was 71.4%, specificity was 98.0% and the threshold was set to 0.434 mm (Table 1). For intermediate and distal phalanges, sensitivity was 86.7% and 62.5%, while specificity was 62.9% and 72.1%, respectively. The threshold for the intermediate models was 0.588 mm and for the distal ones 0.511 mm. For D2 Digit 1 distal phalanx sensitivity was 100.0% and specificity 63.7%, and for Digit 3 proximal phalanx 85.7% and 92.4% respectively. As indicated in Table 2, there was no significant difference in the performance of the two datasets.

ROC analysis results in a cut-off value which combines the best prediction of true pairs with best rejection of non-pairs. It is possible though to calculate, using thousands of simulations, the

Table 2 Sensitivity, specificity, cut-off points and AUC for each paired phalanges

	D1				D2			
	SENS	SPEC	Criterion	AUC	SENS	SPEC	Criterion	AUC
Digit 1 PP	75	96	0.515	0.89	75	94.4	0.514	0.892
Digit 1 DP	93.8	66.1	0.588	0.853	100	63.7	0.586	0.876
Digit 2 PP	71.4	98	0.434	0.842	66.7	95.3	0.435	0.854
Digit 2 MP	86.7	62.9	0.588	0.786	86.7	60.8	0.593	0.781
Digit 2 DP	62.5	72.1	0.511	0.736	43.8	89.1	0.427	0.652
Digit 3 PP	*85.7*	*91.1*	*0.497*	*0.932*	*85.7*	*92.4*	*0.488*	*0.928*
Digit 3 MP	75	91.9	0.443	0.906	68.75	93.02	0.438	0.803
Digit 3 DP	50	98.3	0.386	0.718	56.3	91.6	0.434	0.715
Digit 4 PP	75	81.6	0.526	0.838	78.6	78.2	0.545	0.842
Digit 4 MP	81.3	77.4	0.544	0.84	81.3	74.3	0.54	0.833
Digit 4 DP	81.3	69.3	0.547	0.798	53.3	93.8	0.424	0.764
Digit 5 PP	86.7	68.2	0.619	0.839	73.3	80.4	0.543	0.806
Digit 5 MP	60	89.9	0.51	0.775	53.3	91.7	0.471	0.764
Digit 5 DP	50	81.7	0.521	0.67	42.9	91.5	0.449	0.647

values for different thresholds of specificity and sensitivity and the corresponding cut-off points. Tables 3 and 4 illustrate the adjusted values for a fixed sensitivity of 80, 90, 95, 97.5 and 99% (Table 3) and fixed specificity of 80, 90, 95, 97.5 and 99% (Table 4) for D1 and D2 in an effort to predict the highest number of true pairs and to reject the highest number of non-pairs.

The MTMVs for true pairs were analysed for D1 and D2. Table 5 illustrates mean, standard deviation, minimum and maximum values for both groups. A Wilcoxon paired test was performed using 10.000 Monte Carlo simulations to test whether there were differences between the means of the MTMVs of the true pairs and this test produced negative results.

Interpretation of the results and potential application of the method

To better explain the results of the analysis we will showcase an example using the Digit 3 PP subsample. We analysed 22 left and 15 right Digit 3 PP that can be combined in 330 possible pairs, of which only 14 are true pairs. ROC analysis results in AUC = 0.932 (Fig. 1) with

sensitivity 87.5% and specificity 91.1% for D1 and similar results for D2-87.5% and 92.4% respectively (see Table 4). We will assume that bones were scanned with a CT scanner and we will use D1 for the purpose of this exercise.

For a threshold value of 0.497 mm, all MTMVs equal or less than 0.497 mm will indicate a true pair. In our data this resulted in 40 pairs of which 12 are true pairs. The method correctly identified 12/14 pairs resulting in 85.7% accuracy, but it also identified 26 pairs that are not correct. Similarly, from the remaining 290 comparisons (MTMV > 0.497), all but two do not belong together, thus the method rejects two true pairs and 288 non-pairs (91.1%). To identify all pairs, one must fix specificity to 99% (Table 3), which will raise the MTM threshold value to 0.588 mm. This will result in identifying 95 pairs as matches, even though only 14 are true pairs. At the same time 235 pairs will be rejected. It is worth mentioning that 11 of the 14 true pairs showed the lowest MTMV when compared to each other in contrast with all other comparisons. Mean MTMV for true pairs was 0.4399 ± 0.0853 SD. MTMV values $>0.4399 + 2SD = 0.610$ are excluding

Table 3 Specificity, threshold values and 95% confidence intervals for fixed sensitivity at 80, 90, 95, 97.5 and 99% for D1 and D2

			Sensitivity %				
			80	90	95	97.5	99
Digit 1 PP	D1	Specificity %	84.8	58.13	58.13	32.27	32.27
		95% CI a	80.8–88.3	53.0–63.2	53.0–63.2	27.6–37.3	27.6–37.3
		Threshold	0.588	0.722	0.722	0.889	0.891
	D2	Specificity %	75.2	65.87	65.87	33.33	33.33
		95% CI a	70.5–79.5	60.8–70.7	60.8–70.7	28.6–38.4	28.6–38.4
		Threshold	0.628	0.680	0.681	0.898	0.900
Digit 1 DP	D1	Specificity %	78.13	66.13	66.13	41.87	41.87
		95% CI a	73.6—82.2	61.1–70.9	61.1–70.9	36.8–47.0	36.8–47.0
		Threshold	0.542	0.588	0.588	0.678	0.680
	D2	Specificity %	77.07	68	68	63.73	63.73
		95% CI a	72.5–81.2	63.0–72.7	63.0–72.7	58.6–68.6	58.6–68.6
		Threshold	0.539	0.571	0.571	0.586	0.586

(continued)

Table 3 (continued)

			Sensitivity %				
			80	90	95	97.5	99
Digit 2 PP	D1	Specificity %	66.47	41.84	35.31	35.31	0
		95% CI a	61.2–71.5	36.5–47.3	30.2–40.7	30.2–40.7	0.0–1.1
		Threshold	0.604	0.716	0.753	0.754	1.599
	D2	Specificity %	76.26	42.14	27.89	27.89	27.89
		95% CI a	71.4–80.7	36.8–47.6	23.2–33.0	23.2–33.0	23.2–33.1
		Threshold	0.548	0.714	0.809	0.818	0.818
Digit 2 MP	D1	Specificity %	68.25	46.29	46.29	13.06	13.06
		95% CI a	63.0–73.2	40.9–51.8	40.9–51.8	9.6–17.1	9.6–17.1
		Threshold	0.570	0.645	0.645	0.858	0.861
	D2	Specificity %	64.69	37.09	16.02	16.02	16.02
		95% CI a	59.3–69.8	31.9–42.5	12.3–20.4	12.3–20.4	12.3–20.5
		Threshold	0.572	0.694	0.806	0.806	0.811
Digit 2 DP	D1	Specificity %	57.26	27.37	27.37	3.07	3.07
		95% CI a	52.0–62.4	22.8–32.3	22.8–32.3	1.5–5.4	1.5–5.4
		Threshold	0.556	0.698	0.699	1.485	1.583
	D2	Specificity %	31.01	18.16	18.16	3.63	3.63
		95% CI a	26.2–36.1	14.3–22.5	14.3–22.5	1.9–6.1	1.9–6.2
		Threshold	0.685	0.799	0.808	1.299	1.328
Digit 3 PP	D1	Specificity %	91.14	91.14	76.27	74.37	74.37
		95% CI a	87.4–94.0	87.4–94.0	71.2–80.8	69.2–79.1	69.2–79.1
		Threshold	0.494	0.497	*0.588*	0.595	0.597
	D2	Specificity %	92.41	77.22	77.22	71.2	72.2
		95% CI a	88.9–95.1	72.2–81.7	72.2–81.7	65.9–76.1	65.9–76.2
		Threshold	0.488	0.574	0.578	0.604	0.604
Digit 3MP	D1	Specificity %	82.96	61.17	61.17	55.03	55.03
		95% CI a	78.7–86.7	55.9–66.3	55.9–66.3	49.7–60.3	49.7–60.3
		Threshold	0.500	0.609	0.609	0.644	0.647
	D2	Specificity %	64.25	20.67	20.67	2.79	3.79
		95% CI a	59.0–69.2	16.6–25.2	16.6–25.2	1.3–5.1	1.3–5.2
		Threshold	0.611	1.005	1.011	1.93221123	1.978
Digit 3 DP	D1	Specificity %	32.68	24.02	24.02	10.61	10.61
		95% CI a	27.8–37.8	19.7–28.8	19.7–28.8	7.6–14.3	7.6–14.3
		Threshold	0.644	0.727	0.727	0.896	0.924
	D2	Specificity %	31.84	22.63	22.63	11.45	11.45
		95% CI a	27.0–36.9	18.4–27.3	18.4–27.3	8.3–15.2	8.3–15.2
		Threshold	0.660	0.751	0.752	0.91583532	0.926

(continued)

Table 3 (continued)

			Sensitivity %				
			80	90	95	97.5	99
Digit 4 PP	D1	Specificity %	68.44	55.59	55.59	18.72	18.72
		95% CI a	63.3–73.2	50.3–60.8	50.3—60.8	14.8–23.1	14.8–23.1
		Threshold	0.606	0.662	0.664	0.914	0.915
	D2	Specificity %	77.93	56.7	56.7	22.35	23.35
		95% CI a	73.3–82.1	51.4–61.9	51.4–61.9	18.1–27.0	18.1–27.1
		Threshold	0.545	0.660	0.660	0.896	0.898
Digit 4 MP	D1	Specificity %	77.37	53.91	53.91	23.18	23.18
		95% CI a	72.7–81.6	48.6–59.2	48.6–59.2	18.9–27.9	18.9–27.9
		Threshold	0.544	0.640	0.640	0.823	0.824
	D2	Specificity %	74.3	54.19	54.19	24.02	24.02
		95% CI a	69.4–78.8	48.9–59.4	48.9–59.4	19.7–28.8	19.7–28.8
		Threshold	0.540	0.629	0.631	0.812	0.813
Digit 4 DP	D1	Specificity %	69.27	53.63	53.63	12.85	12.85
		95% CI a	64.2–74.0	48.3–58.9	48.3–58.9	9.6–16.8	9.6–16.8
		Threshold	0.547	0.607	0.608	0.898	0.905
	D2	Specificity %	47.21	46.09	46.09	40.78	40.78
		95% CI a	41.9–52.5	40.8–51.4	40.8–51.4	35.6–46.1	35.6–46.1
		Threshold	0.617	0.622	0.622	0.647	0.648
Digit 5 PP	D1	Specificity %	73.68	68.71	**68.71**	38.3	38.3
		95% CI a	68.7–78.3	63.5–73.6	63.5–73.6	33.1–43.7	33.1–43.7
		Threshold	0.578	0.611	0.619	0.778	0.779
	D2	Specificity %	69.59	38.6	38.6	27.19	27.19
		95% CI a	64.4–74.4	33.4–44.0	33.4–44.0	22.5–32.2	22.5–32.2
		Threshold	0.599	0.768	0.770	0.839	0.839
Digit 5 MP	D1	Specificity %	53.41	41.25	41.25	20.18	20.18
		95% CI a	47.9–58.8	35.9–46.7	35.9–46.7	16.0–24.9	16.0–24.9
		Threshold	0.701	0.763	0.764	0.901	0.902
	D2	Specificity %	55.79	42.43	42.43	5.64	5.64
		95% CI a	50.3–61.2	37.1–47.9	37.1–47.9	3.4–8.7	3.4–8.7
		Threshold	0.664	0.740	0.741	1.079	1.079
Digit 5 DP	D1	Specificity %	44.62	31.01	1.27	1.27	0
		95% CI a	39.1–50.3	26.0–36.4	0.3–3.2	0.3–3.2	0.0–1.2
		Threshold	0.699	0.759	1.120	1.151	1.246
	D2	Specificity %	31.96	14.24	5.38	5.38	0
		95% CI a	26.9–37.4	10.6–18.6	3.2–8.5	3.2–8.5	0.0–1.2
		Threshold	0.731	0.881	1.033	1.040	1.417

Table 4 Sensitivity, threshold values and 95% confidence intervals for fixed specificity at 80, 90, 95, 97.5 and 99% for D1 and D2

			Specificity %				
			80	90	95	97.5	99
Digit 1 PP	D1	Sensitivity %	81.25	75	75	50	43.75
		95% CI a	54.4–96.0	47.6–92.7	47.6–92.7	24.7–75.3	19.8–70.1
		Threshold	0.588	0.586	0.515	0.472	0.454
	D2	Sensitivity %	75	75	68.75	50	50
		95% CI a	47.6–92.7	47.6–92.7	41.3–89.0	24.7–75.3	24.7–75.3
		Threshold	0.627	0.514	0.497	0.488	0.458
Digit 1 DP	D1	Sensitivity %	50	50	43.75	37.5	12.5
		95% CI a	24.7–75.3	24.7–75.3	19.8–70.1	15.2–64.6	1.6–38.3
		Threshold	0.514	0.483	0.431	0.428	0.365
	D2	Sensitivity %	75	43.75	43.75	37.5	6.25
		95% CI a	47.6–92.7	19.8–70.1	19.8–70.1	15.2–64.6	0.2–30.2
		Threshold	0.514	0.473	0.433	0.402	0.366
Digit 2 PP	D1	Sensitivity %	73.33	66.67	66.67	60	53.33
		95% CI a	44.9–92.2	38.4–88.2	38.4–88.2	32.3–83.7	26.6–78.7
		Threshold	0.495	0.488	0.434	0.413	0.406
	D2	Sensitivity %	73.33	73.33	66.67	53.33	26.67
		95% CI a	44.9–92.2	44.9–92.2	38.4–88.2	26.6–78.7	7.8–55.1
		Threshold	0.545	0.503	0.435	0.406	0.330
Digit 2 MP	D1	Sensitivity %	60	40	33.33	26.67	20
		95% CI a	32.3–83.7	16.3–67.7	11.8–61.6	7.8–55.1	4.3–48.1
		Threshold	0.477	0.437	0.415	0.386	0.351
	D2	Sensitivity %	60	46.67	33.33	26.67	13.33
		95% CI a	32.3–83.7	21.3–73.4	11.8–61.6	7.8–55.1	1.7–40.5
		Threshold	0.512	0.440	0.411	0.381	0.354
Digit 2 DP	D1	Sensitivity %	50	43.75	31.25	18.75	18.75
		95% CI a	24.7–75.3	19.8–70.1	11.0–58.7	4.0–45.6	4.0–45.6
		Threshold	0.496	0.429	0.390	0.376	0.357
	D2	Sensitivity %	43.75	43.75	25	12.5	12.5
		95% CI a	19.8–70.1	19.8–70.1	7.3–52.4	1.6–38.3	1.6–38.4
		Threshold	0.502	0.427	0.387	0.380	0.332
Digit 3 PP	D1	Sensitivity %	85.71	85.71	50	35.71	21.43
		95% CI a	57.2–98.2	57.2–98.2	23.0–77.0	12.8–64.9	4.7–50.8
		Threshold	0.585	0.497	0.438	0.396	0.377
	D2	Sensitivity %	85.71	86.71	50	21.43	14.29
		95% CI a	57.2–98.2	57.2–98.3	23.0–77.0	4.7–50.8	1.8–42.8
		Threshold	0.574	0.488	0.428	0.392	0.381

Table 4 (continued)

			Specificity %				
			80	90	95	97.5	99
Digit 3MP	D1	Sensitivity %	75	75	56.25	50	50
		95% CI a	47.6–92.7	47.6–92.7	29.9–80.2	24.7–75.3	24.7–75.3
		Threshold	0.498	0.443	0.420	0.401	0.377
	D2	Sensitivity %	68.75	62.5	50	43.75	31.25
		95% CI a	41.3–89.0	35.4–84.8	24.7–75.3	19.8–70.1	11.0–58.7
		Threshold	0.438	0.438	0.417	0.376	0.366
Digit 3 DP	D1	Sensitivity %	62.5	56.25	50	50	25
		95% CI a	35.4–84.8	29.9–80.2	24.7–75.3	24.7–75.3	7.3–52.4
		Threshold	0.469	0.461	0.461	0.386	0.360
	D2	Sensitivity %	56.25	50	50	43.75	37.5
		95% CI a	29.9–80.2	24.7–75.3	24.7–75.3	19.8–70.1	15.2–64.6
		Threshold	0.434	0.433	0.411	0.382	0.360
Digit 4 PP	D1	Sensitivity %	75	56.25	43.75	31.25	25
		95% CI a	47.6–92.7	29.9–80.2	19.8–70.1	11.0–58.7	7.3–52.4
		Threshold	0.526	0.471	0.415	0.402	0.377
	D2	Sensitivity %	75	62.5	37.5	25	12.5
		95% CI a	47.6–92.7	35.4–84.8	15.2–64.6	7.3–52.4	1.6–38.3
		Threshold	0.518	0.460	0.428	0.384	0.320
Digit 4 MP	D1	Sensitivity %	75	56.25	43.75	25	25
		95% CI a	47.6–92.7	29.9–80.2	19.8–70.1	7.3–52.4	7.3–52.4
		Threshold	0.520	0.481	0.428	0.407	0.336
	D2	Sensitivity %	62.5	56.25	37.5	25	18.75
		95% CI a	35.4–84.8	29.9–80.2	15.2–64.6	7.3–52.4	4.0–45.6
		Threshold	0.515	0.469	0.421	0.343	0.335
Digit 4 DP	D1	Sensitivity %	56.25	50	43.75	31.25	12.5
		95% CI a	29.9–80.2	24.7–75.3	19.8–70.1	11.0–58.7	1.6–38.3
		Threshold	0.480	0.443	0.415	0.401	0.327
	D2	Sensitivity %	56.25	50	31.25	18.75	12.5
		95% CI a	29.9–80.2	24.7–75.3	11.0–58.7	4.0–45.6	1.6–38.3
		Threshold	0.480	0.424	0.407	0.382	0.353
Digit 5 PP	D1	Sensitivity %	66.67	53.33	26.67	20	13.33
		95% CI a	38.4–88.2	26.6–78.7	7.8–55.1	4.3–48.1	1.7–40.5
		Threshold	0.527	0.483	0.412	0.378	0.328
	D2	Sensitivity %	73.33	46.67	33.33	26.67	26.67
		95% CI a	44.9–92.2	21.3–73.4	11.8–61.6	7.8–55.1	7.8–55.1
		Threshold	0.543	0.476	0.407	0.399	0.354

(continued)

Table 4 (continued)

			Specificity %				
			80	90	95	97.5	99
Digit 5 MP	D1	Sensitivity %	60	60	46.67	13.33	13.33
		95% CI a	32.3–83.7	32.3–83.7	21.3–73.4	1.7–40.5	1.7–40.5
		Threshold	0.627	0.510	0.470	0.419	0.358
	D2	Sensitivity %	60	53.33	33.33	26.67	20
		95% CI a	32.3–83.7	26.6–78.7	11.8–61.6	7.8–55.1	4.3–48.1
		Threshold	0.535	0.471	0.429	0.417	0.343
Digit 5 DP	D1	Sensitivity %	50	28.57	28.57	14.29	7.14
		95% CI a	23.0–77.0	8.4–58.1	8.4–58.1	1.8–42.8	0.2–33.9
		Threshold	0.521	0.482	0.419	0.404	0.333
	D2	Sensitivity %	50	42.86	35.71	14.29	7.14
		95% CI a	23.0–77.0	17.7–71.1	12.8–64.9	1.8–42.8	0.2–33.9
		Threshold	0.520	0.449	0.417	0.392	0.308

Table 5 Minimum, maximum, mean and standard deviation of MTMVs for true pairs and Wilcoxon z-scores for paired differences between D1 and D2

|	N	D1				D2				Wilcoxon Z-value	p-value
		Min	Max	Mean	SD	Min	Max	Mean	SD		
Digit 1 PP	16	0.2854	0.8913	0.5038	0.1564	0.2808	0.9003	0.4998	0.1528	−0.414	0.688
Digit 1 DP	16	0.3222	0.6796	0.4768	0.0984	0.3248	0.5863	0.4628	0.0797	−0.414	0.688
Digit 2 PP	15	0.2756	0.7538	0.4602	0.1619	0.2397	0.8178	0.4514	0.1618	−0.170b	0.879
Digit 2 MP	15	0.2807	0.8606	0.4882	0.1474	0.2878	0.8114	0.4848	0.1446	−0.568	0.577
Digit 2 DP	16	0.2914	1.5827	0.5327	0.2999	0.2943	1.3277	0.5603	0.2551	−0.879	0.387
Digit 3 PP	14	0.2942	0.5966	0.4399	0.0853	0.2950	0.6044	0.4408	0.0836	−0.659	0.526
Digit 3 MP	16	0.2636	0.6465	0.4180	0.1088	0.2528	1.9784	0.5551	0.4235	−0.776	0.445
Digit 3 DP	16	0.2948	0.9239	0.4991	0.1791	0.2880	0.9257	0.5022	0.1867	−0.310	0.770
Digit 4 PP	16	0.3055	0.9151	0.4850	0.1588	0.3056	0.8983	0.4858	0.1470	−0.103	0.933
Digit 4 MP	16	0.2652	0.8241	0.4701	0.1389	0.2837	0.8132	0.4706	0.1328	−0.103	0.933
Digit 4 DP	16	0.3071	0.9055	0.4875	0.1435	0.2992	0.6476	0.4781	0.1184	−0.569	0.579
Digit 5 PP	15	0.2990	0.7785	0.4999	0.1217	0.2989	0.8394	0.5096	0.1547	−0.114	0.922
Digit 5 MP	15	0.3375	0.9017	0.5500	0.1672	0.3350	1.0793	0.5442	0.1993	−0.114	0.922
Digit 5 DP	14	0.3326	1.1509	0.5950	0.2076	0.3080	1.0397	0.5823	0.2084	−0.157	0.895
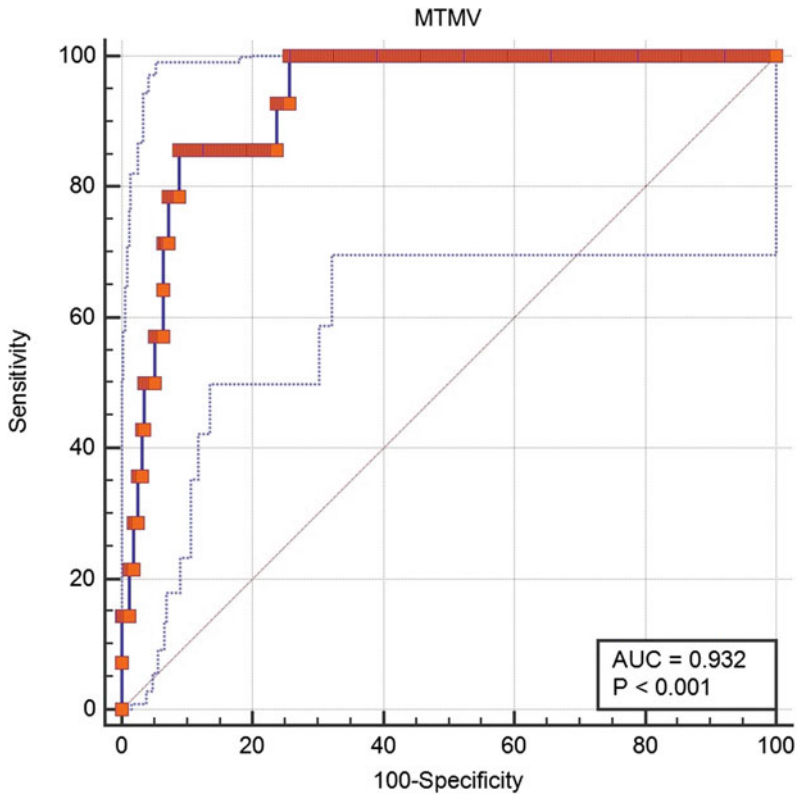
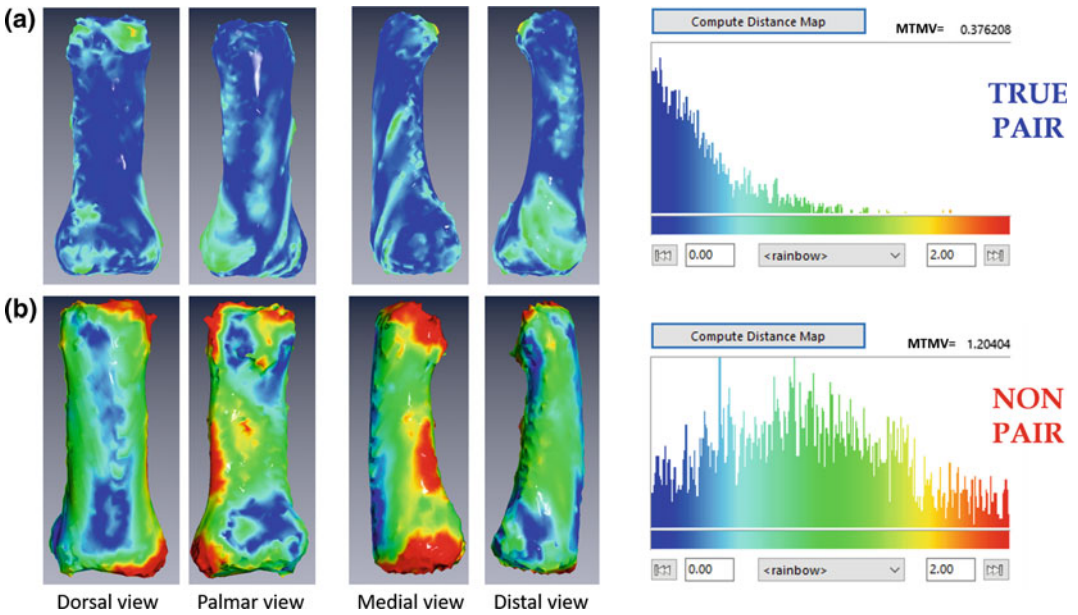**Fig. 1** Area under the curve for Digit 3 proximal phalanx

Fig. 2 A colored distance map applied to two possible pairs of Digit 3 proximal phalanges created with Viewbox 4.1 *beta*, **a** true pair with MTMV = 0.376 and **b** non pair with MTMV = 1.201

correctly 227 non-pairs. MTMV values >0.4399 + 3SD = 0.6958 are classified as non-pairs with great confidence.

Figure 2 illustrates an example of a true match (a, MTMV = 0.376 mm) and of a non match (b, MTVT = 1.201 mm). A colored distance map has been applied in both cases to show true distances between homologous points on the two models. Blue indicates point distances close to zero and red values close to 2 mm.

Thus, one can conclude:

1. MTMV values >0.610 mm are highly indicative of non-pairs in Digit 3 proximal phalanges while >0.700 mm safely indicates non-pairs.
2. Lowest MTMV for a pair of Digit 3 proximal phalanges indicates high possibility of bones to belong together.

Naturally these results need to be confirmed with an independent sample.

4 Discussion

Phalanges are probably the most neglected bones of the human skeleton due to their small size. They have been previously studied for the purpose of creating biometric standards for sex (Smith 1991, 2000; Scheuer and Elkington 1993; Case and Ross 2007; El Morsi and Al Hawary 2013; Mahakkanukrauh et al. 2013), stature (Habib and Kamal 2010) and age estimation (Gilsanz and Ratib 2005) based on both classical osteometry and virtual methods. Yet, there is a substantial lack of population specific standards compared to other bones that are more likely to be recovered in forensic or archaeological settings, such as the skull or long bones. Nevertheless, there are a number of studies examining the hand morphology of modern humans, primates and fossils from an evolutionary and functional perspective (Deane and Begun 2008; Tocheri et al. 2008; Mednikova 2011, 2013; Ward et al. 2014; Almécija et al. 2015; Lorenzo et al. 2015), whereas a few others have examined hand plasticity in relation to activity (Karakostis

et al. 2016). In addition, a few studies focused on providing a detailed methodology for identifying and siding phalanges and exploring directional and functional asymmetry (Case and Heilman 2006; Danforth and Thompson 2008b; Christensen 2009; Garrido Varas and Thompson 2011). To date, no other study has explored potential pair-matching techniques for sorting phalanges in commingled situations.

The current study used 515 hand phalanges belonging to 24 individuals to explore the potential value of the symmetry in these bones in sorting pairs. Left and right bones were CT scanned, modelled and compared to each other using a series of software combining both manual, semi-automated and automated procedures. A single value was used to compare the overall shape similarity between two models (L and MIR) and was then used to differentiate between pairs and non-pairs with the aid of ROC curves. Different threshold values were calculated in order to facilitate either the inclusion of all pairs (99% sensitivity) or the safe exclusion of non-pairs (99% specificity). These calculations allow an interactive use of the ROC curve by the user, depending on the circumstances of the case under investigation. For example, excluding possible matches may be as crucial to a forensic investigation as identifying true matches. In fact, the results of this study indicate that MTMVs > 0.61 for Digit 3 PP (which coincides with 2SD from the mean) is highly indicative of a non-match. This is in agreement with a previous study using the humerus, where all but one true pair were identified using a threshold of the mean plus two standard deviations (Karell et al. 2016). By selecting a threshold of 3 standard deviations from the mean, it is almost certain that one can safely identify non-pairs (see example of Digit 3 PP). This information is very important in a commingled situation as it would reduce significantly the number of DNA comparisons necessary to ascribe the skeletal element to the correct missing person.

The current method is the first semiautomatic 3-dimensional method that attempts to pair-match phalanges. The method was developed in a mixed ancestry and sex sample and these factors do not

seem to affect the performance of the method in accordance with previous studies using MTMVs (Karell et al. 2016). Naturally, the sample needs to be further expanded in size, including surface scans and bones in different taphonomic conditions, so that the preliminary findings presented here can be validated. Nevertheless, the results appear very promising, especially if one takes into account the fact that, although phalanges are the least anatomically complex skeletal elements of the human body, the current method achieves notable success in rejecting a substantial number of possible pairs.

References

Adams BJ, Byrd JE (2006) Resolution of small-scale commingling: a case report from the Vietnam War. Forensic Sci Int 156:63–69. Available from http://www.sciencedirect.com/science/article/pii/S0379073805000952

Akobeng AK (2007) Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr 96:644–647

Almécija S, Smaers JB, Jungers WL (2015) The evolution of human and ape hand proportions. Nat Commun 6:7717. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4510966/

Anastopoulou I, Karakostis FA, Borrini M, Moraitis K (2018a) A statistical method for reassociating human tali and calcanei from a commingled context. J Forensic Sci 63:381–385

Anastopoulou I, Karakostis FA, Moraitis K (2018b) A reliable regression-based approach for reassociating human skeletal elements of the lower limbs from commingled assemblages. J Forensic Sci

Besl PJ, McKay ND (1992) A method for registration of 3-D shapes. IEEE Trans Patt Anal Mach Intell 14:239–256

Bewick V, Cheek L, Ball J (2004) Statistics review 13: receiver operating characteristic curves. Critical care 8:508–512

Case DT, Heilman J (2006) New siding techniques for the manual phalanges: a blind test. Int J Osteoarchaeol 16:338–346. Available from http://dx.doi.org/10.1002/oa.826

Case DT, Ross AH (2007) Sex determination from hand and foot bone lengths*. J Forensic Sci 52:264–270. Available from http://dx.doi.org/10.1111/j.1556-4029.2006.00365.x

Chetverikov D, Svirko D, Stepanov D, Krsek P (2002) The trimmed iterative closest point algorithm. Object Recognit Support by user Interact Serv Robot 3:545–548. Available from http://ieeexplore.ieee.org/document/1047997/

Christensen AM (2009) Techniques for siding manual phalanges. Forensic Sci Int 193:84–87. Available from http://www.sciencedirect.com/science/article/pii/S037907380900396X

Danforth ME, Thompson A (2008a) An evaluation of determination of handedness using standard osteological measurements*. J Forensic Sci 53:777–781. Available from http://dx.doi.org/10.1111/j.1556-4029.2008.00741.x

Danforth ME, Thompson A (2008b) An evaluation of determination of handedness using standard osteological measurements. J Forensic Sci 53:777–781. Available from http://doi.wiley.com/10.1111/j.1556-4029.2008.00741.x

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. Available from http://www.sciencedirect.com/science/article/pii/S004724840800095X

El Morsi DA, Al Hawary AA (2013) Sex determination by the length of metacarpals and phalanges: {X}-ray study on {Egyptian} population. J Forensic Leg Med 20:6–13. Available from http://www.sciencedirect.com/science/article/pii/S1752928X12000935

Garrido-Varas C, Rathnasinghe R, Thompson T, Savriama Y (2015) A new method to pair-match metacarpals using bilateral asymmetry and shape analysis. J Forensic Sci 60:118–123

Garrido Varas C, Intriago Leiva M (2012) Managing commingled remains from mass graves: considerations, implications and recommendations from a human rights case in Chile. Forensic Sci Int 219:e19–e24. Available from http://www.sciencedirect.com/science/article/pii/S0379073811005706

Garrido Varas CE, Thompson TJU (2011) Metric dimensions of the proximal phalanges of the human hand and their relationship to side, position, and asymmetry. HOMO—J Comp Hum Biol 62:126–143. Available from http://www.sciencedirect.com/science/article/pii/S0018442X10001307

Gilsanz V, Ratib O (2005) Hand bone age: a digital atlas of skeletal maturity. Springer, Berlin Heidelberg

Gruss LT, Schmitt D, Gruss LT (2015) The evolution of the human pelvis: changing adaptations to bipedalism, obstetrics and thermoregulation

Habib SR, Kamal NN (2010) Stature estimation from hand and phalanges lengths of Egyptians. J Forensic Leg Med 17:156–160. Available from http://www.sciencedirect.com/science/article/pii/S1752928X09002145

Karakostis FA, Quéré E Le, Vanna V, Moraitis K (2016) Assessing the effect of manual physical activity on proximal hand phalanges using Hellenistic and modern skeletal samples from Greece. {HOMO}—J Comp Hum Biol 67:110–124. Available from http://www.sciencedirect.com/science/article/pii/S0018442X15000979

Karell MA, Langstaff HK, Halazonetis DJ, Minghetti C, Frelat M, Kranioti EF (2016) A novel method for pair-matching using three-dimensional digital models of bone: mesh-to-mesh value comparison. Int J Legal Med 130:1315–1322. Available from http://dx.doi.org/10.1007/s00414-016-1334-3

Kranioti EF, Nathena D, Spanakis K, Bouhaidar R, McLaughlin S, Papadomanolakis A, Karantanas A (2017) Postmortem CT in the investigation of decomposed human remains: advantages and limitations. La Rev Médecine Légale 8:184–185

Kranioti EF, Tzanakis N (2015) Estimation of sex from the upper limb in modern cretans with the aid of ROC-analysis: a technical report. Forensic Res Criminol Int J Forensic Res Criminol Int J 1. Available from http://medcraveonline.com

Lorenzo C, Pablos A, Carretero JM, Huguet R, Valverdú J, Martinón-Torres M, Arsuaga JL, Carbonell E, de Castro JMB (2015) Early Pleistocene human hand phalanx from the Sima del Elefante (TE) cave site in Sierra de Atapuerca (Spain). J Hum Evol 78:114–121. Available from http://www.sciencedirect.com/science/article/pii/S0047248414001857

Lynch J (2017) An automated two-dimensional form registration method for osteological pair-matching. J Forensic Sci 1–7

Mahakkanukrauh P, Khanpetch P, Prasitwattanseree S, Case DT (2013) Determination of sex from the proximal hand phalanges in a Thai population. Forensic Sci Int 226:208–215. Available from http://www.sciencedirect.com/science/article/pii/S0379073813000303

Martínez-Abadías N, Esparza M, Sjøvold T, González-josé R, Santos M, Hernández M (2009) Heritability of human cranial dimensions : comparing the evolvability of different cranial regions 19–35

Mednikova MB (2011) A proximal pedal phalanx of a Paleolithic hominin from denisova cave, Altai. Archaeol Ethnol Anthropol Eurasia 39:129–138. Available from http://www.sciencedirect.com/science/article/pii/S1563011011000444

Mednikova MB (2013) Distal phalanx of the hand of homo from denisova cave stratum 12: a tentative description. Archaeol Ethnol Anthropol Eurasia 41:146–155. Available from http://www.sciencedirect.com/science/article/pii/S1563011013000664

Saukko P, Knight B (2016) Knight's forensic pathology, 4th edn. In: Saukko P, Knight B (eds) CRC Press, LLC, London

Scheuer JL, Elkington NM (1993) Sex determination from metacarpals and the first proximal phalanx. J Forensic Sci 38:769–778

Smith SL (1991) Metrical variation in the thumb, index, and middle finger among four samples of both sexes. Skeletal Radiol 20:353–362

Smith SL (2000) Shape variation of the human pollical distal phalanx and metacarpal. Am J Phys Anthropol 113:329–348. Available from http://doi.wiley.com/10.1002/1096-8644%28200011%29113%3A3%3C329%3A%3AAID-AJPA5%3E3.0.CO%3B2-B

Thomas RM, Ubelaker DH, Byrd JE (2013) Tables for the metric evaluation of pair-matching of human skeletal elements. J Forensic Sci 58:952–956. Available from http://dx.doi.org/10.1111/1556-4029.12133

Tocheri MW, Orr CM, Jacofsky MC, Marzke MW (2008) The evolutionary history of the hominin hand since the last common ancestor of Pan and Homo. J Anat 212:544–562. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2409097/

Varas CEG, Thompson TJU (2011) HOMO—journal of comparative human biology metric dimensions of the proximal phalanges of the human hand and their relationship to side, position, and asymmetry. HOMO-Journal Comp Hum Biol 62:126–143. Available from http://dx.doi.org/10.1016/j.jchb.2010.07.005

Ward CV, Tocheri MW, Plavcan JM, Brown FH, Manthi FK (2014) Early pleistocene third metacarpal from kenya and the evolution of modern human-like hand morphology. Proc Natl Acad Sci 111:121–124. Available from http://www.pnas.org/content/111/1/121.abstract