

The Open University's repository of research publications and other research outputs

## A geometric morphometric approach to the study of variation of shovel-shaped incisors

### Journal Item

#### How to cite:

Carayon, Delphine; Adhikari, Kaustubh; Monsarrat, Paul; Dumoncel, Jean; Braga, José; Duployer, Benjamin; Delgado, Miguel; Fuentes-Guajardo, Macarena; de Beer, Frikkie; Hoffman, Jakobus W.; Oettlé, Anna C.; Donat, Richard; Pan, Lei; Ruiz-Linares, Andres; Tenailleau, Christophe; Vaysse, Frédéric; Esclassan, Rémi and Zanolli, Clément (2019). A geometric morphometric approach to the study of variation of shovel-shaped incisors. *American Journal of Physical Anthropology*, 168(1) pp. 229–241.

For guidance on citations see [FAQs](#).

© 2018 Wiley Periodicals, Inc.

Version: Accepted Manuscript

Link(s) to article on publisher's website:  
<http://dx.doi.org/doi:10.1002/ajpa.23709>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---



**A geometric morphometric approach to the study of  
variation of shovel-shaped incisors.**

Journal:	<i>American Journal of Physical Anthropology</i>
Manuscript ID	AJPA-2018-00075.R2
Wiley - Manuscript type:	Technical Note
Date Submitted by the Author:	n/a
Complete List of Authors:	<p>Carayon, Delphine; Universite Toulouse III Paul Sabatier Faculte des Sciences et d'Ingenierie, Laboratoire AMIS, UMR 5288 CNRS; Universite de Montpellier Faculte d'Odontologie, Service de Prothèses</p> <p>Adhikari, Kaustubh; University College London, Department of Genetics, Evolution and Environment, and UCL Genetics Institute</p> <p>Montsarrat, Paul; Universite Toulouse III Paul Sabatier Faculte des Sciences et d'Ingenierie, Faculté d'Odontologie</p> <p>Dumoncel, Jean; Laboratoire d'Anthropobiologie Moléculaire et d'Imagerie de Synthèse, UMR 5288 CNRS-Université de Toulouse (Paul Sabatier), Braga, José; AMIS UMR 5288 CNRS, Hominid Evolutionary Biology; University of the Witwatersrand, Institute for Human Evolution</p> <p>Duployer, Benjamin; Centre Inter-universitaire de Recherche et d'Ingénierie des Matériaux, UMR 5085 CNRS-Université de Toulouse (Paul Sabatier)</p> <p>Delgado, Miguel; Consejo Nacional de Investigaciones Cientificas y Tecnicas</p> <p>Fuentes-Guajado, Macarena; University College London Research Department of Genetics Evolution and Environment, Department of Genetics, Evolution and Environment</p> <p>de Beer, Frikkie; South African Nuclear Energy Corporation</p> <p>Hoffman, Jakobus; Pelindaba, Radiation Science</p> <p>Oettlé, Anna; University of Pretoria, Department of Anatomy</p> <p>Donat, Richard; Institut National de Recherches Archéologiques Préventives, St Estève</p> <p>Pan, Lei; Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, ; Laboratoire d'Anthropobiologie Moléculaire et d'Imagerie de Synthèse, UMR 5288 CNRS-Université de Toulouse (Paul Sabatier), Ruiz-Linares, Andres; University College London, Biology</p> <p>Tenailleau, Christophe; Centre Inter-universitaire de Recherche et d'Ingénierie des Matériaux, UMR 5085 CNRS-Université de Toulouse (Paul Sabatier)</p> <p>Vaysse, Frédéric; Universite Toulouse III Paul Sabatier Faculte des Sciences et d'Ingenierie, Faculté d'Odontologie</p> <p>Esclassan, Rémi; Universite Toulouse III Paul Sabatier Faculte des Sciences et d'Ingenierie, Faculté d'Odontologie</p> <p>Zanolli, Clément; Universite Toulouse III Paul Sabatier, Laboratoire AMIS,</p>

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

	UMR 5288
Key Words:	shovel-shape incisors, ASUDAS, Procrustes and non-Procrustes superimpositions, virtual anthropology
Subfield: Please select 2 subfields. Select the main subject first.:	Human biology [living humans; behavior, ecology, physiology, anatomy], Primate biology [behavior, ecology, physiology, anatomy]

SCHOLARONE™  
Manuscripts

1  
2  
3 **1 A geometric morphometric approach to the study of variation of shovel-shaped incisors**

4  
5  
6 3 Delphine Carayon<sup>1,2</sup>, Kaustubh Adhikari<sup>3</sup>, Paul Monsarrat<sup>4,5</sup>, Jean Dumoncel<sup>1</sup>, José Braga<sup>1,6</sup>,  
7 4 Benjamin Duployer<sup>7</sup>, Miguel Delgado<sup>8,9</sup>, Macarena Fuentes-Guajardo<sup>3</sup>, Frikkie de Beer<sup>10</sup>,  
8 5 Jakobus W. Hoffman<sup>10</sup>, Anna C. Oettlé<sup>11,12</sup>, Richard Donat<sup>1,13</sup>, Lei Pan<sup>14,15</sup>, Andres Ruiz-  
9 6 Linares<sup>3,16,17</sup>, Christophe Tenailleau<sup>7</sup>, Frédéric Vaysse<sup>1,4</sup>, Rémi Esclassan<sup>1,4</sup>, Clément Zanolli<sup>1</sup>.  
10  
11  
12  
13

14 8 <sup>1</sup>UMR 5288 CNRS, Université Toulouse III - Paul Sabatier, France

15 9 <sup>2</sup>Faculté d'Odontologie, Université Montpellier I, France

16 10 <sup>3</sup>Department of Genetics, Evolution and Environment, and UCL Genetics Institute, University  
17 11 College London, London, United Kingdom

18 12 <sup>4</sup>Faculté d'Odontologie, Université Toulouse III - Paul Sabatier, France

19 13 <sup>5</sup>STROMALab, CNRS ERL 5311, EFS, INP-ENVT, Inserm, Université Toulouse III - Paul  
20 14 Sabatier, France

21 15 <sup>6</sup>Evolutionary Studies Institute and School of Geosciences, University of the Witwatersrand,  
22 16 PO WITS, Johannesburg 2050, South Africa

23 17 <sup>7</sup>Centre Inter-universitaire de Recherche et d'Ingénierie des Matériaux, UMR 5085 CNRS,  
24 18 Université de Toulouse III - Paul Sabatier, France

25 19 <sup>8</sup>Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, Buenos Aires,  
26 20 República Argentina

27 21 <sup>9</sup>División Antropología, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La  
28 22 Plata, La Plata, República Argentina

29 23 <sup>10</sup>Radiation Science Department, South African Nuclear Energy Corporation (Necsa),  
30 24 Pelindaba, South Africa

31 25 <sup>11</sup>Department of Anatomy, University of Pretoria, South Africa

32 26 <sup>12</sup>Department of Anatomy, Sefako Makgatho Health Sciences University, South Africa

33 27 <sup>13</sup>Institut National de Recherches Archéologiques Préventives- St Estève, France

34 28 <sup>14</sup>Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate  
35 29 Paleontology and Paleoanthropology, CAS, Beijing, China

36 30 <sup>15</sup>State Key Laboratory of Palaeobiology and Stratigraphy, Nanjing Institute of Geology and  
37 31 Palaeontology, CAS, Nanjing, China

38 32 <sup>16</sup>Ministry of Education Key Laboratory of Contemporary Anthropology and Collaborative  
39 33 Innovation Center of Genetics and Development, Fudan University, Shanghai, China  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

34 <sup>17</sup>Laboratory of Biocultural Anthropology, Law, Ethics, and Health (Centre National de la  
35 Recherche Scientifique and Etablissement Français du Sang, UMR-7268), Aix-Marseille  
36 University, Marseille, France

38 Corresponding author: Delphine Carayon ([delphinedom.carayon@orange.fr](mailto:delphinedom.carayon@orange.fr))

41 Keywords: shovel-shape incisors; ASUDAS; Procrustes and non-Procrustes superimpositions;  
42 virtual anthropology

1  
2  
3 56 **Abstract**

4 57 **Objectives:** The scoring and analysis of dental non-metric traits are predominantly  
5 58 accomplished by using the Arizona State University Dental Anthropology System  
6 59 (ASUDAS), a standard protocol based on strict definitions and three-dimensional dental  
7 60 plaques. However, visual scoring, even when controlled by strict definitions of features,  
8 61 visual reference and the experience of the observer, includes an unavoidable part of  
9 62 subjectivity. In this methodological contribution, we propose a new quantitative geometric  
10 63 morphometric approach to quickly and efficiently assess the variation of shoveling in modern  
11 64 human maxillary central incisors (UI1).

12 65 **Materials and Methods:** We analyzed 87 modern human UI1s by means of virtual imaging  
13 66 and the ASU-UI1 dental plaque grades using geometric morphometrics by placing  
14 67 semilandmarks on the labial crown aspect. The modern human sample was composed of  
15 68 individuals from Europe, Africa and Asia and included representatives of all seven grades  
16 69 defined by the ASUDAS method.

17 70 **Results:** Our results highlighted some limitations in the use of the current UI1 ASUDAS  
18 71 plaque, indicating that it did not necessarily represent an objective gradient of expression of a  
19 72 non-metric tooth feature. Rating of shoveling tended to be more prone to intra- and inter-  
20 73 observer bias for the highest grades. In addition, our analyses suggest that the observers were  
21 74 strongly influenced by the depth of the lingual crown aspect when assessing the shoveling.

22 75 **Discussion:** In this context, our results provide a reliable and reproducible framework  
23 76 reinforced by statistical results supporting the fact that open scale numerical measurements  
24 77 can complement the ASUDAS method.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Introduction

Teeth display morphological variations of the crown and roots that differ substantially among modern human and fossil groups, some dental characteristics being predominant in certain groups or populations (Turner et al., 1991). As stated by Hillson, "human eyes and brain are unsurpassed in discerning tiny differences between objects compared side by side, but it is much more difficult to define a scheme for recording size and shape in such a way that comparisons can be made between hundreds of such objects" (Hillson, 1996: 68). For this reason, since the 19<sup>th</sup> century, several attempts have been made to classify and assess differences between fossil and extant human populations, at first using detailed descriptive approaches and later elaborating scoring systems (reviewed in Irish and Scott, 2016).

Following the influential early works of Hrdlička (1920) and Dahlberg (1956), who standardized a four grade classification plaque for upper incisor shoveling, some researchers tried to reduce the visual subjectivity by measuring the depth of the lingual fossa. However, they had little success because of issues with the precision of the method (Dahlberg and Mikkelsen, 1947, Carbonell, 1963; Goaz and Miller, 1966; Hanihara, 1969). Later Scott (1973) developed an eight degree scale that was then adapted and integrated by Turner and collaborators (1991) into a formal system for scoring non-metric aspects of dental morphology: the Arizona State University Dental Anthropology System (ASUDAS) (Scott, 1973; Turner et al., 1991; Scott and Turner, 1997). This widely-used standard protocol is based on reference plaster plaques representing the casts of selected teeth showing a gradient of expression of a particular trait (Turner et al., 1991; Scott and Turner, 1997; Scott et al., 2018). Since their initial publication, the number of traits and plaques have increased and some of them have been adapted to the range of variation expressed by fossil hominins (Bailey, 2006; Bailey and Hublin, 2013; Irish et al., 2013; Irish and Scott, 2016). The scoring and analysis of dental non-metric traits currently represents a common diagnostic procedure to highlight ancestry/genetic affinities and investigate human variation in anthropological, paleoanthropological and forensic studies (Turner et al., 1991; Scott and Turner, 1997; Irish, 1998; Irish, 2014a; Irish and Guatelli-Steinberg, 2003; Bailey and Hublin, 2013; Zanolli, 2013; Zanolli et al., 2014; Irish and Scott, 2016). If the observer has been trained by an expert, the ASUDAS approach to morphological characters is easy, fast and reliable, and remains the gold standard technique today (Scott and Irish, 2017; Scott et al., 2018). However, visual scoring, even when controlled by strict definitions of features and the experience of the observer, includes an unavoidable part of subjectivity. In fact, the

1  
2  
3 assessment of shoveling defined by the ASUDAS method has some major limitations inherent  
4 to its definition. The specimens selected to develop the reference grades on the plaque were  
5 chosen by qualitative appreciation, which does not necessarily represent the morphological  
6 variation in a linear way. This can lead to minimal visual difference between some grades of  
7 expression and so to the difficulty experienced by users in classifying the analyzed specimens  
8 with regard to the ASUDAS (especially for beginners). In brief, both the selection of the  
9 reference teeth when creating the ASUDAS method and the comparison of the dental  
10 specimens with the ASUDAS plaques are dependent on observations/palpations and the  
11 experience of the observer (i.e., dependent on operator subjectivity). Nichol and Turner  
12 (1986) have shown that the intra-observer error when assessing the expression of incisor  
13 shoveling is small: 4.1% for more than 1 grade difference and only 2% for presence/absence  
14 differences. However, as mentioned by Scott and Turner (1997), "it will probably never be  
15 possible to attain 100% concordance in replicated observations of tooth crown and root traits,  
16 either by single observers or between observers. The reference plaques developed by  
17 Dahlberg, K. Hanihara, Turner, and others have enhanced observational precision but they  
18 have not been a panacea for the reasons noted above (i.e., threshold expressions, post-eruptive  
19 modifications, surficial noise, varying levels of experience, etc.)" (Scott and Turner, 1997:72).

20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
Incisor shoveling is one of the non-metric features that has received the most attention  
from anthropologists as an indicator of relationships among populations and it is frequently  
used for its taxonomic and phylogenetic relevance (e.g., Scott and Turner, 1997; Bailey and  
Hublin, 2013; Irish et al., 2013, 2014; Carter et al., 2014). This feature can be defined as the  
degree of elevation of the mesial and distal lingual marginal ridges on the lingual surface of  
the maxillary incisors, canines and mandibular incisors, with more pronounced forms  
enclosing a fossa (Hrdlička, 1920; Dahlberg, 1956; Turner et al., 1991, Scott and Turner,  
1997). Shoveling is more marked and variable in the upper central and lateral incisors, the  
former being the polar tooth (Irish and Scott, 2016). Ales Hrdlička (1920) was the first to  
classify the degree of expression of shovel shaped incisors, assess this variation among  
several human populations and describe its occurrence in non-human species (Scott and  
Turner, 1997). Among his findings, he indicated that the prevalence and expression of incisor  
shoveling showed marked geographic variation in modern human populations, being frequent  
and strongly expressed in Asia, with a South to North increasing cline, but less frequent and  
weaker in Africa and Europe (Mizogushi, 1985; Turner, 1990; Kimura et al., 2009). Some  
workers have attempted to quantify the degree of development of the shoveling with an  
interval scale. Dahlberg and Mikkelsen (1947) used a Vernier scale with a modified Boley

1  
2  
3 Gauge to measure the depth of the incisor lingual fossa in millimeters. Hanihara et al. (1975)  
4 measured lingual fossa depth in a Japanese population in order to obtain metrical data to  
5 calculate intrafamilial correlations. Taverne et al. (1979) tried to measure various parts of a  
6 tooth crown surface by an indirect three-dimensional measurement method using  
7 photogrammetry and a Moiré pattern (Mizoguchi, 1985). Also, in a shovel-shaped tooth, the  
8 marginal ridges may extend from the incisal edge to the basal eminence and sometimes, in  
9 very pronounced cases, the ridges can converge on the eminence. In addition, the two  
10 marginal ridges may exhibit different degrees of expression (Mizoguchi, 1985). However,  
11 according to Scott and Turner (1997), the mesial and distal marginal ridges are so strongly  
12 correlated that they can be considered together as a single trait (Scott and Irish, 2017).  
13 Crummett also tried to summarize the main characteristics of incisor shoveling by considering  
14 three aspects: the expression of the marginal ridges, the development of a lingual tubercle at  
15 the lingual base of the crown, from a small swelling to an independent cusp, and the crown  
16 convexity or curvature (Crummett, 1994, 1995). More recently, using X-ray  
17 microtomographic imaging, Denton investigated the relationship of these three aspects  
18 between the external surface of the incisor crown and the enamel-dentin junction in a limited  
19 sample of 10 extant humans (Denton, 2011).

20  
21  
22 Although the expression of dental non-metric features may be sensitive to environmental  
23 or epigenetic factors (Mizoguchi, 2013), it is predominantly determined by genetic factors  
24 (Scott and Turner, 1997; Jernvall et al., 2000; Salazar-Ciudad and Jernvall, 2002, 2010; Park  
25 et al., 2012). To date, the best known genetically-correlated dental trait is incisor shoveling,  
26 which involves a single nucleotide polymorphism (SNP) of the ectodysplasin A receptor gene  
27 (*EDAR*), the most likely target of positive selection in Asian populations resulting in marked  
28 shovel shaped teeth (Kimura et al., 2009, 2012). However, *EDAR* has pleiotropic effects and a  
29 recent study suggested that it was selected in Asian groups for its effect of increasing ductal  
30 branching in the mammary gland, thereby amplifying the transfer of critical nutrients to  
31 infants via the mother's milk (Hlusko et al., 2018). In this case, the dental phenotypic  
32 expression associated with this gene could simply represent a side effect. In any case,  
33 shoveling constitutes a critical marker to discriminate between human groups and assess  
34 ancestry.

35  
36  
37 The objective of this contribution is to propose a new and complementary quantitative  
38 methodological approach to study the concavity of the palatal surface of UI1, used here as a  
39 proxy for the variation of the degree of expression of shoveling. We elaborate a geometric  
40 morphometric (GM) method taking the depth and shape of the labial incisor crown aspect

1  
2  
3 (i.e., two of the three aspects of shoveling: the expression of the marginal ridges and the  
4 curvature of the lingual aspect) into account to assess the degree of UI1 shoveling on a  
5 continuous scale. After comparison with the classical ASUDAS method, we discuss the  
6 implications of implementing such geometric morphometric analyses for the study of the  
7 modern human variability of dental traits and to better track evolutionary trends in hominins.  
8  
9  
10

## 11 **Material and methods**

### 12 *Sample and scanning procedures*

13  
14  
15 Our sample consisted of 87 modern human permanent maxillary central incisors (UI1). It  
16 included specimens of European (n=44), South African (n=30) and Chinese (n=13) ancestry,  
17 as listed in Table 1. Only unworn to moderately worn tooth crowns (reaching maximum stage  
18 2 as defined by Smith, 1984, and corresponding to a thin line of dentine exposure) having no  
19 particular damage or pathology on the labial aspect were included in the analyses. Visual  
20 scoring of the 87 specimens was achieved by two observers (DC and CZ) following the  
21 ASUDAS method (Supporting Information Table 1).  
22

23  
24 We also analyzed the original ASUDAS UI1 shoveling (ASU-UI1) plaque based on  
25 Dahlberg's work (1956) and developed by Turner and collaborators (1991). This plaque  
26 includes seven grades of shoveling expression, from the weakest (grade 0) to the most marked  
27 (grade 6) (Supporting Information Table 2). In a recent revision of the ASUDAS method,  
28 Scott and Irish (2017) described an 8<sup>th</sup> stage for UI1 shoveling (grade 7, defined as any  
29 expression that exceeds grade 6, involving marginal ridges that fold around on themselves,  
30 similar to grade 6 on the UI2 shoveling plaque) but they did not find any suitable example to  
31 put on the plaque. For this reason, we did not consider this last grade here.  
32  
33  
34  
35  
36  
37  
38  
39

40 The 44 European specimens were scanned by X-ray microtomography (micro-CT) at the  
41 CIRIMAT facility of the University of Toulouse with a Phoenix/GE Nanotom 180 instrument,  
42 using the following parameters: 100 kV, 100  $\mu$ A, 0.36° angular step. The virtual records were  
43 reconstructed to a voxel size of 22 to 25  $\mu$ m. The 30 South African teeth and the reference  
44 plaque ASU-UI1 were scanned by X-Ray Micro-CT at the MIXRAD facility of the South  
45 African Nuclear Energy Corporation SOC Limited (Necsa), with a Nikon XT H225-L  
46 instrument by using similar parameters, and reconstructed to a voxel size ranging from 42 to  
47 50  $\mu$ m (Hoffman and de Beer, 2012). The 13 modern human Asian teeth were scanned by X-  
48 Ray micro-CT using similar parameters at the Institute of Vertebrate Paleontology and  
49 Paleoanthropology of Beijing, China, and reconstructed to a voxel size of 31.4  $\mu$ m.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Data were imported into the 3D analytical software Avizo v.8.0. (FEI Visualization  
4 Sciences Group) so that 3D renderings of the tooth external surface could be visualized and  
5 processed. Teeth were first segmented semi-automatically by using a thresholding approach  
6 (Spoor et al., 1993; Fajardo et al., 2002; Coleman and Colbert, 2007) and a surface was  
7 generated from the segmented object. The maximum of curvature was measured on each UI1  
8 crown surface using the "MaxCurvature" module of Avizo. This allowed us to determine the  
9 extreme curvature line of the mesial and distal lingual crests and use these maxima as starting  
10 and ending points of our GM analyses. The cervical best fit plane was defined by placing at  
11 least three landmarks at the most apical points of the cervix on the labial and palatal aspects  
12 (points of maximum curvature on the labial and lingual sides of the cement enamel junction;  
13 Le Cabec et al., 2013). We translated this reference plane to the midpoint between the most  
14 incisal and the most cervical points of the crown (Figure 1a) and then placed 100  
15 semilandmarks along this middle plane following the curve of the lingual aspect of the crown  
16 (Figure 1b).

### 27 *Statistical analysis*

28  
29 Intra-observer error (reliability) of the ASUDAS visual scoring was assessed with respect  
30 to the UI1 reference plaque by intra-class correlation (ICC) using a two-way mixed effects  
31 "absolute agreement" model (Koo and Li, 2017). ICC is generally used to assess the  
32 correlation of various units organized in groups and describes how strongly units in the same  
33 group resemble each other. This analysis was done in order to check for both consistency  
34 (also referred to as precision in the literature; e.g., if a tooth is actually ASUDAS category 2,  
35 but two raters independently assign it to category 5, they are highly consistent with each other  
36 but they have a large bias of 3 units; Schrouf and Fleiss, 1979; Joint Committee for Guides in  
37 Metrology, 2008; Hughes and Hase, 2010) and accuracy (i.e., looking for the degree of  
38 bias/error between observers and our objective landmark-based method, e.g., if grade 3  
39 actually corresponds to grade 3 plaque).

40  
41 To objectively compare the degree of concavity of the labial surface of each incisor from  
42 our sample with the grades of the reference plaque, we performed Procrustes analyses of the  
43 semilandmarks. In the Procrustes method, the original landmarks from all samples are first  
44 superimposed and aligned with one another to produce the Procrustes coordinates.  
45 Subsequently, a principal component analysis (PCA) of the Procrustes coordinates is  
46 performed.

1  
2  
3 The reliability of this computer-based technique was assessed by intra-class correlation  
4 (ICC) of the 100 landmark coordinates among the three operators and 30 samples. Reliability  
5 was higher when the distance between the landmarks assigned by two raters on the same  
6 sample was small. We considered the distance between the landmark and the origin of the 3D  
7 orthonormal reference as outcome, together with the individual X, Y and Z float coordinates.  
8 ICC was obtained after a two-way random effects “absolute agreement” model (Koo and Li,  
9 2017). Levels of agreement between raters were also visually appreciated using Bland-Altman  
10 plots. (This kind of plot, assessing the degree of agreement between two observers, is similar  
11 to a Tukey mean-difference plot.)  
12  
13  
14  
15  
16

17 We also superimposed the curves in a non-Procrustes way, aligning the first and last point  
18 (0 and 100 respectively) of each curve (Figure 2a). This alignment procedure requires an  
19 initial 3D rotation step, which is similar to Procrustes methods, but the subsequent steps are  
20 different from Procrustes. Since each curve lies in an approximate 3D plane (the  
21 semilandmarks are placed along a plane, see Figure 1b), the curve is rotated to approximately  
22 align with the X-Y plane (equation of the 3D plane is obtained by fitting a linear regression on  
23 the X, Y & Z coordinates, as the equation of a 3D plane is  $aX + bY + cZ + d = 0$ ). After this  
24 alignment, the Z coordinate is discarded. The 2D data is then rotated again so that the first and  
25 last points lie on the Y axis (i.e. X coordinate = 0). Finally, they are scaled by a constant  
26 factor on both axes to achieve a fixed Y-axis range of 1, i.e. the first and last points of each  
27 curve now have coordinates (0,0) and (0,1) respectively. This scaling step is a major  
28 difference from the Procrustes method, as the classical Procrustes method scales to have a  
29 centroid size of 1, but here the curves are scaled to have a Y-axis range of 1 to facilitate  
30 comparisons. We then used these aligned coordinates to measure two metrics, the maximum  
31 depth of the lingual aspect with respect to the first and last points of the curves and the hollow  
32 area of the curves (Figure 2b). These metrics are not data-dependent like the previous one (PC  
33 scores have to be calculated in the whole sample, and values change according to the sample  
34 composition) as both the depth and hollow area can be measured directly on the aligned  
35 landmarks. Principal components were also calculated from the non-Procrustes aligned  
36 coordinates (X & Y).  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

50 All statistical analyses and graphic data visualization were performed in MATLAB  
51 R2017b (MATLAB and Statistics Toolbox Release, R2017b) and R 3.4.3 (R Core Team,  
52 2018). The following R packages were used: scatterplot3d (Ligges and Mächler, 2003),  
53 shapes (Dryden, 2017), ade4 (Dray and Dufour, 2007), irr (Gamer et al., 2012) and Bland  
54 Altman Leh (Lehnert, 2015).  
55  
56  
57  
58  
59  
60

1  
2  
3 Most morphological variations in the human dentition vary on a continuous scale.  
4 However, for simplicity of representation, dental anthropological assessment schemes often  
5 use two or more categories into which the range of variation is 'binned' or 'categorized'. For  
6 instance, the amount of melanin pigmentation in the eye is a continuous quantity, but in  
7 traditional analyses it has been categorized into blue vs. brown to represent absence/presence  
8 of melanin, and historically considered to be a Mendelian trait until modern quantitative  
9 analysis showed its complex polygenic nature. Scott and Turner (1997) noted, specifically  
10 with the example of incisor shoveling, that such nonmetric dental traits are possibly 'quasi  
11 continuous' (ordinal or dichotomous) traits, derived from an underlying continuous trait. For  
12 example, while the depth of the incisor crown is a continuous quantity, it can be dichotomized  
13 into absence/presence indicating whether the amount of curvature is below a certain threshold.  
14 In such cases, the underlying continuous variable is called a 'latent variable' corresponding to  
15 the assessed categorical variable. Our analyses suggested that the maximum depth metric was  
16 the most likely candidate for any underlying 'latent' quantitative variable that might be the  
17 basis of the ASUDAS categories for shoveling (see Results below). The results also suggested  
18 that the relationship between the maximum depth variable and the ASUDAS categories was  
19 monotonic but non-linear, i.e. when the latent variable increases the categories also increase,  
20 but the spacing between categories is unequal. Therefore, to 'predict' an objective ASUDAS  
21 category for the 87 modern human specimens, we constructed a prediction function using the  
22 maximum depth values and numerical categories of the ASUDAS specimen teeth. To  
23 preserve nonlinearity, a spline function was fitted on these values, which was then used as the  
24 interpolant to obtain predicted ASUDAS categories from the 3-D measured maximum depth  
25 values on the 87 modern human specimen casts. The predicted ASUDAS categories were  
26 allowed to contain decimals to retain more precision, instead of rounding them off to the  
27 nearest integer category (e.g., 1.67 instead of 2). Similarly, some of the observer-assigned  
28 categories that had an intermediate rating (0\_1 meaning a category between ASUDAS  
29 references 0 and 1) were allowed to retain them (the rating 0\_1 was assigned the middle value  
30 of 0.5, for example). These objective predicted values were compared with subjective  
31 observer-assigned rating values via ICC to obtain accuracy (i.e. unbiasedness) measurements.  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50

## 51 **Results**

52 Our ICC intra- and inter-observer tests on the visual scoring showed highly consistent  
53 assessment of shoveling within and between raters (Table 2). Consistency was high for the  
54 whole sample, but also for each chrono-geographic sub-sample considered in this study. In  
55  
56  
57  
58  
59  
60

1  
2  
3 contrast, the accuracy of the visual assessment performed by the raters, measured using the  
4 ICC between the rater-assigned category and the predicted category was moderate for the  
5 Europeans and Africans. For the Chinese sample, the accuracy of the visual assessment was  
6 very low (further discussion below).  
7  
8

9 We then looked at the results of the landmark-based analyses. Following the standard  
10 procedure in geometric morphometrics, principal component analysis (PCA) was used to  
11 explore the morphospace. Principal components (PCs) were calculated using the Procrustes  
12 coordinates of all landmarks, for all samples including the ASUDAS reference casts. Plotting  
13 the top PCs enabled us to visualize the morphospace, to see how the samples were distributed  
14 with respect to the ASUDAS grades (Figure 3). The first two components (PC1 and PC2) of  
15 the PCA accounted for 87.4% of the total variance (81.5% for PC1 and 5.9% for PC2).  
16  
17  
18  
19

20 PC shape changes could be visualized by plotting the PC loadings. PC loadings for the  
21 Procrustes and non-Procrustes methods were very similar; non-Procrustes PC shape changes  
22 are shown in Supporting Information Figure 1. While PCs are more difficult to interpret than  
23 direct measurements, the PC shape changes give us some idea of the morphological aspects  
24 they capture. The shape of PC1 is roughly proportional to the curve of the lingual aspect, i.e.  
25 the lingual fossa has the highest weight. Therefore the deeper the lingual fossa is from the  
26 baseline, the higher the PC value is. This explains why the PC has such high correlations with  
27 the maximum depth metric (Table 3). And since the deeper the lingual fossa is, the more the  
28 marginal ridges protrude with respect to the fossa, PC1 is also proportional to the ASUDAS  
29 shoveling grade (Table 4).  
30  
31  
32  
33  
34  
35

36 The shape of PC2 gives greatest weight to the corners of marginal ridges, thereby being  
37 proportional to the angle between the labial palate and the marginal ridges. This angle is  
38 called ‘labial convexity’ in Denton (2011), where it is shown that its relationship with the  
39 ASUDAS shoveling grade is not monotonic (largest angles for grades 2-3), which explains  
40 why PC2 is not strongly correlated with it (Table 4).  
41  
42  
43  
44

45 The shape of PC3 seems to reflect the left-right asymmetry present in the shape of the  
46 curve, in particular the asymmetry in the two angles. As asymmetry of the ridges is not  
47 relevant in the definition of the ASUDAS grades, this explains why PC3 is not correlated with  
48 the grades either (Table 4).  
49  
50

51 The shapes of later PCs, such as PC4, are much harder to interpret and, given that they  
52 capture a very small fraction of the variability, they might simply reflect random statistical  
53 variation.  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Along PC1, the European (French contemporary and medieval) and South African  
4 specimens showed a similarly reduced expression of shoveling (expressed here by reduced  
5 lingual depth and a more linear morphology), overlapping with the ASUDAS grades 0 and 1,  
6 while the Chinese material encompassed the grades 2 to 5. Even though the ASUDAS grades  
7 tend to follow a trend along PC1, their distribution is not linearly organized and is  
8 heterogeneous (Figure 3a and Figure 4a). Grades 0-1, 2-3, 4-5 and 6 tended to form four  
9 clusters and grades were not equidistant from one another. There was no visible  
10 discrimination between the chrono-geographic human samples and the ASUDAS grades  
11 along PC3 and PC4, which represented 3.16% and 2.55%, respectively, of the total variance  
12 (Figure 3b). We tested the reproducibility of this Procrustes method. Our results show that the  
13 positioning of the landmarks was highly reproducible, with an ICC >0.990. The graphical  
14 Bland-Altman method confirmed this high level of agreement (Figure 5).

15  
16  
17  
18  
19  
20  
21  
22 When considering the Non-Procrustes analysis, similar results were obtained. In Figure 4,  
23 the histograms showing the distribution of maximum depth (Figure 4b) and hollow area  
24 (Figure 4c) of the specimens and the ASUDAS grades also highlight the non-linear scattering  
25 of the reference grades. Again, grades 0 and 1 are close to each other, while grades 2 to 5 are  
26 grouped together and 6 is alone. In accordance with the knowledge that East Asian  
27 populations show a higher degree of shoveling than the rest, the histogram of Chinese samples  
28 for maximum depth (Figure 4b) shows no overlap with the French and South African samples.  
29 It is also interesting to note that several samples from each population have a maximum depth  
30 value that is intermediate between the ASUDAS categories 1 & 2 (which have a large gap  
31 between them). This might create some difficulties for the observers to assign ASUDAS  
32 ratings to them, and also reduce the distinction between populations when compared via  
33 ASUDAS category frequencies. However, the quantitative maximum depth measurements  
34 provide a complete separation of the Chinese samples from the rest. The hollow area metric  
35 also achieves near-complete separation and shows a similar pattern.

36  
37  
38  
39  
40  
41  
42  
43  
44  
45 Principal components calculated from the non-Procrustes aligned coordinates showed  
46 trends similar to those of the Procrustes PCs. PC1 alone explained 92% of the variation, while  
47 PC2 explained a further 2.7%. The PC loadings are represented as PC shape changes in  
48 Supporting Information Figure 1.

49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60 Table 3 presents the correlation between these various metrics calculated on the 87 human  
specimens plus the 7 ASUDAS reference casts. In addition to the two direct measurements,  
maximum depth and hollow area, PC1 and PC2 are used from both Procrustes and non-  
Procrustes aligned methods. This indicates that the two direct metrics are very similar to each

1  
2  
3 other, so it is sufficient to use either one. It also shows that both direct metric measurements,  
4 maximum depth and hollow area, have very high correlation with PC values obtained from  
5 the two analyses, as expected from the observations above, and therefore can be used in lieu  
6 of Procrustes PC values. The advantage of using direct measurement metrics over Procrustes  
7 PCs is that they are more directly interpretable, and not dependent on the whole dataset.  
8  
9

10  
11 The correlation of ASUDAS category values with these metrics calculated on the  
12 ASUDAS reference casts was also assessed (Supporting Information Table 3). It shows that  
13 both metrics, as well as PC1 from both analyses, are highly correlated with the category  
14 values, while maximum depth has the highest correlation (98%). It also indicates that the  
15 Procrustes and non-Procrustes PCs are very similar to each other. The major feature of the  
16 ASUDAS scale for shoveling is the progressively increasing expression of the marginal ridges  
17 (Scott and Turner, 1997; Scott and Irish 2017). Conversely, more protruding ridges imply a  
18 deeper lingual fossa, and hence the maximal depth metric gives a measure of the development  
19 of the mesial and distal ridges and constitutes an appropriate metric to quantitatively evaluate  
20 the expression of shoveling.  
21  
22  
23  
24  
25  
26

27 When combining the correlation patterns with the observation above that the maximum  
28 depth metric provides a complete separation of the Chinese samples from the rest, we  
29 considered the maximum depth to be the most likely candidate for any underlying ‘latent’  
30 quantitative variable that might be the basis of the ASUDAS categories. Considering that  
31 maximum depth has a 99% correlation with non-Procrustes PC1, which explains 92%  
32 variation of the landmark coordinates, it can be said that the maximum depth metric captures  
33 a large part of the morphological variation of the labial aspect of the U11 crown.  
34  
35  
36  
37

38 To assess the accuracy of the visual ASUDAS assessments made by two observers, we  
39 needed to compare them with objective ‘true’ assessments of the modern human specimens.  
40 To make such objective estimates of ASUDAS categories on the specimens, we started with  
41 the objectively measured maximum depth metric, and employed the previously described  
42 prediction function. The prediction procedure uses the spline interpolation function  
43 constructed entirely on the basis of the ASUDAS reference plaque (Figure 6), and is therefore  
44 free of subjectivity arising from other human raters. The spline model compares the maximum  
45 depth measurements against ASUDAS categories for the ASUDAS reference plaque  
46 elements. Using this objective prediction function on the objectively measured maximum  
47 depth metric for a new specimen provides an objective estimate of the ASUDAS grade for the  
48 new specimen. This estimated grade can then be compared to the subjective ASUDAS ratings  
49 provided by a human observer in order to assess the accuracy of the observer.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 As found in our ICC intra- and inter-observer tests (Table 2), the specimens the observers  
4 assessed as expressing low degrees of shoveling (grades 0-1), namely most of the South  
5 African and French samples, were closer to the predicted 'true' values (using the ASUDAS  
6 reference grades), being less scattered around the spline interpolation curve used for the  
7 prediction (Figure 6). In contrast, observer ratings for those recorded as having marked  
8 shoveling (grade 2 and above) – primarily the Chinese samples – were much more scattered  
9 around the spline curve, indicating that they differed more from the predicted 'true' ASUDAS  
10 grades (Figure 6). This corroborates our observation above that the populations are harder to  
11 separate on the categorical ASUDAS scale than by means of the quantitative maximum depth  
12 metric. Still, no Chinese specimen was visually rated below grade 2, and very few European  
13 and African specimens were recorded as grade 2 or above, which indicates that, despite the  
14 subjective variability, the samples can usually be dichotomized by human raters into high or  
15 low shoveling with reasonable accuracy.  
16  
17  
18  
19  
20  
21  
22  
23

24 When comparing the correlation between the visual assessments (based on the ASUDAS  
25 definition and plaque) and some objectively obtained measurements (e.g., predicted ASUDAS  
26 values, maximum depth, hollow area), we found that the maximum depth and hollow area  
27 metrics correlated highly with visual scoring, even to a higher level than with the predicted  
28 ASUDAS values (Table 4). This could suggest that there is a major unconscious reaction to  
29 the maximum depth aspect when recording shoveling by following the classical ASUDAS  
30 method. In order to test whether the observers were more influenced visually by the depth of  
31 the palatal aspect than by the global morphology, we dichotomized the observer ratings  
32 following the standard protocol for ASUDAS traits (Scott and Irish, 2017, Scott et al., 2018),  
33 which splits the categories of this trait into two broad groups (grades  $\leq 1$  vs. grades  $\geq 2$ ), and  
34 re-ran ICC consistency and accuracy measures (Table 5). This grouping mimicked the  
35 distinction between East Asians and the rest by separating absent/low degrees of shoveling  
36 and marked shoveling. The new ICC values showed high degrees of precision and accuracy,  
37 suggesting that the visual scoring performed by observers was largely successful at separating  
38 low degrees of shoveling from marked shoveling, but not so successful in detecting finer  
39 differences between the ASUDAS casts. This result is relevant as it shows that, while the  
40 ASUDAS method is efficient to distinguish below and above the breakpoint grade 2, it is  
41 more prone to bias when dealing with close grades and limits the possibilities for more  
42 advanced analyses. For example, an important question regarding dental traits is to find which  
43 genetic factors are responsible for a trait's expression. Scott and Turner (1997) used the  
44 example of incisor shoveling to conclude that such nonmetric dental traits possibly arise from  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 an underlying continuous trait which is likely polygenic. They also note that dichotomizing  
4 such traits leads to the loss of a large amount of variation in the trait. As an extreme example,  
5 all Native Americans may have the constant value of 'present' for the shoveling trait, which  
6 causes it to "lose its status as a nonmetric variant as it is present in all individuals." Such loss  
7 of variation, either when constructing a 'quasi continuous' (ordinal) trait from a continuous  
8 underlying trait, or when dichotomizing an ordinal trait, reduces the resolution of the data.  
9 This loss of variation leads to a loss of power in genetic association analysis, where the use of  
10 a continuous trait can entail 'significantly higher power', especially with small sample sizes  
11 (Bhandari et al., 2002). In their literature review, Scott et al. (2018) also note that such  
12 simplifications can create several problems for genetic analyses, e.g. the simplified traits can  
13 "sometimes mimic the segregation patterns of simple Mendelian inheritance where, in reality,  
14 inheritance is complex." (Scott et al., 2018:133). Even in the context of assessing rater  
15 reliability, the higher resolution offered by continuous data provides much better reliability  
16 estimates (Donner and Eliasziw, 1994).

17  
18 In the context of clinical studies, Altman (2006) comments that dichotomized variables  
19 often appear to be more alluring as they simplify the data while retaining the main dichotomy  
20 that is thought to be the crux of the variable, thereby leading to simpler interpretations as well  
21 as higher rater agreement. Yet such deliberate discarding of data causes several problems: loss  
22 of power, increased risk of false positives, underestimation of variation within or between  
23 groups, loss of information about the relationship between the trait and other variables, and  
24 increased confounding with other variables in regression analysis, such as genetic association  
25 analysis. MacCallum et al. (2002) report similar criticisms, and emphasize the problems in  
26 statistical analysis. Using a dichotomous variable means that many statistical procedures are  
27 not applicable, e.g., in a genetic association study, the standard linear regression model, which  
28 allows estimation of effect sizes in absolute units, cannot be used. Dichotomous variables can  
29 only be used with logistic regression, which only estimates effect sizes on an odds ratio scale,  
30 making it much more difficult to interpret the effect of a genetic marker or combine evidence  
31 from multiple studies.

32  
33 The sample-dependent nature of Principal Components means that the PCs depend on the  
34 whole sample composition. We investigated how the PCs could change when the dataset was  
35 limited to European samples, which span only a narrow range of the trait variability (without  
36 any marked shoveling). To assess this, we performed the alignment and PC calculation steps  
37 for both Procrustes and non-Procrustes methods while restricting the dataset to French  
38 samples (contemporary and medieval), and obtained their correlation with the directly  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 measured metrics. For these correlation values, the correlation of PC1 from both analyses  
4 with maximum depth decreased a little (compared to Table 3), especially for the Procrustes  
5 PC (Supporting Information Table 4). This indicates that PCs calculated in this reduced  
6 dataset with a narrower range of variability might have some increased noise or may be  
7 slightly less able to capture the actual underlying metric of lower grades. To verify this, we  
8 looked at the correlation of rater-assigned ASUDAS categories of the French samples with  
9 these PCs and directly measured metrics (Supporting Information Table 5). In contrast to  
10 Table 4, correlations of the rater-assigned ASUDAS categories with maximum depth are  
11 higher than their correlations with PC1 by a larger margin.  
12  
13  
14  
15  
16  
17  
18

### 19 **Discussion**

20 In the literature, the significance of errors due to the role of the observer in the visual scoring  
21 of dental non-metric traits using the ASUDAS method is usually considered as being low  
22 and/or negligible (Scott and Turner, 1997; Bailey and Hublin, 2013; Hillson, 2005). Applied  
23 to decades of research, the classical ASUDAS method has proved to be quite efficient for  
24 inferring biological relationships among modern humans (Scott and Turner, 1997; Scott et al.,  
25 2018), living non-human primates (Pilbrow, 2003) and fossil hominins (Mizoguchi, 1985;  
26 Crummet, 1994; Bailey and Hublin, 2013; Irish et al., 2013; Bailey and Wood, 2007;  
27 Martinon-Torres et al., 2007) – in part because, once the grades are scored, they are  
28 dichotomized into presence/absence to help reduce observer error and because of the current  
29 dichotomous biological distance statistics available (e.g., MMD). There has been a previous  
30 attempt to link morphology and measurements for the UI1 shoveling trait, notably by  
31 considering the depth of the lingual fossa with respect to ASUDAS grades (Hanihara, 2008).  
32 However, this method only considers the maximum depth at the center of the lingual fossa  
33 and does not quantify the shape of this fossa. Thus, it is still possible to develop innovative,  
34 complementary methods. The recent development of quick and efficient methods for  
35 acquiring 3D models of an object (e.g., photogrammetry, laser scanner), together with the  
36 advent of powerful quantitative techniques to assess shape variation (geometric  
37 morphometrics), has opened up new ways to test the reliability (precision and accuracy) of the  
38 ASUDAS method. These methods represent an opportunity to provide objective protocols to  
39 investigate non-metric dental variation. In this preliminary study, we have compared the  
40 classical plaque-based visual scoring assessment with a new 3D geometric morphometric  
41 approach. We propose here a simple, fast method based on geometric morphometrics to  
42 characterize a sample of modern human UI1s using a continuous scale of morphological  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 variation of shoveling. The intra-observer error related to the visual scoring is very low, as  
4 previously demonstrated for ASUDAS plaques (Nichol and Turner, 1986; Scott and Turner,  
5 1997; Scott, 2008; Scott and Irish, 2017). As anticipated, our results highlight some  
6 limitations of the use of the current ASUDAS plaque, indicating that it did not necessarily  
7 represent an objective gradient of expression of a non-metric tooth feature (Figures 3, 4 and  
8 6). Our results also agree with the currently recognized ASUDAS breakpoint between the  
9 recorded absence (grades 0-1) and presence (grades 2-7) of shoveling (Scott and Irish, 2017,  
10 Scott et al., 2018). This method can also distinguish between the French and South African  
11 groups (expressing low degrees of shoveling) and the Chinese sample (being more variable  
12 but mostly showing well-defined shovel-shaped incisors). This is in agreement with the vast  
13 literature on the topic (e.g., Scott and Turner, 1997; Irish and Scott, 2016; Scott and Irish,  
14 2017; Scott et al., 2018) and demonstrates that our method, while confirming the ASUDAS  
15 results, opens a path towards more advanced quantitatively-based assessment for the  
16 distinction of fossil and extant human populations. This modest sample was only used here to  
17 test the method, but by increasing it and incorporating larger chrono-geographic groups,  
18 including fossil hominins, there is a high potential to better understand the evolution of  
19 shovel-shaped incisors. For example, Neanderthals are well-known for their markedly shovel-  
20 shaped incisors and, given the increasing availability of 3D virtual data on their teeth,  
21 paleogenetics techniques and molecular data on tooth morphology (Zanolli et al., 2017), this  
22 new quantitative method is perfectly suited to the investigation of the evolution of UII  
23 shoveling.

24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37 Our protocol integrates the analysis of two different but complementary aspects: the depth  
38 of the lingual surface with respect to the marginal ridges and the shape of the lingual aspect.  
39 This is an important point as our analyses have revealed that visual rating of shoveling tends  
40 to be more prone to intra- and inter-observer bias for the highest grades (even starting at grade  
41 2). In addition, even when the observers are well trained and follow the definition of the UII  
42 shoveling trait (Supporting Information Table 2), when dealing with numerous specimens,  
43 they tend to create a mental image of the ASUDAS categories and then make their judgments,  
44 resulting in a mental scale that is linearly dependent on the maximum depth of the palatal  
45 aspect, while the ASUDAS grades are not distributed linearly for this parameter. This results  
46 in the visually assigned ratings being correlated with the maximum depth rather than with  
47 predicted ASUDAS categories. In this context, our results provide a reliable, reproducible  
48 framework reinforced by statistical results supporting the fact that open scale numerical  
49 measurements can complement the ASUDAS method and provide new information. Of  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 course, similar methods complementing the classic ASUDAS method still need to be  
4 developed for other non-metric dental traits. There are also other possibilities for the  
5 quantitative study of shape variation, with or without landmarks. For example, a surface  
6 deformation-based approach considering a 3D portion of the crown surface (such as the  
7 lingual aspect in the case of UI1 shoveling) could be used to assess the degree of deformation  
8 from one tooth to another and quantify shape variations of the complete set of dental traits  
9 (Durrleman et al., 2012; Durrleman et al., 2014). Thus, although the ASUDAS is a reliable and  
10 efficient tool, it is still possible to complement it with alternative methods.  
11  
12  
13  
14  
15

### 16 17 **Acknowledgements**

18 This study is based on the PhD research program of the first author. It is supported by the  
19 French CNRS. The work of P. Monsarrat is supported by Toulouse University Hospital (CHU  
20 de Toulouse), Université Toulouse III - Paul Sabatier, the Midi-Pyrénées region and the  
21 research platform of the Toulouse Dental Faculty (PLTRO). J. Braga provided access to the  
22 ASUDAS plaque. We thank G. Krüger for granting access to the Pretoria Bone Collection  
23 (PBC) used in this study. For scientific discussion, we are also grateful to F. Duret.  
24  
25  
26  
27  
28  
29

### 30 31 **References**

- 32 Altman, D.G. (2006). The cost of dichotomising continuous variables. *The British Medical*  
33 *Journal*, **332**, 1080.  
34  
35 Bailey, S.E., & Wood, B.A. (2007). Trends in postcanine occlusal morphology within the  
36 hominin clade: the case of *Paranthropus*. In S.E. Bayley & J.-J. Hublin (Eds.), *Dental*  
37 *Perspectives on Human Evolution: State of the Art Research in Dental Anthropology* (pp.  
38 33-52). Dordrecht: Springer.  
39  
40 Bailey, S.E., & Hublin J.-J. (2013). What does it mean to be dentally “modern”? In G.R. Scott  
41 & J.D. Irish (Eds.), *Anthropological Perspectives on Tooth Morphology, Genetics,*  
42 *Evolution, Variation* (pp. 222-249). Cambridge: Cambridge University Press.  
43  
44 Bhandari, M., Lochner, H., & Tornetta P. (2002). Effect of continuous versus dichotomous  
45 outcome variables on study power when sample sizes of orthopaedic randomized trials are  
46 small. *Archives of Orthopaedic and Trauma Surgery*, **122**, 96-98.  
47  
48 Carbonell, V.M. (1963). Variations in the frequency of shovel-shaped incisors in different  
49 populations. In D.R. Brothwell (Ed.), *Dental Anthropology* (pp. 211-234). New York:  
50 Pergamon Press.  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Carter, K., Worthington, S., & Smith, T. (2014). Non-metric dental traits and hominin  
4 phylogeny. *Journal of Human Evolution*, **69**, 123-128.
- 5  
6 Coleman, M.N., & Colbert, M.W. (2007). CT thresholding protocols for taking measurements  
7 on three-dimensional models. *American Journal of Physical Anthropology*, **133**, 723-725.
- 8  
9 Crummett, T. (1994). *The evolution of shovel shaping: regional and temporal variation in*  
10 *human incisor morphology*. Ph.D. Dissertation, University of Michigan, Ann Arbor.
- 11  
12 Crummett, T. (1995). The three dimensions of shovel-shaping. In J. Moggi-Cecchi (Ed.),  
13 *Aspects of Dental Biology: Palaeontology, Anthropology and Evolution* (pp. 305-313).  
14 Florence: International Institute for the Study of Man.
- 15  
16  
17 Dahlberg, A.A. (1956). *Materials for the Establishment of Standards for Classification of*  
18 *Tooth Characteristics, Attributes, and Techniques in Morphological Studies of the*  
19 *Dentition*. Chicago: Zoller Laboratory of Dental Anthropology, University of Chicago.
- 20  
21  
22 Dahlberg, A.A. & Mikkelsen, O. (1947). The shovel-shaped character in the teeth of the Pima  
23 Indians. *American Journal of Physical Anthropology*, **5**, 234-235.
- 24  
25  
26 Denton, L.C. (2011). *Shovel-Shaped Incisors and the Morphology of the Enamel-Dentin*  
27 *Junction: An Analysis of Human Upper Incisors in Three Dimensions*. Master dissertation,  
28 Colorado State University, Fort Collins.
- 29  
30  
31 Donner, A., & Eliasziw, M. (1994). Statistical implications of the choice between a  
32 dichotomous or continuous trait in studies of interobserver agreement. *Biometrics*, **50**, 550-  
33 555.
- 34  
35  
36 Dray, S. & Dufour, A.B. (2007). The ade4 package: implementing the duality diagram for  
37 ecologists. *Journal of Statistical Software*, **22**, 1-20.
- 38  
39  
40 Dryden, I.L. (2017). shapes package. R Foundation for Statistical Computing, Vienna,  
41 Austria. Contributed package. Version 1.2.3. URL <http://www.R-project.org>.
- 42  
43  
44 Durrleman, S., Prastawa, M., Korenberg, J.R., Joshi, S., Trouvé, A., & Gerig, G. (2012).  
45 Topology preserving atlas construction from shape data without correspondence using  
46 sparse parameters. In N. Ayache, H. Delingette, P. Golland & K. Mori (Eds.), *Proceedings*  
47 *of Medical Image Computing and Computer Aided Intervention* (pp. 223-230). Nice:  
48 Springer.
- 49  
50  
51 Durrleman, S., Prastawa, M., Charon, N., Korenberg, J.R., Joshi, S., Gerig, G., & Trouvé, A.  
52 (2014). Morphometry of anatomical shape complexes with dense deformations and sparse  
53 parameters. *NeuroImage*, **101**, 35-49.
- 54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Fajardo, R.J., Ryan, T.M., & Kappelman J. (2002). Assessing the accuracy of high-resolution  
4 X-ray computed tomography of primate trabecular bone by comparisons with histological  
5 sections. *American Journal of Physical Anthropology*, **118**, 1-10.
- 6  
7  
8 Gamer, M., Lemon, J., & Puspendra Singh, I. (2012). irr: Various Coefficients of Interrater  
9 Reliability and Agreement. R package version 0.84. [https://CRAN.R-](https://CRAN.R-project.org/package=irr)  
10 [project.org/package=irr](https://CRAN.R-project.org/package=irr).
- 11  
12 Goaz, P.W., & Miller, M.C. (1966). A preliminary description of the dental morphology of  
13 the Peruvian Indians. *Journal of Dental Research*, **45**, 106-119.
- 14  
15  
16 Hanihara, K. (1969). Mongoloid dental complex in the permanent dentition. *Proceedings of*  
17 *the VIII<sup>th</sup> International Congress of Anthropological and Ethnological Sciences*, **1**, 298-  
18 300.
- 19  
20  
21 Hanihara, K., Masuda, T., Tanaka, T. (1975). Family studies of the shovel trait in the  
22 maxillary central incisor. *Journal of the Anthropological Society of Nippon*, **83**, 107-12.
- 23  
24 Hanihara, K. (2008). Morphological variation of major human populations based on  
25 nonmetric dental traits. *American Journal of Physical Anthropology*, **136**, 169-182.
- 26  
27 Hillson, S.W. (1996). Dental anthropology. Cambridge: Cambridge University Press.
- 28  
29 Hillson, S.W. (2005). Teeth. Cambridge: Cambridge University Press.
- 30  
31 Hlusko, L.J., Carlson, J.P., Chaplin, G., Elias, S.A., Hoffecker, J.F., Huffman, M., Jablonski,  
32 N.G., Monson, T.A., O'Rourke, D.H., Pilloud, M.A., & Scott, G.R. (2018). Environmental  
33 selection during the last ice age on the mother-to-infant transmission of vitamin D and fatty  
34 acids through breast milk. *Proceedings of the National Academy of Sciences of the United*  
35 *States of America* (in press, [doi.org/10.1073/pnas.1711788115](https://doi.org/10.1073/pnas.1711788115)).
- 36  
37  
38 Hoffman, J.W., & de Beer, F.C. (2012) Characteristics of the Micro-Focus X-ray  
39 Tomography Facility (MIXRAD) at Necsas in South Africa. In *Proceedings of the 18th*  
40 *World Conference on Nondestructive Testing* (pp. 1-12). Durban: South African Institute  
41 for Non-Destructive Testing.
- 42  
43  
44 Hrdlička, A. (1920). Shovel-shaped teeth. *American Journal of Physical Anthropology*, **3**,  
45 429-465.
- 46  
47  
48 Hughes, I., & Hase, T. (2010). *Measurements and Their Uncertainties: A Practical Guide to*  
49 *Modern Error Analysis*. Oxford: Oxford University Press.
- 50  
51  
52 Irish, J.D. (1998). Ancestral dental traits in recent Sub-Saharan Africans and the origins of  
53 modern humans. *Journal of Human Evolution*, **34**, 81-98.
- 54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Irish, J.D., & Guatelli-Steinberg, D. (2003). Ancient teeth and modern human origins: An  
4 expanded comparison of African Plio-Pleistocene and recent world dental samples.  
5 *Journal of Human Evolution*, **45**, 113-144.  
6  
7 Irish, J.D., Guatelli-Steinberg, D., Legge, S.S., Ruitter, D.J., & Berger, L.R. (2013). Dental  
8 morphology and the phylogenetic “place” of *Australopithecus sediba*. *Science*, **340**,  
9 1233062.  
10  
11 Irish, J.D. (2014). Dental nonmetric variation around the world: Using key traits in  
12 populations to estimate ancestry in individuals. In G.E. Berg & S.C. Ta'ala (Eds.),  
13 *Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black*  
14 *and White* (pp. 165-190). Boca Raton: CRC Press-Taylor and Francis Group.  
15  
16 Irish, J.D., Guatelli-Steinberg, D., Legge, S.S., Ruitter, D.J., & Berger, L.R. (2014). Response  
17 to “Non-metric dental traits and hominin phylogeny” by Carter et al., with additional  
18 information on the Arizona State University Dental Anthropology System and  
19 phylogenetic “place” of *Australopithecus sediba*. *Journal of Human Evolution*, **69**, 129-  
20 134.  
21  
22 Irish, J.D., & Scott, G.R. (2016). *A Companion to Dental Anthropology*. Chichester: Wiley  
23 Blackwell.  
24  
25 Jernvall, J., Keränen, S.V.E., & Thesleff, I. (2000). Evolutionary modification of development  
26 in mammalian teeth: Quantifying gene expression patterns and topography. *Proceedings of*  
27 *the National Academy of Sciences of the United States of America*, **97**, 14444–14448.  
28  
29 Joint Committee for Guides in Metrology. (2008). *The International Vocabulary of*  
30 *Metrology, Basic and General Concepts and Associated Terms*.  
31  
32 Kimura, R., Yamaguchi, T., Takeda, M., Kondo, O., Toma, T., Haneji, K., Hanihara, T.,  
33 Matsukusa, H., Kawamura, S., Maki, K., Osawa, M., Ishida, H., & Oota, H. (2009). A  
34 common variation in EDAR is a genetic determinant of shovel shaped incisors. *The*  
35 *American Journal of Human Genetics*, **85**, 528-535.  
36  
37 Koo, T.K., Li, M.Y. (2017). A guideline of selecting and reporting intraclass correlation  
38 coefficients for reliability research. *Journal of Chiropractic Medicine*, **15**, 155-163.  
39  
40 Le Cabec, A., Gunz, P., Kupczik, K., Braga, J., & Hubin, J.J. (2013). Anterior tooth root  
41 morphology and size in Neanderthals: taxonomic and functional implications. *Journal of*  
42 *Human Evolution*, **64**, 169-193.  
43  
44 Lehnert, B. (2015). BlandAltmanLeh package. R Foundation for statistical computing,  
45 Vienna, Austria. Contributed package. Version 0.3.1. Available from: [http://www.R-](http://www.R-project.org)  
46 [project.org](http://www.R-project.org).  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Ligges, U., & Mächler, M. (2003). Scatterplot3d - an R Package for Visualizing Multivariate  
4 Data. *Journal of Statistical Software*, **8**, 1-20.
- 5  
6 MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of  
7 dichotomization of quantitative variables. *Psychological Methods*, **7**, 19-40.
- 8  
9 Martínón-Torres, M., Bermúdez de Castro, J.M., Gómez-Robles, A., Arsuaga, J.L., Carbonell,  
10 E., Lordkipanidze, D., Manzi, G., & Margvelashvili, A. (2007). Dental evidence on the  
11 hominin dispersals during the Pleistocene. *Proceedings of the National Academy of*  
12 *Sciences of the United States of America*, **104**, 13279-13282.
- 13  
14  
15  
16 MATLAB and Statistics Toolbox Release. (R2017ba). The MathWorks, Inc., Natick,  
17 Massachusetts, United States.
- 18  
19 Mizogushi, Y. (1985). Shoveling: a statistical analysis of its morphology. *The University*  
20 *Museum, the University of Tokyo Bulletin*, **26**.
- 21  
22 Mizogushi, Y. (2013). Significant among-population associations found between dental  
23 characters and environmental factors. In G.R. Scott & J.D. Irish (Eds.), *Anthropological*  
24 *Perspectives on Tooth Morphology, Genetics, Evolution, Variation* (pp. 108-125).  
25 Cambridge: Cambridge University Press.
- 26  
27  
28  
29 Nichol, C.R., & Turner, C.G., II (1986). Intra- and interobserver concordance in classifying  
30 dental morphology. *American Journal of Physical Anthropology*, **69**, 299-315.
- 31  
32 Park, J.H., Yamaguchi, T., Watanabe, C., Kawaguchi, A., Haneji, K., Takeda, M., Kim, Y.I.,  
33 Tomoyasu, Y., Watanabe, M., Oota, H., Hanihara, T., Ishida, H., Maki, K., Park, A.B., &  
34 Kimura, R. (2012). Effects of an Asian-specific nonsynonymous EDAR variant on  
35 multiple dental traits. *Journal of Human Genetics*, **57**, 508-514.
- 36  
37  
38  
39 Pilbrow, V. (2003). *Dental variation in African apes with implications for understanding*  
40 *patterns of variation in species of fossil apes*. PhD Dissertation. New York: University of  
41 New York.
- 42  
43  
44 R Core Team. (2017). R Foundation for statistical computing. Vienna, Austria. Version 3.4.3.  
45 Available from: <http://www.R-project.org>.
- 46  
47 Salazar-Ciudad, I. & Jernvall, J. (2002). A gene network model accounting for development  
48 and evolution of mammalian teeth. *Proceedings of the National Academy of Sciences of the*  
49 *United States of America*, **99**, 8116-8120.
- 50  
51 Salazar-Ciudad, I. & Jernvall, J. (2010). A computational model of teeth and the  
52 developmental origins of morphological variation. *Nature*, **464**, 583-586.
- 53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Scott, G.R (1973). *Dental Morphology: a genetic study of American White Families and*  
4 *Variation in living Southwest Indians*. PhD dissertation, Department of Anthropology,  
5 Arizona State University, Tempe.  
6  
7 Scott, G.R. (2008). Dental Morphology. In A. Katzenburg & S. Saunders (Eds.), *Biological*  
8 *Anthropology of the Human Skeleton, 2<sup>nd</sup> Edition* (pp. 265-298). New York: Wiley-Liss.  
9  
10 Scott, G.R., & Irish, J.D. (2017). *Human Tooth Crown and Root Morphology. The Arizona*  
11 *State University Dental Anthropology System*. Cambridge: Cambridge University Press.  
12  
13 Scott, G.R., & Turner, C.G. (1997). *The Anthropology of Modern Human Teeth: Dental*  
14 *Morphology and its Variation in Recent Human Populations, 2<sup>nd</sup> Edition*. Cambridge:  
15 Cambridge University Press.  
16  
17 Scott, G.R., Turner, C.G., Townsend, G.C., & Martínón-Torres, M. (2018). *The Anthropology*  
18 *of Modern Human Teeth. Dental Morphology and Its Variation in Recent and Fossil Homo*  
19 *sapiens*. Cambridge: Cambridge University Press.  
20  
21 Shroud, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability.  
22 *Psychological Bulletin*, **86**, 420-428  
23  
24 Smith, H. B. (1984). Patterns of molar wear in hunter-gatherers and agriculturalists. *American*  
25 *Journal of Physical Anthropology*, **63**, 39-56.  
26  
27 Spoor, C.F., Zonneveld, F.W., & Macho, G.A. (1993). Linear measurements of cortical bone  
28 and dental enamel by computed tomography: Applications and problems. *American*  
29 *Journal of Physical Anthropology*, **91**, 469-484.  
30  
31 Taverne, P.P., Amesz-Voorhoeve, W.H.M., Leertouwer, H.L. (1979). A photogrammetric  
32 method which provides geometric data in dental morphology. *Zeitschrift fur Morphologie*  
33 *und Anthropologie*, **70**, 163-173.  
34  
35 Turner, II C.G. (1990). Major features of sundadonty and sinodonty, including suggestions  
36 about East Asian microevolution, population history, and Late Pleistocene relationships  
37 with Australian aboriginals. *American Journal of Physical Anthropology*, **82**, 295-317.  
38  
39 Turner, II C.G., Nichol, C.R., & Scott, G.R. (1991). Scoring procedures for key  
40 morphological traits of the permanent dentition: the Arizona State University Dental  
41 Anthropology System. In M.A. Kelley & C.S. Larsen (Eds.), *Advances in Dental*  
42 *Anthropology* (pp. 13-31). New York, Wiley-Liss.  
43  
44 Zanolli, C. (2013). Additional evidence for morpho-dimensional tooth crown variation in a  
45 new Indonesian *H. erectus* sample from the Sangiran Dome (Central Java). *PLoS One*, **8**,  
46 e67233.  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- 1  
2  
3 Zanolli, C., Bondioli, L., Coppa, A., Dean, M.C., Bayle, P., Candilio, F., Capuani, S., Dreossi,  
4 D., Fiore, I., Frayer, D.W., Libsekal, Y., Mancini, L., Rook, L., Medin, T., Tuniz, C., &  
5 Macchiarelli, R. (2014). The late Early Pleistocene human dental remains from Uadi Aalad  
6 and Mulhuli-Amo (Buia), Eritrean Danakil: macromorphology and microstructure. *Journal*  
7 *of Human Evolution*, **74**, 96-113.  
8  
9  
10 Zanolli, C., Hourset, M., Esclassan, R., & Mollereau, C. (2017). Neanderthal and Denisova  
11 tooth protein variants in present-day humans. *PLoS One*, **12**, e0183802.  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Captions for figures

**Figure 1.** (a). Position of the reference plane located at the midpoint of the crown, which was used to place the semilandmarks along the curve of the lingual aspect of the crown from the mesial to the distal side (b). The maxima of the extreme curvature line were used as starting and ending points of our GM analysis

**Figure 2.** Illustration of the maximum depth and hollow area (a) used in the non-Procrustes analyses (semilandmark curves aligned with their first and last points). The ASUDAS reference plaque teeth curves are superimposed following this non-Procrustes approach, showing the non-linear variation in shape from grade 0 to grade 6 (b).

**Figure 3.** Principal component analysis (PCA) of the Procrustes-registered shape coordinates of the 100 semilandmarks used as proxy to assess UI1 shoveling. A: PC1 vs. PC2; B: PC3 vs. PC4.

**Figure 4.** Histograms showing the frequency of Procrustes PC1 (a), maximum depth (b) and hollow area (c) values for the 87 modern human specimens and the distribution of the ASUDAS reference grades (black vertical lines).

**Figure 5.** Bland-Altman visualization for agreement of the ASUDAS visual observations and Procrustes coordinates of the semilandmark curves. The different agreements between the two raters were plotted for X, Y and Z float coordinates. Among the 100 landmarks, only the first, mid and last landmarks are drawn.

**Figure 6.** Plots of the maximum depth against the visual ASUDAS scoring of the first observer's tests (VS1 T1: a; VS1 T2: b) and the second rater (VS2: c). The black dots correspond to the values of the ASUDAS reference plaque and they were joined up via spline interpolation. The symbols represent the chrono-geographic origin as indicated in the legend of the graphs.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

**Table 1.** List of 87 modern human UI1 elements considered in this study.

<b>Time period</b>	<b>Geographic origin</b>	<b>depository</b>
contemporary (n=13)	France	MHNT <sup>1</sup>
medieval(n=31)	France	INRAP <sup>2</sup>
contemporary (n=13)	China	IVPP <sup>3</sup>
contemporary (n=30)	South Africa	PBC <sup>4</sup>

<sup>1</sup>MHNT: Natural History Museum of Toulouse; <sup>2</sup>INRAP: French National Institute for Preventive Archeological Research; <sup>3</sup>IVPP: Institute of Vertebrate Paleontology and Paleoanthropology of Beijing; <sup>4</sup>PBC: Pretoria Bone Collection of the Department of Anatomy of the University of Pretoria, representing individuals of various ancestries including Nbele, N Sotho, Swazi and Zulu.

**Table 2.** Intra-Class Correlation (ICC) values for intra-observer (VS1 T1 vs. T2) and inter-observer (VS1 vs. VS2) consistency measures and accuracy (comparing the three ratings with their predicted values from the ASUDAS reference plate teeth). Shades of color represent the degree of consistency and accuracy, with darker green corresponding to the highest degrees and white to the lowest degree.

	consistency (precision)		accuracy (unbiasedness)		
	VS1 T1 vs T2	VS1 vs VS2	VS1 T1 vs predicted	VS1 T2 vs predicted	VS2 vs predicted
<b>all samples</b>	0.985	0.987	0.759	0.749	0.772
<b>South African</b>	0.946	0.986	0.602	0.565	0.602
<b>French contemporary</b>	0.900	0.721	0.496	0.570	0.494
<b>French medieval</b>	0.971	0.993	0.550	0.592	0.559
<b>Chinese</b>	0.976	0.975	0.000	0.000	0.000

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

**Table 3.** Correlations between various measurements obtained from the coordinates of all samples (87 human specimens plus the 7 ASUDAS reference casts). Maximum depth and hollow area are the directly measured metrics based on the non-Procrustes alignment method. Principal components were obtained from both Procrustes and non-Procrustes methods. Shades of color represent the degree of correlation, with darker green corresponding to the highest degrees and white to the lowest degree. Red indicates negative correlation.

		direct metrics		Procrustes PC		non-Procrustes PC	
		maximum depth	hollow area	PC1	PC2	PC1	PC2
direct metrics	maximum depth	-	0.98	0.98	0.08	0.99	-0.08
	hollow area	0.98	-	0.96	0.20	1.00	0.06
Procrustes PC	PC1	0.98	0.96	-	0.00	0.97	-0.21
	PC2	0.08	0.20	0.00	-	0.16	0.86
non-Procrustes PC	PC1	0.99	1.00	0.97	0.16	-	0.00
	PC2	-0.08	0.06	-0.21	0.86	0.00	-

**Table 4.** Correlation of observer-assigned ASUDAS categories with the predicted ASUDAS value, the first four PCs of the Procrustes analysis, the maximum palatal depth and the hollow area. Shades of color represent the degree of correlation, with darker green corresponding to the highest degrees and white to the lowest degree. Red indicates negative correlation.

	VS1 T1	VS1 T2	VS2
<b>predicted ASUDAS</b>	0.776	0.769	0.788
<b>PC1</b>	0.836	0.831	0.850
<b>PC2</b>	0.153	0.169	0.177
<b>PC3</b>	-0.106	-0.097	-0.082
<b>PC4</b>	0.055	0.028	0.056
<b>maximum depth</b>	0.840	0.833	0.852
<b>hollow area</b>	0.832	0.828	0.848

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47

**Table 5.** ICC values for intra-observer (VS1 T1 vs. T2) and inter-observer (VS1 vs. VS2) consistency measures and accuracy (comparing the three ratings with their predicted values from the ASUDAS reference plate teeth) when grades 0-1 and 2-6 are fused into two groups (2-category split). Shades of color represent the degree of consistency and accuracy, with darker green corresponding to the highest degrees and white to the lowest degree.

	consistency (precision)		accuracy (unbiasedness)		
	VS1 T1 vs T2	VS1 vs VS2	VS1 T1 vs predicted	VS1 T2 vs predicted	VS2 vs predicted
<b>all samples</b>	0.945	0.964	0.812	0.827	0.842

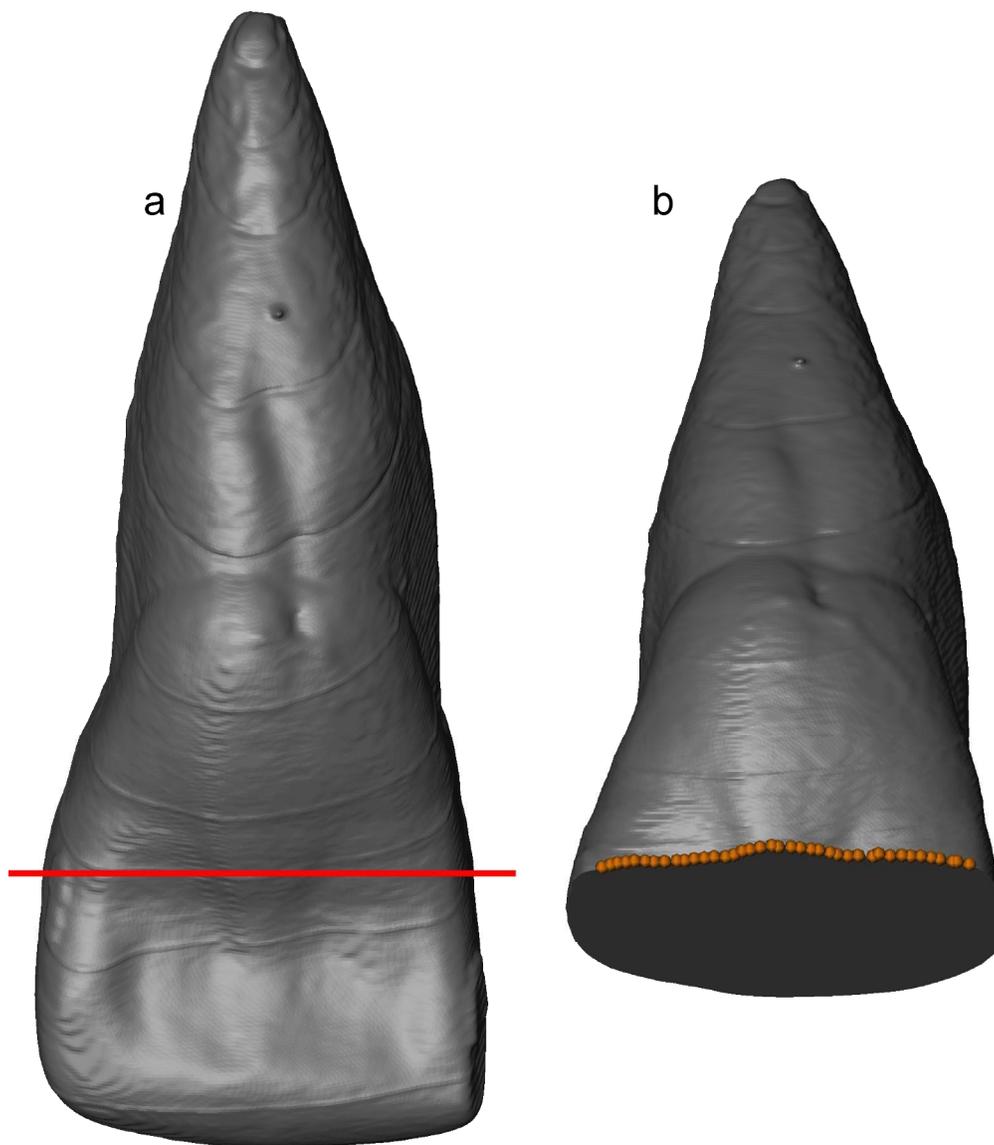


Figure 1. (a). Position of the reference plane located at the midpoint of the crown, that is used to put the semilandmarks along the curve of the lingual aspect of the crown from mesial to distal side. (b). We use the maxima of the extreme curvature line as starting and ending points of our GM analysis.

356x407mm (300 x 300 DPI)

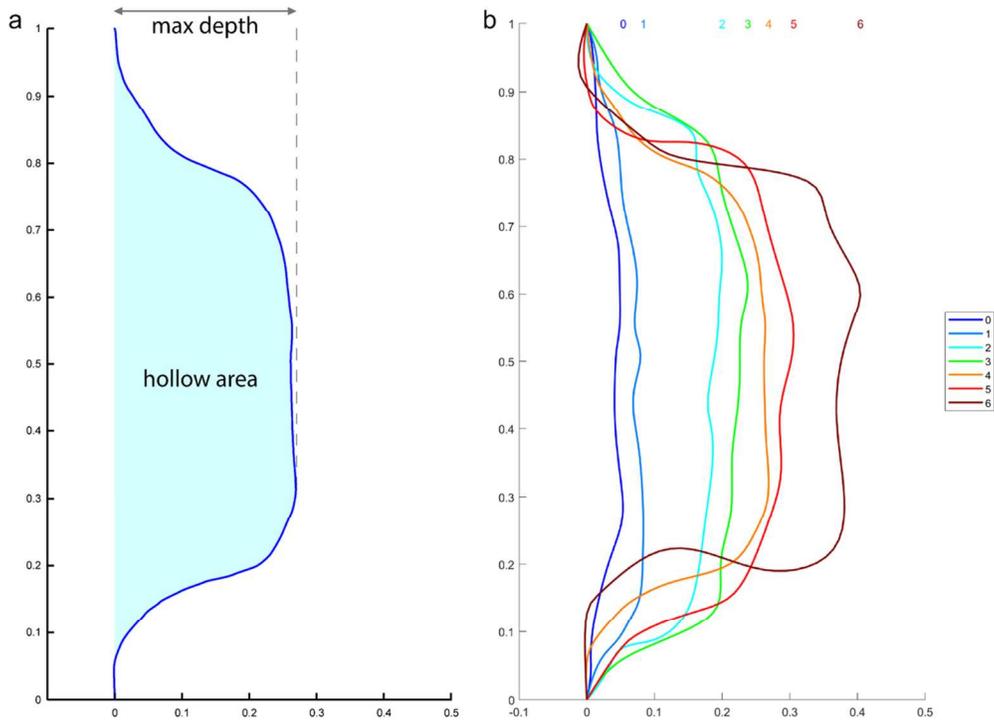


Figure 2. Illustration of the maximal depth and hollow area (a) used in the non-Procrustes analyses (semilandmark curves aligned with their first and last points). The ASUDAS reference plaque teeth curves are superimposed following this non-Procrustes approach, showing the non-linear variation in shape from grade 0 to grade 6 (b).

106x76mm (300 x 300 DPI)

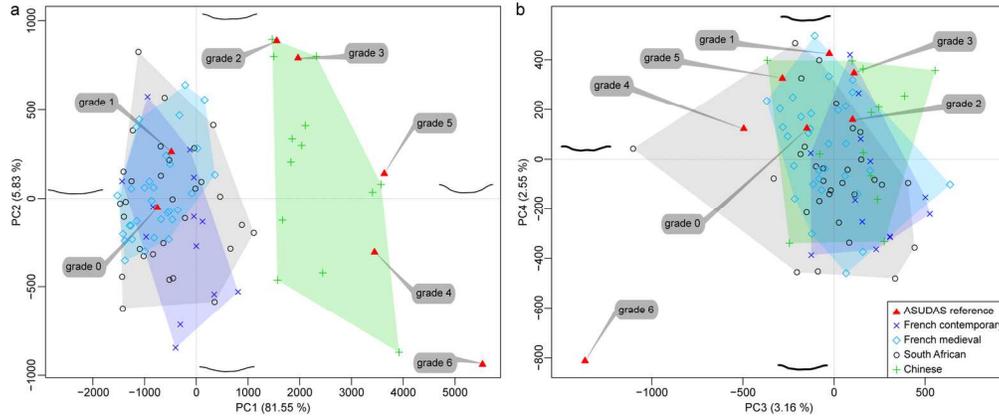


Figure 3. Principal component analysis (PCA) of the Procrustes-registered shape coordinates of the 100 semilandmarks used as proxy to assess UI1 shoveling. A: PC1 vs. PC2; B: PC3 vs. PC4.

154x64mm (300 x 300 DPI)

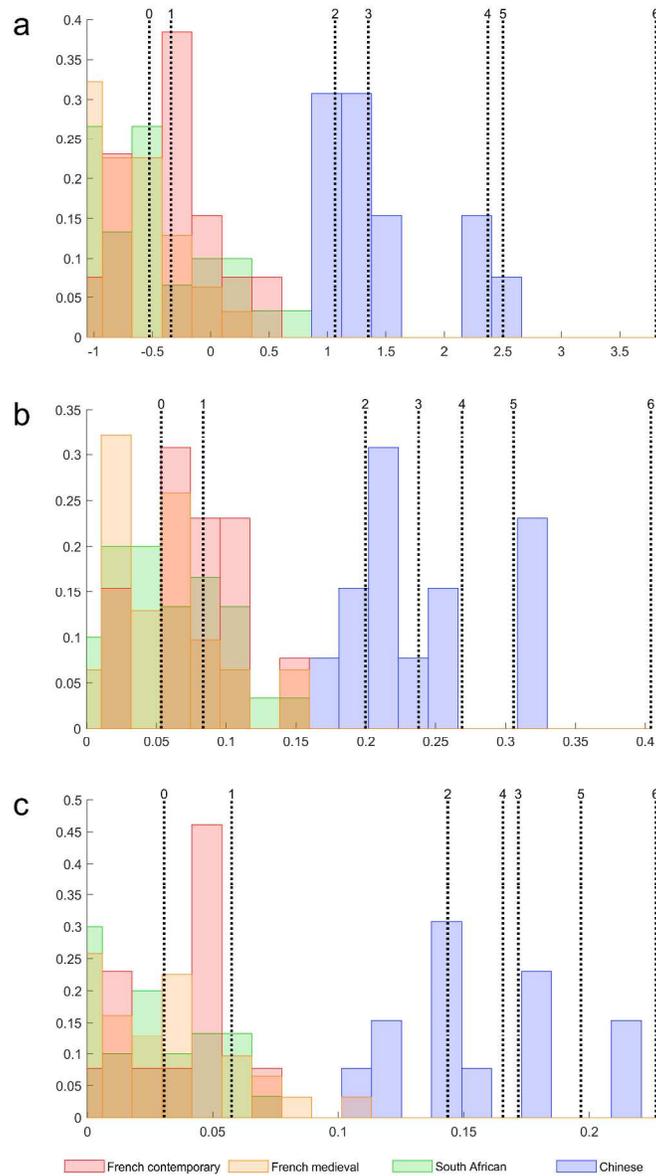


Figure 4. Histograms showing the frequency of Procrustes PC1 (a), maximum depth (b) and hollow area (c) values for the 87 modern human specimens and the distribution of the ASUDAS reference grades (black vertical lines).

187x319mm (300 x 300 DPI)

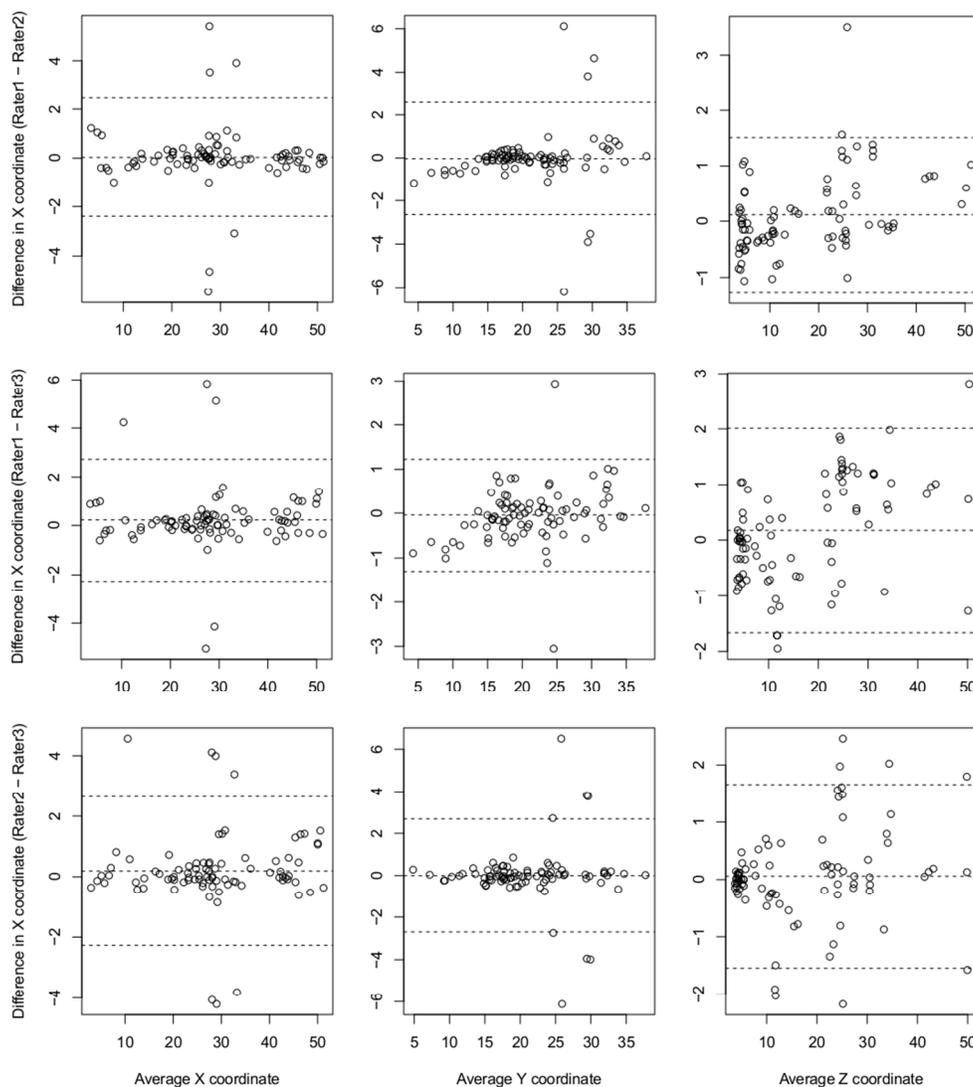


Figure 5. Bland-Altman visualization for agreement of the ASUDAS visual observations and Procrustes coordinates of the semilandmark curves. The different agreements among the pair of raters were plotted for X, Y and Z float coordinates. Among the 100 landmarks, only first, mid and last landmarks were drawn.

92x103mm (300 x 300 DPI)

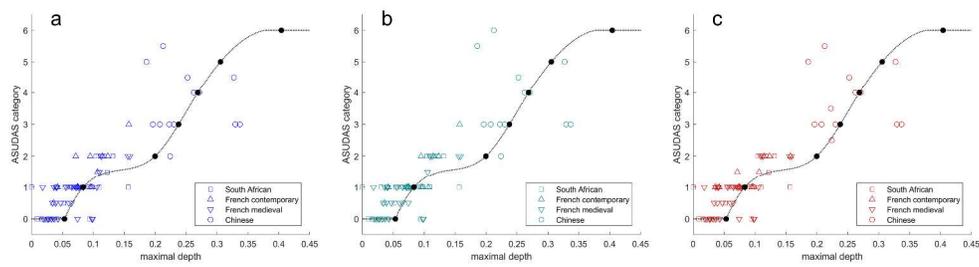


Figure 6. Plots of the maximum depth against the visual ASUDAS scoring of the first observer's tests (VS1 T1: a; VS1 T2: b) and the second rater (VS2: c). The black dots correspond to the values of the ASUDAS reference plaque and they joined between them via spline interpolation. The symbols represent the chrono-geographic origin as indicated in the legend of the graphs.

411x111mm (300 x 300 DPI)