Taxonomic differences in deciduous lower first molar crown outlines of Homo sapiens and Homo neanderthalensis

Taxonomic differences in deciduous lower first molar crown outlines of *Homo sapiens* and *Homo neanderthalensis*

SE Bailey[1,2],R Sorrentino[3,5], G Mancuso[5], J-J Hublin[2,3], S Benazzi[5,2]

[1]Department of Anthropology, New York University, New York USA, [2] Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Leipzig, GERMANY, [3]Collège de France, 70005 Paris, France, [4]Department of Biological Geological and Environmental Sciences – BiGeA, University of Bologna, Bologna, ITALY, [5]Deparment of Cultural Heritage, University of Bologna, Ravenna, ITALY

Corresponding author: Shara Bailey, Department of Anthropology, New York University, 25 Waverly Place, New York, NY 10003 sbailey@nyu.edu

Abstract

Recent studies have demonstrated that the outline shapes of deciduous upper and lower second molars and the deciduous upper first molar are useful for diagnosing hominin taxa – especially *Homo neanderthalensis* and *H. sapiens*. Building on these studies, we use geometric morphometric methods to assess the taxonomic significance of the crown outline of the lower first deciduous molar (dm_1). We test whether the crown shape of the dm_1 distinguishes *H. neanderthalensis* from *H. sapiens* and explore whether dm_1 crown shape can be used to accurately assign individuals to taxa. Our fossil sample includes 3 early *H. sapiens*, 7 Upper Paleolithic *H. sapiens* and 13 *H. neanderthalensis* individuals. Our recent human sample includes 103 individuals from Africa, Australia, Europe, South America and South Asia. Our results indicate that *H. neanderthalensis* dm_1s cluster fairly tightly and separate well from those of Upper Paleolithic *H. sapiens*. However, we also found that the range of shapes in the recent human sample completely overlaps the ranges of all fossil samples. Consequently, results of the quadratic discriminant analysis based on the first 8 PCs representing more than 90% of the variation were mixed. Lower dm1s were correctly classified in 87.3% of the individuals: the combined *H. sapiens* sample had greater success (90.2%) in assigning individuals than did the *H. neanderthalensis* sample (61.5%). When the analysis was run removing the highly variable recent human sample, accuracy increased to 84.6% for *H. neanderthalensis* and 57.1% of Upper Paleolithic *H. sapiens* were classified correctly by using the first four PCs (70.3%). We conclude that caution is warranted when assigning isolated dm_1 crowns to taxa: while an assignment to *H. neanderthalensis* has a high probability of being correct, assignment to Upper Paleolithic *H. sapiens* is less certain.

Key Words: *Homo sapiens*, Neanderthals, Tooth shape, Deciduous molars, Geometric morphometrics

1. Introduction

Before we can test evolutionary hypotheses explaining patterns in, and distribution of, morphological variation in our fossil relatives we must first be able to accurately identify hominin species from a fragmentary fossil record. Recent studies have demonstrated this need by showing the importance of accurately associating a culture with the species that made it (Benazzi et al., 2011a; Benazzi et al., 2015). Correctly identifying isolated dental remains has also shed important light on the timing of dispersals of our species (Benazzi et al., 2011b). The ability to accurately assign isolated skeletal and dental elements to taxa may also result in larger fossil sample sizes, which provide greater power to statistical tests aimed as testing the significance of differences among taxa.

Skeletons recovered from the Late Pleistocene, especially during the European Upper Paleolithic, are often incomplete and fragmentary (Churchill and Smith, 2000). Complicating matters is the fact that fragmentary skeletal elements often are morphologically undiagnostic and may be unusable unless they preserve ancient hominin DNA. Dental elements, on the other hand, are more frequently recovered and, due to their durable enamel, are often complete.

Although tooth size alone is not very informative for diagnosing Late Pleistocene taxa (Bailey and Hublin, 2005), tooth crown and root morphology has proven to be quite useful, especially in distinguishing *Homo neanderthalensis* (hereafter: Neanderthals) from *H. sapiens* during the periods in which they overlapped in time and space (Bailey et al., 2009; Been et al., 2017; Benazzi et al., 2011b, 2014; Fabbri et al., 2016; Hublin et al., 2020; Kupczik and Hublin, 2010; Le Cabec et al., 2013 ; Margherita et al., 2016). When complete dentitions are found and

crowns are relatively unworn, assigning specimens to taxa is fairly straightforward because Neanderthals have diagnostic combinations of dental characters (Bailey, 2002a; 2002b, 2006). Even incomplete dentitions can be diagnostic if the appropriate teeth and/or characters are preserved (Bailey et al., 2009). However, while many tooth crowns are found complete, they often suffer from wear that obscures or eliminates minor morphological features on the crown (e.g., occlusal crests and small accessory cusps).

Early studies of molar crown shape relied on the position of, and relationships between, cusp tips, which required relatively unworn teeth (Bailey, 2004; Morris, 1981). More recently, methods of assessing crown shape (e.g., Elliptical Fourier Analysis - EFA, semi-landmark-based methods) from crown outlines have allowed for the inclusion of both worn and unworn molar crowns in analyses (Benazzi et al., 2012). Studies using these methods have shown that crown outlies of permanent molars are quite useful for partitioning out variation and assigning specimens to taxa (Bailey and Lynch, 2005; Benazzi et al., 2011a; Gómez-Robles et al., 2007; Gómez-Robles et al., 2008, 2011).

The small size and thin enamel of deciduous molars make them especially prone to loss of surface information through attrition, especially in paleoanthropological and archaeological samples that predate the advent of processed food. For this reason, the crown outline is particularly useful for assessing shape differences among groups. Over the past decade several studies have confirmed that the outlines of postcanine deciduous crowns can be used to accurately assign individuals to taxa (Bailey et al., 2014b, 2016; Fornai et al., 2016; Moroni et al., 2018a).

In hominins, the deciduous second molar (dm2 or dp4[1]) is remarkably similar to the permanent first molar (M1) in both crown outline and morphology (Fig. 1A). While about 15% smaller in size than the M1 (Bailey et al., 2014a), within individuals the dm2 preserves the same number of primary cusps; and the number and expression of accessory features are highly correlated between the two (Edgar and Lease, 2007; Kieser, 1984; Paul et al., 2017). Because the dm2 forms early during ontogeny (Liversidge and Molleson, 2004) it is presumed to be little influenced by environmental variation. Moreover, studies have shown it to be less variable in size and morphology than the deciduous first molar (Farmer and Townsend, 1993; Liversidge and Molleson, 1999; Margetts and Brown, 1978). Thus, it is perhaps not surprising that just like the M1, the dm2 has proven to discriminate between Neanderthals and *H. sapiens* quite well (Bailey et al., 2014a, 2015; Benazzi, 2012; Moroni et al., 2018a).

In contrast to the dm2, the dm1 can be more premolar-like than molar-like in form, at least in later *Homo* (Fig. 1B). The dm1 often preserves fewer cusps, with the distal aspects of both upper and lower dm1 reduced compared to the dm2. The dm[1] may even be bicuspid (preserving only mesial cusps) in some *H. sapiens* groups. Like the dm[1], the distal cusps of the dm1 may be completely missing, preserving only the protoconid and metaconid. This variation in cusp number and expression is reflected in the crown's shape.

[FIGURE 1A and 1B ABOUT HERE]

[1] Here we follow terminology in the dental anthropological literature, which refers to this tooth as a molar. We are aware that in the paleontological literature this tooth is referred to as a premolar.

An earlier study of dm^1 shape of Neanderthals and *H. sapiens* resulted in 96.3% accuracy in separating the two groups (Benazzi et al., 2011b). The current study builds on our previous studies of the diagnostic utility of deciduous molar shape for taxonomic affiliation by examining variation of the dm_1 (dp_3). We analyze the crown shapes of Neanderthals and early, Upper Paleolithic and recent *H. sapiens*, applying geometric morphometric (GM) methods to crown outlines taken from digital occlusal images. Based on our previous research, we expect that the dm_1 will distinguish Neanderthals from *H. sapiens* with a high degree of accuracy (80% or higher). Based on results of our earlier study showing that the dm_2 and M_1 were slightly less diagnostic than the dm^2 and M^1 (Bailey et al., 2016), we expect this may also to be the case for the dm_1. The ability of the dm_1 to discriminate among taxa will rely, at least in part, on the amount of variation within each group. At a broader level, knowing the degree of variability within groups may allow us to test hypotheses about the evolutionary forces, or the relaxation of such forces, driving this variation.

If the dm_1 crown outline proves to discriminate well between Neanderthals and *H. sapiens*, it will add to the tools available for assessing isolated teeth and assigning them to fossil taxa. If, unlike the dm^1 (Benazzi et al., 2011b), the dm_1 crown outline cannot accurately assign teeth to taxa, future work will focus on exploring the possible reasons why the lower molars are less distinctive than the upper molars.

2.0 Materials

2.1. Samples

The materials used in this study include occlusal photographs of dm_1s from 126 recent and fossil *H. sapiens* and Neanderthals (Table 1). Our recent *H. sapiens* (RHS) sample includes

103 individuals representing Africa, Australia, Europe, South America and South Asia. Deciduous teeth are scarce in the fossil record and our comparative fossil sample, while small, includes nearly all relevant fossil dm_1s available for study: 3 early *H. sapiens* (EHS), 7 Upper Paleolithic *H. sapiens* (UPHS) and 13 Neanderthals. We assigned specimens to taxa based on assignments made in the published literature. These assignments were based on a combination of criteria including: cranial morphology, age, cultural association, and/or their association with taxonomically diagnostic adult human remains.

We included only complete and undamaged crowns in our samples. With one exception (Die Kelders 6291), these crowns ranged in status from unworn to moderately worn (three or more small dentine patches, stages 1–4; Molnar, 1971). Figure 2 illustrates the single crown with stage 5 wear (see Methods below for how worn outlines were reconstructed). Even in moderatly worn crowns it was primarily the distal aspect that required correction.We did not consider sex as a variable in this study due to the difficulty in assigning sex to fossil individuals, especially those represented by isolated teeth.

[TABLE 1 ABOUT HERE]

We arbitrarily chose to use the left dm_1 to represent each individual. If the left side was not represented or was damaged, we used the right side and mirror-imaged the crown using Adobe PhotoShop® before the analysis. Although the left and right sides may be asymmetrical in size and/or shape, studies have shown that dental asymmetry occurs randomly with regard to side. This phenomenon is known as fluctuating asymmetry (Van Valen, 1962). To date we know of no study quantifying the differences in crown shapes between left and right antimeres.

However, we assume that crown shape asymmetry is randomly distributed — as it is for tooth size and dental nonmetric traits, which influence crown shape (see Scott and Turner, 1997 for review).

2.2. Methods of data collection and analysis

All but seven occlusal images were taken using a Canon EOS Rebel XT digital 8 MP camera equipped with a macro lens (see Supplementary Online Material [SOM] Table S1). All images were taken from original skeletal and fossil materials (i.e., no casts were used). Photographic images of the fossils were taken by SEB. Some images of recent humans (primarily the African samples) were taken by Caroline Souday (see acknowledgements) under the supervision of SEB. Individual teeth were oriented so that the cervical border was perpendicular to the camera's optical axis. A bubble device was used to level the camera and each image included a similarly leveled millimeter scale that was placed at approximately the same height as the cusp tips. Bailey et al. (2004) have shown that inter-observer error due to differences in image orientation and camera equipment is low (2.4%–4.5%) and not significantly greater than intra-observer error.

In seven cases (SOM Table S1) occlusal images were acquired from microtomographic (µCT) image data of original specimens performed by the Department of Human Evolution of the Max Planck Institute for Evolutionary Anthropology. In those cases, either an industrial µCT system or a desktop system was used, and the subsequent voxel resolutions ranged from 14 to 70 µm. The image stacks of each tooth were filtered to improve tissue grayscale homogeneity and then segmented into enamel and dentine components manually with Avizo® v.9 (Thermo Fisher Scientific). The crown surface was extracted as a 3D digital

surface model (.ply format). The models of the µCT scans were opened in Avizo® v.9 and then manipulated in 3D space so that the cervical border was perpendicular to the optical axis in both mesiodistal and buccolingual directions (Benazzi et al., 2009). Aviso® v.9 was used to add an appropriate scale and then a screen shot of the occlusal surface (analogous to taking a digital photograph) was taken and saved as a .jpg file. A recent study has shown that there is no significant difference between crown outlines obtained from photos and 3D digital models (Buti, 2013).

Screen shots and digital images were imported into Adobe Photoshop®. Backgrounds were removed and image contrast was adjusted to provide a clear distinction between the crown outline and the background. Finally, each image was scaled to approximately the same size and resolution (300 dpi).

Even in moderately worn dm_1s, interproximal wear sometimes distorted the distal aspect of the crown outline. Less often, the mesial aspect was also affected. In these cases, the outline was reconstructed by estimating the original mesial and/or distal borders (see Bailey, 2004; Gómez-Robles et al., 2007; Wood and Abbott, 1983; Wood and Engleman, 1988). These estimations were based on the buccolingual extent of the wear facet and the overall contour of the tooth (Fig. 2); all estimations were made by SEB.

[FIGURE 2 ABOUT HERE]

The occlusal images of the dm_1s were imported in Rhino 4.0 Beta CAD environment (Robert McNeel & Associates, Seattle, WA), placed on the xy-plane of the Cartesian coordinate system, and rotated along the z-axis to have its lingual aspect parallel to the x-axis. Then, for

each tooth the crown outline was manually digitized using the curve function. The outlines were centered on their centroid, and equiangularly spaced radial vectors emanating from their superimposed centroids (the first radius parallel to the y-axis and buccally directed) intersected the outlines. Ultimately 24 pseudolandmarks were identified for each outline (Fig. 3; Benazzi et al., 2011a). Finally, the pseudolandmark configurations were scaled to unit centroid size (i.e., Procrustes shape coordinates) and variation in crown outline shape was explored by principal components analysis (PCA) of the matrix of shape coordinates (Bailey et al., 2014a, b, 2016; Benazzi et al., 2011b; Benazzi et al., 2012; Lacy et al., 2018; Moroni et al., 2018b).

[FIGURE 3 ABOUT HERE]

We conducted two separate PCAs. The first analysis included all samples to examine variation among fossil and recent groups. The second analysis used only the recent *H. sapiens* sample to investigate the role of geographic origin in the variation observed.

To identify potentially significant differences in crown shape of the dm_1 between groups, permutation tests (n = 10,000) were conducted using the first three PCs. These tests compared the distance between two group means to the distances obtained by random assignment of observations to this groups (using Morpho v. 2.8 in R). Values were considered significant at $p < 0.05$. Because Neanderthal molars are, on average, slightly larger than those of *H. sapiens* and because size and shape may be related, we also conducted an analysis examining the relationship between shape variables (PCs) and size allometry (logarithm of crown base area). This analysis was investigated by Procrustes ANOVA with permutation procedures (n = 1,000) using the R package geomorph v. 3.2.1 (Adams and Otárola-Castillo, 2013).

The Shapiro-Wilks test was used to assess the normality of distribution of Procrustes shape coordinates for each group in the sample (Ghasemi and Zahediasl, 2012). Fligner-Killeen's test was performed to test the homogeneity of variances across the groups, rejecting the null hypothesis H0 (variances homogeneity) if $p < 0.05$. Since both assumption of normality and homogeneity of variance were violated, we used leave-one-out cross-validation Quadratic Discriminant Analysis (QDA) to test how well crown shape discriminates taxa (see Results for details). The QDA used the first eight PCs representing about 90% of the variation in the comparison of *H. sapiens* (fossil and recent) and Neanderthals. Whereas, considering the small sample size of UPHS ($n = 7$), the QDA used the first four PCs (70.3%) in the comparison among recent *H. sapiens,* UPHS and Neanderthals, as well as between UPHS and Neanderthals. The number of PCs used for QDA was chosen in order to find the minimum optimal combination of variables (i.e., PCs) within the sensible cutoff in the range of 70% to 90% of variation (Jolliffe, 2002; Sorrentino et al., 2020). Posterior probabilities were calculated using equal prior probability of 0.5. The data were processed and analyzed through software routines written in R v. 3.4.3 (R Core Team, 2017).

3. Results

3.1. Principal components analysis

Figure 4 illustrates the results of the PCA. The first three principal components account for about 60% of the variance (PC1 = 31.6%, PC2 = 15.5%, and PC3 = 12.4%; Fig. 4a). Allometry is responsible for only 2.1% of overall crown variation ($F = 2.72$, $R^2 = 0.021$, df = 1, $p < 0.05$) considering the whole sample; and it remains similar (2.3%) when excluding EHS ($F = 2.86$, $R^2 = 0.023$, df = 1, $p < 0.05$) in Procrustes ANOVA. The contribution of allometry

increases to 10.8% in the comparison of Neanderthals and UPHS ($F = 2.19$, $R^2 = 0.108$, df = 1, $p > 0.05$), but the effects of shape variation due to size allometry are not significant in this case. It is, therefore, unlikely that size is a significant driver of shape differences between the two groups.

[FIGURE 4 ABOUT HERE]

The range of variation in recent humans is wide and spans all four quadrants of the PCA plot. With the exception of two *H. sapiens* individuals (Die Kelders 6291 and La Madeleine) all fossil individuals, regardless of taxon, fall within the RHS range. Recent humans appear to be distributed randomly but it is possible that their distribution reflects the geographic range sampled in this study. The results of a PCA exploring the RHS distribution further by grouping RHS samples by geographic region are provided in Figure 5 and discussed below (3.3 Recent human variation).

[FIGURE 5 ABOUT HERE]

In Figures 4b and 5 positive PC1 scores represent a relatively rectangular crown shape, whereas negative PC1 scores reflect a more trapezoidal shape with a mesiobuccal projection related to the tuberculum molare. Along PC2, positive scores reflect an asymmetrical crown with a somewhat reduced trigonid portion and unreduced talonid, while negative PC2 scores are associated with a somewhat triangular shape with a reduction in the talonid portion of the crown.

3.2. Fossil hominin variation

The three EHS individuals are variable for PC1. However, none have particularly high negative PC1 scores, indicating the absence of a strong mesiobuccal projection (i.e., tuberculum molare). All three individuals have negative scores for PC2, which reflect relatively large mesial cusps. The three EHS individuals fall closer to the range of Neanderthals than they do to the range of UPHS. All of the UPHS individuals possess negative PC1 scores, which reflect the presence of a prominent tuberculum molare. Along PC2 UPHS individuals have mainly positive scores (or low negative scores), indicating crowns with a relatively wider talonid than trigonid. Neanderthal individuals have both positive and negative PC1 scores and mainly negative PC2 scores. Along PC1 the Neanderthal dm_1 scores range from moderately positive to moderately negative, reflecting the observation that some possess a strong tuberculum molare, while others are more rectangular and/or symmetrically shaped. Table 2 presents the results of a permutation test of the significance of differences among groups. Significant differences are obtained between the UPHS sample and all the other groups ($p < 0.05$). Significant differences are also found between Neanderthals and the UPHS and RHS samples ($p < 0.05$), but not between the Neanderthal and the EHS samples ($p > 0.05$).

[TABLE 2 ABOUT HERE]

The PCA plots in Figure 4 shows that the UPHS and Neanderthal samples are less variable than the RHS sample despite their wider temporal sampling, although small sample sizes may play a role this result. In fact, the two fossil groups separate quite well in shape space (especially in the 3D plot of the first three PCs: Fig 4a), with only one individual falling in the range of both. Figure 6 provides the mean dm_1 crown shapes of UPHS and Neanderthals. As

suggested from the PCA plots, the mean shape of UPHS reflects the marked mesiobuccal projection frequently observed in that sample, whereas the mean shape in Neanderthals reflects the wider range of expression in this feature.

[FIGURE 6 ABOUT HERE]

3.3. Recent human variation

Figure 5 provides a PCA plot of the geographic subgroups within the RHS sample. Figure 7 illustrates the wide range of shape variation within the subgroups. Procrustes ANOVA showed no significant effects (1.8%) of crown variation due to size allometry in the RHS sample ($F = 1.84$, $R_2 = 0.018$, df = 1, $p > 0.05$). Table 3 presents the results of the permutation test of significant differences among recent human subgroups in which the two Australian individuals were not included. With the exception of South America, all subgroups span the four quadrants of the PCA graph. Significant differences were obtained between the Sub Saharan African and European ($p < 0.05$), South American ($p < 0.05$) and South Asian ($p < 0.05$) subgroups. Significant differences were also found between South American and European ($p < 0.05$) and South Asian ($p < 0.05$) subgroups. The North African subsample differs significantly from the European ($p < 0.05$) and South Asia ($p < 0.05$) subgroups. Even though significant differences were found, Figure 5 suggests the geographic patterning to the variation is not very strong. Among the recent geographic subgroups, the South American sample shows the narrowest distribution: individuals have positive and negative PC1 scores but only positive PC2 scores.

[FIGURE 7 and TABLE 3 ABOUT HERE]

3.4. Quadratic Discriminant Functions Analysis

Shapiro-Wilks tests show that the distribution of Procrustes shape coordinates of the RHS violate the assumption of normality ($W = 0.945$, $p < 0.05$), whereas UPHS ($W = 0.879$, $p > 0.05$ and Neanderthals ($W = 0.894$, $p > 0.05$) do not. The variances of the groups are not homogeneous ($\chi2 = 555.7$, df = 3, $p < 0.05$), even if EHS are excluded ($\chi2 = 321$, df = 2, $p < 0.05$). Furthermore, Fligner-Killeen's test shows different variance between RHS and Neanderthals ($\chi2 = 8.65$, df = 1, $p < 0.05$), RHS and UPHS ($\chi2 = 112.79$, df = 1, $p < 0.05$), and between UPHS and Neanderthals ($\chi2 = 4.8$, df = 1, $p < 0.05$).

Results of the QDA are provided in Tables 4 and 5. When grouped according to taxon (*H. neanderthalensis* and *H. sapiens*), individuals were correctly assigned 87.3% of the time (Table 4). The classification for *H. sapiens* was better (90.2%) than it was for Neanderthals (61.5%). When *Homo sapiens* was separated into fossil and recent groups and reanalyzed, RHS were correctly classified 76.7% of the time, but only 42.9% of the UPHS individuals and 53.8% of the Neanderthals classified correctly (Table 5). EHS was not considered in this second analysis, due to its small sample size.

[TABLES 4 and 5 ABOUT HERE]

To explore the effect of the recent human variation on our results, and because our primary goal was to ascertain whether dm_1 shape can accurately distinguish between Neanderthals and fossil *H. sapiens*, we re-ran the QDA focusing only on Neanderthal and UPHS groups. Doing this increased the accuracy substantially (Table 6). Correct assignment to the

Neanderthal group rose to 84.6% while correct assignment to UPHS increased to 57.1%, with two Neanderthals (Bruniquel and Roc du Marsal) and three UPHS individuals (Estelas, Isturitz and Solutre) misclassified.

[INSERT TABLE 6 ABOUT HERE]

Discussion

Results of the present study are in agreement with previous ones, which have demonstrated that there are significant differences between the deciduous molar crown shapes of UPHS and Neanderthals. As was the case for other deciduous molars, we found that assessment of the dm_1 shape provides a relatively accurate method for identifying Neanderthal individuals. However, and in contrast to our previous studies, the success rate in classifying UPHS based on dm_1 shape is substantially lower. This leads us to conclude that a dm_1 assigned to 'Neanderthal' is very likely to be correct, but a dm_1 assigned to 'UPHS' is less certain to be correct.

We are somewhat surprised at the mediocre classification accuracy for the UPHS individuals, especially given that in the PCA the UPHS and Neanderthal samples appear to be well separated in shape space (Fig. 4). We believe that our QDA results reflect, at least in part, the choice of PCs and the variance for the QDA. We chose a number of PCs (4) that was both less than the smallest group size (n = 7) and also accounted for at least 70% of the variance. Re-running the QDA with five and six PCs (accounting for a slightly higher amount of variation) did not improve the results. Re-running the QDA with only the first three PCs (which are illustrated in Fig. 4a) led to better classificatory results, but the first three PCs accounted for only 60% of the variance. Therefore, we do not have confidence in those results. Since both number of PCs and variance are affected by the size of samples used, we believe that small sample size is

responsible, at least in part, for the lower classification accuracy indistinguishing Neanderthals and UPHS in this study compared to previous ones (e.g., Bailey et al 2016).

Results from the present study are consistent with those of our previous studies, which found that lower molars are less powerful in discriminating *H. sapiens* (both fossil and recent) and Neanderthal groups than are the upper molars. The first study using dm^1 crown shape to distinguish Neanderthals from UPHS showed the method to be successful 96% of the time (Benazzi et al 2011a). In the same study, the shape of the dm^2 proved to be 100% accurate at discriminating individuals from these two groups. In a follow-up study that included a wide geographic range of recent *H. sapiens*, the accuracy of the dm^2 was only slightly lower (97%; Bailey et al 2014). Subsequent studies that assessed the lower dentition suggested that dm_2 shape was also a powerful discriminator of Neanderthals and UPHS and recent European *H. sapiens,* but it was slightly less accurate (92%) than the upper deciduous molars (Benazzi et al., 2012). And in a study comparing dm_2 and M_1 shapes, Bailey et al. (2016) confirmed that both lower molars discriminated between these two species less successfully than the upper molars. The results of the present study show that the dm_1 is the least powerful in terms of discriminating Neanderthals from *H. sapiens.*

The mediocre discriminatory power of the dm_1 in the present study is at least somewhat related to the wide range of shape variation in recent humans (more than has been observed in the other deciduous molars) and the greater similarity of EHS dm_1 shape to that of Neanderthals, at least as far as can be determined with this small EHS sample. A previous study of the dm_2 and M_1 (Bailey et al., 2016) also suggested that EHS and Neanderthal dm_2 shapes do not differ significantly. However, in that study EHS specimens plotted well within the variation of both RHS and UPHS groups, which makes the dissimilarity between EHS and UPHS dm_1 shapes

found in this study somewhat surprising. The similarities between EHS and Neanderthals may suggest that the lower dentition has undergone less change in EHS than it has in UPHS and RHS. Additional specimens from the Middle Pleistocene would help clarify the polarity of dm_1 crown shapes and confirm that this is the case.

Conclusions

Based on the recent series of studies of molar crown shapes, we conclude that the lower deciduous molars and the lower permanent M1 are less reliable than the upper molars for discriminating between Neanderthals and *H. sapiens*. Although our results for the dm_1 are somewhat mediocre over all, from a practical standpoint we can say that crown shape of the dm_1 is useful for identifying Neanderthals in a Late Pleistocene European context. Unfortunately, we would hesitate to use the dm_1 to identify *H. sapiens* from the same time period/region because the success rate is not much better than chance. In addition, we would not recommend using the dm_1 crown shape to discriminate between these two groups where they co-occur in the Near East, since the early *H. sapiens* dm_1 crown outline does not differ significantly from that of Neanderthals. In sum, the dm_1 crown shape is only of limited use for assigning isolated teeth to taxa.

Acknowledgements

SEB is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 724046 – SUCCESS, http://www.erc-success.eu). SEB's data collection was supported by the Leakey Foundation and the Max Planck Institute. Caroline Souday (CS) took photographs of the African and some of the

European samples under SEB's supervision while employed as a New York University postdoctoral fellow.

REFERENCES CITED

Adams, D.C., Otárola-Castillo, E., 2013. Geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution 4, 393-399.

Bailey, S.E., 2002a. A closer look at Neanderthal postcanine dental morphology. I. The mandibular dentition. Anatomical Record 269, 148-156.

Bailey, S.E., 2002b. Neandertal dental morphology: implications for modern human origins, Ph.D. Dissertation, Arizona State University.

Bailey, S.E., 2004. A morphometric analysis of maxillary molar crowns of Middle-Late Pleistocene hominins. Journal of Human Evolution 47, 183-198.

Bailey, S.E., 2006. Beyond shovel shaped incisors: Neandertal dental morphology in a comparative context. Periodicum Biologorum 108, 253-267.

Bailey, S.E., Benazzi, S., Buti, L., Hublin, J.-J., 2016. Allometry, merism, and tooth shape of the lower second deciduous molar and first permanent molar. American Journal of Physical Anthropology 159, 93-105.

Bailey, S.E., Benazzi, S., Hublin, J.-J., 2014a. Allometry, Merism and tooth shape of the upper deciduous M2 and permanent M1. American Journal of Physical Anthropology 154, 104-114.

Bailey, S.E., Benazzi, S., Souday, C., Astorino, C., Paul, K., Hublin, J.-J., 2014b. Taxonomic differences in deciduous upper second molar crown outlines of *Homo sapiens*, *Homo neanderthalensis* and *Homo erectus*. Journal of Human Evolution 72, 1-9.

Bailey, S.E., Hublin, J.-J., 2005. Who made the Early Aurignacian? A reconsideration of the Brassempouy dental remains. Bulletins et mémoires de la Société d'anthropologie de Paris 17, 115-121.

Bailey, S.E., Lynch, J.M., 2005. Diagnostic differences in mandibular P4 shape between Neandertals and anatomically modern humans. American Journal of Physical Anthropology 126, 268-277.

Bailey, S.E., Pilbrow, V.C., Wood, B.A., 2004. Interobserver error involved in independent attempts to measure cusp base areas of *Pan* M1s. Journal of Anatomy 205, 323-331.

Bailey, S.E., Weaver, T.D., Hublin, J.-J., 2009. Who made the Aurignacian and other early Upper Paleolithic industries? Journal of Human Evolution 57, 11-26.

Been, E., Hoversc, E., Ekshtain, R., Malinski-Buller, A., Aghna, N., Barash, A., Bar-Yosef, Y., Mayer, D., Benazzi, S., Hublin, J.-J., Leven, L., Greenbaum, N., Mitki, N., Oxilia, G., Porat, N., Roskin, J., Soudack, M., Yeshurun, R., Shahack-Gross, R., Nir, N., Stahlschmidt, M.C., Rak, Y., Barzilai, O., 2017. The first Neanderthal remains from an open-air Middle Palaeolithic site in the Levant. Scientific Reports 7, 1-8.

Benazzi, S., 2012. The first modern Europeans. Journal of Anthropological Sciences 90, 3-6.

Benazzi, S., Bailey, S.E., Peresani, M., Mannino, M.A., Romandini, M., Richards, M.P., Hublin, J.-J., 2014. Middle Paleolithic and Uluzzian human remains from Fumane Cave, Italy. Journal of Human Evolution 70, 61-68.

Benazzi, S., Coquerelle, M., Fiorenza, L., Bookstein, F., Katina, S., Kullmer, O., 2011a. Comparison of dental measurement systems for taxonomic assignment of first molars. American Journal of Physical Anthropology 144, 342-354.

Benazzi, S., Douka, K., Fornai, C., Bauer, C.C., Kullmer, O., Svoboda, J., Pap, I., Mallegni, F., Bayle, P., Coquerelle, M., Condemi, S., Ronchitelli, A., Harvati, K., Weber, G.W., 2011b. Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. Nature 479, 525-528.

Benazzi, S., Fantini, M., De Crescenzio, F., Persiani, F., Gruppioni, G., 2009. Improving the spatial orientation of human teeth using a virtual 3D approach. Journal of Human Evolution 56, 286-293.

Benazzi, S., Fornai, C., Buti, L., Toussaint, M., Mallegni, F., Ricci, S., Gruppioni, G., Weber, G.W., Condemi, S., Ronchitelli, A., 2012. Cervical and crown outline analysis of worn Neanderthal and modern human lower second deciduous molars. American Journal of Physical Anthropology 149, 537-546.

Benazzi, S., Slon, V., Talamo, S., Negrino, F., Peresani, M., Bailey, S.E., Sawyer, S., Panetta, D., Vicino, G., Starnini, E., Mannino, M.A., Salvadori, P.A., Meyer, M., Pääbo, S., Hublin, J.-J., 2015. The makers of the Protoaurignacian and implications for Neandertal extinction. Science 348, 793-796.

Buti, L., 2013. Nuove prospettive di indagine tassonomica di denti decidui usurati attraverso analisi di immagine e tecnologie tridimensionali. Ph.D. Dissertation, Università di Firenze.

Churchill, S., Smith, F., 2000. Makers of the early Aurignacian of Europe. Yearbook of Physical Anthropology 43, 61-115.

Edgar, H.J.H., Lease, L.R., 2007. Correlations between deciduous and permanent tooth morphology in a European sample. American Journal of Physical Anthropology 133, 726-734.

Fabbri, P.F., Panetta, D., Sarti, L., Martini, F., Salvadori, P., Caramella, D., Fedi, M., Benazzi, S., 2016. Middle Paleolithich human deciduous incisor from Grotta del cavallo, Italy. American Journal of Physical Anthropology 161, 506-512.

Farmer, V., Townsend, G., 1993. Crown size variability in the deciduous dentition of South Australian children. American Journal of Human Biology 5, 681-690.

Fornai, C., Benazzi, S., Gopher, A., Barkai, R., Sarig, R., Bookstein, F.L., Hershkovitz, I., Weber, G.W., 2016. The Qesem Cave hominin material (part 2): a morphometric analysis of dm2-QC2 deciduous lower second molar. Quaternary International 398, 175-189.

Ghasemi, A., Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. International Journal of Endocrinology and Metabolism 10, 486-489.

Gómez-Robles, A., Martinón-Torres, M., Bermúdez De Castro, J.M., Margvelashvili, A., Bastir, M., Arsuaga, J.L., Pérez-Pérez, A., Estebaranz, F., Martínez, L.M., 2007. A geometric morphometric analysis of hominin upper first molar shape. Journal of Human Evolution 55, 627-638.

Gómez-Robles, A., Martinón-Torres, M., Bermúdez de Castro, J.M., Prado-Simón, L., Arsuaga, J.L., 2011. A geometric morphometric analysis of hominin upper premolars. Shape variation and morphological integration. Journal of Human Evolution 61, 688-702.

Gómez-Robles, A., Martinón-Torres, M., Bermúdez de Castro, J.M., Prado, L., Sarmiento, S., Arsuaga, J.L., 2008. Geometric morphometric analysis of the crown morphology of the lower first premolar of hominins, with special attention to Pleistocene *Homo*. Journal of Human Evolution 55, 627-638.

Hublin, J-J., Sirakov, N., Aldeias, V., Bailey, S., Bard, E., Delvigne, V., Endarova, E., Fagault, Y., Fewlass, H., Hajdinjak, M., Kromer, B., Krumov, I., Marreiros, J., Martisius, N., Paskulin, L., Sinet-Mathiot, V., Meyer, M., Pääbo, S., Popov, V., Rezek, Z., Svoboda, S., Skinner, M., Smith, G., Spasov, R., Talamo, S., Tuna, T., Wacker, L., Welker, F., Wilcke, A., Zahariev, N., McPherron, S., Tsanova, T., 2020. Initial Upper Paleolithic *H. sapiens* from Bacho Kiro Cave, Bulgaria. Nature. 581, 299-302. doi.org/10.1038/s41586-020-2259-z

Jolliffe, I.T. 2002. Principal Component Analysis. Springer, New York.

Kieser, J.A., 1984. An analysis of the Carabelli trait in the mixed deciduous and permanent human dentition. Archives of Oral Biology 29, 403-406.

Kupczik, K., Hublin, J.-J., 2010. Mandibular molar root morphology in Neanderthals and Late Pleistocene and recent *Homo sapiens*. Journal of Human Evolution 59, 525-541.

Lacy, S.A., Bailey, S., Benazzi, S., Delage, C., 2018. Newly Recognized Human Dental Remains at Les Fadets (Lussac-les-Châteaux, Vienne, France). Bulletins et mémoires de la Société d'anthropologie de Paris 30, 180-191.

Le Cabec, A., Gunz, P., Kupczik, K., Braga, J., Hublin, J.-J., 2013 Anterior tooth root morphology and size in Neanderthals: taxonomic and functional implications. Journal of Human Evolution 64, 169-193.

Liversidge, H.M., Molleson, T., 1999. Deciduous tooth size and morphogenetic fields in children from Christ Church, Spitalfields. Archives of Oral Biology 44, 7-13.

Liversidge, H.M., Molleson, T., 2004. Variation in crown and root formation and eruption of human deciduous teeth. American Journal of Physical Anthropology 123, 172-180.

Margetts, B., Brown, T., 1978. Crown diameters of the deciduous teeth in Australian Aboriginals. American Journal of Physical Anthropology 48, 493-502.

Margherita, C., Talamo, S., Wiltschke-Schrotta, K., Senck, S., Oxilia, G., Sorrentino, R., Mancuso, G., Gruppioni, G., Lindner, R., Hublin, J.-J., Benazzi, S., 2016. A reassessment of the preseumed Torrener Bärenhöhle's Paleolithic human tooth. Journal of Human Evolution 93, 120-125.

Molnar, S., 1971. Human tooth wear, tooth function and cultural variability. American Journal of Physical Anthropology 34, 27-42.

Moroni, A., Ronchitelli, A., Simona, A., Aureli, D., Bailey, S.E., Boscato, P., Boschin, F., Capecchi, G., Crezzini, J., Douka, K., Marciani, G., Panetta, D., Ranaldo, F., Ricci, S., Scaramucci, S., Spagnolo, V., Benazzi, S., Gambassini, P., 2018b. Grotta del Cavallo (Apulia – Southern Italy). The Uluzzian in the mirror. Journal of Anthropological Sciences 96, 125 - 160.

Morris, D.H., 1981. Maxillary first premolar angular differences between North American Indians and non-North American Indians. American Journal of Physical Anthropology 54, 431-433.

Paul, K.S., Astorino, C.M., Bailey, S.E., 2017. The Patterning Cascade Model and Carabelli's trait expression in metameres of the mixed human dentition: exploring a morphogenetic model. American Journal of Physical Anthropology 162, 3-18.

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

Scott, G.R., Turner, C.G., II, 1997. The Anthropology of Modern Human Teeth. Dental Morphology and its Variation in Recent Human Populations. Cambridge University Press, Cambridge.

Sorrentino R, Belcastro MG, Figus C, Stephens NB, Turley K, Harcourt-Smith W, Ryan T, Benazzi S. 2020. Exploring sexual dimorphism of the modern human talus through geometric morphometric methods. PLoS ONE 15(2): e0229255.

Van Valen, L., 1962. A study of fluctuating asymmetry. Evolution 16, 125-142.

Wood, B.A., Abbott, S.A., 1983. Analysis of the dental morphology of Plio-Pleistocene hominids. I. Mandibular molars: crown area measurements and morphological traits. Journal of Anatomy 136, 197-219.

Wood, B.A., Engleman, C.A., 1988. Analysis of the dental morphology of Plio-Pleistocene hominids. V. Maxillary postcanine tooth morphology. Journal of Anatomy 161, 1-35.

Figure Legends

Figure 1. Comparison of A) upper and lower left dm2 and M1, and B) upper and lower left dm1 and dm2 (all images represent the same recent *H. sapiens* from Peru). In both photos, upper is on the left, lower is on the right. For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 2. Illustration showing the most worn crown (stage 5 wear: Molnar, 1971) in our sample and how minor corrections were made to the outline before analysis (Early *H. sapiens* Die Kelders 6291). For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 3. Illustration showing methods for acquisition of pseudolandmarks on the left dm_1 of the Kebara 1 Neanderthal. For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 4. Results of the Principal Components Analysis: all samples. The range of variation in recent *H. sapiens* encompasses that of nearly all fossil samples, whereas the fossil samples are more tighly constrained along the first three PCs. Center plot: PC1 against PC2. Upper left: PC1, PC2 and PC3. N, Neanderthal; EHS, Early *Homo sapiens*; RHS, Recent *Homo sapiens*; UPHS, Upper Paleolithic *Homo sapiens*. For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 5. Results of the Principal Components Analysis of recent *H. sapiens* grouped by geographic origin. With the exception of the South American sample, which has only positive PC2 scores, there appears to be no geographic patterning to ldm1 shape based on the first two principal components. For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 6. Comparison of mean shapes between Neanderthals (left) and Upper Paleolithic *H. sapiens* (right). Right arrow indicates mesiobuccal expansion (tuberculum molare) in Upper Paleolithic *H. sapiens.* Left arrow indicates more equal sized buccal and lingual cusps in Neanderthals. For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 7. Variation of left dm_1 crown shape within recent *H. sapiens* geographic populations. For orientation: B = buccal, L = lingual, M = mesial, D = distal.

Figure 1

Click here to access/download;Figure;Bailey et al Fig 1 dm1dm2M1 comparisons.tif

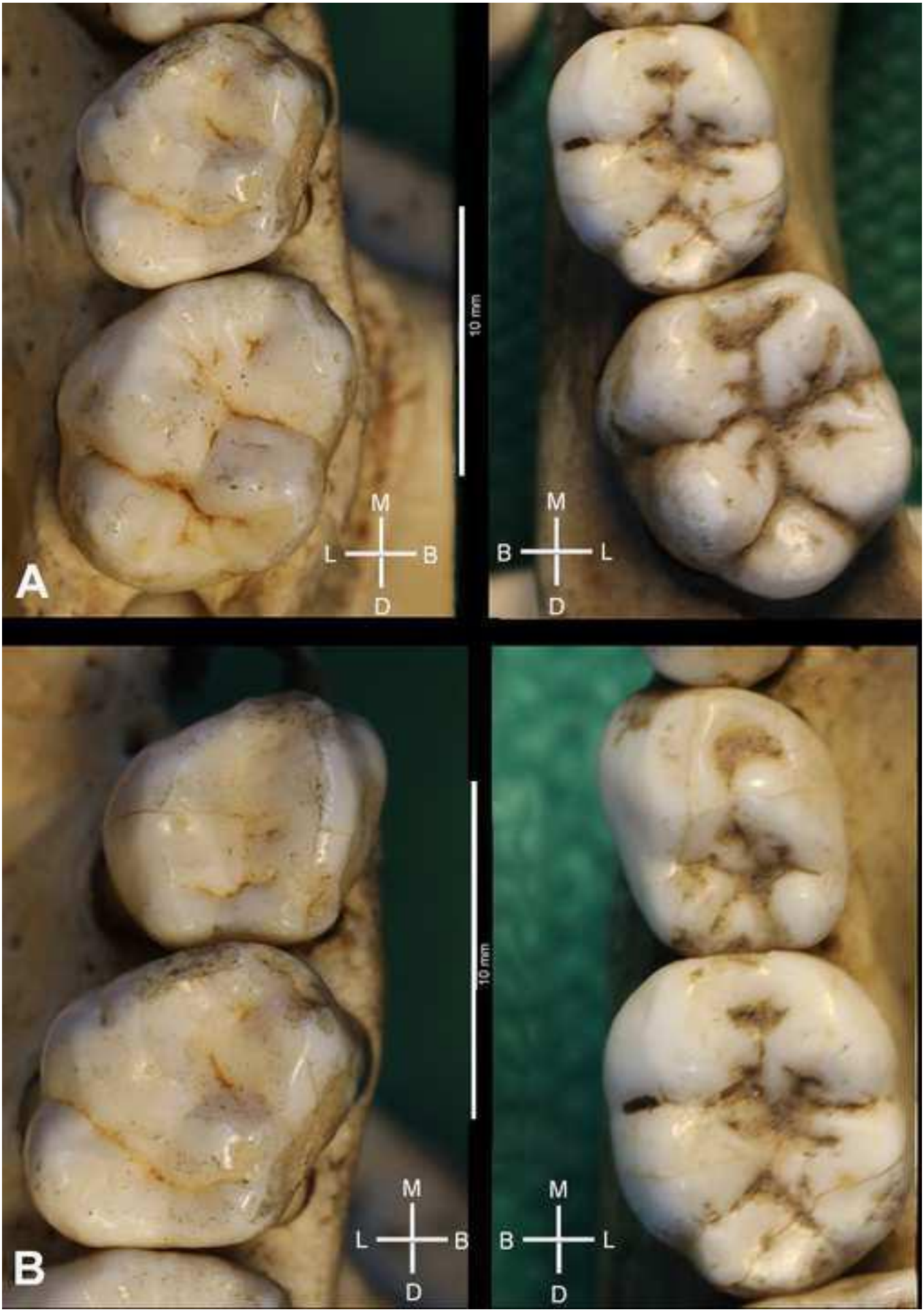

Figure 2

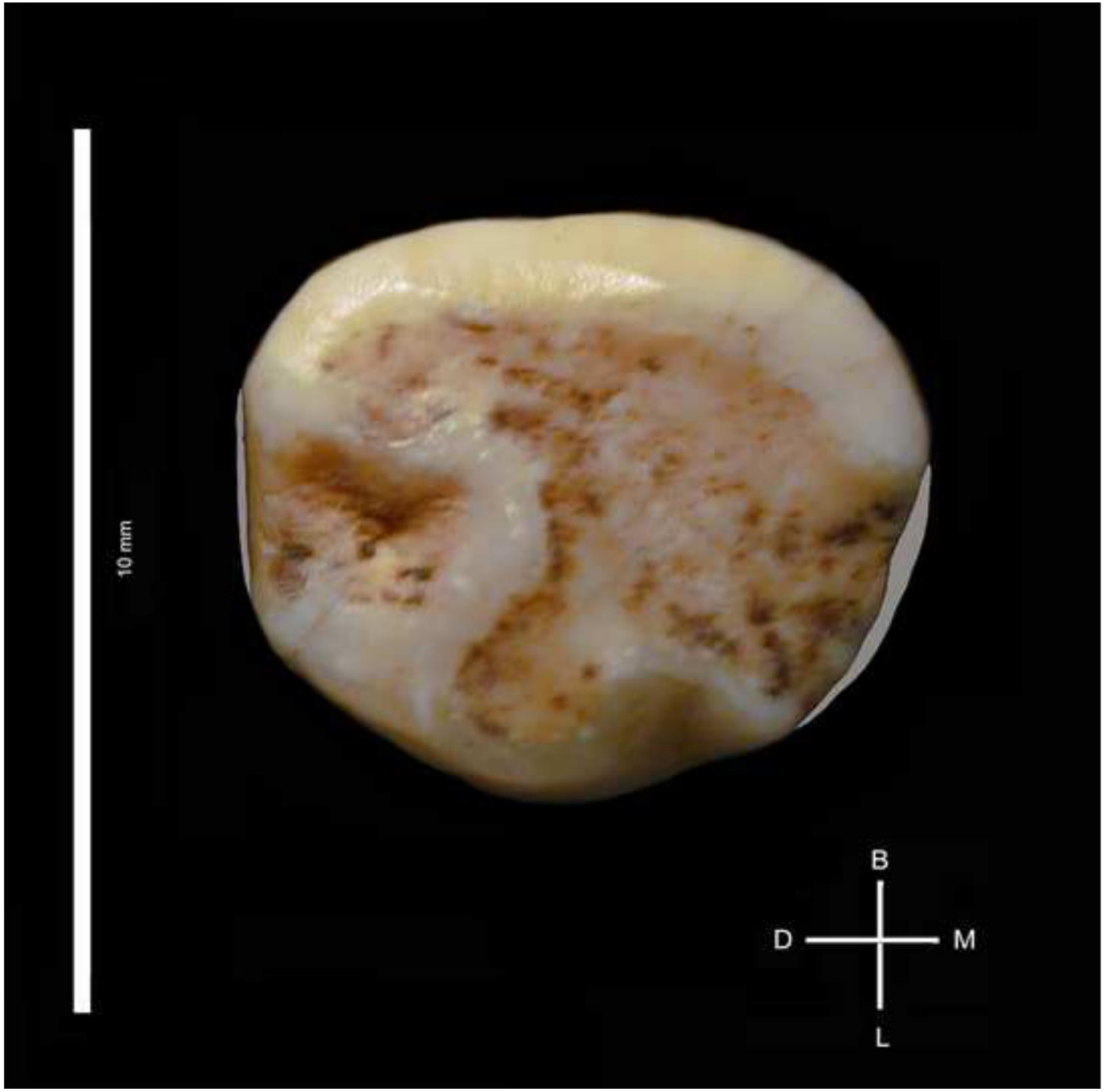

Figure 3

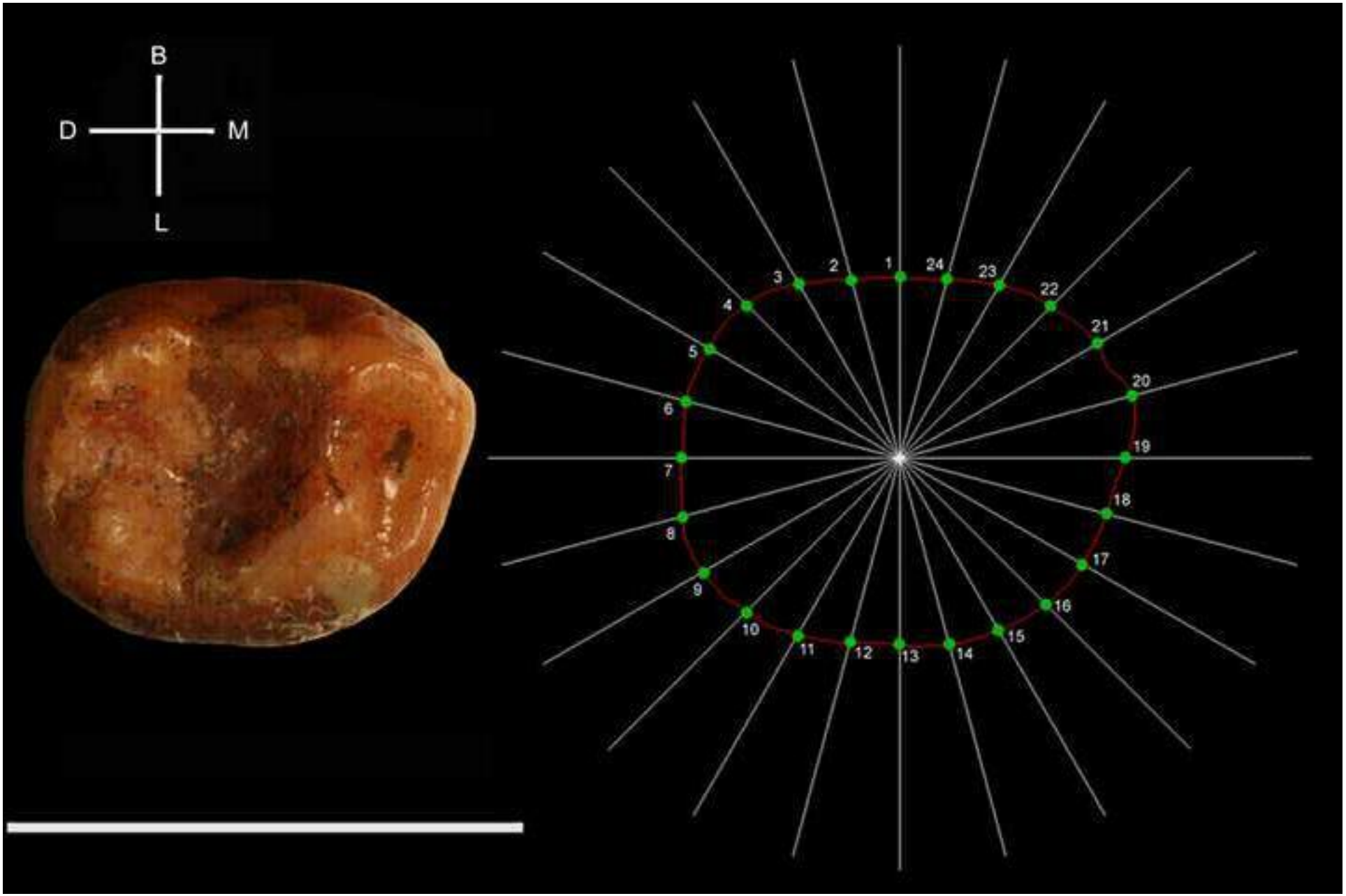

Figure 4

Click here to access/download;Figure;Bailey et al Fig 4 PCA all symbols 7-20.tif

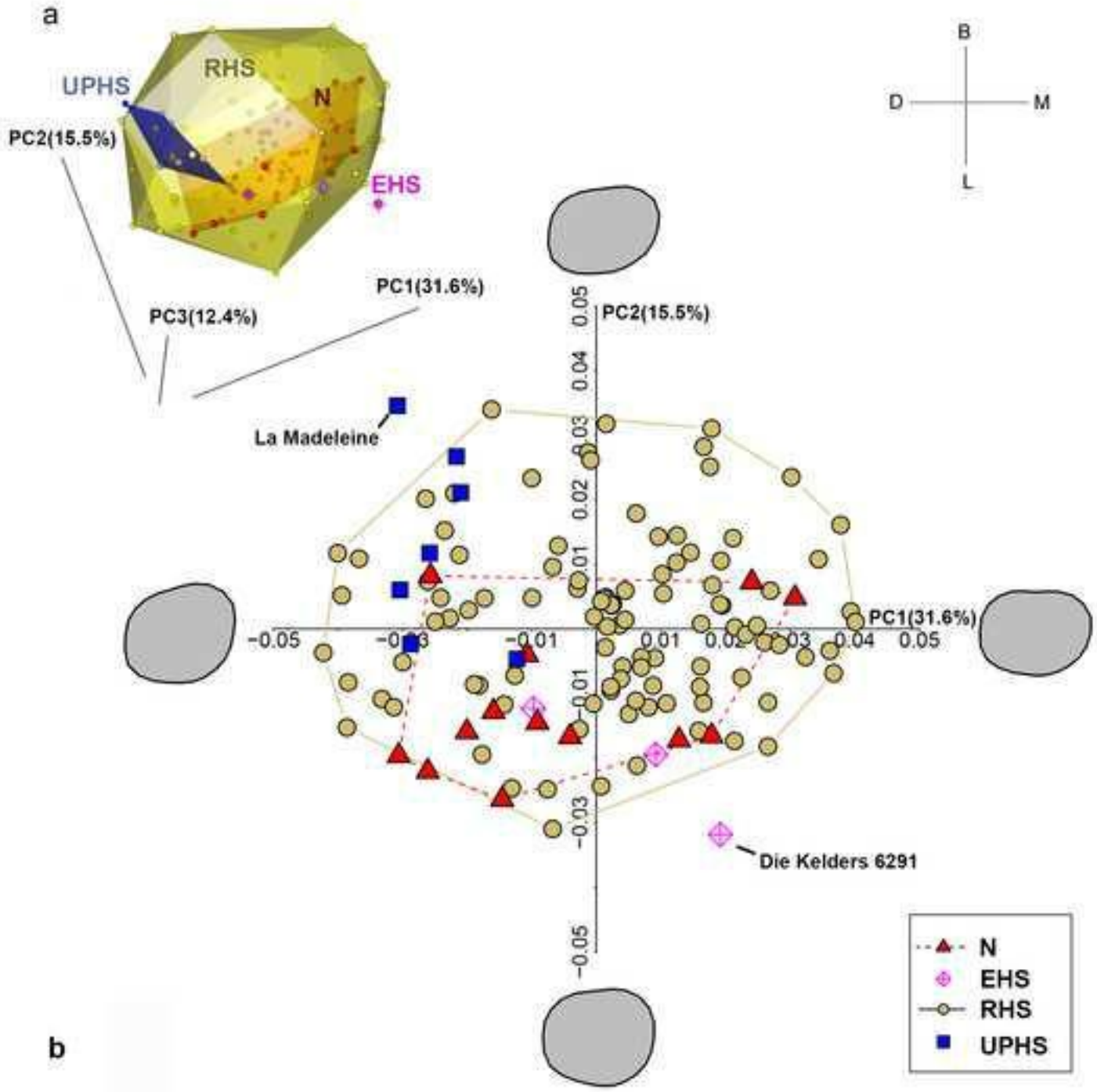

Figure 5

Click here to access/download;Figure;Bailey et al Fig 5 PCA recent symbols.tif

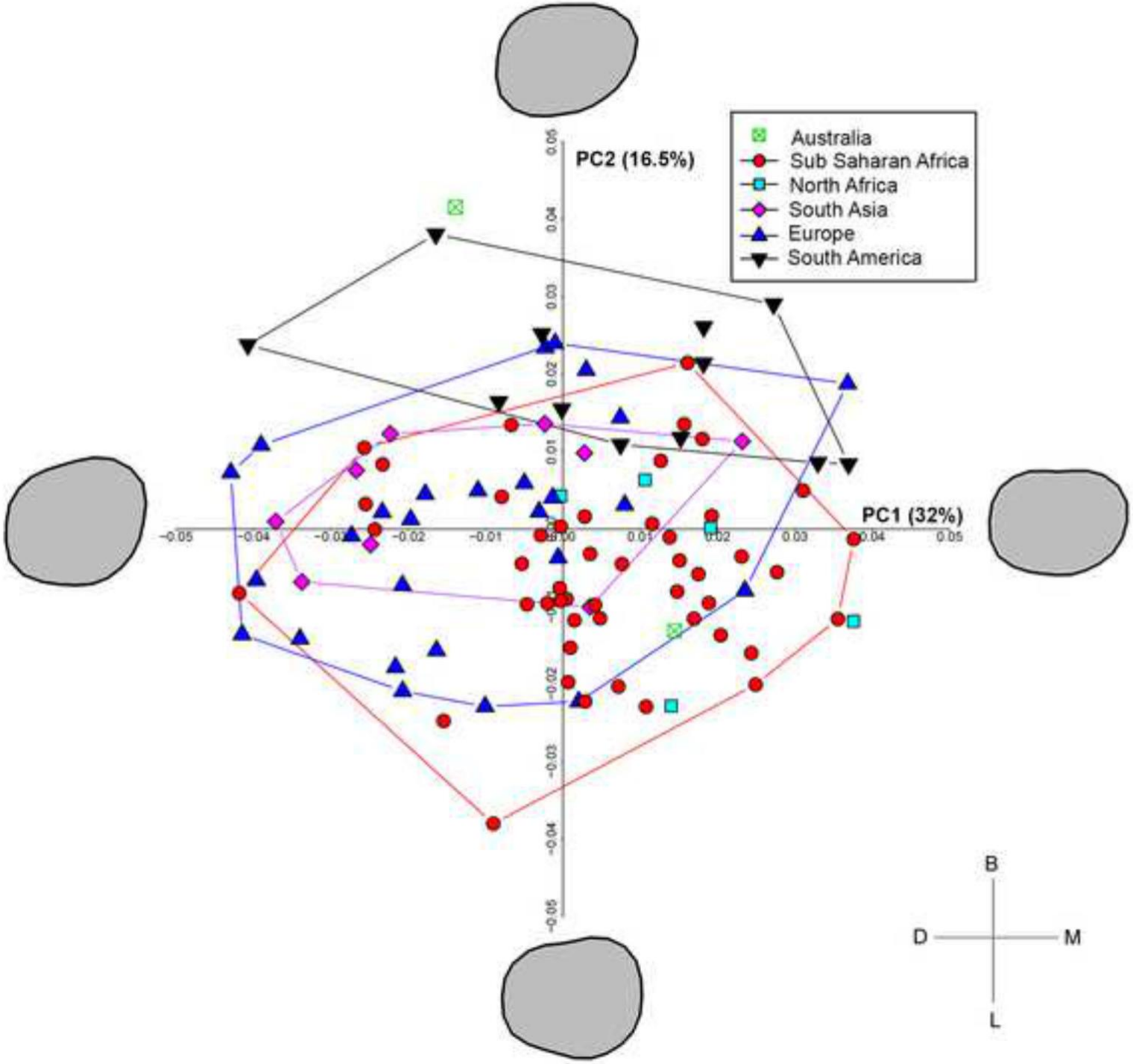

Figure 6

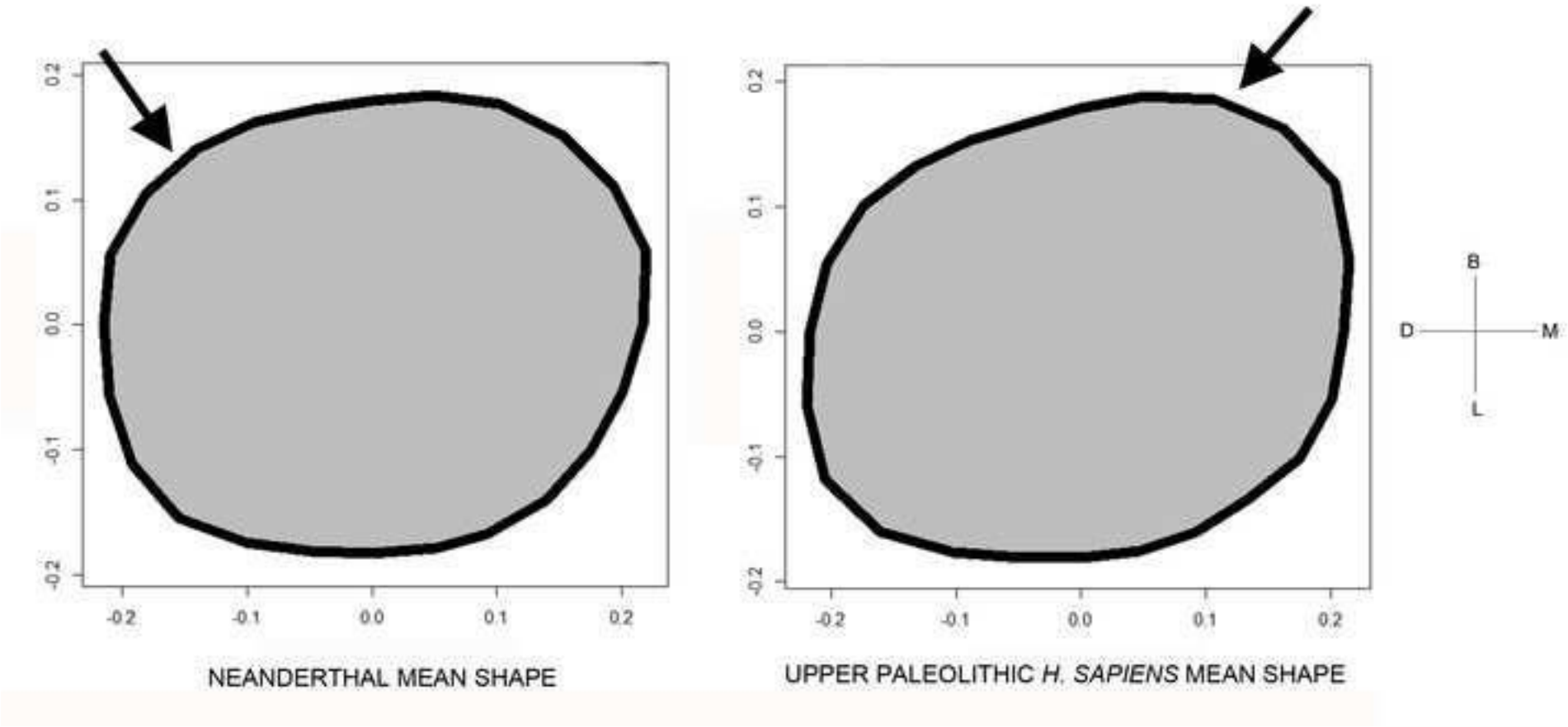

Figure 7

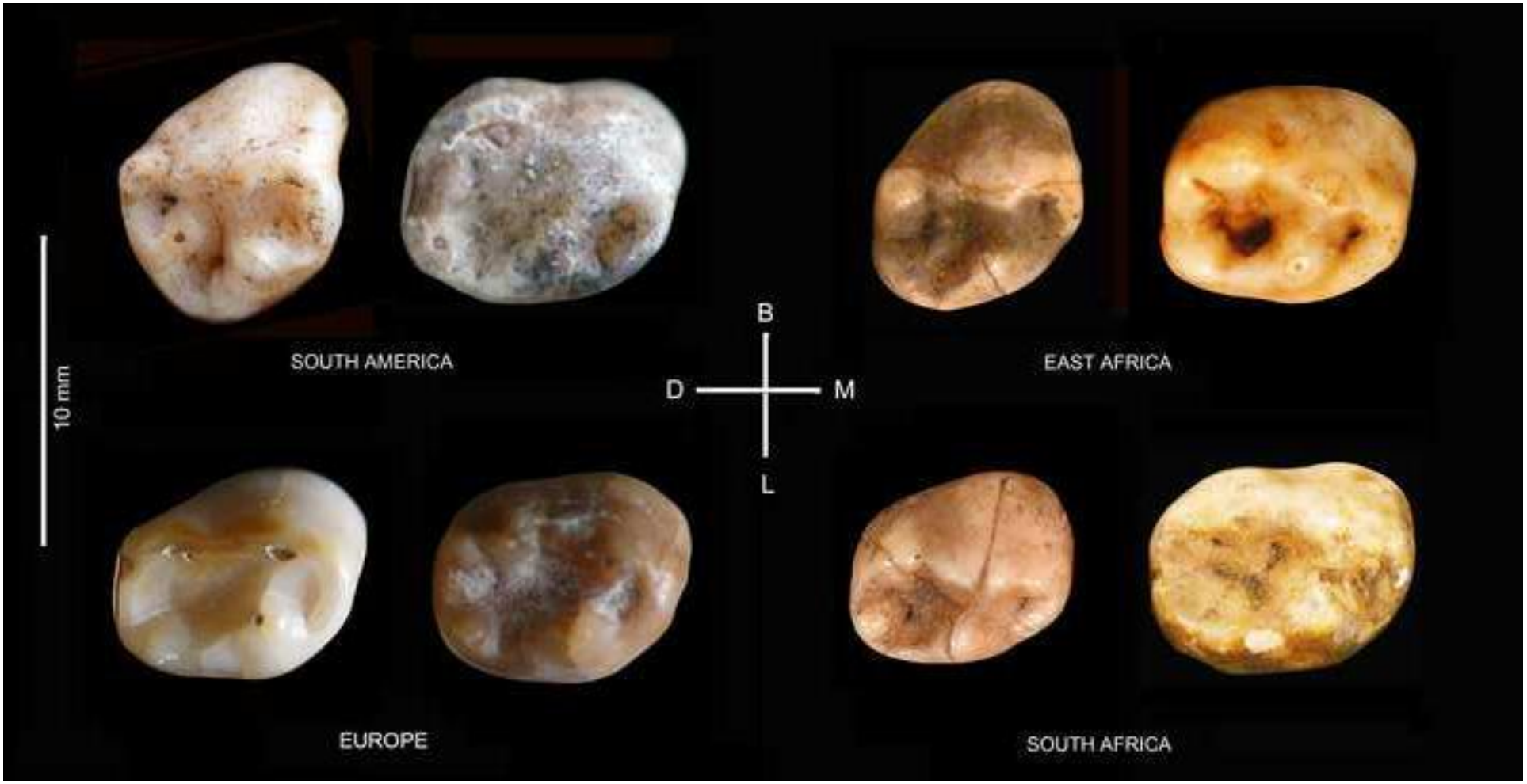

Table 1
Materials used in this study.[a]

	No.	Sites sampled
Early *H. sapiens*	3	Die Kelders, Qafzeh
Upper Paleolithic *H. sapiens*	7	Balla Barlang, Estelas, La Grotte du Figuier, Isturitz, Lagar Velho, Abri de la Madeleine, Roche de Solutré
H. neanderthalensis	13	Archi, Arcy-sur-Cure, Barakai Cave, Bruniquel, Combe Grenal, Engis, Kebara, La Ferrassie, La Chaise, Riparo del Molare, Peche de l'Azé, Roc de Marsal, Mezmaiskaya
Recent *H. sapiens*	103	Africa, Asia, Australia, Europe, South America

[a] See SOM for sources of materials.

Table 2
Permutation tests of differences in crown shape of the dm_1 between fossil and recent human samples. [a]

	Early *H. sapiens*	Neanderthal	Recent *H. sapiens*
H. neanderthalensis (n=13)	0.502		
Recent *H. sapiens* (n=103)	0.182	**0.010**	
Upper Paleolithic *H. sapiens* (n=7)	**0.002**	**0.002**	**0.001**

[a]Significant differences ($p < 0.05$) are in bold.

Table 3
Permutation tests of differences in crown shape of the dm_1 between recent human geographic subsamples.[a-b]

	Europe	North Africa	South America	South Asia
Europe (n=28)				
North Africa (n=5)	**0.025**			
South America (n=12)	**0.001**	0.054		
South Asia (n=9)	0.884	**0.040**	0**.032**	
Sub-Saharan Africa (n=49)	**0.001**	0.884	**0.001**	**0.032**

[a]The Australia sample is excluded in the permutation test because of its small (n=2) sample size'
[b]Significant differences ($p < 0.05$) are in bold.

Table 4
Results of quadratic discriminant functions assignments (fossil and recent *H. sapiens* combined) based on crown shape of the dm_1 by using 8 PCs (accounting for 90.7% of the variation).

	H. neanderthalensis	*H. sapiens*	% correct
H. neanderthalensis (n=13)	8	5	61.5
H. sapiens (n=113)	11	102	90.2

Table 5

Results of quadratic discriminant functions assignments (fossil and recent *H. sapiens* separated) based on crown shape of the dm_1 using 4 PCs (accounting for 70.3% of the variation). Early *H. sapiens* are excluded due to small sample size.

	H. neanderthalensis	*H. sapiens*	Upper Paleolithic *H. sapiens*	% correct
H. neanderthalensis (*n* = 13)	7	4	2	53.8
Recent *H. sapiens* (*n* = 103)	4	79	16	76.7
Upper Paleolithic *H. sapiens* (*n* = 7)	0	4	3	42.9

Table 6

Results of quadratic discriminant functions assignments (Upper Paleolithic *H. sapiens* and *H. neanderthalensis* only) by using 4 PCs (accounting for 70.3% of the variation). Early *H. sapiens* are excluded due to small sample size.

	H. neanderthalensis	Upper Paleolithic *H. sapiens*	% correct
H. neanderthalensis (*n* = 13)	11	2	84.6
Upper Paleolithic *H. sapiens* (*n* = 7)	3	4	57.1