3D geometric morphometrics of thorax-pelvis covariation and its potential for predicting the thorax morphology: A case study on Kebara 2 Neandertal

Nicole Torres-Tamayo[a,b,*], Stefan Schlager[c], Daniel García-Martínez[a,d], Juan Alberto Sanchis-Gimeno[b], Shahed Nalla[e], Naomichi Ogihara[f], Motoharu Oishi[g], Sandra Martelli[h], Markus Bastir[a]

[a] *Paleoanthropology Group, Museo Nacional de Ciencias Naturales (CSIC), J.G. Abascal 2, 28006, Madrid, Spain*

[b] *GIAVAL Research Group, Department of Anatomy and Human Embryology, University of Valencia, Av. Blasco Ibanez, 15. E-46010, Valencia, Spain*

[c] *Biological Anthropology, Faculty of Medicine, University of Freiburg, Hebelstr 29, D-79104, Freiburg, Germany*

[d] *Centro Nacional de Investigación sobre la Evolución Humana (CENIEH), Avenida de La Sierra de Atapuerca 3, 09002, Burgos, Spain*

[e] *Department of Human Anatomy and Physiology, Faculty of Health Sciences, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Gauteng, South Africa, Johannesburg, South Africa*

[f] *Department of Biological Science, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan*

[g] *Laboratory of Anatomy 1, School of Veterinary Medicine, Azabu University, Kanagawa 252-5201, Japan*

[h] *UCL Centre for Integrative Anatomy (CIA), Department of Cell and Developmental Biology, Faculty of Life Sciences, Gower Street WC1E 6BT, London, UK*

*Corresponding author.

E-mail address: nicole_bio_ucm@hotmail.com (N. Torres-Tamayo).

Acknowledgments

The authors thank the members of the Virtual Morphology Lab (MNCN-CSIC) and Lucía Peña López for helpful and critical discussions on previous versions of the manuscript. The authors are grateful to Caroline VanSickle, Jeremy DeSilva, Karen Rosenberg, Gary Sawyer, Blaine Maley, Tara Chapman, and Antonio Rosas for providing access to the casts and 3D models of Kebara 2 reconstructions used in this study. Special thanks go to Asier Gómez-Olivencia for uploading the 3D virtual model of his reconstruction of Kebara 2 ribcage to the online open access repository Figshare, for kindly accepting to revise this manuscript and for helpful comments and discussions. Lastly, the authors thank Miss Abigail O'Connell (Hull York Medical School, York) for generously accepting to revise the English grammar of the last version of this manuscript. This paper is dedicated to the memory of Dennis E. Slice for his contribution to the development of geometric morphometrics methods and software. This research was supported by the Ministerio de Economía, Industria y Competitividad of Spain (AEI/FEDER, EU, grant number CGL-2015-63648-P).

3D geometric morphometrics of thorax-pelvis covariation and its potential for predicting thorax morphology: A case study on Kebara 2 Neandertal

ABSTRACT

The skeletal torso is a complex structure of outstanding importance to understanding human body shape evolution, but reconstruction usually entails an element of subjectivity as each researcher applies their own anatomical expertise to the process. Among different fossil reconstruction methods, 3D geometric morphometric techniques have been increasingly used in the last decades. Two-block partial least squares (2B-PLS) analysis has shown great potential for predicting missing elements by exploiting the covariation between two structures (blocks) in a reference sample: one block can be predicted from the other one based on the strength of covariation between blocks. The first aim of this study is to test whether this predictive approach can be used for predicting thorax morphologies from pelvis morphologies within adult *Homo sapiens* reference samples with known covariation between the thorax and the pelvis. The second aim is to apply this method to Kebara 2 Neandertal (Israel, ~60 ka) to predict its thorax morphology using two different pelvis reconstructions as predictors. We measured 134 fixed landmarks, 720 curve semilandmarks and 160 surface semilandmarks on n = 60 3D virtual torso models segmented from CT scans. We conducted three 2B-PLS analyses between the thorax (block 1) and the pelvis (block 2) based on the *H. sapiens* reference samples, after performing generalized Procrustes superimposition on each block separately. Comparisons of these predictions in full shape space by means of Procrustes distances show that the male-only predictive model yields the most reliable predictions within modern humans. Additionally, Kebara 2 thorax predictions based on this model concur with the thorax morphology proposed for Neandertals. The method presented here does not aim to replace other techniques, but to rather complement them through the quantitative prediction of a virtual 'scaffold' to articulate the thoracic fossil elements, thus extending the potential of missing data estimation beyond the methods proposed in previous works.

Keywords: Partial least squares; *Homo neanderthalensis*; Ribcage; Prediction; Fossil

1. Introduction

The skeletal torso is a complex structure of huge importance to understanding the evolution of human body shape and its paleobiological implications (Ruff, 1991; Schmid, 1991; Carretero et al., 2004; Holliday, 2012; Jungers et al., 2016; Bastir et al., in press). Torso reconstruction in fossil hominins is crucial to shed light on their body plans, particularly in view of how tightly it is linked to body mass and stature (Ruff et al., 1997; Arsuaga et al., 1999, 2015; Holliday, 2012; Carretero et al., 2004; Simpson et al., 2008; Ruff, 2010), and potentially body energetic demands (Franciscus and Churchill, 2002; Churchill, 2006, 2014; Froehle and Churchill, 2009; García-Martínez et al., 2018a, b, c; Gómez-Olivencia et al., 2018). However, torso reconstruction may pose challenges because its morphology depends on the interaction not only within the thorax (ribs, vertebrae and sternum), the lumbar spine, and the pelvis, but also between them and their associated soft tissues, which are not generally preserved in the fossil record. Fortunately, as thorax and pelvis morphologies covary (Middleton, 2015; Torres-Tamayo et al., 2018), these anatomical parts can inform us about each other and previous researchers used this information to elucidate complete hominin torso morphologies (Schmid, 1983, 1991; Jellema et al., 1993; Walker and Ruff, 1993; Tague and Lovejoy, 1998; Sawyer and Maley, 2005; Simpson et al., 2008; Berge and Goularas, 2010; Schmid et al., 2013; Latimer et al., 2016; Brassey et al., 2018; Gómez-Olivencia et al., 2018; Laudicina et al., 2019).

1.1. Reconstruction methods of fossil trunk elements

In some of the aforementioned studies, reconstructions of skeletal trunk elements have been made through classical methodological approaches using different materials, such as wax or plaster (e.g., Schmid, 1983; Walker and Ruff, 1993; Sawyer and Maley, 2005; Simpson et al., 2008). However, virtual reconstruction methods have become increasingly popular in the past few years and are an excellent alternative to these classic techniques as they reduce the risk of damaging the original fossils during handling (Zollikofer and Ponce de León, 2005; Gunz et al., 2009; Weber and Bookstein, 2011; Bastir et al., 2019a). Among these techniques, geometric morphometrics (GM) has become an important tool for aiding in fossil reconstruction, and a 3D approach, in addition to the inclusion of sliding semilandmarks, have considerably improved the estimation of missing data (Slice, 2005, 2007; Benazzi et al., 2009; Gunz et al., 2009; Mitteroecker and Gunz, 2009; O'Higgins et al., 2011; Gunz and Mitteroecker, 2013; García-Martínez et al., 2014; Brassey et al., 2018; Schlager et al., 2018). In turn, these methods have been the basis for virtual reconstructions of different trunk fossil elements (Berge and Goularas, 2010; García-Martínez et al., 2014, 2018d; Claxton et al., 2016; Brassey et al., 2018; Rmoutilová et al., 2019).

Among the missing data estimation techniques are those using multiple multivariate regressions for statistical reconstruction based on a reference sample (Bookstein et al., 2003; Gunz et al., 2004; Weber and Bookstein, 2011; Stelzer et al., 2018). In this approach, multiple variables (e.g., a set of 3D coordinates on an anatomical structure) are regressed on all other variables in a reference sample of complete specimens, and missing values (3D coordinates) are predicted by the generated linear regression model. Gunz et al. (2004) presented this method for reconstructing incomplete human crania, and demonstrated that regression-based reconstruction was more accurate than the thin plate spline (TPS) warping and mean substitution methods. This technique has been recently employed for fossil reconstruction by Stelzer et al. (2018), who demonstrated that dental arcades of extinct hominins can be reliably predicted using the covariation between upper jaws and lower jaws of a reference sample composed of extant hominoids with previously demonstrated known associations (Stelzer et al., 2017).

Two-block partial least squares (2B-PLS) regression method has also shown great potential for predicting missing elements. This technique was firstly applied to investigate covariation in shape data by Rohlf and Corti (2000), and subsequent studies have used this method to statistically assess the covariation between two or more different sets of shape variables (Bookstein et al., 2003; Bastir, et al., 2005; Mitteroecker and Bookstein, 2007; Mitteroecker et al., 2012; Klingenberg and Marugán-Lobón, 2013; Adams and Collyer, 2016; Arlegi et al., 2018; Neaux et al., 2018; Scott et al., 2018; Torres-Tamayo et al., 2018). Gunz et al. (2009) referred to the 2B-PLS analysis as a method to reduce high-dimensionality of data in regression-based predictive analysis while exploiting the morphological integration between the known and missing parts. Furthermore, it has also shown great potential for predicting missing elements: one block can be predicted from the other one based on the strength of covariation between blocks in a reference sample (Schlager, 2013; Archer et al., 2018; Bastir et al., 2019b; Torres-Tamayo et al., 2019). By studying stone tools, Archer et al. (2018) demonstrated that a flake body could be accurately predicted from the platform body based on the covariation between these two structures. In testing whether this predictive analysis could be applied to an organism, Bastir et al. (2019b) successfully predicted anatomically connected lumbar spines from isolated lumbar vertebrae in modern humans, thus validating this method for the first time in an anatomical system.

1.2. Kebara 2 thorax reconstructions

One of the best fossil examples to which classic and virtual techniques have been applied is the Kebara 2 specimen (*Homo neanderthalensis*, ~60 ka; Valladas and Valladas, 1991). This well-preserved young adult, presumed to be an adult male, was found in Kebara Cave (Israel) in 1983 and has been used as a key specimen in describing Neandertal postcranial anatomy (Rak and Arensburg, 1987; Vandermeersch, 1991; Been et al., 2010; Gómez-Olivencia et al., 2009, 2013, 2017, 2018; García-Martínez et al., 2014; Chapman et al., 2017). The thorax, which is usually badly preserved mainly due to the fragility of the costal skeletal elements, has been the focus of several works that characterized isolated thoracic elements of Kebara 2 and helped to lay the foundation to describe the thoracic morphology of this specimen (Arensburg, 1991; Gómez-Olivencia et al., 2009; García-Martínez et al., 2014, 2018a, b, c; Been et al., 2017; Chapman et al., 2017).

The first approach to a Neandertal skeleton reconstruction was performed by Sawyer and Maley (2005), who based their reconstruction of La Ferrassie 1 on recovered material from other Neandertal skeletons. Among these specimens, these authors used the thoracic vertebrae, ribs and sternum of Kebara 2 Neandertal to reconstruct the thorax, using clay and epoxy paste to fill missing areas. This thorax exhibits a dome-shaped upper part and a markedly flaring lower part which, together with notable rib declination, results in an anteroposterior flattened bell-shaped reconstruction. According to Sawyer and Maley (2005: 30), "Although the ribcage and pelvis are visually compelling and convincing in demonstrating the relative difference in Neanderthals and modern humans, the introduction of some degree of artistic license makes it difficult to comment on the significance of these differences". As part of this artistic license, the authors altered the first and second rib lengths of their thorax reconstruction to better match the somewhat larger La Ferrassie 1 shoulder girdle. Even so, the bell-shaped thorax reconstruction performed by Sawyer and Maley (2005) has been considered as the only approach to Kebara 2 thorax morphology until recently.

Gómez-Olivencia et al. (2018) combined published data as well as anatomical expertise to reconstruct the thorax of Kebara 2 using 3D virtual techniques. These authors first reconstructed the thoracic vertebral column of this specimen by slightly modifying a previous reconstruction undertaken by Been et al. (2017), then correcting taphonomic and reconstruction deformation of the isolated ribs and using mirror images when necessary, and virtually articulating the ribs and sternal elements. This reconstruction was analyzed using both traditional morphometrics and 3DGM within a comparative framework composed of 16 thoraces of adult male *H. sapiens* individuals. These authors found that Kebara 2 showed a wider lower thorax with a pronounced invagination of the thoracic spine and a more horizontal orientation of the ribs compared to modern human males of similar stature.

The reconstructions produced by Sawyer and Maley (2005) and Gómez-Olivencia et al. (2018) reinforce the widely accepted morphology of the Neandertal thorax: this species showed a large costal skeleton, especially in the mid-thoracic ribs, and more dorsally oriented transverse processes in the mid-thoracic spine that leads to a wider lower thorax in the mid-lower segment (Franciscus and Churchill, 2002; Sawyer and Maley, 2005; Gómez-Olivencia et al., 2009, 2018, 2019; García-Martínez et al., 2014, 2017, 2018a, b, c; Bastir et al., 2015, 2017; Gómez-Olivencia et al., 2018, 2019). However, achieving these reconstructions has proved difficult as the thorax is a complex structure composed of different elements that need to be articulated, and this requires vast anatomical expertise. In vivo, these structures are anatomically connected through the costovertebral and the costochondral joints (Graeber and Nazim, 2007; Beyer et al., 2014), so thorax morphology relies on anatomical relationships between these elements. However, these cartilaginous joints are generally not preserved in the fossil record, and thus any thorax reconstruction relies on how ribs are articulated to their corresponding thoracic vertebrae. As a consequence, there could be more than one way to align the thoracic elements depending on each researcher's criteria, which could affect rib orientation and declination, so modifying the anteroposterior, mediolateral and craniocaudal diameters of the articulated thorax. These different researcher's assumptions might also be based on different interpretations of the body shapes of extinct hominins, which is then reflected in their reconstructions (Schmid, 1983; Bonmatí et al., 2010; Arsuaga et al., 2015; Latimer et al., 2016; Gómez-Olivencia et al., 2018; Haeusler et al., 2019). This reinforces the necessity of developing new and complementary approaches for such reconstructions.

1.3. Aims of this work

This work aims to test whether thorax morphologies can be predicted from pelvic morphologies using 3DGM techniques and 2B-PLS analysis working as a predictive method (Schlager, 2013; Archer et al., 2018; Bastir et al., 2019b; Torres-Tamayo et al., 2019). This method was first tested on a comparative sample composed of living humans with known associations between thorax and pelvis shape and then applied to Kebara 2 Neandertal specimen, thoracic and pelvic morphologies of which are well-known thanks to previous reconstructions of these two anatomical structures (Rak and Arensburg, 1987; Sawyer and Maley, 2005; Gómez-Olivencia et al., 2018). We used two different pelvic predictors and three different thoraco-pelvic covariation models based on an adult living *H. sapiens* reference sample. The resulting six thorax predictions were compared with each other and with the two thorax reconstructions of Kebara 2 Neandertal previously published (Sawyer and Maley, 2005; Gómez-Olivencia et al., 2018). Lastly, the limitations of this predictive method are discussed.

2. Materials and methods

2.1. Sample

A total of 60 thoracoabdominopelvic CT scans of adult living *H. sapiens* individuals (sample composition: n = 31 Spanish, n = 29 South-African; see details in Supplementary Online Material [SOM] Table S1) were segmented using the open-source software 3D Slicer v. 4.8 (Kikinis et al., 2014), which applied the marching cubes algorithm to render 3D triangular meshes (Lorensen and Cline, 1987). The use of the Spanish human sample for research purposes was approved in the context of the mutual scientific collaboration agreement between ASCIRES ERESA (Exploraciones Radiológicas Especiales Sociedad Anónima) and the Universitat de València, and is in accordance with the Declaration of Helsinki (Goodyear et al., 2007). The use of the South-African human sample was approved by the Human Research Ethics Committee Medical Clearance Certificate NO. M130844. The resulting 3D virtual torso models were postprocessed (i.e., hole filling, surface smoothing, mesh simplification) in the open source software Meshlab (Cignoni et al., 2008) to obtain optimal models for data analysis.

3D virtual models of fossil reconstructions used in this study (Figs. 1 and 2) include: (1) the cast of Kebara 2 thorax reconstruction of Sawyer and Maley (2005; see Fig. 1a); (2) Kebara 2 thorax reconstruction published by Gómez-Olivencia et al. (2018), available from Figshare (https://doi.org/10.6084/m9.figshare.7012256; see Fig. 1b); (3) the cast of Kebara 2 pelvis reconstruction made by Rak and Arensburg (1987; Fig. 2a), scanned at the Anthropology Department of the University of Delaware (Newark, USA) using a Creaform Go!SCAN scanner; and (4) the cast of the Neandertal pelvis reconstruction made by Sawyer and Maley (2005; Fig. 2b), scanned at the Departamento de Paleobiología (Museo Nacional de Ciencias Naturales, Madrid, Spain) using an Artec Spider 3D scanner.

2.2. Data measurement and intraobserver measurement error

The 3D models were measured in license-free version of Viewbox 4.0 software (Halazonetis, 2018) by the same researcher to avoid interobserver measurement error. The protocol of digitization was composed of 134 fixed landmarks, 720 curve semilandmarks and 160 surface semi-landmarks (Fig. 3; Table 1). Intraobserver measurement error was assessed by measuring a random specimen three times on three different days by the same researcher. Then, Procrustes distances (i.e., Euclidean distances between two configurations of Procrustes coordinates, used as a metric of shape difference following Mitteroecker and Gunz, 2009) between sets of Procrustes shape coordinates were calculated between ten random specimens and these three repetitions. For an acceptable intraobserver error, we expected that the highest Procrustes distance between two repetitions would be lower than the lowest Procrustes distance between two random specimens (SOM Table S2; Morecroft et al., 2010).

Later procedures were performed in R software v. 3.6.3 (R Development Core Team, 2017) mostly using the 3DGM R-package Morpho v. 2.6 (Schlager, 2017). All of the analyzed data and code are available in the Open Science Framework (OSF) repository (https://osf.io/efqjt/?view_only=2f2d3c9d59814b63a2d1084e0e528fa6). Since a random specimen was used as a template to digitize the sample, and to avoid that any feature inherent to this individual was transferred to the rest of the specimens, the torso configurations were reslid against a previously calculated torso mean configuration. This iterative process minimized the bending energy of each specimen with respect to this mean, reducing the uncertainty of the semilandmarks location (Gunz et al., 2005; Gunz and Mitteroecker, 2013).

2.3. Partial least squares regression

We subdivided the dataset into the thorax (block 1, 677 landmarks and semilandmarks) and the pelvis (block 2, 337 landmarks and semilandmarks; Fig. 3) and performed a full generalized Procrustes analysis (GPA) of the raw coordinates of each block separately to remove any variation related to a different position, size and orientation of the configurations (Gower, 1975; O'Higgins, 2000). Then, we used the 2B-PLS analysis of Procrustes shape coordinates to assess thorax-pelvis covariation (Rohlf and Corti, 2000; Bookstein et al., 2003; Torres-Tamayo et al., 2018) as follows.

2B-PLS analysis computes the variance-covariance matrix that comprises the within-block variance-covariance matrices of the thorax (x, y, z) and the pelvis (x', y', z'), and the covariance matrix between the two blocks (Zelditch et al., 2012). Based on the singular value decomposition (SVD) of the corresponding subset of the joint covariance matrix, 2B-PLS analysis projects the data into a latent space (generating latent variables [LV]) that models the covariance between the two blocks. Analogous to a principal component analysis (PCA) that computes the axes of the main variation, this can be interpreted as the space modeling the joint covariance between two sets of variables, with the dimensions in the latent space explaining the overall covariance. They can also be sorted according to their relevance in explaining the covariance, which is encoded in the respective singular value in the aforementioned SVD. After projection into that latent space, the linear correlation between those newly obtained pairs of (latent) variables (one for each set or block of shape variables in each latent dimension) can be assessed in n different dimensions. For each dimension, corresponding p-values can be computed for testing against the null hypothesis of complete independence between the two blocks.

Working as a predictive method, the 2B-PLS analysis decomposes that part of the common covariance matrix that encodes the covariation between those blocks of variables to obtain the basis vectors for each set that are spanning that latent space (or feature space). After projecting the predictor and response into that feature space, the resulting latent variables can be regressed onto each other and the result can then be rotated back into the original space to obtain the resulting shape. That ensures that those linear combinations can be identified that are known to maximize the linear association between both sets of variables.

When dealing with high dimensional predictors where the amount of variables is close to or exceeds sample size (as is the case here), it is necessary to reduce the number of variables to avoid overfitting (Gunz et al., 2009). Overfitting means that the resulting regression model will work almost perfectly on the training data but very poorly when applied to predict data that was not used for training that specific model. To reduce the number of variables, one could simply perform a PCA on both sets of variables to extract those linear combinations contributing most to the total predictor's variance, thus restricting the predictor to a subspace only composed of the first few principal components (PCs) that are chosen based on a percentage criterion. This approach is normally used when multiple multivariate regression is the preferred method to predict missing/partial structures (Stelzer et al., 2018). Thus, the matrix of scores of the subset of PCs computed from the response is regressed on the matrix of scores of the subset of PCs computed from the predictor in a multiple multivariate regression model (Stelzer et al., 2018). However, by choosing only the first few principal components based on a percentage criterion one might miss shape variations that are relatively unimportant regarding the overall variability of the predictor but which could strongly covary with the response (Schlager, 2013). As we aim at taking into account those linear combinations maximizing covariation between predictor and response, the partial least squares regression (PLSR; Martens and Naes, 1992; Wold et al., 2001; Abdi, 2007), and more specifically the 2B-PLS analysis (Rohlf and Corti, 2000) is preferred here over the multiple multivariate regression of PC scores. The validity of the usage of partial least-squares regression (PLSR) for predicting missing/partial shapes has been previously shown (Schlager, 2013; Archer et al., 2018; Bastir et al., 2019b; Torres-Tamayo et al., 2019).

To avoid overfitting in the 2B-PLS analysis, only the LVs yielding the minimum prediction error (mean square error of the Procrustes distance between each original specimen and its prediction) were used as projection matrices (Schlager, 2013; Archer et al., 2018). To find the number of LV that yields the minimum prediction error, we performed leave-one-out cross-validation analyses (LOOCV), i.e., a separate 2B-PLS analysis was carried out for each prediction excluding that particular individual for which the prediction was being calculated. Thus, prediction of specimen 1 was calculated from its predictor but excluding that specimen from the predictive model to avoid self-inference, and the same was done for the rest of the specimens. Hence, individual specimens did not unduly influence the training sample used to develop the predictive models.

2.4. Validation of human predictive models

A reference sample composed of adult living *H. sapiens* individuals was used to predict the thorax morphology of Kebara 2. While it is true that the choice of training data results in a taxon-specific bias, using a model generated from modern human data likely provides suitable information given the overall genetically, temporal and morphological similarity of *H. sapiens* and *H. neanderthalensis*. Firstly, we analyzed the reference sample to demonstrate that this predictive approach yields reliable results within modern human samples with known associations between thorax and pelvis shape prior to moving on to the Kebara 2 predictions. For this purpose, we calculated three predictive models as follows:

1. Model A: 2B-PLS analysis on $n = 27$ male *H. sapiens* individuals. This model could simulate the scenario when we know the genus and the sex of the fossil (Kebara 2

Neandertal is a presumed male specimen) and we have a reference sample composed of specimens of the same genus and sex as those of the fossil.

2. Model B: 2B-PLS analysis on n = 60 *H. sapiens* individuals (n = 27 males and n = 33 females). This model could simulate the scenario when we know that the fossil belongs to genus *Homo* but we do not know its sex so we have to use a reference sample composed of specimens of the same genus and both sexes.
3. Model C: 2B-PLS analysis on n = 33 female *H. sapiens* individuals. This model simulates the scenario when we know the genus and the sex of the fossil and we only have a reference sample composed of specimens of the same genus as that of the fossil but of the opposite sex. As Kebara 2 Neandertal is a presumed male specimen, a female-only predictive model was calculated as an exclusion model to avoid male-only model bias.

Once the most suitable LVs number for predictions was identified in each predictive model via LOOCV, predictive power of each model was first validated within each reference sample as a precursor to the analyses that include the Kebara 2 specimen. This validation was performed in two steps: first, thorax morphology was predicted from pelvis morphologies for each specimen using the three models, and Procrustes distances between actual thoraces and predicted thoraces were computed for each model; second, Procrustes distances between every possible pair within each reference sample were also calculated to investigate the actual intraspecific variation if each sample. We used Shapiro-Wilk tests to check for normality (SOM Table S3) and Mann-Whitney U tests to compare these Procrustes distances within each model. Reliable predictive models will show Procrustes distances between actual and predicted thoraces below the actual variation within the corresponding reference sample, and the best predictive model will show the lowest Procrustes distance between actual and predicted thoraces when the three models are compared (Stelzer et al., 2018). The files containing these Procrustes distances can be found in OSF (https://osf.io/efqjt/?view_only=2f2d3c9d59814b63a2d1084e0e528fa6).

Additionally, we computed the mean Procrustes distance and the 95% CI between every possible pair within the complete *H. sapiens* sample (*n* = 60) for the thorax and the pelvis to have a quantitative reference of the intraspecific morphological variation of our entire sample. Procrustes distance between the two thorax reconstructions made by Sawyer and Maley (2005) and by Gómez-Olivencia et al. (2018), and Procrustes distance between the two pelvis reconstructions made Rak and Arensburg (1987) and Sawyer and Maley (2005) were calculated to quantify how different/similar these reconstructions are in relation to the intraspecific human variation. The files containing these Procrustes distances can be found in OSF (https://osf.io/efqjt/?view_only=2f2d3c9d59814b63a2d1084e0e528fa6).

2.5. Kebara 2 thorax predictions

Kebara 2 thorax predictions were calculated based on Model A, Model B and Model C using the pelvis reconstructions made by Rak and Arensburg (1987) and by Sawyer and Maley (2005) as predictors (Fig. 2). Hereafter we will refer to thorax predictions based on Model A as A/RA (using the pelvis reconstruction of Rak and Arensburg, 1987 as a predictor) and A/SM (using the pelvis reconstruction of Sawyer and Maley, 2005 as a predictor) and the same applied to all predictions: B/RA, B/SM, C/RA and C/SM. 3D models of these thorax predictions were deposited in OSF (https://osf.io/efqjt/?view_only=2f2d3c9d59814b63a2d1084e0e528fa6).

These thorax predictions were compared with thorax reconstructions previously published for Kebara 2 (Sawyer and Maley, 2005; Gómez-Olivencia et al., 2018; Fig. 1):

for visualization of morphological similarities and differences, we superimposed these reconstructions and predictions via GPA (Gower, 1975; O'Higgins, 2000); for quantification, Procrustes distances in full shape space between these configurations were calculated.

2.6. Data sharing statement

The data supporting the findings of this study are available in the SOM and in OSF. Because of ethical and legal reasons, the CT scans and 3D models of the human sample used in this study must remain confidential and cannot be shared. The Spanish CT data are deposited in GIAVAL Research Group, Department of Anatomy and Human Embryology, University of Valencia, Av. Blasco Ibanez, 15. E-46010, Valencia, Spain. The South-African CT data are deposited in the Department of Human Anatomy and Physiology, Faculty of Health Sciences, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Gauteng, South Africa, Johannesburg, South Africa For requesting access to these data, please, contact J.A.S.-G. and S.N., respectively.

3. Results

3.1. Human predictive models

Table 2 shows mean Procrustes distances between actual and predicted thoraces for each model, as well as mean Procrustes distances between every possible pair within each reference sample. Further details of each predictive model are described below.

<u>Model A</u> Table 3 shows the covariance explained by the first five LVs (more than 85% of the total covariance) of the 2B-PLS analysis in *n* = 27 male *H. sapiens* individuals. This covariance exceeded what one might expect by chance in the first four LVs. Among these LVs, LOOCV analysis yielded that three LVs were optimal for thorax prediction in Model

A (Fig. 4a). Mean Procrustes distance between actual and predicted thoraces (mean Procrustes distance = 0.0824; Table 2) is significantly smaller than the actual variation showed by the *n* = 27 *H. sapiens* males (mean Procrustes distance = 0.1138; U = 1506; $p < 0.05$; Table 2; Fig. 4b). Almost 50% of covariance is explained by the first LV (Table 3): males showing wide lower thoraces with more horizontal ribs also have relative wide upper pelves (negative PLS scores), and males that have narrow lower thoraces with more declined ribs also show narrow upper pelvis (positive PLS scores; Fig. 5a; SOM Fig. S1).

<u>Model B</u> Table 4 shows the covariance explained by the first five LVs (more than 85% of the total covariance) of the 2B-PLS analysis of *n* = 60 *H. sapiens*. As only the first two LVs showed covariance that exceeded what is expected by chance, LOOCV analysis showed that two LVs were optimal for thorax predictions based on Model B (Fig. 4c). Mean Procrustes distance between actual and predicted thoraces (mean Procrustes distance = 0.1021; Table 2) is significantly smaller than that shown by the actual variation of *n* = 60 *H. sapiens* (mean Procrustes distance = 0.1259; U = 30430; $p < 0.05$; Table 2; Fig. 4d). Singular values of the first LV explain more than 60% of the total covariance (Table 4), and shape changes associated with PLS axes in the first LV (Fig. 5b; SOM Fig. S2) showed thoracopelvic covariation patterns driven by sexual dimorphism: males (negative PLS scores) have relatively wider lower thoraces than their narrow upper pelves and females show the opposite trend (positive PLS scores).

<u>Model C</u> Table 5 shows that the first five LVs explain more than 80% of the total covariance of the 2B-PLS analysis in *n* = 33 female *H. sapiens* individuals, although this covariance is nonsignificant in the first three LVs (almost 70% of the total covariance). Accordingly, LOOCV analysis shows a high prediction error in this model (Fig. 4e) but a relative minimum prediction error when five LVs are collected. Thus, we used five LVs for thorax predictions based on this model. Mean Procrustes distance between actual and

predicted thoraces (mean Procrustes distance = 0.1185; Table 2) is not significantly different from the actual variation showed by *n* = 33 female *H. sapiens* individuals (mean Procrustes distance = 0.1307; U = 7024; *p* = 0.07; Table 2; Fig. 4f). As singular values of the first LV explained covariance that was not significantly different from that expected by chance (Table 5), it was not possible to describe and interpret shape changes associated to the PLS axes of this model in biological terms.

Lastly, Procrustes distance calculations in full shape space within the entire *H. sapiens* reference sample (*n* = 60) yielded a mean Procrustes distance of 0.1258 with a 95% CI of 0.1242–0.1275 for the thorax and a mean Procrustes distance of 0.0840 with a 95% CI of 0.0832–0.0848 for the pelvis. Quantitative shape comparisons between the thorax reconstructions made by Sawyer and Maley (2005) and by Gómez Olivencia et al. (2018) by means of Procrustes distance in full shape space yielded a Procrustes distance of 0.1251 that is within the 95% CI of Procrustes distances between every possible pair within *n* = 60 *H. sapiens* thorax sample. In turn, quantitative shape comparisons between the pelvis reconstructions performed by Rak and Arensburg (1987) and by Sawyer and Maley (2005) yielded a Procrustes distance of 0.0602 that is below the mean and below the 95% CI of Procrustes distances between every possible pair within *n* = 60 *H. sapiens* pelvis sample.

3.3. Kebara 2 thorax predictions

<u>Thorax predictions based on Model A</u> Figure 6 shows Kebara 2 thorax predictions based on Model A, as well as the superimposition of these configurations with previous reconstructions. A/RA (Fig. 6a) and A/SM (Fig. 6d) show very similar thoracic morphologies: both predictions have a lower part of the thorax that is relatively wider than the upper part, indicating that the different predictors used in this study do not

substantially influence prediction results of this model. This was quantitatively confirmed by means of Procrustes distances: Procrustes distance between A/RA and A/SM (Procrustes distance = 0.0239) is considerably below the mean Procrustes distance and below the 95% CI of Procrustes distances between every possible pair within *n* = 60 *H. sapiens* thorax sample (mean Procrustes distance = 0.1258; 95% CI = 0.1242–0.1275). A/RA and A/SM present notable differences when compared to the reconstruction made by Sawyer and Maley (2005; Fig 6b, e). Although similar in upper thoracic widths, this reconstruction shows wider lower thoracic widths (ninth-eleventh ribs), more declined ribs and thus less anteroposterior depth than A/RA and A/SM. Conversely, A/RA and A/SM are very similar to the reconstruction produced by Gómez-Olivencia et al. (2018; see Fig. 6c, f) in rib declination, anteroposterior depth and in the relative width of the lower part of the thorax (sixth-eleventh ribs). Main differences are shown in the upper part of the thorax, with A/RA and A/SM showing relatively longer upper ribs (first pair) and thus a relatively wider upper thorax than the reconstruction made by Gómez-Olivencia et al. (2018). Table 6 shows that Procrustes distance comparisons in full shape space confirm morphological visualizations: A/RA and A/SM estimates are considerably more similar to the reconstruction made by Gómez-Olivencia et al. (2018) than to the one made by Sawyer and Maley (2005).

<u>Thorax predictions based on Model B</u> Figure 7 shows that B/RA (Fig. 7a) and B/SM (Fig. 7d) have very similar morphologies. Both predictions show a narrow thorax and marked rib declination, indicating that the different predictors used in this study do not substantially influence prediction results of this model. This was quantitatively confirmed by means of Procrustes distances: Procrustes distance between B/RA and B/SM (Procrustes distance = 0.0101) is considerably below the mean Procrustes distance and below the 95% CI of Procrustes distances between every possible pair within *n* = 60 *H.*

sapiens thorax sample. Contrary to the findings in Model A, B/RA and B/SM are more similar to Sawyer and Maley (2005) reconstruction (Fig. 7b, e) than to Gómez-Olivencia et al. (2018) reconstruction (Fig. 7c, f); this is mostly due to the marked rib declination and similar anteroposterior depths present in these predictions and also in Sawyer and Maley (2005) reconstruction. Also, B/RA and B/SM show similar relative upper thorax width as to that of Sawyer and Maley (2005) thorax reconstruction, although the lower thorax is notably wider in this reconstruction than in the predictions. When compared to Gómez-Olivencia et al. (2018) thorax reconstruction, differences are notable as well, as this reconstruction shows a wider lower part of the thorax than B/RA (Fig. 7c) and B/SM (Fig. 7f). In addition, these predictions also exhibit marked rib declination compared to Gómez-Olivencia et al. (2018) thorax reconstruction. Consequently, B/RA and B/SM show considerably shorter anteroposterior depth than this reconstruction. Table 6 shows that Procrustes distance comparisons in full shape space confirms morphological trends: B/RA and B/SM are slightly more similar to Sawyer and Maley (2005) reconstruction than to Gómez-Olivencia et al. (2018) reconstruction, although these predictions are overall fairly different from previous reconstructions.

<u>Thorax predictions based on Model C</u> Figure 8 shows that predictions based on Model C are not realistic in terms of anatomy. C/RA (Fig. 8a) and C/SM (Fig. 8d) have very similar morphologies with a narrow thorax that is artificially anteroposterior flattened because of the extremely declined ribs, indicating that the different predictors used in this study do not substantially influence predictions results of this model. This was quantitatively confirmed by means of Procrustes distances: Procrustes distance between C/RA and C/SM (Procrustes distance = 0.0329) is considerably below the mean Procrustes distance and below the 95% CI of Procrustes distances between every possible pair within $n = 60$ *H. sapiens* thorax sample. However, C/RA and C/SM are quite different from the

reconstruction made by Sawyer and Maley (2005) (Fig. 8b, e) and by Gómez-Olivencia et al. (2018) (Fig. 8c, f) both in rib declination and overall thorax width, as it is shown in the superimposition of these configurations. Table 6 shows the inconsistency of these predictions with previous reconstructions in terms of Procrustes distance in full shape space.

4. Discussion

In the present study, we aimed to apply the 2B-PLS analysis as a predictive method (Schlager, 2013; Archer et al., 2018; Bastir et al., 2019b; Torres-Tamayo et al., 2019) to test whether thorax morphologies can be predicted from pelvis morphologies within genus *Homo*. For this purpose, we first validated the method in three modern human reference samples following Stelzer et al. (2018) to then predict the thorax morphology of the well-known specimen Kebara 2 Neandertal. The purpose of this section is to discuss the implications of the method in terms of: (1) validity of human predictive models, (2) structures used as predictors, (3) choice of reference samples for predictive models, and (4) limitations of this technique.

4.1 Validity of human predictive models

Model A (Fig. 5a) shows that male *H. sapiens* displays morphological correspondence between lower thorax and upper pelves in our sample, with relatively wide lower thoraces and upper pelves that lead to relatively wide torsos in negative values and the opposite trend in positive values. This model shows the smallest prediction error and the smallest mean Procrustes distance between actual and predicted thoraces when the three predictive models are compared (Fig. 4g, h; Table 2) so it is the best-case scenario of this study as it yields the most reliable results within humans.

Model B (Fig. 5b) shows that *H. sapiens* displays a pattern of thoracopelvic covariation driven by sexual dimorphism, with a clear mean difference between females and males: for a given thorax shape there is a specific pelvis shape, irrespective of being female or male, but there are thoracic and pelvic shapes more likely to be found only in males or in females. Consequently, females show relatively narrower (lower) thoraces than their wide (upper) pelves and males show the opposite trend. Model B shows the second smallest prediction error and the second smallest mean Procrustes distance between actual and predicted thoraces when the three predictive models are compared (Fig. 4g, h; Table 2). This means that a model using a modern human reference sample composed of both males and females represents an intermediate-case scenario in this study.

In turn, the female-only model (Model C; Fig. 5c) represents the worst-case scenario in this study because of three reasons. First, ~70% of the covariance explained by this model is not significantly different from that expected by chance (Table 5). Second, prediction error is higher than that of Model A and Model B (Fig. 4g). And third, morphological differences between actual and predicted thorax configurations do not significantly differ from intraspecific differences in females, showing by far the lowest predictive power among the three models used in this study (Table 2; Fig. 4h). Weaker patterns of integration between the thorax and the pelvis in females when compared to males are not unexpected, as Middleton (2015) found that integration of the thorax with both the false pelvis and the lumbar region (not quantified in the present study) is stronger in males than in females. However, it is quite surprising that covariance showed by Model C is not significantly different from that expected by chance, which might have a biological cause, but more likely relates to sample size/composition. Further research using larger female *H. sapiens* samples to investigate potential underlying sample issues.

4.2. Predictors

The original pelvic remains recovered from Kebara Cave are composed of a well-preserved right innominate bone, a taphonomically distorted left innominate bone, and an almost complete and well-preserved sacrum (Rak and Arensburg, 1987). Unlike other Neandertal pelvic remains that need to be extensively reconstructed, e.g. Tabun C1 (Ponce de León et al., 2008; Weaver and Hublin, 2009), the well preserved nature of Kebara 2 allows for the retention of much of the anatomical and biological information of the original Neandertal individual who lived ~60 ka. This reduces the uncertainty usually associated with pelvis reconstruction, so that Kebara 2 is a suitable fossil specimen to apply this predictive method in a paleoanthropological context.

Among the Neandertal pelvis reconstructions published to date (Rak and Arensburg, 1987; Sawyer and Maley, 2005; Ponce de León et al., 2008; Weaver and Hublin, 2009), there have been two approaches to the Kebara 2 pelvis. Rak and Arensburg (1987; Fig. 2a) mirrored the right hipbone to obtain a virtually complete pelvis of this specimen, while Sawyer and Maley (2005; Fig. 2b) put together pelvic fragments of three different Neandertal individuals: the right ilium, sacrum and left ischium belong to Kebara 2, most of the left ilium and left ischium are from Feldhofer 1 and the pubic bones are from La Ferrassie 1. Although these two reconstructions are not identical (Fig. 2c), they present many similarities: pelvic inlet, anterosuperior iliac spine orientation and sacrum position are overall the same, with some differences in the inferior pubic ramus and in the orientation of the posterior part of the ilium. These morphological similarities are confirmed by the Procrustes distance between these reconstructions (Procrustes distance = 0.0602), which is below the mean and below the 95% confidence interval of the Procrustes distance between every possible pair within n = 60 *H. sapiens* pelvis sample.

Here we used these two pelvis reconstructions as predictors to show how the resulting predictions can vary depending on the predictor used.

It is tempting to assume that the reconstruction made by Rak and Arensburg (1987) should be a 'better' predictor than the reconstruction made by Sawyer and Maley (2005) as the latter is a composite of fragments from three different Neandertals. Although that is probably true, in this particular study we are not in a position to say that these predictors are 'good' or 'bad' because to make such a statement we would require comparisons of our predictions with the nonexistent original thorax of Kebara 2. Based on our results, we could say that one predictor yields morphologies 'more consistent' or 'less consistent' with that accepted or expected for a specific taxon (e.g., for Neandertals) than the other predictor according to previous studies (Arensburg, 1991; Franciscus and Churchill, 2002; Gómez-Olivencia, 2009, 2017, 2018; García-Martínez et al., 2014, 2018a, b, c; Chapman et al., 2017). However, when resulting predictions using each predictor are compared within each model (e.g. A/RA vs. A/SM) they do not differ substantially, as Procrustes distances are considerably below the mean Procrustes distance and the 95% CI representing thorax variation within our sample. This means that our study does not show a clear signal of one predictor yielding more consistent morphologies with previous reconstructions than the other predictor, as results using different predictors show very similar Procrustes distance with both reconstructions (Table 6). This also implies that these pelvis reconstructions made by previous authors are consistent with each other, as supported by the Procrustes distance between them.

The pelvis is composed of three bones that need to be articulated. Even in the case that a fossil pelvis was found complete and undistorted, this articulation process may entail an element of subjectivity, as demonstrated by previous authors on several pelvis reconstructions of *Australopithecus afarensis* (Lovejoy, 1979; Schmid, 1983; Häusler and

Schmid, 1995). Because of this subjectivity, it would be interesting to use further reconstructions of Kebara 2 pelvis to keep track of the uncertainties due to the predictor utilized, as when using different reference specimens for TPS-based estimation (Gunz et al., 2009). The uncertainty resulting from the predictor's reconstruction may be reduced only in the unlikely case of having a complete and undistorted fossil predictor not requiring articulation. An example of predictors that do not need articulation was shown by Stelzer et al. (2018), who found success in predicting upper dental arcades from lower dental arcades and vice versa in extant hominoids and extinct hominins. Other than potential missing parts and/or taphonomic distortions, dental arcades do not need to be articulated, so uncertainty due to this process is reduced compared to structures that do need to be assembled. In summary, our study indicates that the main morphological differences between predictions are due to different predictive models rather than distinct predictors as we detail below.

4.3. Predictive models

One of the most difficult decisions in reference-based statistical reconstructions is the choice of the reference sample to build predictive models, as the statistical reconstruction is computed using the sample covariance matrix and thus relies on the covariation among the present coordinates (Gunz et al., 2009). Stelzer et al. (2018) used multivariate regression models in a predictive way and demonstrated that a model using individuals of the same genus or species as the target specimen yielded the most reliable results (Stelzer et al., 2018). However, the vast majority of works on primate covariation patterns are carried out on the cranium (Bookstein et al., 2003; Bastir and Rosas, 2006, Bastir et al., 2005, Mitteroecker and Bookstein, 2007; Bruner et al., 2017; Profico et al., 2017; Neaux et al., 2018; Stelzer et al., 2018). As a consequence, little is known about thoracopelvic covariation in extant hominoids, which could potentially serve as reference sample.

Middleton (2015) analyzed integration patterns within and between individual trunk elements in *H. sapiens* and *Pan troglodytes*, with the finding that these taxa share overall similar patterns of integration but different magnitude. In the same line, Torres-Tamayo et al. (in press) applied 3DGM to 3D torso models to assess and compare patterns of thoracopelvic covariation in *H. sapiens* and *P. troglodytes*. They found that these taxa share some aspects of the thoracopelvic covariation, supporting the findings of Middleton (2015). However, Torres-Tamayo et al. (in press) also suggested that other aspects of this covariation, especially those related to sexual dimorphism and allometry, might be species-specific, i.e., might be present in *H. sapiens* but not in *P. troglodytes* and vice versa. If these two taxa share some aspects of thoracopelvic covariation, these common features might have been maintained since the last common ancestor of *Pan-Homo*, i.e., in the human lineage. In line with this suggestion, *H. sapiens* and *H. neanderthalensis* might have shown covariation commonalities, but also their own species-specific thoracopelvic covariation features related to their own paleobiology and/or bioenergetic demands that might be different between these two taxa (Ruff et al., 1997; Arsuaga et al., 1999, 2015; Franciscus and Churchill, 2002; Carretero et al., 2004; Churchill, 2006, 2014; Simpson et al., 2008; Froehle and Churchill, 2009; Ruff, 2010; Holliday, 2012; García-Martínez et al., 2018a, b, c; Gómez-Olivencia et al., 2018; Bastir et al., in press). However, Neandertal fossils recovered to date do not suffice neither to confirm nor to refute the existence of potential shared aspects of the thoracopelvic covariation between *H. sapiens* and this taxon. Considering this lack of fossil evidence to properly assess the phenotypic thoracopelvic covariation in Neandertals, our best bet here is using extant humans as reference to model the thoracopelvic covariation of this closely related extinct taxon.

Model A yields Kebara 2 thorax predictions with the lower part being wider than the upper part (Fig. 6). In addition, this model also yields predictions with less rib declination and great anteroposterior depths. Even though Sawyer and Maley (2005) reconstruction also shows a wider lower thoracic part than the upper part (bell-shaped thorax), this reconstruction is characterized by marked rib declination and flattened thorax. Consequently, predictions based on Model A are not fully consistent with Sawyer and Maley (2005) thorax reconstruction (Fig. 6b, e; Table 6). On the contrary, because the reconstruction made by Gómez-Olivencia et al. (2018) possesses a mediolaterally expanded lower part, less rib declination and great anteroposterior depth, predictions based on Model A are largely consistent with this reconstruction (Fig. 6c, f; Table 6).

Predictions based on Model B are relatively narrow in both the upper and the lower part of the thorax, so the relative lower widths of previous reconstructions considerably exceeds those of predictions based on Model B (Fig. 7). These predictions also show marked rib declination and relatively shorter anteroposterior depths, so they are slightly more similar to Sawyer and Maley (2005; Fig. 7b, e) reconstruction than to Gómez-Olivencia et al. (2018; Fig. 7c, f) reconstruction in these features (Table 6). However, Model B is discarded as a valid model for Neandertals because it yields narrow thoracic configurations which in principle would not be consistent with the thoracic morphology expected for this taxon.

Model C does not show significant covariation patterns within our female *H. sapiens* sample (Fig. 5c) and it is highly unreliable for making any kind of prediction. Consequently, thorax predictions based on this model do not have anatomical realism (Fig. 8), so Model C is rejected as a valid model to predict Neandertal thorax morphologies. As the target fossil specimen of this study is assumed to be male, this constraint should not affect our study as it is designed. However, we must emphasize the clear limitations of this method had our purpose been predicting the thorax morphology of a Neandertal female, e.g., Tabun C1 (McCown and Keith, 1939; Ponce de León et al., 2008; Weaver and Hublin, 2009). Therefore, larger female *H. sapiens* samples are necessary before predictive attempts are made for Neandertal females.

This study is based on predictive models that were validated exclusively within modern humans, and the predictions carried out are premised on the assumption that Neandertals had the same thoracopelvic covariation pattern as modern humans. This might not be true. Neandertal thoracic and pelvic morphologies overall represent the plesiomorphic condition for the genus *Homo*, with modern humans being derived in many aspects, such as mediolaterally narrower thorax and pelvis morphologies, shortened dorsoventral thoracic depth, and shortened and thicker pubic bones (Arsuaga et al., 1999, 2015; Carretero et al., 2004; Rosenberg, 2007; Gómez-Olivencia et al., 2009, 2018; Bastir et al., 2017, in press). Such derived features result in morphological differences between modern humans and Neandertals that could lead to different patterns of thorax-pelvis covariation between these two taxa. Although Model A shows predictions largely consistent with an independent thorax reconstruction of Kebara 2 Neandertal made by Gómez-Olivencia et al. (2018) using a different methodological approach, we must clearly state that only if modern humans have thoracopelvic covariation patterns like those of Neandertals, then these findings would be confirmable. In addition, the different predictive models utilized here mainly account for potential influences of sexual dimorphism, not for other sources of variation that might influence modern humans and Neandertals differently, such as body mass, thermoregulation, intra- and interpopulation variation, etc. While the sample is composed of individuals from both Europe and South

Africa, and thus reflects aspects of thoracopelvic covariation of two modern human populations, this does not suffice to model modern human variation in thorax-pelvis covariation. Thus, an ideal scenario would include a large and diverse human sample populations showing a wide range of morphological thoracopelvic covariation.

Lastly, one might consider the inclusion of more extant hominoids to model thorax-pelvis covariation. This might prove beneficial when dealing with more plesiomorphic taxa such as *Australopithecus*, since australopiths were habitual bipedal hominins retaining plesiomorphic thoracic features (Stern and Susman, 1983; Stern, 2000; Ward, 2002; Schmid et al., 2013; Arias-Martorell et al., 2015; Kappelman et al., 2016; Latimer et al., 2016). In such a case, a *Homo*-based reference sample might not be enough for modeling this thorax-pelvis covariation, and a reference sample composed of several hominoid species would provide a wider range of morphological covariation, which is needed before predictive attempts are made for *Australopithecus.* In our case, however, the disparate thoracic and pelvic shape of extant hominoids when compared to Neandertals would very likely result in obfuscation, rather than in improvement, of the resulting shape model.

4.4. Limitations of the method

As we have outlined above, it is important to bear in mind that the mathematical predictions mainly depend on two variable factors: (1) the structures used as predictors, since different pelvis reconstructions of the same specimen can lead to different thorax morphologies, and (2) the reference sample(s) used to build the predictive models, which is influenced by the sample size, composition, and the factors driving the main patterns of covariation. In this particular case, we used a modern human reference sample to predict the thoracic morphology of a Neandertal, assuming that both taxa show similar patterns of thorax-pelvis integration. A further limitation of this method is that the resulting mathematical 'scaffold' might not reflect important features present in the original fossil specimen, such as deviations of spinal curvatures, as is the case with the nonpathological thoracic scoliosis described for Kebara 2 (Been et al., 2017; Gómez-Olivencia et al., 2018). Likewise, anatomical details such as muscle attachment points or tuberosities could be underestimated using this method. Therefore, this method should be combined with other geometric and statistical methods to reconstruct individual fossil elements. In no case should the 2B-PLS prediction method replace previous techniques, but rather should complement them. And more importantly, the biological realism and consistency of the results should always prevail over reference-based reconstruction methods.

5. Conclusions

This work shows the potential of a 3DGM method to predict thorax morphologies from pelvis morphologies based on a statistical model that exploits the covariation between the thorax and the pelvis. Prediction results are less sensitive to the structures used as predictors than to the reference sample chosen to calculate the predictive models. Therefore, careful reflection about the choice of the reference sample is necessary before any mathematical predictions. In this study, the male-only predictive model (Model A) yields the most reliable predictions within modern humans. However, cross-validations in fossil specimens are not possible to perform because of the small sample sizes. For this reason, when this method is applied to Kebara 2, it is difficult to evaluate whether thorax predictions yielded by each predictive model are reliable or not, as there is no true Kebara 2 thorax to compare them to. Assuming that Neandertals showed thorax configurations with the lower part being mediolaterally wider than the upper part as previously proposed, our study suggests that only Model A yields thorax predictions that match such proposed morphology. When anteroposterior depths and rib declination are also considered, predictions based on Model A would more closely resemble the reconstruction made by Gómez-Olivencia et al. (2018) than the one made by Sawyer and Maley (2005).

The method proposed here aims to reduce this subjectivity by allowing the calculation of a mathematical prediction that works as a 'scaffold' to articulate the individual fossil elements. In this particular case, the 2B-PLS analysis exploits the covariation between the thorax and the pelvis to return a statistical 'scaffold' that works as a 3D virtual template upon which the thoracic vertebrae and the costal skeleton can be articulated. Therefore, although anatomical experience is still required to articulate the fossils to the resulting 'scaffold', this method aims to reduce the subjectivity involved in fossil articulation based on the researcher's preconceptions. Future studies will inquire the effect of pooling extant primates for generating an overarching shape model and its performance in estimating missing/partial structures in extinct species.

References

Abdi, H., 2003. Partial least squares (PLS) regression. In: Lewis-Beck M., Bryman, A., Futing, T. (Eds.), Encyclopedia of Social Sciences Research Methods. Thousand Oaks, Sage, pp. 792-795.

Adams, D.C., Collyer, M.L., 2016. On the comparison of the strength of morphological integration across morphometric datasets. Evolution 70, 2623-2631.

Archer, W., Pop, C.M., Rezek, Z., Schlager, S., Lin, S.C., Weiss, M., Dogandžić, T., Desta, D., McPherron, S.P., 2018. A geometric morphometric relationship predicts stone flake shape and size variability. Archaeological and Anthropological Sciences 10, 1991-2003.

Arensburg, B., 1991. The vertebral column, thoracic cage and hyoid bone. In: Bar-Yosef, O., Vandermeersch, B. (Eds.), Le Squelette Moustérien de Kebara 2. Éditions du CNRS, Paris, pp. 113–147

Arias-Martorell, J., Potau, J.M., Bello-Hellegouarch, G., Pérez-Pérez, A., 2015. Like father, like son: assessment of the morphological affinities of AL 288–1 (*A. afarensis*), Sts 7 (*A. africanus*) and Omo 119–73–2718 (*Australopithecus sp.*) through a three-dimensional shape analysis of the shoulder joint. PLoS One 10, e0117408.

Arlegi, M., Gómez-Robles, A., Gómez-Olivencia, A., 2018. Morphological integration in the gorilla, chimpanzee, and human neck. American Journal of Physical Anthropology 166, 408-416.

Arsuaga, J.L., Lorenzo, C., Carretero, J.M., Gracia, A., Martínez, I., García, N., Bermúdez de Castro, J.M., Carbonell, E., 1999. A complete human pelvis from the Middle Pleistocene of Spain. Nature 399, 255-258.

Arsuaga, J.L., Carretero, J.-M., Lorenzo, C., Gómez-Olivencia, A., Pablos, A., Rodríguez, L., García-González, R., Bonmatí, A., Quam, R.M., Pantoja-Pérez, A., Martínez, I., Aranburu, A., Gracia-Téllez, A., Poza-Rey, E., Sala, N., García, N., Alcázar de Velasco, A., Cuenca-Bescós, G., Bermúdez de Castro, J.M., Carbonell, E., 2015. Postcranial morphology of the middle Pleistocene humans from Sima de los Huesos, Spain. Proceedings of the National Academy of Sciences USA 112, 11524-11529.

Bastir, M., García-Martínez, D., Estalrrich, A., García-Tabernero, A., Huguet, R., Ríos, L., Barash, A., Recheis, W., Rasilla, M., Rosas, A., 2015. The relevance of the first ribs of the El Sidrón site (Asturias, Spain) for the understanding of the Neandertal thorax. Journal of Human Evolution 80, 64-73.

Bastir, M., García-Martínez, D., Ríos, L., Higuero, A., Barash, A., Martelli, S., García-Tabernero, A., Estalrrich, A., Huguet, R., de la Rasilla, M., Rosas, A., 2017. Three-dimensional morphometrics of thoracic vertebrae in Neandertals and the fossil evidence from El Sidrón (Asturias, Northern Spain). Journal of Human Evolution 108, 47-61.

Bastir, M., García-Martínez, D., Torres-Tamayo, N., Palancar, C.A., Fernández-Pérez, F.J., Riesco-López, A., Osborne-Márquez, P., Ávila, M., López-Gallo, P., 2019a. Workflows in a Virtual Morphology Lab: 3D scanning, measuring, and printing. Journal of Anthropological Sciences 97, 1-28.

Bastir, M., García-Martínez, D., Torres-Tamayo, N., Palancar, C.A., Beyer, B., Barash, A., Villa, C., Sanchis-Gimeno, J.A., Riesco-López, A., Torres-Sánchez, I., García-Río, F., Been, E, Gómez-Olivencia, A., Haeusler, M., Williams, S.A., Spoor, F. in press. Rib cage anatomy in *Homo erectus* suggests a recent evolutionary origin of modern human body shape. Nature Ecology & Evolution.

Bastir, M., Rosas, A., Sheets, H. D., 2005. The morphological integration of the hominoid skull: a partial least squares and PC analysis with implications for European Middle Pleistocene mandibular variation. In: Slice, D.E. (Ed.), Modern Morphometrics in Physical Anthropology. Springer, Boston, pp. 265-284.

Bastir, M., Rosas, A., 2006. Correlated variation between the lateral basicranium and the face: a geometric morphometric study in different human groups. Archives of Oral Biology 51, 814-824.

Been, E., Gómez-Olivencia, A., Kramer, P.A., Barash, A., 2017. 3D reconstruction of spinal posture of the Kebara 2 Neanderthal. In: Marom, A., Hovers, E. (Eds.), Human Paleontology and Prehistory: Contributions in Honor of Joel Rak. Springer, Cham, pp. 239-251.

Been, E., Peleg, S., Marom, A., Barash, A., 2010. Morphology and function of the lumbar spine of the Kebara 2 Neandertal. American Journal of Physical Anthropology 142, 549-557.

Benazzi, S., Stansfield, E., Milani, C., Gruppioni, G., 2009. Geometric morphometric methods for three-dimensional virtual reconstruction of a fragmented cranium: the case of Angelo Poliziano. International Journal of Legal Medicine 123, 333-344.

Berge, C., Goularas, D., 2010. A new reconstruction of Sts 14 pelvis (*Australopithecus africanus*) from computed tomography and three-dimensional modeling techniques. Journal of Human Evolution 58, 262-272.

Beyer, B., Sholukha, V., Dugailly, P.M., Rooze, M., Moiseev, F., Feipel, V., Jan, S.V.S., 2014. In vivo thorax 3D modelling from costovertebral joint complex kinematics. Clinical Biomechanics 29, 434-438.

Bonmatí, A., Gómez-Olivencia, A., Arsuaga, J.L., Carretero, J.M., Gracia, A., Martínez, I., Lorenzo, C., Bermúdez de Castro, J.M., Carbonell, E., 2010. Middle

Pleistocene lower back and pelvis from an aged human individual from the Sima de los Huesos site, Spain. Proceedings of the National Academy of Sciences USA 107, 18386-18391.

Bookstein, F.L., Gunz, P., Mitteroecker, P., Prossinger, H., Schæfer, K., Seidler, H., 2003. Cranial integration in *Homo*: singular warps analysis of the midsagittal plane in ontogeny and evolution. Journal of Human Evolution 44, 167-187.

Brassey, C.A., O'Mahoney, T.G., Chamberlain, A.T., Sellers, W.I., 2018. A volumetric technique for fossil body mass estimation applied to *Australopithecus afarensis*. Journal of Human Evolution 115, 47-64.

Bruner, E., Pereira-Pedro, A. S., Bastir, M., 2017. Patterns of morphological integration between parietal and temporal areas in the human skull. Journal of Morphology 278, 1312-1320.

Carretero, J.M., Arsuaga, J.-L., Martínez, I., Quam, R.M., Lorenzo, C., Gracia, A., Ortega, A.I., 2004. Los humanos de la Sima de los Huesos (Sierra de Atapuerca) y la evolución del cuerpo en el género *Homo*. In: Baquedano, E. (Ed.), Homenaje a Emiliano Aguirre. Museo Arqueológico Regional, Alcalá de Henares, pp. 120-136.

Chapman, T., Beyer, B., Sholukha, V., Semal, P., Feipel, V., Louryan, S., Jan, S.V.S., 2017. How different are the Kebara 2 ribs to modern humans? Journal of Anthropological Sciences 95, 1-20.

Churchill, S.E., 2006. Bioenergetic perspectives on Neanderthal thermoregulatory and activity budgets. In: Harvati, H., Harrison, T. (Eds.), Neanderthals revisited: new approaches and perspectives, Springer, Dordrecht, pp. 113-133.

Churchill, S.E., 2014. Thin on the ground: Neandertal biology, archeology, and ecology. John Wiley & Sons, New York.

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G., 2008. Meshlab: an open-source mesh processing tool. In: Scarano, V., De Chiara, R., Erra, U (Eds.), Eurographics Italian Chapter Conference (2008), pp. 129-136.

Claxton, A.G., Hammond, A.S., Romano, J., Oleinik, E., DeSilva, J.M., 2016. Virtual reconstruction of the *Australopithecus africanus* pelvis Sts 65 with implications for obstetrics and locomotion. Journal of Human Evolution 99, 10-24.

Franciscus, R.G., Churchill, S.E., 2002. The costal skeleton of Shanidar 3 and a reappraisal of Neandertal thoracic morphology. Journal of Human Evolution 42, 303-356.

Froehle, S.E., Churchill, S., 2009. Energetic competition between Neandertals and anatomically modern humans. PaleoAnthropology 2009, 96-116.

García-Martínez, D., Barash, A., Recheis, W., Utrilla, C., Torres-Sánchez, I., García-Río, F., Bastir, M., 2014. On the chest size of Kebara 2. Journal of Human Evolution 70, 69-72.

García-Martínez, D., Bastir, M., Huguet, R., Estalrrich, A., García-Tabernero, A., Ríos, L., Cunha, E., Rasilla, M., Rosas, A. 2017. The costal remains of the El Sidrón Neanderthal site (Asturias, northern Spain) and their importance for understanding Neanderthal thorax morphology. Journal of Human Evolution 111, 85-101.

García-Martínez, D., d'Angelo del Campo, M.D., González, A.M., Cambra-Moo, Ó., Barash, A., Bastir, M., 2018a. Reevaluation of 'endocostal ossifications' on the Kebara 2 Neanderthal ribs. Journal of Human Evolution 122, 33-37.

García-Martínez, D., Radovčić, D., Radovčić, J., Cofran, Z., Rosas, A., Bastir, M., 2018b. Over 100 years of Krapina: New insights into the Neanderthal thorax from the study of rib cross-sectional morphology. Journal of Human Evolution 122, 124-132.

García-Martínez, D., Torres-Tamayo, N., Torres-Sánchez, I., García-Río, F., Rosas, A., Bastir, M., 2018c. Ribcage measurements indicate greater lung capacity in Neanderthals and Lower Pleistocene hominins compared to modern humans. Communications Biology 1, 117.

García-Martínez, D., Riesco, A., Bastir, M., 2018d. Missing elements estimation in sequential anatomical structures: the case of the human thoracic vertebrae and its potential application to the fossil record. In: Rissech, C., Lloveras, Ll., Nadal, J., Fullola, J.M. (Eds.), Geometric Morphometrics: Trends in Biology, Paleobiology and Archaeology. Universitat de Barcelona, Barcelona, pp. 93-99.

Gómez-Olivencia, A., Eaves-Johnson, K.L., Franciscus, R.G., Carretero, J.M., Arsuaga, J.L., 2009. Kebara 2: new insights regarding the most complete Neandertal thorax. Journal of Human Evolution 57, 75-90.

Gómez-Olivencia, A., Been, E., Arsuaga, J.L., Stock, J.T., 2013. The Neandertal vertebral column 1: The cervical spine. Journal of Human Evolution 64, 608-630.

Gómez-Olivencia, A., Arlegi, M., Barash, A., Stock, J.T., Been, E., 2017. The Neandertal vertebral column 2: The lumbar spine. Journal of Human Evolution 106, 84-101.

Gómez-Olivencia, A., Barash, A., García-Martínez, D., Arlegi, M., Kramer, P., Bastir, M., Been, E., 2018. 3D virtual reconstruction of the Kebara 2 Neandertal thorax. Nature Communications 9, 4387.

Gómez-Olivencia, A., Holliday, T., Madelaine, S., Couture-Veschambre, C., Maureille, B., 2019. The costal skeleton of the Regourdou 1 Neandertal. Journal of Human Evolution 130, 151-171.

Gower, J.C., 1975. Generalized procrustes analysis. Psychometrika 40, 33-51.

Graeber, G.M., Nazim, M., 2007. The anatomy of the ribs and the sternum and their relationship to chest wall structure and function. Thoracic Surgery Clinics 17, 473-489.

Gunz, P., Mitteroecker, P., 2013. Semilandmarks: a method for quantifying curves and surfaces. Hystrix 24, 103-109.

Gunz, P., Mitteroecker, P., Bookstein, F.L., 2005. Semilandmarks in three dimensions. In: Slice, D.E. (Ed.), Modern Morphometrics in Physical Anthropology. Kluwer Academic, New York, pp. 73-98.

Gunz, P., Mitteroecker, P., Bookstein, F.L., Weber, G.W., 2004. Computer aided reconstruction of incomplete human crania using statistical and geometrical estimation methods. In: Campana, S., Scopigno, R., Carpentiero, G., Cirillo, M. (Eds.), Enter the Past: Computer Applications and Quantitative Methods in Archaeology. Archaeopress, Oxford, pp. 96–98.

Gunz, P., Mitteroecker, P., Neubauer, S., Weber, G.W., Bookstein, F.L., 2009. Principles for the virtual reconstruction of hominin crania. Journal of Human Evolution 57, 48-62.

Halazonetis, D., 2018. Viewbox 4 - Cephalometric Software. http://dhal.com/viewboxindex.htm.

Haeusler, M., Trinkaus, E., Fornai, C., Müller, J., Bonneau, N., Boeni, T., Frater, N., 2019. Morphology, pathology, and the vertebral posture of the La Chapelle-aux-Saints Neandertal. Proceedings of the National Academy of Sciences USA 116, 4923-4927.

Häusler, M., Schmid, P., 1995. Comparison of the pelves of Sts 14 and AL 288-1: implications for birth and sexual dimorphism in australopithecines. Journal of Human Evolution, 29, 363-383.

Holliday, T.W., 2012. Body size, body shape, and the circumscription of the genus *Homo*. Current Anthropology 53(S6), S330-S345.

Jellema, L.M., Latimer, B., Walker, A., 1993. The rib cage. In: Walker, A., Leakey, R. (Eds.), The Nariokotome *Homo erectus* Skeleton. Harvard University Press, Cambridge, pp. 294–325.

Jungers, W.L., Grabowski, M., Hatala, K.G., Richmond, B.G., 2016. The evolution of body size and shape in the human career. Philosophical Transactions of the Royal Society B 371, 20150247.

Kappelman, J., Ketcham, R.A., Pearce, S., Todd, L., Akins, W., Colbert, M.W., Faseha, M., Maisano, J.A., Witzel, A., 2016. Perimortem fractures in Lucy suggest mortality from fall out of tall tree. Nature 537, 503-507.

Kikinis, R., Pieper, S. D., Vosburgh, K. G., 2014. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support. In: Jolesz, F.A. (Ed.), Intraoperative imaging and image-guided therapy. Springer, New York, pp. 277-289

Klingenberg, C.P., 2016. Size, shape, and form: concepts of allometry in geometric morphometrics. Development Genes and Evolution 226, 113-137.

Klingenberg, C.P., Marugán-Lobón, J., 2013. Evolutionary covariation in geometric morphometric data: analyzing integration, modularity, and allometry in a phylogenetic context. Systematic Biology 62, 591-610.

Latimer, B. M., Lovejoy, C. O., Spurlock, L., Haile-Selassie, Y., 2016. The thoracic cage of KSD-VP-1/1. In: Haile-Selassie, Y., Su, D.F. (Eds.), The Postcranial Anatomy of *Australopithecus afarensis*. Springer, Dordrecht, pp. 143-153.

Laudicina, N.M., Rodríguez, F., DeSilva, J.M., 2019. Reconstructing birth in *Australopithecus sediba*. PLoS One 14, e0221871.

Lorensen, W.E., Cline, H.E., 1987. Marching cubes: A high resolution 3D surface construction algorithm. Computer Graphics 21, 163-169.

Martens, H., Naes, T., 1992. Multivariate Calibration. John Wiley & Sons, New York.

McCown, T. D., Keith, A., 1939. The Stone Age of Mount Carmel, Vol. 2: The Fossil Human Remains from the Levalloiso-Mousterian. Clarendon Press, Oxford.

Middleton, E. R., 2015. Ecogeographic influences on trunk modularity in recent humans Ph.D. Dissertation, New York University.

Mitteroecker, P., Gunz P., 2009. Advances in geometric morphometrics. Evolutionary Biology 36, 235-247.

Mitteroecker, P., Bookstein, F., 2007. The conceptual and statistical relationship between modularity and morphological integration. Systematic Biology 56, 818-836.

Mitteroecker, P., Gunz, P., Neubauer, S., Müller, G., 2012. How to explore morphological integration in human evolution and development? Evolutionary Biology 39, 536-553.

Morecroft, L., Fieller, N., Dryden, I. Evison, M., 2010. Shape variation in anthropometric landmarks in 3D. In: Evison, M.P., Vorder Bruegge, R.W. (Eds.), Computer-Aided Forensic Facial Comparison. CRC Press, Boca Raton, pp. 35-52.

Neaux, D., Sansalone, G., Ledogar, J. A., Ledogar, S. H., Luk, T. H., Wroe, S., 2018. Basicranium and face: assessing the impact of morphological integration on primate evolution. Journal of Human Evolution 118, 43-55.

O'Higgins, P., 2000. The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. Journal of Anatomy 197, 103-120.

O'Higgins, P., Cobb, S. N., Fitton, L. C., Gröning, F., Phillips, R., Liu, J., Fagan, M. J., 2011. Combining geometric morphometrics and functional simulation: an emerging toolkit for virtual functional analyses. Journal of Anatomy 218, 3-15.

Ponce de León, M.S.P., Golovanova, L., Doronichev, V., Romanova, G., Akazawa, T., Kondo, O., Ishida, H., Zollikofer, C., 2008. Neanderthal brain size at birth provides insights into the evolution of human life history. Proceedings of the National Academy of Sciences USA 105, 13764-13768.

Profico, A., Piras, P., Buzi, C., Di Vincenzo, F., Lattarini, F., Melchionna, M., Veneziano, A., Raia, P., Manzi, G., 2017. The evolution of cranial base and face in Cercopithecoidea and Hominoidea: Modularity and morphological integration. American Journal of Primatology 79, e22721.

Rak, Y., Arensburg, B., 1987. Kebara 2 Neanderthal pelvis: first look at a complete inlet. American Journal of Physical Anthropology 73, 227-231.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rmoutilová, R., Bruzek, J., Ledevin, R., Gomez-Olivencia, A., Couture-Veschambre, C., Holliday, T., Madelaine, S., Velemínská, J., Maureille, B., 2019. New insight on the Neandertal pelvis: Virtual reconstruction of the pelvis of Regourdou 1 specimen. American Journal of Physical Anthropology 168(S68), 205-205.

Rosenberg, K.R., 2007. Neandertal pelvic remains from Krapina: peculiar or primitive? Periodicum Biologorum 109, 387-392.

Rohlf, F.J., Corti, M., 2000. Use of two-block partial least-squares to study covariation in shape. Systematic Biology 49, 740-753.

Ruff, C. B., Trinkaus, E., Holliday, T. W., 1997. Body mass and encephalization in Pleistocene *Homo*. Nature 387, 173-176.

Ruff, C., 2010. Body size and body shape in early hominins–implications of the Gona pelvis. Journal of Human Evolution 58, 166-178.

Ruff, C.B., 1991. Climate and body shape in hominid evolution. Journal of Human Evolution 21, 81-105.

Sawyer, G.J., Maley, B., 2005. Neanderthal reconstructed. The Anatomical Record 283, 23-31.

Schlager, S., 2013. Soft-tissue reconstruction of the human nose. Population differences and sexual dimorphism. Ph.D. Dissertation, Universitatsbibliothek Freiburg.

Schlager, S., 2017. Morpho and Rvcg – shape analysis in R: R-packages for geometric morphometrics, shape analysis and surface manipulations. In: Zheng, G., Li, S., Székely, G. (Eds.), Statistical Shape and Deformation Analysis. Elsevier Academic Press, San Diego, pp. 217-256.

Schlager, S., Profico, A., Di Vincenzo, F., Manzi, G., 2018. Retrodeformation of fossil specimens based on 3D bilateral semi-landmarks: Implementation in the R package "Morpho". PLoS One 13, e0194073.

Schmid, P., 1983. Eine rekonstruktion des skelettes von AL 288-1 (Hadar) und deren konsequenzen. Folia Primatologica 40, 283-306.

Schmid, P., 1991. The trunk in australopithecines. In: Coppens, Y., Senut, B. (Eds.), Origine(s) de la Bipedie chez les Hominides. Editions du CNRS, Paris, pp. 225-234.

Schmid, P., Churchill, S.E., Nalla, S., Weissen, E., Carlson, K.J., de Ruiter, D.J., Berger, L. R., 2013. Mosaic morphology in the thorax of *Australopithecus sediba*. Science 340, 1234598.

Schultz, A.H., 1961. Vertebral column and thorax. In: Hofer, H., Schultz, A. H., Stark, D. (Eds.), Primatologia, Handbuch der Primatenkunde. Karger, Basel, pp. 1-46.

Scott, N. A., Strauss, A., Hublin, J. J., Gunz, P., Neubauer, S., 2018. Covariation of the endocranium and splanchnocranium during great ape ontogeny. PLoS One 13, e0208999.

Senck, S., Bookstein, F.L., Benazzi, S., Kastner, J., Weber, G.W., 2015. Virtual reconstruction of modern and fossil hominoid crania: consequences of reference sample choice. The Anatomical Record 298, 827-841.

Simpson, S.W., Quade, J., Levin, N. E., Butler, R., Dupont-Nivet, G., Everett, M., Semaw, S., 2008. A female *Homo* erectus pelvis from Gona, Ethiopia. Science 322, 1089-1092.

Slice, D.E. (Ed.), 2005. Modern Morphometrics in Physical Anthropology. Kluwer Academic/Plenum Publishers, New York.

Slice, D.E., 2007. Geometric morphometrics. Annual Review of Anthropology 36, 261-281.

Steegmann Jr, A.T., Cerny, F.J., Holliday, T.W., 2002. Neandertal cold adaptation: physiological and energetic factors. American Journal of Human Biology 14, 566-583.

Stelzer, S., Gunz, P., Neubauer, S., Spoor, F., 2017. Hominoid arcade shape: Pattern and magnitude of covariation. Journal of Human Evolution 107, 71-85.

Stelzer, S., Gunz, P., Neubauer, S., Spoor, F., 2018. Using the covariation of extant hominoid upper and lower jaws to predict dental arcades of extinct hominins. Journal of Human Evolution 114, 154-175.

Stern, J.T., 2000. Climbing to the top: A personal memoir of *Australopithecus afarensis*. Evolutionary Anthropology 9, 113–133

Stern Jr, J.T., Susman, R.L., 1983. The locomotor anatomy of Australopithecus afarensis. American Journal of Physical Anthropology 60, 279-317.

Tague, R.G., Lovejoy, C.O., 1998. AL 288-1—Lucy or Lucifer: gender confusion in the Pliocene. Journal of Human Evolution 35, 75-94.

Torres-Tamayo, N., García-Martínez, D., Nalla, S., Barash, A., Williams, S.A., Blanco-Pérez, E., Mata-Escolabo, F., Sanchis-Gimeno, J.A., Bastir, M., 2018. The torso integration hypothesis revisited in *Homo sapiens*: Contributions to the understanding of hominin body shape evolution. American Journal of Physical Anthropology 167, 777-790.

Torres-Tamayo, N., Martelli, S., García-Martínez, D., Schlager, S., Sanchis-Gimeno, J.A., Nalla, S., Ogihara, N., Oishi, M., Bastir, M., in press. Assessing thoraco-pelvic covariation in *Homo sapiens* and *Pan troglodytes*: A 3D geometric morphometric approach. American Journal of Physical Anthropology.

Torres-Tamayo, N., Martelli, S., Lois-Zlolniski, S., Palancar, C.A., García-Martínez, D., Nalla, S., Barash, A., Sanchis-Gimeno, J.A., Villa, C., Hossein Khonsari, R., Schlager, S., Bastir, M., 2019. A new predictive method for quantitative 3D reconstruction of lumbar spine morphology in extinct hominins. Proceedings of the 8th European Society for the Study of Human Evolution 8, 187

Vandermeersch, B., 1991. La ceinture scapulaire et les membres supérieures. Le squelette mousterien de Kébara 2, In: Bar-Yosef, O., Vandermeersch, B. (Eds.), Le Squelette Moustérien de Kébara 2. Éditions du CNRS, Paris, pp. 157-178.

Valladas, H., Valladas, G., 1991. Datation par la thermoluminescence de silex chauffés des grottes de Kébara et de Qafzeh. In: Bar-Yosef, O., Vandermeersch, B. (·ds.), Le Wquelette Moustérien de Kébara 2. Éditions du CNRS, Paris, pp. 43-47.

Walker, A., Ruff, C., 1993. The reconstruction of the pelvis. In: Walker, A., Leakey, R. (Eds.), The Nariokotome *Homo erectus* Skeleton. Harvard University Press, Cambridge, pp. 221–233.

Ward, C.V., 2002. Interpreting the posture and locomotion of *Australopithecus afarensis*: where do we stand?. American Journal of Physical Anthropology 119(S35), 185-215.

Weaver, T.D., Hublin, J.J., 2009. Neandertal birth canal shape and the evolution of human childbirth. Proceedings of the National Academy of Sciences USA 106, 8151-8156.

Weber, G.W., Bookstein, F.L., 2011. Virtual Anthropology: a Guide to a New Interdisciplinary Field. Springer Verlag, Vienna.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometrics and intelligent laboratory systems 58, 109-130.

Zollikofer, C.P., Ponce de León, M., 2005. Virtual Reconstruction: A Primer in Computer-Assisted Paleontology and Biomedicine. Wiley-Interscience, Hoboken.

Captions to the figures

Figure 1. 3D models of previous thorax reconstructions used in this study: A) thorax reconstruction made by Sawyer and Maley (2005); B) thorax reconstruction made by Gómez-Olivencia et al. (2018); C) superimposition of the two thorax reconstructions. First row: anterior view; second row: coronal view; third row: left lateral view.

Figure 2. 3D models of previous pelvis reconstructions used in this study: A) pelvis reconstruction made by Rak and Arensburg (1987); B) pelvis reconstruction made by Sawyer and Maley (2005); C) superimposition of the two pelvis reconstructions. First row: anterior view; second row: coronal view; third row: left lateral view.

Figure 3. 3D landmarks (red), curve semilandmarks (green) and surface semilandmarks (purple) digitized on the 3D virtual torso models: A) frontal view; B) right lateral view. The 12th rib level was not included in the template of digitization due to its high morphological variability (and sometimes absence) in humans. Modified from Torres-Tamayo et al. (2018).

Figure 4. Results of predictive models validation within modern human samples. A, C, E, G) Results of leave-one-out cross-validated mean square error (MSE) in Model A (A), Model B (C), Model D (E), and in the three models for comparison purposes (G); results of non-cross-validated analyses are also shown. B, D, F, H) Probability distributions of Procrustes distances between actual and predicted thoraces within each reference sample (solid lines), and probability distributions of Procrustes distances between every possible pair of thoraces within each reference sample (dashed lines) in Model A (B), Model B (D), Model C (F), and in the three models for comparison purposes (H).

Figure 5. Predictive models used in this study as depicted in plots of the scores (first latent variable) of the two-block partial least squares analysis (left) and associated shape deformations (frontal and left lateral view) at ±1 SD of each axis' scores distribution (right): A) Model A (n = 27 adult males of *Homo sapiens*); B) Model B (n = 60 adult *H. sapiens*); C) Model C (n=33 adult females of *H. sapiens*). Only the first latent variable is shown, because it explains almost or more than 50% of the total covariance in each model.

Figure 6. Kebara 2 thorax predictions based on Model A. A–C) Prediction result using Rak and Arensburg (1985) pelvis reconstruction as predictor (A), superimposition of this prediction and Sawyer and Maley (2005) thorax reconstruction (B), and superimposition

of this prediction and Gómez-Olivencia et al. (2018) thorax reconstruction (C). D–F) Prediction result using Sawyer and Maley (2005) pelvis reconstruction as predictor (D), superimposition of this prediction and Sawyer and Maley (2005) thorax reconstruction (E), and superimposition of this prediction and Gómez-Olivencia et al. (2018) thorax reconstruction (F). First row: anterior view; Second row: coronal view; Third row: left lateral view.

Figure 7. Kebara 2 thorax predictions based on Model B. A–C) Prediction result using Rak and Arensburg (1985) pelvis reconstruction as predictor (A), superimposition of this prediction and Sawyer and Maley (2005) thorax reconstruction (B), and superimposition of this prediction and Gómez-Olivencia et al. (2018) thorax reconstruction (C). D–F) Prediction result using Sawyer and Maley (2005) pelvis reconstruction as predictor (D), superimposition of this prediction and Sawyer and Maley (2005) thorax reconstruction (E), and superimposition of this prediction and Gómez-Olivencia et al. (2018) thorax reconstruction (F). First row: anterior view; Second row: coronal view; Third row: left lateral view.

Figure 8. Kebara 2 thorax predictions based on Model C. A–C) Prediction result using Rak and Arensburg (1985) pelvis reconstruction as predictor (A), superimposition of this prediction and Sawyer and Maley (2005) thorax reconstruction (B), and superimposition of this prediction and Gómez-Olivencia et al. (2018) thorax reconstruction (C). D–F) Prediction result using Sawyer and Maley (2005) pelvis reconstruction as predictor (D), superimposition of this prediction and Sawyer and Maley (2005) thorax reconstruction (E), and superimposition of this prediction and Gómez-Olivencia et al. (2018) thorax

reconstruction (F). First row: anterior view; Second row: coronal view; Third row: left lateral view.