Journal Pre-proofs

Short Communications

Taphonomic characterisation of tooth marks of extinct Eurasian carnivores through Geometric Morphometric

José Yravedra, Lloyd Austin Courtenay, Darío Herranz-Rodrigo, Gonzalo Linares-Matás, Juan José Rodríguez-Alba, Verónica Estaca-Gómez, Carmen Luzón, Alexia Serrano-Ramos, Miguel Ángel Maté-González, José Antonio Solano, Diego González-Aguilera, Juan Manuel Jiménez-Arenas

Short Communication

Taphonomic characterisation of tooth marks of extinct Eurasian carnivores through Geometric Morphometric

José Yravedra[1,2,3,4,5*], Lloyd Austin Courtenay[6,1*], Darío Herranz-Rodrigo[2,1], Gonzalo Linares-Matás[7], Juan José Rodríguez-Alba[1], Verónica Estaca-Gómez[1], Carmen Luzón[8], Alexia Serrano-Ramos[8], Miguel Ángel Maté-González[6,9], José Antonio Solano[10], Diego González-Aguilera[6], Juan Manuel Jiménez-Arenas[10,5*]

[1] Department of Prehistory, Ancient History and Archaeology, Complutense University of Madrid, Madrid 28040, Spain

[2] C.A.I. Archaeometry and Archaeological Analysis, Complutense University, Madrid 28040, Spain

[3] Grupo de Investigación Ecosistemas Cuaternarios, Complutense University, Madrid 28040, Spain

4 Grupo de Investigación Arqueología Prehistórica, Complutense University, Madrid 28040, Spain

[5] Museo Primeros Pobladores de Europa 'Josep Gibert', Orce 18858, Spain

[6] Department of Cartographic and Land Engineering, Higher Polytechnic School of Avila, University of Salamanca, Avila 05003, Spain

[7] School of Archaeology, St. Hugh's College, University of Oxford, Oxford 01865, UK

[8] History and Arts Doctoral Program, University of Granada, Granada 18071, Spain

[9] Escuela Técnica Superior de Ingenieros en Topografía, Geodesia y Cartografía. Universidad Politécnica de Madrid, Madrid 28031, Spain

[10] Department of Prehistory and Archaeology, University of Granada, Granada 18071, Spain

* Corresponding authors, E-mail: joyravedra@hotmail.com, jyravedr@ucm.es (José Yravedra); ladc1995@gmail.com; jumajia2007@gmail.com

`

Determining the cause and nature of the postmortem processes that living organisms experience is one of the main common issues faced by forensic experts, zooarchaeologists, palaeontologists, and other specialists. Carnivores are among the most destructive agents that can interact with a corpse, since their feeding behaviour can lead to very extensive alterations, complicating the diagnostic identification of which carnivore species was responsible for the death of an individual, a livestock unit, or the formation of a fossil assemblage. Even though some currently available techniques enable forensic experts to undertake a differential diagnosis of carnivore agency from corpse examination, these are very difficult to apply when skeletal parts are all that remains. Nevertheless, a computational taphonomic approach can help identify which carnivore could have generated the tooth marks present on bone surfaces, and thus aid in the reconstruction of their forensic biography.

Recent studies [1] have successfully classified the tooth marks generated by 8 different carnivores— wolf, fox, lycaon, hyaena, tiger, jaguar, lion—with over 90% accuracy, thus achieving a methodological innovation that enables us to identify a carnivore from tooth mark data. Nonetheless, these studies only focused on extant carnivores, and questions remain as to whether they can be used to successfully characterise extinct carnivores in fossil assemblages.

To extend the usefulness of these methodological solutions, we have selected as a case-study the palaeontological site of Venta Micena 3 (VM3; Orce, Granada, Spain), a fossil assemblage believed to have been primarily generated by the large extinct hyaena *Pachycrocuta brevirostris* [2], a species that

lived in Eurasia during the Early and Middle Pleistocene (Supp. File 1 online). There are several other Eurasian localities with faunal assemblages accumulated by this extinct carnivore, such as Dmanisi (Georgia) Le Vallonet (France) or Zhoukoudian (China).

The VM3 sample used for this paper involves 28 tooth marks from 10 different long bone diaphyses of medium- and large-sized ungulates, randomly selected from the collection hosted at the Museum of Granada and the local Museum of Orce. Carnivores generate several types of tooth marks, and previous studies have demonstrated that pits tend to be the more reliable for diagnosis [3]. Our comparative reference framework includes a sample of 483 tooth pits caused by different carnivores, described in (see Supp. File 2 from Ref. [1]).

Our methodological approach is based on the analysis of three-dimensional digital models of the tooth marks, generated using the Structured Light Surface Scanning DAVID SLS-3. Following the construction of 3D models of the bones, tooth pits were digitized using a landmark configuration consisting of 25 landmarks. 5 fixed Type II landmarks located on the exterior and interior of each pit, and a 5×5 semilandmark patch [4]. Of the 5 fixed landmarks, LM1 and LM2 mark the maximal length (l) of each pit. For the correct orientation of the pit, LM1 is placed farthest away from the perpendicular axis marking the maximal width, and LM2 is thus placed on the opposite extremity. LM3 and LM4 are then placed along this perpendicular axis marking the left (LM3) and right (LM4) maximum extremities, while LM5 is placed at the deepest point of the pit. The semilandmark patch is then positioned over the entirety of the pit, so as to capture the internal morphology of the mark and its walls. Landmarks produced by the semilandmark patch that overlap with the 5 fixed landmarks are then deleted leaving a final 25 landmark model (Fig. S1 online). Finally landmarks are slid across the surface of the model using the R (v.4.0.4.) programming language, so as to minimize bending energy [5]. Once digitized, landmarks were formatted as a morphologika file and imported into an R environment (v.4.0.4). Landmarks were first subjected to a Generalized Procrustes Analysis (GPA), so as to normalize data and project landmarks into new superimposed feature space [5]. This procedure was performed excluding the scaling procedure, so as to analyze pits in form space. The use of form, as opposed to shape, was chosen considering prior observations on the weight that size has on morphological variability [1,3,4]. Once superimposed, landmark configurations were analysed in terms of the Procrustes distances between each other, and Centroid Size (CS) distributions.

For Procrustes distances and CS, distributions were first analysed for homogeneity using Shapiro-Wilk tests. All following statistical tests were then conditioned by these results, using traditional statistical approaches where homogeneity was found to be present and robust statistical approaches otherwise [4]. From this perspective, descriptive statistics were either performed using the mean or median measurement for central tendency (for Gaussian and non-Gaussian data respectively), while distribution variability was measured in terms of the standard deviation or the Median Absolute Deviation (MAD). From a different perspective, univariate statistical tests were either performed using a linear ANOVA model, or the Kruskal-Wallis test.

For multivariate analyses, dimensionality reduction via Principal Component Analysis (PCA) was performed [1,5]. The PC scores representing up to 99% of morphological variance were then selected, and used for further statistical processing. Multivariate Analyses of Variance (MANOVA) were used to assess for differences in form feature space, using either the Hotelling-Lawley or Wilk's Lambda test statistic. Similarly, the Mahalanobis distances from each fossil individual to the multivariate distribution for each modern day carnivore were calculated. Finally, Thin Plate Splines (TPS) were also calculated [6], using a Delauney 2.5D Triangulation algorithm to facilitate the visualization of landmark configuration patterns. All statistics were performed in the R (v.4.0.4) programming language.

To support the observations made using statistical approaches in geometric morphometrics, computational learning algorithms were also trained, following the procedures recommended by Ref.

[1]. This methodological approach consists first in the augmentation of data via unsupervised algorithms [1], followed by the training of supervised classification algorithms that can then be used to predict class labels for each of the fossil tooth pits [1]. For data augmentation, a multivariate Monte Carlo Markov Chain was used to simulate the morphological characteristics of 100 tooth marks per sample. This was performed so as to balance data set sizes, as well as provide enough information for the supervised algorithms to learn from Ref. [1]. The quality of augmented data was then evaluated by calculating the statistical similarities with the original data [4, with the final augmented dataset being calculated to be highly equivalent to the original data ($|d| = 0.01$, $P = 3.6 \times 10^{-62}$).

Once augmented, Support Vector Machines (SVM) and Neural Support Vector Machines (NSVM) were trained [1,7]. SVMs were trained using a k-fold cross-validated approach ($k = 10$), and a *Radial Basis Function* kernel. Optimal configuration of the kernel was computed using Bayesian Optimisation algorithms [8]. NSVMs were trained using typical deep learning approaches [9], first by training a Laplacian Random Fourier Function based neural network [10], and then replacing the final activation layer with a linear SVM [1]. NSVM was trained in batches of 32 for 1000 epochs, using a triangular cyclic learning rate, and the adam optimizer. Additional tuning of the SVM activation layer was also performed using Bayesian approaches [8].

Both SVM and NSVM were trained on 80:20% train: test sets, and then used to predict labels and label probabilities for each of the fossil individuals. The summary of the two trained algorithm performance on test sets is provided in Table S1 (online). SVMs were programmed in the R programming language (v.4.0.4), while NSVMs were programmed in Python (v.3.7.4). For more details see Ref. [1].

As mentioned above, *P. brevirostris* has been considered the main accumulating and modifying agent in the VM3 site. For this reason, once marks had been classified, the VM3 marks observed to be produced by hyaenids were separated to perform a more in depth characterisation of *P. brevirostris*. This characterization was performed using the same methodological procedure as the geometric morphometric analyses described above. Nevertheless, these analyses were complemented with two one-sided equivalency tests (TOST), according to Cohen's d, so as to calculate the magnitude of similarities between samples [7]. For homogeneous distributions, Welch's t-statistic was used, while non-parametric approaches employed the use of Yuen's trimmed robust t-statistic. Finally, TPS were used to warp mean configurations of hyena and *Pachycrocuta* to a 3D model [6], so as to calculate the distance between the faces of each mesh and quantify differences between the mean configurations. Distance calculations were computed using the nearest neighbor distance from a reference mesh to a warped mesh, using as a reference mesh the 3D model corresponding to the median individual of one of the groups [11].

All hypothesis tests were evaluated using Bayesian calibrations of P-values. Under this premise, the false positive risk (FPR) was calculated for each P-value [12], using the Sellke-Berger approach [13], for the definition of null hypothesis (H_0) and alternative hypothesis (H_a) ratios. Where necessary, FPR was also used to derive probability of H_0 values ($p(H_0)$), providing a means to calibrate P values over 0.3681 [1,3]. Unless specified otherwise, prior probabilities in support of H_a were set at 0.5, indicating complete randomness, as recommended by Ref. [12]. In light of these calibrations, P-values were thus evaluated using a robust value of 0.003 (3σ) as a threshold for more conclusive results. This P-value can be considered to have and FPR of 4.5 +/− [1.2, 15.9] %, using priors of 0.5 +/− [0.2, 0.8] [1].

After applying this method, the first step in the corroboration of carnivore agency at VM3 is to compare the tooth pits from VM3 with the reference dataset. Initial comparisons find the VM3 tooth sample to consist of relatively large tooth pits (CS = 8.32 mm), slightly larger than those of the modern day spotted hyena (CS = 8.16 mm), but lower than those of larger felids such as the lion (CS = 12.29 mm) [1,3]. Statistically, these similarities are reflected with the greatest approximations appearing when comparing VM3 with hyaenids ($\chi^2 = 0.02$, $P = 0.88$, $p(H_0) = 76.6\%$), while

differences are present from all other samples ($\chi^2 > 7.5$, $P < 0.006$, $p(H_0) = 7.7\%$). Procrustes distances, on the other hand, reveal VM3 to be similar to large felids ($P = 0.80$, $p(H_0) = 67.3\%$), and hyaenids ($P = 0.60$, $p(H_0) = 54.6\%$), while presenting vague similarities with ursids ($P = 0.04$, $p(H_0) = 25.9\%$). On all accounts, the VM3 sample appears to remain separate from canids ($P = 0.004$, $p(H_0) = 5.7\%$). Finally, multivariate analyses of both size and shape (Fig. 1a) confirm VM3 to be notably different from both ursids and canids ($P > 0.001$, $p(H_0) = 1.8\%$), while similarities are still clear when compared with hyaenids and large felids ($P > 0.123$, $p(H_0) = 41.2\%$).

When using SVM and NSVM algorithms to classify these traces, both algorithms confirm with an average confidence of 94%, the presence of 20 hyaenid tooth marks, 6 large felid tooth marks, and a single tooth mark associated with the genus *Canis* (Table S2 and Fig. S2 online). When analyzing SVM and NSVM performance, both algorithms produce similar classification results, with NSVM proving the most confident classifier in the majority of cases. Additionally, CL produced labels coincide in the majority of Procrustes distance associations, with only two pits presenting Procrustes distances that contradict CL class labels. In each of these cases, the large Procrustes distances can be attributed to abnormally high CS values for a Hyaenidae, however, in form feature space, CL algorithms are still confident in attributing these marks to that of a large hyena. The final indeterminable pit, however, is found to have an abnormally large Procrustes distance that does not reveal a conclusive association to any of the carnivores used in the present comparative sample.

Visualization of the corresponding feature space (Fig. 1a) effectively reveal the proximity of the 6 classified pits to the large felid sample (Procrustes $D = 2.65$, Mahlanobis $d = 0.33$), as opposed to their association with hyaenids (Procrustes $D = 5.16$, Mahlanobis $d = 3.25$). While the hyaenids from VM3 are observed to present a large spread across form feature space, their association with Hyaenids can also be confirmed (Procrustes $D = 3.67$, Mahlanobis $d = 1.15$), as opposed to their association with felids (Procrustes $D = 5.78$, Mahlanobis $d = 1.36$). Thin plate splines derived from this data additionally reveal the VM3 felids to produce deeper more elongated pits, like those of lions, while the VM3 hyaenids can be characterized by producing a wide range of tooth mark morphologies, from superficial small pits to large deeper pits, as will be explained in detail in continuation.

Exploration of form feature space in more dimensions reveals PC3 (3.81% variation) to be represented mostly by the asymmetry of pit morphologies (Fig. S2 online), with constrictions in the lateral edges of each pit between LM1 and LM3. This characteristic is mostly observed in felids, with hyaenids present a generally more ovular and symmetric pit morphology.

When considering the alterations per bone specimen (Table S2 online), it can be observed that only a single bone presents the intervention of both large felids and hyaenids (Sp. Nº 10). Similarly, the single *Canis* tooth mark is associated to a hyaenid tooth mark on the same bone (Sp. Nº 3) (Table S2 online).

The tooth pits produced by *Pachycrocuta*, as seen through the present sample, can be described as morphologically similar to the tooth pits produced by modern day *Crocuta crocuta*. This is apparent both in terms of morphology and size.

As would be expected [5], allometry is revealed to be a variable of great importance when considering hyaenid tooth pit morphology ($F = 325.6$, residuals^2 = 0.758, effect size = 5.08, $P = 0.001$, FPR = 1.8%). Nevertheless, shape-size relationships according to groups reveal the present sample of *Pachycrocuta* and hyaenid tooth marks to be of similar size and shape ($F = 1.5985$, residuals^2 = 0.003, effect size = 1.06, $P = 0.149$, FPR = 43.5%). From this perspective, the present sample reveals *P. brevirostris* to have a CS of 6.4 +/− 2.2 mm (Mean +/− 1st Std. Dev.; Shapiro-Wilk $w = 0.9$, $P = 0.1$). This value is 0.5 mm smaller than the comparative hyaenid sample used here, with *Crocuta crocuta* also presenting a larger spread; 6.9 +/− 4.0 mm (Median +/− MAD; $w = 0.9$, $P = 8.7 \times 10^{-6}$). This measurement groups *P. brevirostris* in the larger group of carnivore species alongside lion, jaguar,

hyena and *Lycaon* [1]. Through this sample, *P. brevirostris* can also be described to produce similar sized pits to hyena [1], and much larger pits than other studied hyaenid species [14]. Nevertheless, it is important to note that the small sample size presented here ($n = 20$) is likely to exaggerate some similarities in tooth pit size (TOST $t = -3.1$, $P = 0.002$). This results in an 11.9% probability of this observation being a false positive (corrected prior probability = 0.2).

Analysis of morphological traits in form space reveal both hyaenids to produce a large variety of different tooth pit morphologies, with a relatively large spread when compared with other carnivores (Fig. 1a). When comparing both hyaenids separately (Fig. 1b), notable overlap can be observed ($t = 3.076$, $P = 0.002$, FPR = 2.8%). In this feature space, the majority of morphological information is explained by shifts in the position of LM5, and the sliding landmarks that mark the base of the tooth pit (PC1, 84.64% morphological variance). From this perspective, changes can be seen not in depth, but more in the point where the maximal depth of each tooth pit is located. Moreover, changes in the elongation of each pit can also be observed, with a combination of PC1 and PC2 (90.06% morphological variance) describing more elongated tooth marks. This can be seen in greater length: width ratios towards the positive portion of this region, and the variability in distance from LM2 and LM5. Nevertheless, as noted in other hyaenid samples [1,3], both modern day and fossil hyaenids present great morphological variance, with the ability to produce a combination of small, large, deep and superficial pits within the same sample.

In light of each of these observations, and when considering the weight size has over the morphology, hyaenids can be considered close to three other species, including lions, jaguars and *Lycaon* [1,3]. Nevertheless, each of these species have been noted to produce much deeper pits (Fig. S2 online), with jaguars and *Lycaon* also creating more elongated pits along the LM1-LM2 axis. The pits observed in VM3 follow this trait, appearing much more superficial than those typically produced by other large carnivores (Fig. 1a, b).

Finally, when projecting the mean configuration of hyena pits onto the median *Pachycrocuta* 3D model (Fig. 1c), warp analysis reveals a difference of as little as 0.02 +/− 0.04 mm between the 3D models (Fig. 1e). Visualizing these changes, it can be confirmed that depth is not necessarily a conditioning factor in tooth pit morphological variation, while the majority of deformations are concentrated as mild changes along the wall of each pit (Fig. 1d). These variations are likely resulting from *Pachycrocuta*'s hyperdeveloped mastication muscles—particularly the masseter and the pterygoid—in comparison to modern hyaenas, which enabled them to exert a greater premolar bite force despite having a similar cranial and dental configuration [15]. This study applies a series of recently-developed computational techniques for taphonomic research [1,3,4], and presents the first 3D reconstruction and morphometric characterisation of tooth marks present in the Early Pleistocene site of VM3. Our investigation confirms that the faunal assemblage from this palaeontological locality was accumulated and modified by the extinct Eurasian giant hyaena *Pachycrocuta brevirostris*, confirming previous interpretations of the assemblage [2]. Our results also enable the creation of a reference framework that will allow the identification of the taphonomic impact of *Pachycrocuta brevirostris* in Eurasian bone assemblages from the three-dimensional analysis of tooth mark morphometrics.

References

1. Courtenay, L.A., Herranz-Rodrigo, D., González-Aguilera, D., et al. Developments in data science solutions for carnivore tooth pit classification. Sci Rep, 2021, 11, 10209
2. Arribas, A., Palmqvist, P. Taphonomy and palaeoecology of an assemblage of large mammals: Hyaenid activity in the Lower Pleistocene site at Venta Micena (Orce, Guadix-Baza Basin, Granada, Spain). Geobios, 1998, 31, 3-47.

3. Courtenay, L.A., Herranz-Rodrigo, D., Yravedra, J., et al., 3D Insights into the Effects of Captivity on Wolf Mastication and Their Tooth Marks; Implications in Ecological Studies of Both the Past and Present. Animals, 2021. 11, 2323.

4. Courtenay, L. A., Herranz-Rodrigo, D., Huguet, R., et al., Obtaining new resolutions in carnivore tooth pit morphological analyses: a methodological update for digital taphonomy. PLoS ONE, 2020, 15: e0240328.

5. Bookstein, F.L., Morphometric tools for landmark data. Cambridge: Cambridge University Press. 1991.

6. Bookstein, F. L. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989, 11, 567-585.

7. Lakens, D. Equivalence Tests: a Practical Primer for T Tests, Correlations and Meta Analyses. Soc Psychol Pers Sci, 2017, 8, 355-362

8. Shahriari, B., Swersky, K., Wang, Z., et al., Taking the human out of the loop: A review of Bayesian optimization. Proc. IEEE, 2016. 104, 148–175.

9. Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge: MIT Press, 2016.

10. Rahimi A, Recht B. Random features for large-scale kernel machines. NeurIPS. 2007, 20, 1-8.

11. Rohlf F.K. Shape statistics: Procrustes superimpositions and tangent spaces. J Classif, 1999, 16, 197-223.

12. Colquhoun D. The False Positive Risk: A proposal concerning what to do about P-values. Am. Stat. 2019, 73: 192–201.

13. Benjamin D.J., Berger J.O. Three recommendations for improving the use of P-values. Am Stat, 2019, 73: 186–191.

14. Arriaza M. C., Aramendi J., Maté-González M.A., et al., Characterizing leopard as taphonomic agent throught the use of micro-photogrammetric reconstruction of tooth marks and pit to score ratio. Hist Biol, 2019. 33, 176-185.

15. Mutter R.J., Berger L.R., Schmidt P. New evidence of the giant hyaena, Pachycrocuta brevirostris (Carnivora, Hyaenidae), from the Gladysvale Cave Deposit (Plio-Pleistocene, John Nash Nature Reserve, Gauteng, South Africa). Palaeontologia Africana, 2001, 37: 103-113.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

Author Contributions

José Yravedra contributed in the conceptualization, formal analysis, investigation, resources, writing (original draft, review and editing), supervision, project administration and funding acquisition for this study. Lloyd Austin Courtenay contributed in the conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing (original draft, review and editing) and visualization of the present study. Darío Herranz-Rodrigo contributed in the investigation and data curation for the present study. Gonzalo Linares-Matás, Juan José Rodríguez-Alba, Veróncia Estaca-Gómez, and Miguel Ángel Maté-González, contributed in the reviewing and editing of the manuscript. Carmen Luzón contributed in the Investigation of the present study. Alexia Ramos was in charge of supervising the present study. José Antonio Solano was in charge of investigation, supervision, as well as writing (review and editing). Diego González-Aguilera provided resources for the present study and was in charge of supervision. Juan Manuel Jiménez-Arenas contributed in the writing (review and editing), project administration and funding acquisition of the present study.

Appendix . Supplementary data

Caption Figures and Tables

Figure 1 (a) Mean principal component analysis in form feature space characterizing the central configuration of each modern carnivore sample with each of the VM3 tooth marks and their classification label. Predicted form deformations via thin plate splines are depicted on each extremity of their corresponding PC score. Visualization of form deformations are visualized using a 2.5D Triangulation algorithm. (b) Principal component analysis in form feature space characterizing the central configuration of Crocuta crocuta and the VM3 *Pachycrocuta* tooth mark sample. Predicted form deformations via thin plate splines are depicted on each extremity of the graph. Form deformations are visualized using a 2.5D Triangulation algorithm. (c) Results when warping 3D models in a comparison between hyena and *Pachycrocuta* samples. The median *Pachycrocuta* pit used for projecting the mean configuration of hyena onto the 3D model. (d) Colour map indicating areas of higher distances between the reference and the warped 3D model, with a histogram of these distances presented in (e).

José Yravedra is the director of the Archaeological Analysis Unity of the Centro de Asistencia a la Investigación of Earth Science and Archaeometrý (UCM), and profesor in the Prehistory, Ancient History and Archeology Department of Complutense University. He has been the director of several research projects including Olduvai Gorge and other Iberian sites, and currently is studying the taphonomy of Project ORCE and he is pioneer in the application of photogrammetry and geometric morphometrics in taphonomy.

Lloyd Austin Courtenay is a Ph.D. student in the University of Salamanca, and has been working in the TIDOP research group, of the Department of Cartographic and Terrain Engineering, since 2019. He is an active participant in multiple research teams from diverse fields of science, with the majority of his work focusing on early human evolution in both Europe and Africa. His specialties are in robust statistics, programming, and data science.

Juan Manuel Jiménez-Arenas received his Ph.D. degree in history from University of Granada (Spain) in 2006. Later, he completed research stays at University of Zurich (Switzerland), and at the Catalan Institute of Human Paleoecology and Social Evolution (Tarragona, Spain). He is assistant professor at the University of Granada since 2014. His research interests focus on the paleocological context of the first human settlement in Western Eurasia (ProjectORCE) and human variability.
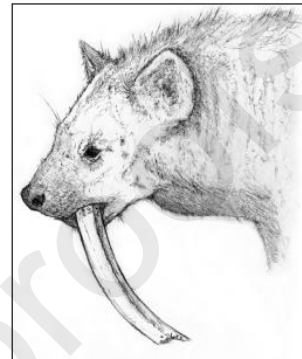
¿ Who ate it?

Tooth Marks
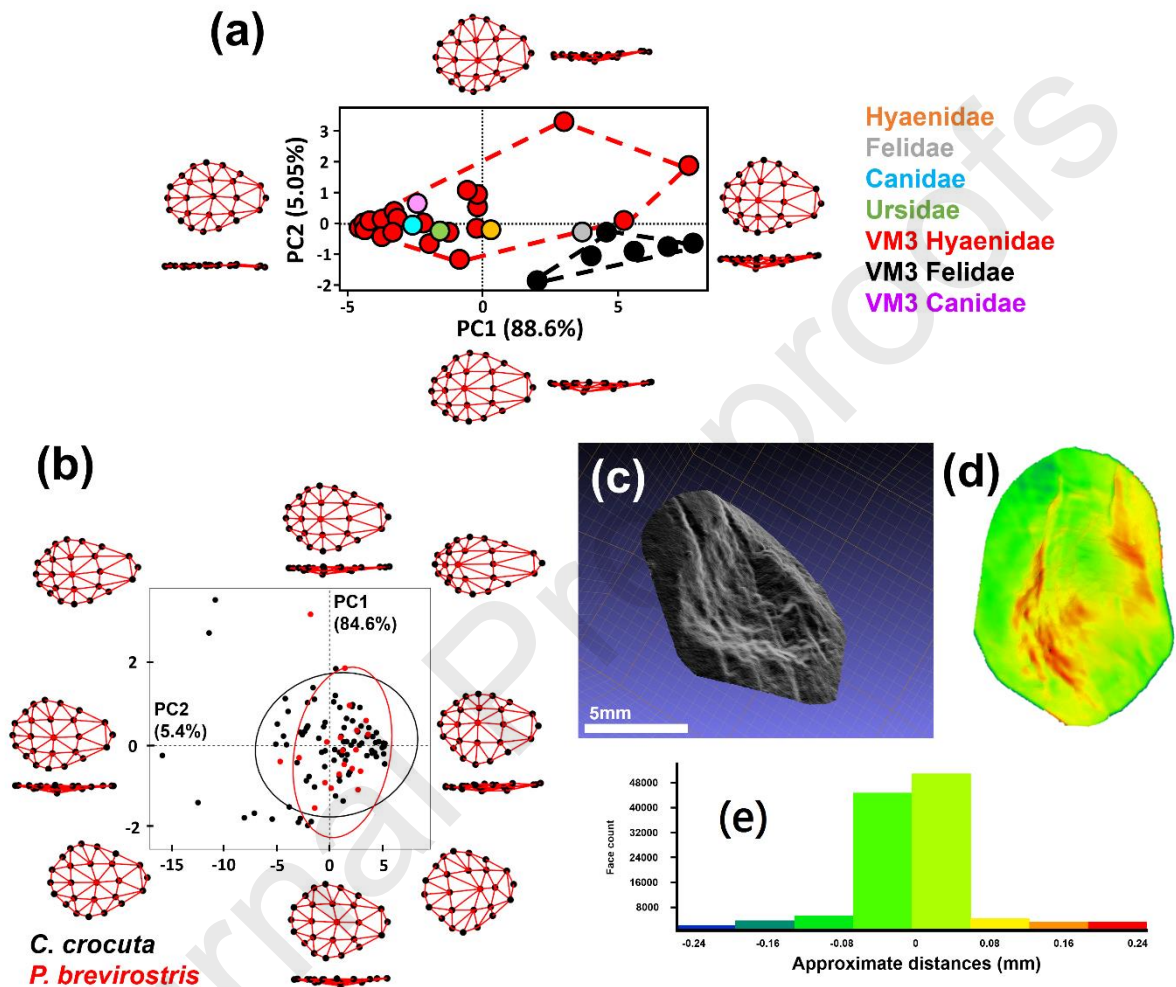
New Methods for taphonomical-forensic Data Analysis
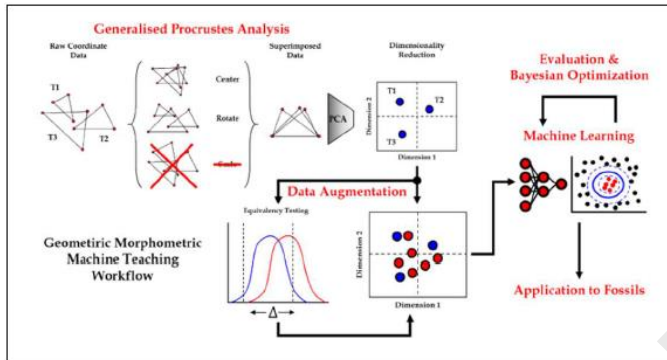
Results

Pachycrocuta brevirostris