



OPEN Bone diagenesis and stratigraphic implications from Pleistocene karst systems

Héctor Del Valle^{1,2}✉, Alejandro B. Rodríguez-Navarro³, Abel Moclán^{4,5}, Paula García-Medrano^{1,2,6,7} & Isabel Cáceres^{1,2}

Bone diagenesis is a complex process that modifies bone components in response to burial conditions. These modifications help to understand deposit formation and classify fossils by stratigraphy. The combined techniques of X-ray diffraction with Rietveld refinement and infrared spectroscopy were used to study the bone diagenetic processes along the complete stratigraphic sequence of Galería site (Sierra de Atapuerca, Spain). Eleven chemometric indices considering the different bone components (phosphates, carbonates, organic phase), together with the apatite unit cell parameters and cell volume were evaluated by 9 machine learning algorithms for bone diagenesis/stratigraphic classification. The results showed differences along the stratigraphic sequence due to changes in the apatite structure chemistry (i.e., F⁻ and OH⁻), producing a gradual shift of the unit cell volume (from 531.9 to 526.1 Å³) from GII to GIV associated with coupled dissolution–precipitation processes. Two diagenetic pathways are indicated: The lowest unit (GII) is characterized by leaching and carbonate loss in bone, suggesting an acidic and wet burial environment with the formation of authigenic phosphate minerals. The uppermost units (GIII–GIV) show bone apatite undergoing F⁻ and CO₃ incorporation, suggesting a slightly alkaline and drier environment. These differences enabled the development of classification models to understand deposit formation dynamics and also recontextualize dissociated fossil bones.

Keywords Infrared spectroscopy, X-ray diffraction, Apatite, Fossil bones, Machine learning

Bone is a complex composite biomaterial that fulfils several mechanical, physiological, and chemical regulatory functions, such as mineral ion homeostasis or cell storage^{1–4}. Bone is largely composed of an organic phase and a mineral phase. The mineral phase, made up of carbonated apatite nanocrystals⁵, can be described as Ca_{10–x}[(PO₄)_{6–x}(CO₃)_x](OH)_{2–x}·nH₂O⁶, whose non-stoichiometric composition includes multiple ionic substitutions, such as CO₃^{2–}, Na⁺, F⁻, Zn²⁺, Sr²⁺, and Ba²⁺⁶. Apatite has hexagonal symmetry (space group P6₃/m) and contains 4–8% carbonate, substituting PO₄^{3–} groups (B-type) or OH⁻ groups (A-type) in the crystal lattice⁷. The organic phase of bone is mainly composed of type I collagen together with non-collagenous proteins, proteoglycans, osteocalcin, and glycoproteins^{5,8}. Proteins are intimately associated with the mineral part of the bone, controlling mineral formation and many of its properties^{5,8}.

The modification of bone remains over time begins with the death and deposition of an organism. From this moment, the remains are affected by various processes. These can be divided into two main phases within the framework of taphonomic studies. The first is biostratinomy, encompassing all processes that occur before the sedimentation and burial of the remains. The second is fossil diagenesis, which includes the physical and chemical processes that occur after the sedimentation and burial. All taphonomic processes are framed by the environment in which they develop, thus being linked to their taphosystem^{9,10}. Bone diagenetic processes are influenced by their burial environment. However, it is important to consider that many biostratinomic modifications (e.g., bone heating, digestion, or weathering) may also affect diagenetic processes (e.g., dissolution-

¹Institut Català de Paleoeologia Humana i Evolució Social (IPHES-CERCA), Zona Educacional 4, Campus Sescelades URV (Edifici W3), 43007 Tarragona, Spain. ²Departament d'Història i Història de l'Art, Universitat Rovira i Virgili, Avinguda de Catalunya 35, 43002 Tarragona, Spain. ³Departamento de Mineralogía y Petrología, Universidad de Granada, 18002 Granada, Spain. ⁴Laboratoire PALEVOPRIM, Université de Poitiers and CNRS, Poitiers, France. ⁵Institute of Evolution in Africa (IDEA), University of Alcalá de Henares, Covarrubias 36, 28010 Madrid, Spain. ⁶UMR 7194 HNHP (MNHN-CNRS-UPVD), Département Homme et Environnement, Muséum National d'Histoire Naturelle, Paris, France. ⁷Department Britain, Europe and Prehistory, British Museum, London, UK. ✉email: hectorvalleblanco@gmail.com; hdelvalle@iphes.cat

(re) precipitation^{11–13}, altering the initial conditions of the fossil-diagenetic phase even within the same burial context. Due to the complexity of these processes, it is necessary to study them using complementary techniques such as histological, mineralogical, structural, and chemical analyses to fully characterize and understand them. While the taphonomic processes in the Sierra de Atapuerca sites have been thoroughly investigated from a biostratigraphic perspective^{14–19}, their fossil-diagenetic aspects remain less explored (e.g., Del Valle et al.¹³).

In this study, we focus on the diagenetic processes studied from infrared spectroscopy (ATR-FTIR)^{20,21} and X-ray diffraction (XRD) based on the Galería site. Galería is one of the sites that makes up the archaeopaleontological complex of the Sierra de Atapuerca (Burgos, Spain), which is located to the north of the northern plateau at an altitude of 1085 masl. It is composed of several cave systems. Specifically, Galería, along with Gran Dolina and Sima del Elefante, is located on the western side of the Sierra within a railway trench (Fig. 1). It measures approximately 14 m high, 18 m wide, and over 12 m deep. The stratigraphic sequence belonging to the Middle Pleistocene is composed of six lithostratigraphic units named from GI to GV²² (Figs. 1 and 2). Detailed information on the site and each stratigraphic unit is provided in Supplementary Information 1 (SI 1).

The aims of this study are (1) to characterize bone diagenetic processes throughout a complete Pleistocene stratigraphic sequence using ATR-FTIR and XRD and (2) to determine whether differences in bone preservation exist among the stratigraphic units. This could be used to reconstruct the burial history and environmental conditions at the deposition site. Once the differences were understood, classification methods (3) were proposed to assign fossil bones to their corresponding geological units based on their chemical properties. This approach involves Principal Component Analysis (PCA) to evaluate the contribution of the variables, combined with supervised Machine Learning (ML) models for classification. These models are particularly significant as they provide a framework to recontextualize archaeological fossils which currently lack stratigraphic context.

Material and methods

A total of 150 fossil bone samples have been analysed for this study (Table S1) from each Galería unit and subunit (30 by unit/subunit) containing archaeopaleontological remains (GIIa, GIIb, GIIIa, GIIIb, and GIV). Each bone sample corresponds to cortical bone from anatomical elements of large and medium-sized fossil mammals (Table S1).

Infrared spectroscopy and chemometrics

Fourier transform infrared spectroscopy (FTIR) is a powerful tool for understanding changes in the mineral and organic components at the molecular level. FTIR analysis was performed in attenuated total reflection (ATR) mode (Fig. 2). The advantages of using ATR-FTIR over the transmission mode for archaeological bones are evident as it is fast, requires minimal and fast preparation, a small amount of sample, and allows for the reduction of analytical errors^{20,21}. In addition, it allows semi-quantitative chemometric analyses to be carried out between different bone remains or archaeological assemblages.

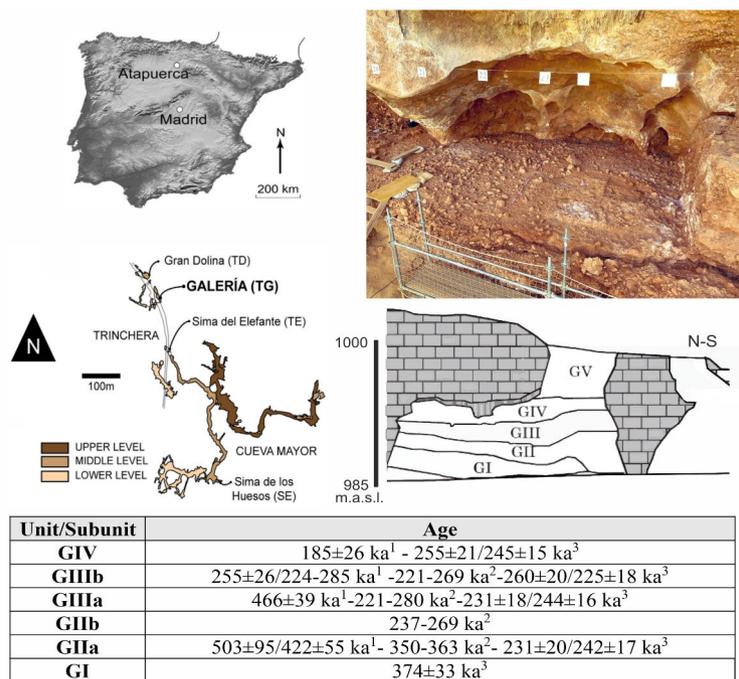


Fig. 1. Location of the Galería site within the Las Torcas karst system (modified from^{23,24}. Schematic karstic infill of the Galería site (modified from²⁵ and age (1:²⁶, 2:²⁷; 3:²³). Photograph of the current excavation surface on the top of GII (photograph by M. Guillén/IPHES-CERCA).

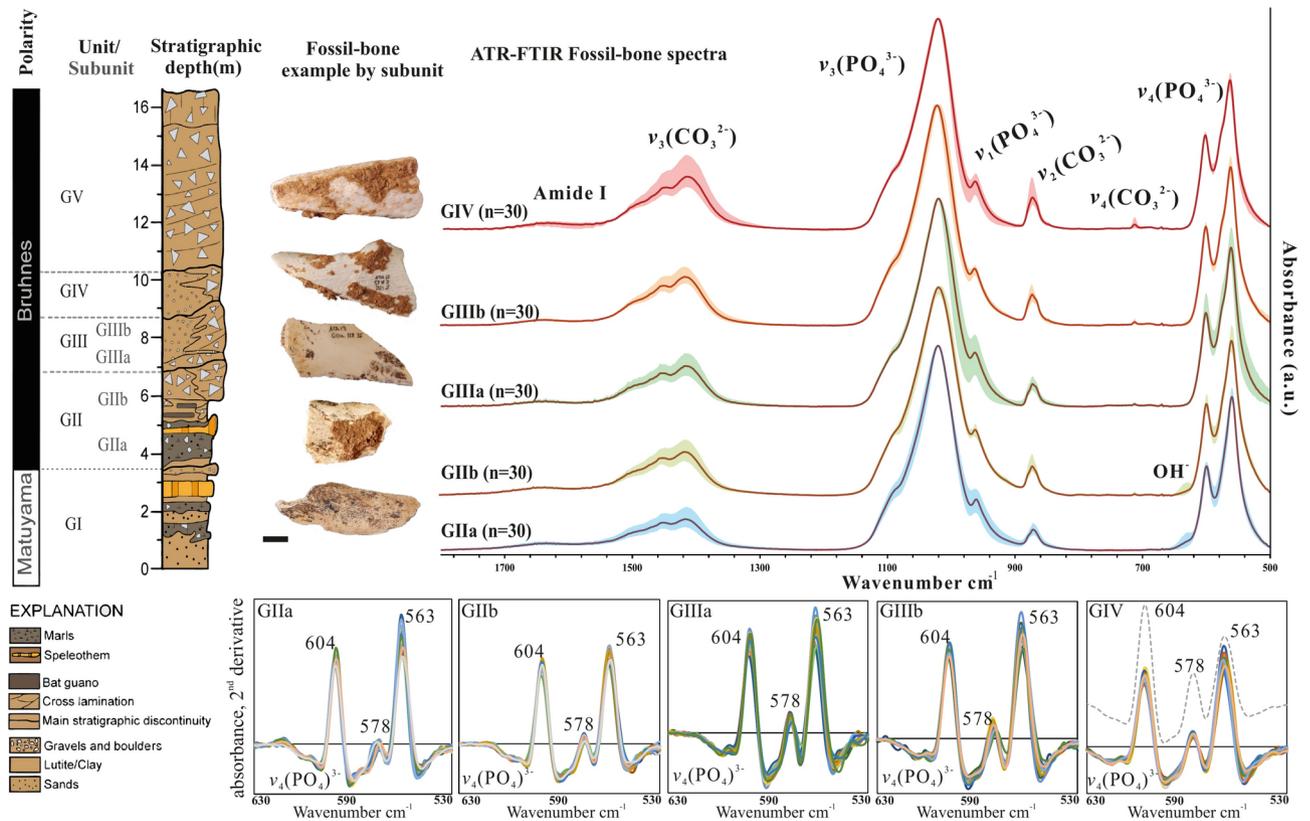


Fig. 2. Average normalised ATR-FTIR spectra set for bone samples from each stratigraphic unit. Images of fossil bone examples of each unit and the stratigraphic column of the Galería site are shown alongside the spectra (modified from²⁴). Below, detail on the second-derivative spectra of the region $\nu_4(\text{PO}_4^{3-})$ for each unit/subunit. The grey dashed spectrum in the GIV chart shows an example of the same bands for fluorapatite.

Each bone sample was processed following the methodology proposed by Kontopoulos et al.²⁸ and Lebon et al.²⁹. The spectra were collected on a Jasco FT/IR-6800 spectrometer with a resolution of 2 cm^{-1} and 64 scans in the range $4000\text{--}400\text{ cm}^{-1}$ at the SRCiT URV (University of Rovira i Virgili's Scientific and Technical Service, Tarragona, Spain). Around 1 mg of sample powder was pressed on the surface of a diamond crystal with a $50\text{--}20\text{ }\mu\text{m}$ particle size²⁸. The anvil pressure on the ATR crystal was adjusted to obtain a 0.5 absorbance for the $\nu_3(\text{PO}_4^{3-})$ band around 1010 cm^{-1} ²⁹. Three spectra were analysed for each sample. Spectra analysis was performed using Spectra Manager™ 2.15.00 software (JASCO).

Overlapping peaks were resolved and their integrated areas were measured using specifically designed curve fitting software³⁰. From the peak areas, the following compositional parameters were determined to define bone material properties: (1) $\nu_1, \nu_3(\text{PO}_4^{3-})$ band area ratio ($800\text{--}1200\text{ cm}^{-1}$).

For additional information, the second derivative was calculated using a Savitzky-Golay algorithm (2nd derivative order and 3 polynomial order) with 5 points of window for $\nu_3(\text{PO}_4^{3-})$ and $\nu_4(\text{PO}_4^{3-})$ spectral regions^{31,32}.

After processing the spectra, the chemometric indices were obtained. The indices are shown in Table 1. These were chosen to describe the different taphonomic modifications in the main bone components (phosphates, carbonates, and organic phase).

More specifically, the indices that analyse phosphates' crystallinity and behaviour in the spectrum include the infrared splitting factor (IRSF), which is the crystallinity index proposed by Weiner and Bar-Yosef³³ and is calculated from the $\nu_4(\text{PO}_4^{3-})$ band ($\text{IRSF} = 600\text{ cm}^{-1} + 560\text{ cm}^{-1} / 590\text{ cm}^{-1}$). The highly crystalline phosphate index (HCP) and the poorly crystalline phosphate index (PCP), which are calculated from the area of the 1030 cm^{-1} and 960 cm^{-1} bands in the deconvoluted spectrum as presented by Rodríguez-Navarro et al.³⁴, provide information on the degree of crystallinity and the order or disorder of phosphates in the crystal lattice. An increase in the intensity of the HCP indicates that the bone mineral has become more stable and structurally ordered, while a high PCP value reflects a less crystalline bone mineral. To understand the presence and behaviour of non-apatitic phosphates in different contexts, the band areas in the region between 1100 and 1200 cm^{-1} , attributed to HPO_4^{2-} in non-apatitic environments^{34–36}, were used based on the deconvolution of the $\nu_1, \nu_3(\text{PO}_4^{3-})$ region. The FWHM of the $\nu_3(\text{PO}_4^{3-})$ region around 1010 cm^{-1} was used to understand the order and strain in the crystal lattice, considering particle size and sample preparation²⁸.

Indices	Peaks (cm ⁻¹)	Description	Baseline correction (cm ⁻¹)	References
Amide I/PO ₄	1640 / 1010	Amide I to phosphate	1710–1590	29,37
			1150–890	
IRSF	(560 + 600) / 590	Infrared splitting factor. Indicate crystal size and order in the matrix	660/640 470/420	33
C/P	1410 / 1010	Carbonate to phosphate	1590–1290	38
C/C	1455 / 1410	A + B type carbonates band area in deconvoluted spectrum		34,39
FWHM(ν_3 PO ₄)	1010	Full width at half-maximum of the ν_3 (PO ₄) band, indicating lattice order	1150–890	21,28
OH/PO ₄	630 / 605	OH ⁻ groups libration	660/640 470/420	40
Calcite/PO ₄	712 / 1010	Calcite to phosphate	730–700	41
HPO ₄ 1118	Acid phosphate band area in deconvoluted spectra at 1118 cm ⁻¹			34–36
HPO ₄ 1145	Acid phosphate band area in deconvoluted spectra at 1145 cm ⁻¹			35,36
HCP	Highly crystalline phosphate, band area at 1030 cm ⁻¹ in deconvoluted spectra			34
PCP	Poorly crystalline phosphate, band area at 960 cm ⁻¹ in deconvoluted spectra			34

Table 1. Chemical parameters calculated from the analysis of bone through ATR-FTIR and used in this study.

X-ray diffraction (XRD)

The mineralogy and crystallinity of fossil bone were characterized by X-ray diffraction (XRD) using an Xpert Pro X-ray powder diffractometer (Panalytical) at the Universidad de Granada (Spain). Bone powder samples (approx. 0.2 g) were measured in reflection mode in a standard sample holder using copper radiation and a 5–120° 2Theta angular range (0.013° steps; 100 s integration time). Rietveld refinement analysis of XRD profiles was performed with TOPAS 5.0 software (Bruker). The Rietveld refinement was applied to calculate the apatite unit cell parameters and volume^{31,42}. Mineral phases fitted were hydroxylapatite, calcite, and quartz. For hydroxylapatite, unit cell parameters were refined. Other relevant parameters (zero error) were also refined, and we got a goodness of fit (Rwp) between 5 and 7%.

Statistical analysis

Statistical analyses were performed using R software version 4.2.2. The results were expressed as mean ± standard deviation (Table S2). The chemometric indices were first previsualised through violin plot diagrams and examined for normality using the Shapiro–Wilk test and Levene's test to test homoscedasticity. Based on the results, the non-parametric Kruskal–Wallis test or Welch's ANOVA test were used to compare the difference assemblage. After that, a Wilcoxon rank sum test for pairwise comparison was applied to discern the differences between assemblages. *P*-values ≤ 0.05 were considered statistically significant. Subsequently, a PCA, which is a linear combination of the original variables, was used to observe the distribution and importance of the variables. A PCA allows us to better understand which variables are more important in the distribution of the samples as well as allows us to better understand the trends of the diagenetic trajectories⁴³.

Additionally, a correlation matrix was constructed to examine the associations between the variables. This is a valuable technique for managing multicollinearity in the dataset. If two or more variables exhibit a high correlation (values greater than 0.75 or -0.75), it suggests that there is considerable redundancy in the information they provide for the classification models. For enhanced understanding of the data, the parameters for a modern adult pig's cortical femur (*Sus scrofa domestica*) are provided herein: Amide I/PO₄ = 0.168 ± 0.000; IRSF = 3.235 ± 0.021; C/P = 0.231 ± 0.001; C/C = 0.736 ± 0.026; FWHM(ν_3 PO₄) = 114.80 ± 2.42; HPO₄1118 = 0.023 ± 0.000; HPO₄1145 = 0; HCP = 0.098 ± 0.002; PCP = 0.073 ± 0.002; Cell volume = 532.5; a-axis = 9.437; c-axis = 6.904.

Classification based on machine learning algorithms

Machine Learning (ML) models are used to identify patterns and relationships in data with the goal of making predictions or classifications (among other tasks) based on new information. A model is the result of training a dataset using an algorithm, which is the process or set of steps that adjusts the model's parameters. While the algorithm defines how learning occurs from the data, the model encapsulates the learned knowledge and is used to perform specific tasks. When these methods are used as classification tools, they are very robust. Moreover, with respect to the type and size of samples supervised, ML models encounter fewer issues than other traditional statistical methods or unsupervised ML methods. This makes them very useful in addressing archaeological problems⁴⁴.

The application of ML to biostratigraphic problems has already been addressed in several studies on surface modifications, bone breakage, and skeletal profiles^{44–47}. Similarly, they have been applied to diagenetic problems in both archaeology and forensics^{13,48,49}, demonstrating important results in the field of classification and prediction.

In this study, we used 9 algorithms that have been identified as the most powerful classifiers available, allowing us to compare the variability of results and samples: NNET (neural networks), SVM (support vector machines), KNN (k-nearest neighbour), RF (random forest), DTC5.0 (decision trees using the C5.0 algorithm), LDA (linear discriminant analysis), MDA (mixture discriminant analysis), PLS (partial least squares), NB (naive Bayes). Due to the differences in the way each algorithm processes the data set, different taphonomic aspects can be understood. Therefore, it is necessary to consider those algorithms that give inferior results when

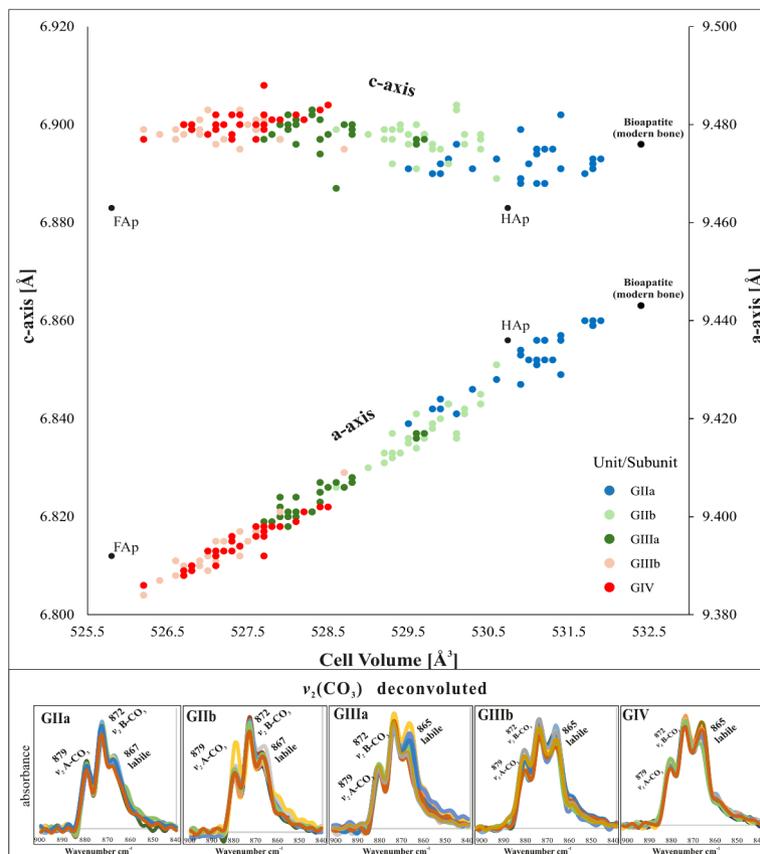


Fig. 3. >Dependence of the apatite crystallographic a-axis and c-axis values on the unit cell volume of fossil bones from the Galería site's stratigraphical sequence. An adult pig femur (*Sus scrofa domestica*) is used as a modern bone reference (Bioapatite). FAp (Fluorapatite reference), HAp (hydroxylapatite reference). Refer to Fig. S2 to view the diffractograms. Below, the deconvolution of the $\nu_2(\text{CO}_3)$ region in the FTIR spectrum for each unit is shown. 879 cm^{-1} refers to type-A carbonates, 872 cm^{-1} to type-B carbonate substitution. Note that the carbonate band at 865 cm^{-1} is more defined in the bones bearing F^- (GIII and GIV units).

	GIIa	GIIb	GIIIa	GIIIb	GIV
IRSF	4.803 ± 0.386	4.362 ± 0.281	4.242 ± 0.209	4.186 ± 0.144	4.161 ± 0.134
C/P	0.138 ± 0.024	0.188 ± 0.027	0.194 ± 0.026	0.204 ± 0.022	0.226 ± 0.056
Calcite/ PO_4	–	0.002 ± 0.005	0.003 ± 0.007	0.003 ± 0.008	0.013 ± 0.002
OH/PO_4	0.125 ± 0.078	0.019 ± 0.051	–	–	–
HPO_4_{1118}	0.017 ± 0.002	0.017 ± 0.002	0.020 ± 0.002	0.021 ± 0.001	0.024 ± 0.002
HPO_4_{1145}	0.270 ± 0.065	0.136 ± 0.043	0.082 ± 0.040	0.117 ± 0.020	0.066 ± 0.028
C/C	0.977 ± 0.093	0.843 ± 0.052	0.764 ± 0.048	0.785 ± 0.060	0.725 ± 0.048
PCP	0.028 ± 0.007	0.034 ± 0.008	0.028 ± 0.009	0.022 ± 0.007	0.018 ± 0.007
HCP	0.170 ± 0.018	0.148 ± 0.009	0.151 ± 0.006	0.170 ± 0.010	0.181 ± 0.011
$\text{FWHM}(\nu_3\text{PO}_4)$	69.437 ± 4.111	72.962 ± 2.889	70.788 ± 4.052	68.850 ± 2.719	66.318 ± 2.768
Amide I/ PO_4	0.009 ± 0.002	0.010 ± 0.002	0.010 ± 0.001	0.011 ± 0.001	0.013 ± 0.002
Cell volume (Å^3)	530.9 ± 0.7	529.7 ± 0.5	528.5 ± 0.9	527.1 ± 0.5	527.4 ± 0.6
a-axis (Å)	9.431 ± 0.007	9.417 ± 0.006	9.405 ± 0.008	9.393 ± 0.005	9.394 ± 0.004
c-axis (Å)	6.892 ± 0.003	6.897 ± 0.003	6.899 ± 0.003	6.899 ± 0.002	6.900 ± 0.002

Table 2. Summary of fossil bone property values (mean \pm SD) for each unit/subunit. The value units are dimensionless in the case of ratios and area calculations in FTIR, except for $\text{FWHM}(\nu_3\text{PO}_4)$, which is expressed in cm^{-1} .

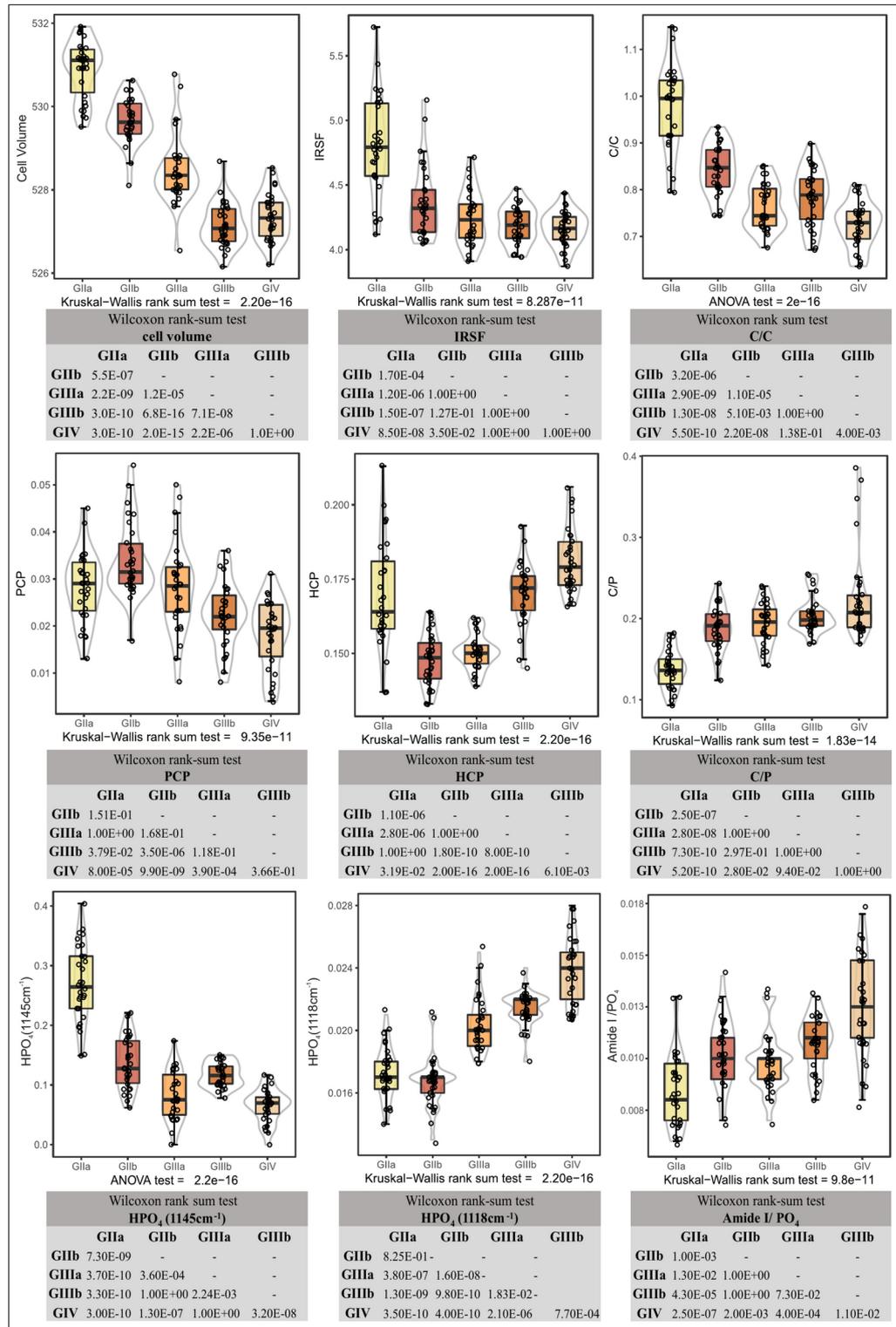


Fig. 4. Box plot of diagenetic parameters determined from fossil bone and pairwise comparison between stratigraphical sequence. The rest of the parameters are shown in Fig. S1 and Table S2.

interpreting an assemblage. The training of all these classification models was performed using the “caret” library in R language⁵⁰.

The operation was based on dividing the set of samples into 70% to be used for training and the remaining 30% for testing. For this purpose, the data are first centred and scaled. The accuracy value obtained checks the precision of the classification models, a value between 0 and 1. The closer to 1, the closer to a 100% correct classification of the samples. The kappa value is an index that reflects the possibility of a prediction occurring only

by chance. This value ranges from -1 to 1 , where 1 indicates perfect classification, -1 represents a completely inverted model that systematically misclassifies all labels, and 0 corresponds to a classification that is entirely random. Cross-validation has been used with ten folds and 10 repetitions to improve the quality of the analyses, overcoming the model's overfitting. In order to create accurate models, we used the function “tuneLength” from the ‘caret’ library 45 to create 10 different random models with different hyperparameter configurations of all the algorithms (except LDA, as it does not allow the configuration of the hyperparameters to be changed), which can be compared through kappa agreement index. In addition, the values of sensitivity (true positive rate), specificity (true negative rate), and balanced accuracy (precision of each category analysed) are considered. Note that the final configuration of the best-fit model (i.e. the selected hyperparameters) has been included in the supplementary files for all algorithms, along with the complete script used in this study (Table S3).

Results

The main chemometric and mineralogical results of the analysis of fossil bone samples obtained by FTIR and XRD are summarized in Table 2, Figs. 4, S1, S2, and Table S2. The infrared spectra primarily showed bands attributed to phosphate and carbonate groups, with minor contributions from amides. Specifically, peaks associated with carbonates were observed at 1450 cm^{-1} ($\nu_3\text{CO}_3$), 1415 cm^{-1} ($\nu_3\text{CO}_3$), $872\text{--}878\text{ cm}^{-1}$ ($\nu_2\text{CO}_3$) and 712 cm^{-1} ($\nu_4\text{CO}_3$), while peaks at 1030 cm^{-1} ($\nu_3\text{PO}_4$), 960 cm^{-1} ($\nu_3\text{PO}_4$), 604 cm^{-1} and 565 cm^{-1} are attributable to phosphate bands ($\nu_4\text{PO}_4$). Finally, peaks at 1118 cm^{-1} and 1145 cm^{-1} , observed in the deconvoluted spectra, are attributed to $\nu_3(\text{PO}_4)$ bands associated with HPO_4 (Fig. S3). The $\sim 1640\text{ cm}^{-1}$ peak is attributable to Amide I and the 630 cm^{-1} peak is associated with hydroxyl groups (exclusively in unit GII). Furthermore, using the 2nd derivative analysis of the $\nu_4\text{PO}_4$ region, after Savitzky–Golay smoothing, allowed us to confirm that bone mineral was mainly carbonate hydroxylapatite for unit GII, while samples from GIIIA onwards presented F-containing phases^{31,42,51–53} (Figs. 2 and 3). Higher intensities of these bands can be observed in units GIII and GIV compared to GII, suggesting a mix of crystalline phases in the bone assemblages of units GIII and GIV (Fig. 2). Similarly, the apatite unit cell parameters (determined by XRD; Table S1; Fig. 3) demonstrate a decrease in the unit cell volume in relation to the increased incorporation of F^- in the apatite crystal lattice within the bone assemblage of each subunit. The increase of F^- in the bone mineral remains occurs towards the top of the stratigraphic sequence. Fossil bones found in stratigraphic units GIIIB and GIV have a higher mixed phase compared to subunits GIIA and GIIb, which have values closer to hydroxylapatite than fluorapatite. The ranges and averages for each unit are expressed in Fig. 4 and Table 2. A progressive decrease in a-axis and unit cell volume values is observed as the stratigraphic sequence increases until GIIIB is reached. The results of the Wilcoxon sum test for apatite unit cell volume and a-axis show differences between all stratigraphic units (see Fig. 4 and Table S2) except GIIIB and GIV.

Regarding the infrared data, the results of the several indices used for crystallinity and phosphate phase (IRSF, HCP, PCP, HPO_4 , $\text{FWHM}(\nu_3\text{PO}_4)$) showed significant differences across all fossil bone assemblages (Fig. 4 and Table 2). Although the results suggested a tendency for apatite crystallinity (IRSF) to increase as the remains get older in the stratigraphic sequence, the Wilcoxon test showed that only the GII unit exhibits dissimilarities in contrast to the other units. On the other hand, GIIIA, GIIIB, and GIV demonstrate no differences through the IRSF (Fig. 4). Statistically significant differences (Kruskal–Wallis = $2.20\text{E}-16$) were found through highly crystalline phosphates (HCP), excluding the comparison between GIIb–GIIIA (Wilcoxon test = 1) and GIIA and GIIIB (Wilcoxon test = 1). The variations in poorly crystalline phosphates (PCP) are less noticeable (Kruskal–Wallis = $9.35\text{E}-11$). However, statistically significant differences were found among the subunits GIIA–GIIb (Wilcoxon test = $1.10\text{E}-06$), GIIIA–GIIb (Wilcoxon test = 0.000), as well as GIIIB–GIV (Wilcoxon test = 0.006). The $\text{FWHM}(\nu_3\text{PO}_4)$, which considers the width of the phosphate band at $\sim 1010\text{ cm}^{-1}$, is related to the order/disorder of the crystalline lattice²⁸, with a value of approximately $\sim 115\text{ cm}^{-1}$ in modern bones. The results indicated that the GIIIB and GIV units differ from the others, and there are no statistically significant differences between GIIA–GIIIA–GIIb (Wilcoxon test = 1), and GIIIA–GIIb (Wilcoxon test = 0.496).

Finally, acid phosphate, also known as monohydrogen phosphate (HPO_4) in bone mineral is a factor to consider in bone diagenesis, as the substitution of phosphate groups with HPO_4 induces distortions in the apatite crystal lattice^{54,55}. Two positions for acid phosphate ions have been recorded around 1118 cm^{-1} and 1145 cm^{-1} for the $\nu_3\text{PO}_4$ domain^{35,36,56}. The HPO_4 content around 1118 cm^{-1} is correlated with the a-axis parameter ($r^2 = -0.74$) and the signal of poorly crystalline phosphates ($r^2 = 0.69$) (Fig. 5). The HPO_4 ions at 1145 cm^{-1} are correlated with the substitution of type-A and type-B carbonates ($r^2 = 0.86$) and with the a-axis parameter ($r^2 = 0.76$). The HPO_4 content showed statistically significant differences for all sets. The case of HPO_4 (1118 cm^{-1}) allowed for the differentiation of all the sets except GIIA–GIIb (Wilcoxon test = $8.25\text{E}-01$). Furthermore, the case of HPO_4 (1145 cm^{-1}) allowed for the differentiation of all the sets except GIIIA–GIV (Wilcoxon test = 1).

Regarding carbonate content in bone mineral, the indices carbonate to phosphate (C/P) and the A+B carbonates ratio (C/C) have enabled the differentiation of assemblages and the observation of distinct dynamics among geological units (Fig. 4). The bone assemblage of the GIIA subunit exhibits a loss of carbonate (C/P 0.14 ± 0.02). Starting from the GIIb subunit, the content becomes more stable (C/P 0.19 ± 0.03) and only allows for differentiation between GIIb and GIV (Wilcoxon test = 0.028). The C/C index, which relates type-A carbonates (substitution of OH^- groups) and B-type carbonates (substitution of PO_4^{3-}), has revealed differences among all bone assemblages (Welch's ANOVA test = $2\text{e}-16$) (Fig. 4). The C/C ratio decreases as the stratigraphic sequence progresses (Table 2), with a higher content of A-type carbonates found in the bone remains of the GIIA subunit. Additionally, there is a band development around 865 cm^{-1} in the fossil bone remains of the GIII and GIV units (Fig. 3).

Calcite, as a separate mineral phase, is detected through the vibrational mode $\nu_4\text{CO}_3$ at 712 cm^{-1} . The results indicate a high content of calcite in the bone assemblage of the GIV unit (ranging from 0 to 0.073 with an average of 0.013 ± 0.021), as compared to the bone assemblages of GIIIA (ranging from 0 to 0.022 with an average of

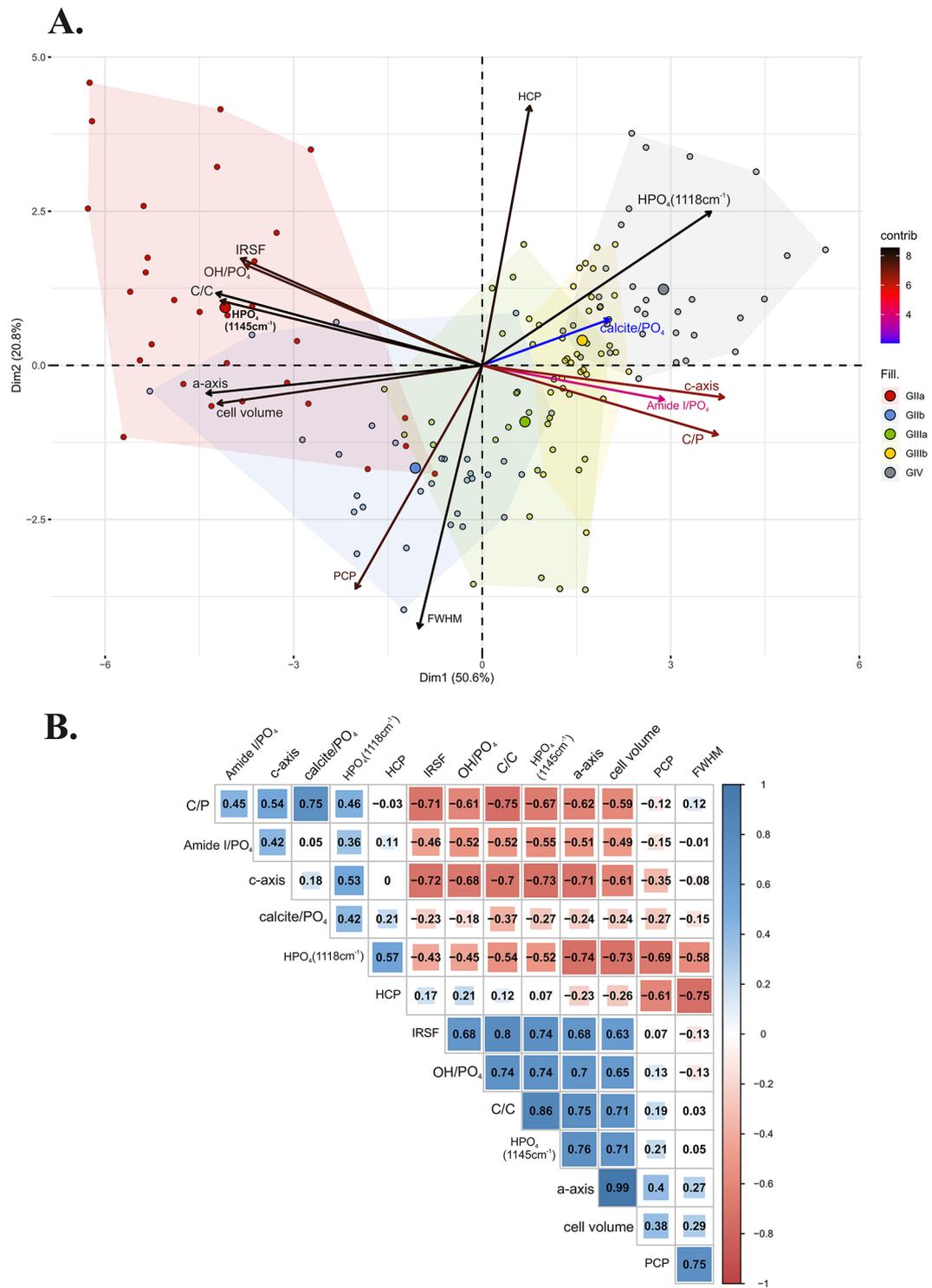


Fig. 5. (A) Principal component analysis (PCA) showing the dispersion of the bone samples according to all the diagenetic related chemometric parameters used in this study. The darker a vector is, the more it contributes to the spatial distribution. (B) Correlation matrix for all parameters used.

0.003 ± 0.007), GIIb (ranging from 0 to 0.032 with an average of 0.004 ± 0.002), GIIb (ranging from 0 to 0.019 with an average of 0.002 ± 0.001), and the absence of calcite in the remains of GIIa (Fig. S1).

Finally, OH⁻ liberation is common in the fossil bone remains of the GIIa subunit (23 out of 30 with an average of 0.13 ± 0.08) and some in the GIIb subunit (4 out of 30 with an average of 0.02 ± 0.05). This however does not occur in the rest of the sequence. This index is correlated with C/C (r² = 0.74), HPO₄ (1145 cm⁻¹) (r² = 0.74), and

the a-axis ($r^2=0.71$), which may indicate an arrangement of carbonate ions and acid phosphate ions along the a-axis, especially in the remains of the GIIa subunit.

Concerning the organic matter contents, the Amide I/PO₄ ratio falls below 0.02 ± 0.00 , indicating minimal organic matter preservation. Nevertheless, the findings demonstrate statistically significant distinctions in the Amide I amount between units GIV and GIIa (Wilcoxon test = $2.50E-07$). Conversely, no statistically significant differences were observed among subunits GIIb, GIIa, and GIIb. Notably, it is within unit GIV that fossil bones have retained a higher Amide I content (0.013 ± 0.002) (Fig. 4).

Principal components analysis results

The PCA explained 71.4% of the variance with two principal components, revealing the distribution of samples around two dynamics (Fig. 5a, Table S2). On one side, subunits GIIa and GIIb are positioned on the left half of the plot, associated with indices indicating more defined chemical and structural changes of bone mineral compared to units GIII and GIV (IRSF, C/C, HPO₄-1145 cm⁻¹, a-axis, and unit cell volume contributed to their distribution). On the other side, subunits GIIa, GIIb, and unit GIV are positioned on the right half of the plot, linked primarily to carbonate content indices C/P, HPO₄-1118 cm⁻¹, and HCP. This illustrates two trajectories within this stratigraphic sequence: one related to hydroxylapatite precipitation and carbonate dissolution, and another one related to the absorption of exogenous carbonates and fluoride ions. The means of each group are distributed differently and the means of GIIb and GIV are closer to each other. The variables contributing to the distinct distribution of GIIb compared to GIIa and GIIa compared to GIIb are FWHM(ν_3 PO₄) and PCP, indices related to the crystallinity of bone mineral.

Machine learning classification results

All the results of the ML analysis are presented in supplementary files (Table S3). The classification analyses of bone samples using ML algorithms yielded consistent results across the different methods (Table 3). As demonstrated in the PCA and the correlation matrix of variables, a strong relationship (≥ 0.75) exists among different variables. Furthermore, several variables express the same information in the component space (Fig. 5). Therefore, there is multicollinearity in the results that may lead to data overfitting⁵⁷. Taking this into account, we removed variables with a correlation ≥ 0.75 , either positive or negative, those that contribute little to the component space, and exclusive variables, i.e., those not present in all fossil bone assemblages (Amide I/PO₄, IRSF, C/P, calcite/PO₄, OH⁻ liberation, C/C, FWHM(ν_3 PO₄), c-axis, a-axis). Thus, the models considered the variables cell volume, HCP, PCP, HPO₄ (1118 cm⁻¹) and HPO₄ (1145 cm⁻¹). The least accurate results (Tables 3 and S3) were obtained with the PLS algorithm (accuracy = 0.69; kappa = 0.61) and DTC5.0 (accuracy = 0.76; kappa = 0.69). In contrast, the models that best classified the assemblages were NNET (accuracy = 0.89; kappa = 0.86) and MDA (accuracy = 0.87; kappa = 0.83). The remaining algorithms yielded similar results: SVM (accuracy = 0.84; kappa = 0.81), NB (accuracy = 0.84; kappa = 0.81), RF and LDA (accuracy = 0.82; kappa = 0.78), and KNN (accuracy = 0.80; kappa = 0.75). In other words, neural networks and the mixture discriminant analysis models were better classifiers.

Therefore, the most important variables for classification are therefore related to the behaviour of phosphates with the apatite crystal lattice parameters. In both the NNET and MDA models, the importance of variables lies between the apatite unit cell volume and HPO₄ ions. The performance of each algorithm is presented in Table S3, which includes sensitivity, specificity, precision, and balanced accuracy values in relation to the results for each unit and subunit. It is observed that the difficulty in classifying the fossil bone assemblages lies in distinguishing between the remains of GIIb and GIV, a result that is also evident through the Wilcoxon rank-sum test (Fig. 4; Table S2). The main differences observed between GIIb and GIV lie in the arrangement of HPO₄ ions.

Discussion

A taphosystem consists of a preserved association and its external environment¹⁰. In the case of bone diagenesis in karstic taphosystems, bones are modified by processes stemming from the composition of sedimentary fill, pH, Eh, environmental conditions surrounding caves, soil hydrology, cave hydrological regime, taphon interactions, and the prior taphonomic history before fossil diagenesis^{58,59}. Consequently, common features in the preservation mode of remains in karstic taphosystems may be evident, indicating a diagenetic pathway. Furthermore, differences in the preservation mode of remains can occur along a stratigraphic sequence due to the variability of these mentioned characteristics throughout the site's formation. In the case of the Galería site, the obtained results suggest that the diagenetic history differs for each unit (GII, GIII, and GIV). Additionally, differences have been observed between subunits (GIIa-GIIb, GIIa-GIIb). These differences arise during the formation of the stratigraphic sequence, resulting in a diagenetic history that will be discussed for each unit and subunit (Fig. 6).

	NNET	SVM	KNN	RF	MDA	NB	LDA	PLS	DTC5.0
Accuracy	0.89	0.84	0.80	0.82	0.87	0.84	0.82	0.69	0.76
Kappa	0.86	0.81	0.75	0.78	0.83	0.81	0.78	0.61	0.69
Accuracy lower	0.76	0.71	0.65	0.68	0.73	0.71	0.68	0.53	0.61
Accuracy upper	0.96	0.94	0.90	0.92	0.95	0.94	0.92	0.82	0.87

Table 3. Accuracy and kappa values provided by the ML algorithms used to classify the different units from Galería site.

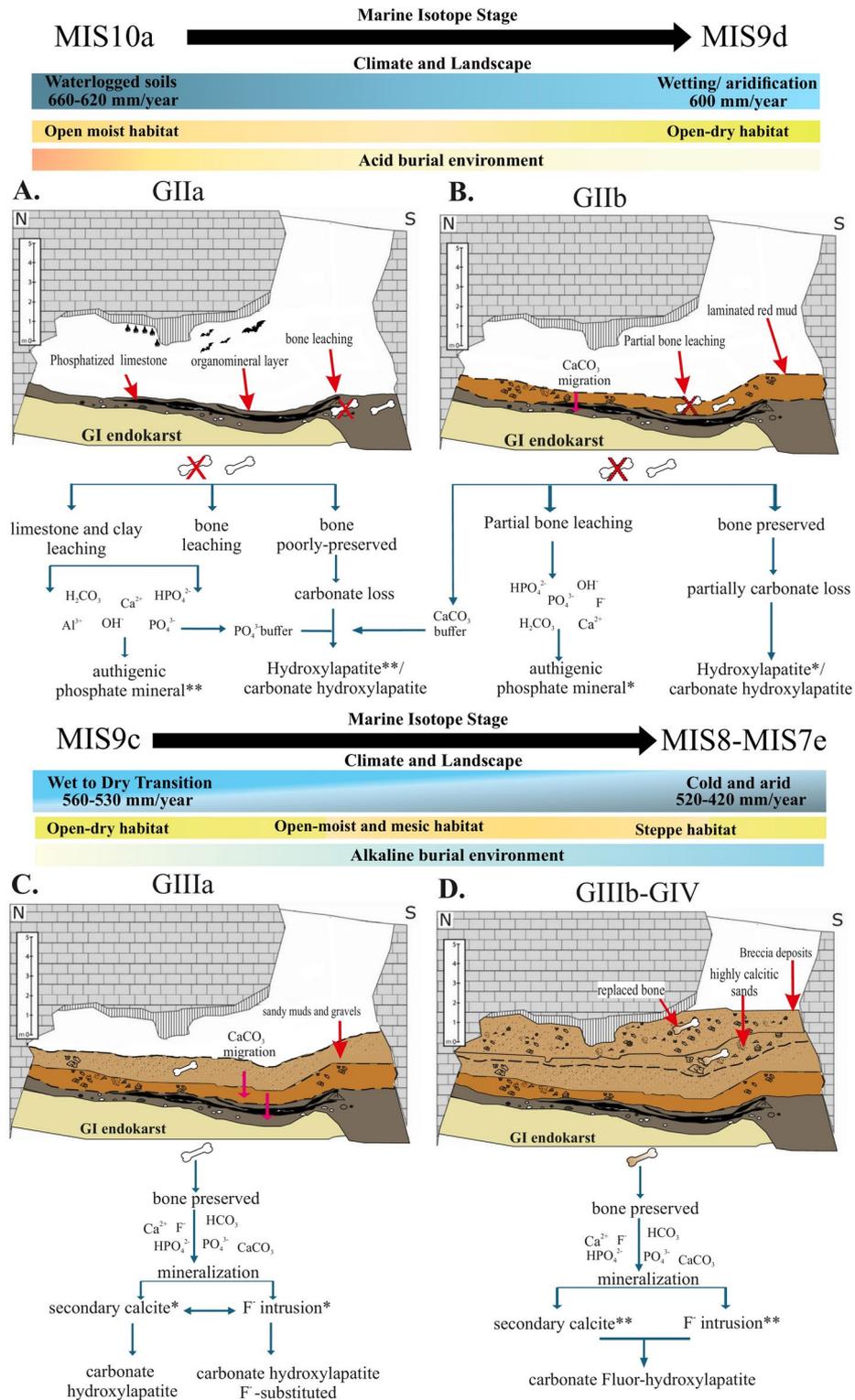


Fig. 6. Conceptual model by unit/subunit with climatic and environmental evolution over time. * = moderate process; ** = intense process. Schematic section modified from⁶⁰. (A) *GIIa subunit*: Under this acidic environment with organomineral layers, dissolution–precipitation processes lead to the formation of authigenic phosphate minerals and the loss of carbonate from the bone remains, bringing the mineral phase closer to hydroxylapatite. (B) *GIIb subunit*: As calcite precipitates, the environment becomes more alkaline, which helps to mitigate the leaching of bone material. (C) *GIIIa subunit*: Under increasingly alkaline conditions, calcite precipitation occurs, along with a replacement by F-containing apatite phases. (D) *GIIIb-GIV units*: The intensity of fluorine incorporation, along with calcite precipitation, increases. Enhanced calcite precipitation further promotes greater fluorine incorporation.

From bone mineral to hydroxylapatite in GII unit

The GIIa subunit, framed in MIS 9–10 with a weighted mean age of 313 ± 14 ka by ESR/US method²⁷ (combined electron spin resonance and U-series dating), is lithologically defined by layered mudstone beds. This is a mollisol, a soil rich in organic matter with a high water-holding capacity, where acid leaching leads to the alteration of limestone clasts^{22,61,62}. The fine-grained sediments, enriched with organic matter, exhibit significant post-depositional alteration, with evidence of in situ leaching, formation of humate microfacies, high biological activity, rhizolites, and authigenic phosphate minerals^{23,62}.

Considering the sediment composition, crandallite ($\text{CaAl}_3(\text{PO}_4)_2(\text{OH})_5 \cdot \text{H}_2\text{O}$), an authigenic mineral formed from phosphate-rich solutions, is also present²², indicating a shift in pH from alkaline or near-neutral to acidic^{63–65}. Regarding the paleoenvironment and climatic context of this subunit, Bógalo et al.⁶¹ were able to interpret that the hydrological regime corresponds to 620–660 mm/year. Meanwhile, other studies have observed that it represents humid conditions with an open habitat and waterlogged soils inside the cave⁶⁶. The GIIb subunit, framed in MIS9 and dating back to 237–269 ka by ESR-U-series²⁷, presents a milder hydrological regime of around 600 mm/year with periods of wetting and aridification⁶¹. Similarly, Vallverdú⁶² aligns with the alternation of wet and dry periods for this subunit, defined as vertisol and inceptisol, where a lower content of organic matter is recorded. Thus, in the GII unit, an acidic environment develops, which prevents carbonation (acid leaching) and affects the limestone clasts, altering them. There are processes from which the bone remains are not exempt. Pérez-González et al.²² and Vallverdú⁶², pointed out that in the GIIa subunit, there is an absence of carbonates (2%), and as the stratigraphic sequence advances toward GIIb, there is a more pronounced presence (18–32%). Similarly, this process affects the bones, significantly dissolving their carbonate portion in the GIIa subunit and more mildly in the GIIb subunit (Fig. 6a,b), where some fossil bone remains even contain calcite. Thus, it is observed that the preserved remains in GIIa have undergone structural carbonate loss (Fig. 4), while the dissolved remains have contributed to the formation of authigenic phosphate minerals. The faunal remains also suggest significant bone dissolution and loss. Out of the 2704 bone remains recovered from the GII unit, only 2.7% belong to subunit GIIa⁶⁷. In subunit GIIb, there is a proportion of 83.4% ($n = 2632$) of bone remains compared to 16.6% ($n = 523$) of lithic industry, a trend that persists throughout the sequence. However, in the GIIa subunit, 19.2% ($n = 72$) of bone remains were found compared to 80.8% ($n = 303$) of lithic industry⁶⁷. While the difference between subunits in the lithic assemblage is not particularly large, there is a difference in the proportion of bone remains. Clearly, significant acidic leaching processes have occurred in GIIa, resulting in the loss of bone material in a subunit where much anthropic activity would be expected due to the full functioning of the natural trap as lithic material appears^{67–71}.

As the burial history continues, the environmental conditions in GIIb become increasingly dry, alternating between periods of wetting and aridification⁶¹. Additionally, the diagenesis affecting the bone assemblage in GIIb is less pronounced, with reduced leaching compared to GIIa. Sedimentologically, the GIIb subunit is characterized by layers of red clay, where the limestones exhibiting less alteration and reduced dissolution compared to those in GIIa (Fig. 6a,b). Our results indicate that the deterioration of the carbonate component of bones leads to a precipitation of hydroxylapatite phase. This dissolution-(re) precipitation process is more pronounced in GIIa than in GIIb (as evident from the values of cell volume, *c*-axis, HCP, IRSE, C/P in Figs. 3 and 4). Even for the fossil bone remains in the GIIa subunit, slight OH^- groups liberation is detected, which show an advanced deterioration of the fossil bone remains⁷². The loss of carbonate in the apatite crystal structure results in a pure hydroxylapatite phase, and a shoulder at 630 cm^{-1} can be obtained⁷³. This process may also be related to the incorporation of HPO_4 ions and pH changes⁷⁴. Our findings regarding bone preservation suggest a trajectory characterized by acid leaching/dissolution, decarbonation, and reprecipitation into a more stable apatite phase in the GII unit. The dissolution and (re)precipitation can lead to the formation of authigenic minerals, which may remain stable during subsequent episodes of environmental changes. These processes have been experimentally documented in the early stages of guano-driven diagenesis⁷⁵. The dissolution of bone material stops, enabling the fossilization of remains once the pH is buffered by the migration of CaCO_3 from the overlying units (beginning in GIIb, see Fig. 6b). This buffering facilitates a transition from an acidic environment to a mildly acidic environment where hydroxylapatite precipitates along with other authigenic phosphate minerals. In this way, the leaching processes are reduced in GIIb, allowing for better bone preservation and less limestone alteration.

Fluorine replacement in GIII and GIV units

A notable shift in the burial history and conditions at the deposition site is observed beginning with the GIII unit. From this unit onward, bone diagenesis and fossilization are dominated by the incorporation of fluorine and calcite (see Fig. 6). GIIIa is situated within MIS9, with a chronology of 231 ± 18 ka TT-OSL (thermally transferred optically stimulated luminescence) and 244 ± 16 ka pIR-IR225 (post-infrared infrared stimulated luminescence)²³, mainly made up of allochthonous sediments, predominantly containing rounded pebbles. From this unit onward, a transition from wet to dry environments is observed, progressing from GIIIa to GIIIb, reflecting cold and arid conditions by the time GIV is reached^{24,61,66}. This should reflect differences in the burial environment and, consequently, in the fossilization of the bone assemblages. According to our results, the main differences between these subunits occur in the substitution of carbonates within the apatite lattice and the arrangement of HPO_4 ions, which appear to be related to the substitution of A-type and B-type carbonates in the apatite structure. Therefore, in the fossil bone assemblage of GIIIb, there is a higher predominance of A-type carbonates compared to the GIIIa assemblage (Figs. 3 and 4). From this lithostratigraphic unit onwards, leaching is not as pronounced as it is in the GII unit, a fact that is also observed in the lack of chemical alteration in the limestones⁶². Thus, it can be observed that the burial environment is more alkaline than in GII, leading to the precipitation of calcite, which is also reflected in the bone assemblages. Moreover, the information provided by the Savitzky-Golay 2nd derivative algorithm and the results from the dependence of *a*-axis and cell volume (Figs. 2 and 3) have allowed us to observe an F^- substitution in the fossil bone remains of units GIII and GIV.

Therefore, the diagenetic pathway is characterized by coupled dissolution-(re) precipitation processes^{76,77}, where OH⁻ groups are replaced by F⁻, resulting in the formation of distinct fluorine-containing apatite phases. (Fig. 6c).

Both the GII and GIII units hold key significance for the interpretation of Middle Pleistocene human groups in Europe. This importance stems from the presence of highlighted Acheulean lithic set⁷⁸ as well as accumulated faunal records crucial for understanding subsistence dynamics^{68,79}. Therefore, thorough control of fossilization and preservation pathways is essential for aiding in the resolution of zooarchaeological and taphonomic data. It is also essential for the acquisition of paleoproteomics, isotopes, paleogenetics data, or dating methods^{80–84}.

Continuing with the GIV unit, framed in MIS 8–7, the sedimentary dynamics observed in the GIII unit persist, although breccia sediments become more predominant towards the end of the sequence. Several proxies have indicated cold and arid conditions for this event^{24,61,66,68}. Regarding cave usage, there is no longer evidence of its recurrent utilization by human groups. Diagenetic dynamics, like sedimentation, closely resemble those observed in the GIIb subunit. However, differences have been identified between the GIIb and GIV fossil assemblages, in which with GIV exhibits a higher content of highly crystalline phosphates and HPO₄ at 1118 cm⁻¹. Nevertheless, the similarity in their fossilization dynamics can be attributed to rapid sedimentation during the latter part of the sequence, as discussed in previous studies^{22,85}. If we consider the behaviour of carbonates in the fossil assemblages, we also identify significant contributions. Concerning the ν₃CO₃ region for the GIII and GIV units, it is worth noting that the 865 cm⁻¹ band, attributed to an environment characterized by labile amorphous calcium carbonates^{42,86,87}, exhibits greater definition compared to unit GII, where it is positioned at 867 cm⁻¹ and is less well-defined (Fig. 3). This result may suggest that labile carbonates are favoured by ionic substitutions of F⁻ and OH⁻, as previously proposed³¹. However, it should be noted that this region is ambiguous, and ionic interactions can be highly complex⁸⁸, making it challenging to discern the contribution of carbonate ions. Furthermore, calcite precipitation within the bone tissue promotes micro-fracturing and increased porosity, facilitating F⁻ replacement⁸⁹, as observed in the GIV unit. Finally, following the formation of the GIV unit, the GV unit completes the cavity's infilling, enabling the fossilization of bone remains within the described contexts.

Overview of the burial history

The Galería site reveals different diagenetic processes that evolved in response to changing burial environments along the stratigraphic sequence (Fig. 6). Crystalline phases began to evolve with the loss of organic material during early diagenesis and were refined during late diagenesis through dissolution-(re)precipitation processes.

In the GII unit (MIS 10a–MIS 9d), acidic and wet conditions dominated, particularly in GIIa, where leaching caused significant loss of bone material and led to the formation of authigenic phosphates. The preserved bone remains exhibit a hydroxylapatite phase with dissolved carbonate (Fig. 6a). As conditions progressed into GIIb, drier environments facilitated pH buffering through CaCO₃ precipitation, which stabilized the carbonate hydroxylapatite phase in the bone mineral (Fig. 6b), improving bone preservation.

In the most recent units, the burial environments of the GIII and GIV units shifted to slightly alkaline conditions. These units are characterized by calcite precipitation and fluorine replacement into the bone apatite. The dissolution-(re) precipitation processes intensified from GIIIa (MIS 9c) to GIIIb, resulting in the formation of fluorine-containing apatite phases (Fig. 6c,d). The GIIIb and GIV units (MIS 8–MIS 7e) share such a closely related diagenetic trajectory that it suggests rapid sedimentation, ultimately contributing to the complete infilling of the cave.

Machine Learning-based classification of fossil bones across stratigraphic units

Considering the observed burial history, the diagenetic parameters have enabled the establishment of classification models for the bone assemblages along the stratigraphy. These diagenetic differences allowed machine learning algorithms to classify the assemblages with an accuracy of 0.89 and 0.87 using neural networks (NNET) and mixture discriminant analysis (MDA) algorithms, respectively (Tables 3 and S3). The variables that contribute most to this classification are related to the behaviour of phosphates and the unit cell structure. The evaluation metrics demonstrate the effectiveness of the NNET model with high performance in GIIa and GIIIa (balanced accuracy of 1.00 and 0.96, respectively). Phosphate bands (HCP, HPO₄ at 1145 and 1118 cm⁻¹) and apatite unit cell volume were key contributors to the classification. Although GIIIb showed lower sensitivity (0.67), the neural network (NNET) model, optimized with a hidden layer size of 5 and a decay value of 0.1, proved robust and efficient, highlighting the relevance of chemical and structural diagenetic changes in distinguishing units.

At the nanoscale, fossil bone assemblages homogenize as they seek stability within burial environments. This homogenization enables the tracing of common preservation trajectories for fossil bone remains^{13,43}. This remains valid even when analysing cortical bone from different mammalian taxa, which, as shown in previous studies^{48,59,90} exhibit compositional variability in vivo. In any event, we can observe that the karstic taphosystem developed at the Galería site differentially preserves its fossil assemblage according to the stratigraphy. While it is true that the wide range of some variables may be attributed to spatial variability at different points within the stratigraphic units, this issue can be addressed in future studies. These machine learning models will allow us to address issues of looting and fossil bone decontextualization, at least for sites in karstic systems.

Conclusion

This study investigates the different preservation states of fossil bone remains in the Galería deposit. XRD and FTIR chemometric parameters were used to characterize the fossil assemblages within the several stratigraphic units of the Galería site. It was observed that these units showed different patterns of fossilization and preservation in the chemistry and mineralogy of the fossil bone remains. This differential preservation is attributed to different burial environment conditions during the deposit formation.

The oldest unit, the GII unit, is characterised by leaching, carbonate loss, and dissolution–precipitation processes, bringing the bone remains closer to the hydroxylapatite phase. This suggests an acidic and wet burial

environment where the dissolution of bone remains leads to the formation of authigenic phosphate minerals such as crandallite. The GIII and GIV units are characterised by an incorporation of F⁻ and CO₃ into the crystal structure. The results suggest a slightly alkaline burial environment with lower moisture than in the lower part of the sequence.

These differences have enabled the development of supervised Machine Learning classification models for fossil bone remains at the Galería site. The ML algorithms (NNET and MDA) identified apatite crystal lattice parameters and phosphate molecules as the most relevant factors for differentiating fossil bones across the stratigraphic sequence. These models provide a valuable tool for recontextualizing fossil remains and understanding diagenetic processes, with potential applications in other karstic systems.

Data availability

All data and Rcode are presented in the main text and the Supplementary Information.

Received: 21 July 2024; Accepted: 3 February 2025

Published online: 14 February 2025

References

- Rho, J. Y., Kuhn-Spearing, L. & Zioupos, P. Mechanical properties and the hierarchical structure of bone. *Med. Eng. Phys.* **20**, 92–102 (1998).
- Weiner, S., Traub, W. & Wagner, H. D. Lamellar bone: Structure-function relations. *J. Struct. Biol.* **126**, 241–255 (1999).
- Currey, J. D. The structure and mechanics of bone. *J. Mater. Sci.* **47**, 41–54 (2012).
- Schmahl, W. W., Kocsis, B., Toncala, A. & Grupe, G. J. Mineralogic characterisation of archaeological bone. In *Isotopic Landscapes in Bioarchaeology* (eds Grupe, G. & McGlynn, G. C.) 91–110 (Springer Berlin Heidelberg, 2016). https://doi.org/10.1007/978-3-662-48339-8_5.
- Stock, S. R. The Mineral-Collagen interface in bone. *Calcif. Tissue Int.* **97**, 262–280 (2015).
- Pasteris, J. D., Yoder, C. H. & Wopenka, B. Minerals in the human body: Molecular water in nominally unhydrated carbonated hydroxylapatite: The key to a better understanding of bone mineral. *Am. Mineral.* **99**, 16–27 (2014).
- Grunenwald, A. et al. Revisiting carbonate quantification in apatite (bio)minerals: A validated FTIR methodology. *J. Archaeol. Sci.* **49**, 134–141 (2014).
- Macías-Sánchez, E. et al. Spherulitic crystal growth drives mineral deposition patterns in collagen-based materials. *Adv. Funct. Mater.* <https://doi.org/10.1002/adfm.202200504> (2022).
- Fernández-López, S. R. & Fernández-Jalvo, Y. The limit between biostratigraphy and fossil diagenesis. *Curr. Top. Taphon. Foss.* 27–37 (2002).
- Fernández-López, S. R. Taphonomic alteration and evolutionary taphonomy. *J. Taphon.* **4**, 111–142 (2006).
- Hedges, R. E. M., Millard, A. R. & Pike, A. W. G. Measurements and relationships of diagenetic alteration of bone from three archaeological sites. *J. Archaeol. Sci.* **22**, 201–209 (1995).
- Hedges, R. E. M. Bone diagenesis: An overview of processes. *Archaeometry* **44**, 319–328 (2002).
- Del Valle, H. et al. ATR-FTIR to distinguish Holocene fumier facies. A perspective from bone diagenesis at El Mirador cave (Sierra de Atapuerca, Spain). *J. Archaeol. Sci.* **141**, 105582 (2022).
- Fernández-Jalvo, Y., Díez, J. C., Cáceres, I. & Rosell, J. Human cannibalism in the early Pleistocene of Europe (Gran Dolina, Sierra de Atapuerca, Burgos, Spain). *J. Hum. Evol.* **37**, 591–622 (1999).
- Cáceres, I., Lozano, M. & Saladié, P. Evidence for bronze age cannibalism in El Mirador cave (Sierra de Atapuerca, Burgos, Spain). *Am. J. Phys. Anthropol.* **133**, 899–917 (2007).
- Blasco, R. et al. Earliest evidence for human consumption of tortoises in the European Early Pleistocene from Sima del Elefante, Sierra de Atapuerca, Spain. *J. Hum. Evol.* **61**, 503–509 (2011).
- Saladié, P. et al. The role of carnivores and their relationship to hominin settlements in the TD6-2 level from Gran Dolina (Sierra de Atapuerca, Spain). *Quat. Sci. Rev.* **93**, 47–66 (2014).
- Sala, N., Arsuaga, J. L., Martínez, I. & Gracia-Téllez, A. Breakage patterns in Sima de los Huesos (Atapuerca, Spain) hominin sample. *J. Archaeol. Sci.* **55**, 113–121 (2015).
- Rodríguez-Hidalgo, A. et al. Human predatory behavior and the social implications of communal hunting based on evidence from the TD10.2 bison bone bed at Gran Dolina (Atapuerca, Spain). *J. Hum. Evol.* **105**, 89–122 (2017).
- Beasley, M. M., Bartelink, E. J., Taylor, L. & Miller, R. M. Comparison of transmission FTIR, ATR, and DRIFT spectra: Implications for assessment of bone bioapatite diagenesis. *J. Archaeol. Sci.* **46**, 16–22 (2014).
- Hollund, H. I., Ariese, F., Fernandes, R., Jans, M. M. E. & Kars, H. Testing an alternative high-throughput tool for investigating bone diagenesis: FTIR in attenuated total reflection (ATR) mode. *Archaeometry* **55**, 507–532 (2013).
- Pérez-González, A. et al. Géologie de la Sierra de Atapuerca et stratigraphie des remplissages karstiques de Galería et Dolina (Burgos, Espagne). *Anthropologie* **105**, 27–43 (2001).
- Demuro, M. et al. New luminescence ages for the Galería Complex archaeological site: Resolving chronological uncertainties on the Acheulean record of the Sierra de Atapuerca, Northern Spain. *PLoS ONE* **9**, (2014).
- Núñez-Lahuerta, C., Galán, J., Cuenca-Bescós, G., García-Medrano, P. & Cáceres, I. A bird assemblage across the MIS 9/8 boundary: The Middle Pleistocene of Galería (Atapuerca). *Quat. Sci. Rev.* **293**, 107708 (2022).
- Campana, I. et al. Reconstructing depositional environments through cave interior facies: The case of Galería Complex (Sierra de Atapuerca, Spain). *Geomorphology* **440**, (2023).
- Berger, G. W. et al. Luminescence chronology of cave sediments at the Atapuerca paleoanthropological site Spain. *J. Hum. Evol.* **55**, 300–311 (2008).
- Falguères, C. et al. Combined esr/u-series chronology of acheulean hominid-bearing layers at Trincheras Galería site, Atapuerca, Spain. *J. Hum. Evol.* **65**, 168–184 (2013).
- Kontopoulos, I., Presslee, S., Penkman, K. & Collins, M. J. Preparation of bone powder for FTIR-ATR analysis: The particle size effect. *Vib. Spectrosc.* **99**, 167–177 (2018).
- Lebon, M., Reiche, I., Gallet, X., Bellot-Gurlet, L. & Zazzo, A. Rapid quantification of bone collagen content by ATR-FTIR spectroscopy. *Radiocarbon* **58**, 131–145 (2016).
- Rodríguez-Navarro, A. B. et al. Influence of physical activity on tibial bone material properties in laying hens. *J. Struct. Biol.* **201**, 36–45 (2018).
- Stathopoulou, E. T., Psycharis, V., Chryssikos, G. D., Gionis, V. & Theodorou, G. Bone diagenesis: New data from infrared spectroscopy and X-ray diffraction. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **266**, 168–174 (2008).
- Querido, W., Ailavajhala, R., Padalkar, M. & Pleshko, N. Validated approaches for quantification of bone mineral crystallinity using transmission Fourier transform infrared (FT-IR), attenuated total reflection (ATR) FT-IR, and Raman spectroscopy. *Appl. Spectrosc.* **72**, 1581–1593 (2018).

33. Weiner, S. & Bar-Yosef, O. States of preservation of bones from prehistoric sites in the Near East: A survey. *J. Archaeol. Sci.* **17**, 187–196 (1990).
34. Rodriguez-Navarro, A. B., Romanek, C. S., Alvarez-Lloret, P. & Gaines, K. F. Effect of in ovo exposure to PCBs and Hg on clapper rail bone mineral chemistry from a contaminated salt marsh in coastal Georgia. *Environ. Sci. Technol.* **40**, 4936–4942 (2006).
35. Rey, C., Shimizu, M., Collins, B. & Glimcher, M. J. Resolution-enhanced Fourier transform infrared spectroscopy study of the environment of phosphate ion in the early deposits of a solid phase of calcium phosphate in bone and enamel and their evolution with age: 2. Investigations in the v3 PO₄ domain. *Calcif. Tissue Int.* **49**, 383–388 (1991).
36. Kourkoumelis, N. & Tzaphlidou, M. Spectroscopic assessment of normal cortical bone: Differences in relation to bone site and sex. *ScientificWorldJournal*. **10**, 402–412 (2010).
37. Trueman, C. N. G., Behrensmeyer, A. K., Tuross, N. & Weiner, S. Mineralogical and compositional changes in bones exposed on soil surfaces in Amboseli National Park, Kenya: Diagenetic mechanisms and the role of sediment pore fluids. *J. Archaeol. Sci.* **31**, 721–739 (2004).
38. Wright, L. E. & Schwarcz, H. P. Infrared and isotopic evidence for diagenesis of bone apatite at Dos Pilas, Guatemala: Palaeodietary implications. *J. Archaeol. Sci.* **23**, 933–944 (1996).
39. Thompson, T. J. U., Gauthier, M. & Islam, M. The application of a new method of Fourier transform infrared spectroscopy to the analysis of burned bone. *J. Archaeol. Sci.* **36**, 910–914 (2009).
40. Thompson, T. J. U., Islam, M. & Bonniere, M. A new statistical approach for determining the crystallinity of heat-altered bone mineral from FTIR spectra. *J. Archaeol. Sci.* **40**, 416–422 (2013).
41. Dal Sasso, G. et al. Bone diagenesis variability among multiple burial phases at Al Khiday (Sudan) investigated by ATR-FTIR spectroscopy. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **463**, 168–179 (2016).
42. Margariti, E. et al. A geochemical approach to fossilization processes in Miocene vertebrate bones from Sahabi NE Libya. *J. African Earth Sci.* **149**, 1–18 (2019).
43. Smith, C. I., Nielsen-Marsh, C. M., Jans, M. M. E. & Collins, M. J. Bone diagenesis in the European Holocene I: Patterns and mechanisms. *J. Archaeol. Sci.* **34**, 1485–1493 (2007).
44. Domínguez-Rodrigo, M. Successful classification of experimental bone surface modifications (BSM) through machine learning algorithms: A solution to the controversial use of BSM in paleoanthropology?. *Archaeol. Anthropol. Sci.* **11**, 2711–2725 (2019).
45. Arriaza, M. C. & Domínguez-Rodrigo, M. When felids and hominins ruled at Olduvai Gorge: A machine learning analysis of the skeletal profiles of the non-anthropogenic Bed I sites. *Quat. Sci. Rev.* **139**, 43–52 (2016).
46. Courtenay, L. A. et al. Combining machine learning algorithms and geometric morphometrics: A study of carnivore tooth marks. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **522**, 28–39 (2019).
47. Moclán, A. et al. Identifying the bone-breaker at the Navalmañillo Rock Shelter (Pinilla del Valle, Madrid) using machine learning algorithms. *Archaeol. Anthropol. Sci.* **12**, (2020).
48. Zupanič Pajnič, I., Leskovar, T. & Jerman, I. ATR-FTIR spectroscopy and chemometric complexity: Unfolding the intra-skeleton variability. *J. Chemom.* **36**, 1–15 (2022).
49. Pal Chowdhury, M. et al. Machine learning ATR-FTIR spectroscopy data for the screening of collagen for ZooMS analysis and mtDNA in archaeological bone. *J. Archaeol. Sci.* **126**, 105311 (2021).
50. Kuhn, M. et al. Caret: Classification and regression training. (2020).
51. Shemesh, A. Crystallinity and diagenesis of sedimentary apatites. *Geochim. Cosmochim. Acta* **54**, 2433–2438 (1990).
52. Geiger, S. B. & Weiner, S. Fluoridated carbonatoapatite in the intermediate layer between glass ionomer and dentin. *Dent. Mater.* **9**, 33–36 (1993).
53. Trueman, C. N., Privat, K. & Field, J. Why do crystallinity values fail to predict the extent of diagenetic alteration of bone mineral?. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **266**, 160–167 (2008).
54. LeGeros, R. Z. Apatites in biological systems. *Prog. Cryst. Growth Charact. Mater.* **4**, 1–45 (1981).
55. Wopenka, B. & Pasteris, J. D. A mineralogical perspective on the apatite in bone. *Mater. Sci. Eng. C* **25**, 131–143 (2005).
56. Lebon, M. et al. New parameters for the characterization of diagenetic alterations and heat-induced changes of fossil bone mineral using Fourier transform infrared spectrometry. *J. Archaeol. Sci.* **37**, 2265–2276 (2010).
57. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning, Springer Texts in Statistics* (Springer New York, 2013). <https://doi.org/10.1007/978-1-4614-7138-7>.
58. Karkanas, P., Bar-Yosef, O., Goldberg, P. & Weiner, S. Diagenesis in prehistoric caves: The use of minerals that form in situ to assess the completeness of the archaeological record. *J. Archaeol. Sci.* **27**, 915–929 (2000).
59. Kendall, C., Eriksen, A. M. H., Kontopoulos, I., Collins, M. J. & Turner-Walker, G. Diagenesis of archaeological bone and tooth. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **491**, 21–37 (2018).
60. Bermejo, L. et al. Karst features interpretation using ground-penetrating radar: A case study from the Sierra de Atapuerca, Spain. *Geomorphology* **367**, 107311 (2020).
61. Bógalo, M. F. et al. High-resolution late Middle Pleistocene paleoclimatic record from the Galería Complex, Atapuerca archaeological site, Spain - An environmental magnetic approach. *Quat. Sci. Rev.* **251**, 106721 (2021).
62. Vallverdú, J. Soil-stratigraphy in the cave entrance deposits of Middle Pleistocene age at the Trincheras del Ferrocarril sites (Sierra de Atapuerca, Spain). *Quat. Int.* **433**, 199–210 (2017).
63. Mestour, B. Contribution à l'étude des mécanismes de mise en place et d'évolution du remplissage de la Baume-Bonne (Quinson, Alpes-de-Haute-Provence, France): Apports de l'analyse minéralogique et micromorphologique des dépôts phosphatés et des dépôts carbonatés. (Archéologie et Préhistoire. Museum national d'histoire naturelle - MNHN PARIS, 1996).
64. Karkanas, P., Rigaud, J. P., Simek, J. F., Albert, R. M. & Weiner, S. Ash bones and guano: A study of the minerals and phytoliths in the sediments of Grotte XVI, Dordogne France. *J. Archaeol. Sci.* **29**, 721–732 (2002).
65. Sokol, E. V. et al. Phosphate record in pleistocene-holocene sediments from Denisova cave: Formation mechanisms and archaeological implications. *Minerals* **12**, 553 (2022).
66. Expósito, I., Burjachs, F. & Allué, E. Filling in the gaps: The contribution of non-pollen palynomorphs to knowledge about the local environment of the Sierra de Atapuerca caves during the Pleistocene. *Quat. Int.* **433**, 224–242 (2017).
67. Cáceres, I. Tafonomía de yacimientos antrópicos en karst. Complejo Galería (Sierra de Atapuerca, burgos), Vanguard cave (Gibraltar) y Abric Romani (Capellades, Barcelona). (Universitat Rovira i Virgili, 2002).
68. Cáceres, I. et al. *El yacimiento de Galería (Sierra de Atapuerca, Burgos, España): Un enclave para la obtención de recursos cárnicos en el Pleistoceno Medio. 1ª Reunión de científicos sobre cubiles de hiena (y otros grandes carnívoros) en los yacimientos arqueológicos de la Península Ibérica* **13**, 157–165 (2010).
69. García-Medrano, P. et al. The earliest Acheulean technology at Atapuerca (Burgos, Spain): Oldest levels of the Galería site (GII Unit). *Quat. Int.* **353**, 170–194 (2014).
70. García-Medrano, P., Ollé, A., Mosquera, M., Cáceres, I. & Carbonell, E. The nature of technological changes: The Middle Pleistocene stone tool assemblages from Galería and Gran Dolina-subunit TD101 (Atapuerca, Spain). *Quat. Int.* **368**, 92–111 (2015).
71. García-Medrano, P., Cáceres, I., Ollé, A. & Carbonell, E. The occupational pattern of the Galería site (Atapuerca, Spain): A technological perspective. *Quat. Int.* **433**, 363–378 (2017).
72. Weiner, S. *Microarchaeology: Beyond the Visible Archaeological Record* (Cambridge University Press, 2010). <https://doi.org/10.1017/CBO9780511811210>.
73. Shaw, C. L. An evaluation of the infrared 630 cm⁻¹ OH libration band in bone mineral as evidence of fire in the archaeological record. *J. Archaeol. Sci. Rep.* **46**, 103655 (2022).

74. Combes, C., Rey, C. & Mounic, S. Identification and evaluation of HPO₄ ions in biomimetic poorly crystalline apatite and bone mineral. *Key Eng. Mater.* **192–195**, 143–146 (2001).
75. McAdams, C., Morley, M. W. & Roberts, R. G. The acid test: An experimental microarchaeological study of guano-driven diagenesis in tropical cave sediments. *J. Archaeol. Sci. Rep.* **37**, 102947 (2021).
76. Putnis, C. V. & Putnis, A. A mechanism of ion exchange by interface-coupled dissolution-precipitation in the presence of an aqueous fluid. *J. Cryst. Growth* **600**, 126840 (2022).
77. Kral, A. G., Geisler, T., Wiedenbeck, M., Guagliardo, P. & Tütken, T. Phosphate uptake is an essential process for rapid bone mineralization during early diagenesis – evidence from bone alteration experiments. *Geochim. Cosmochim. Acta* **375**, 173–185 (2024).
78. García-Medrano, P., Moncel, M. H., Maldonado-Garrido, E., Ollé, A. & Ashton, N. The Western European acheulean: Reading variability at a regional scale. *J. Hum. Evol.* **179**, 103357 (2023).
79. Hugué, R. et al. Le gisement de Galería (Sierra de Atapuerca, Burgos, Espagne): Un modèle archéozoologique de gestion du territoire au Pléistocène. *Anthropologie* **105**, 237–257 (2001).
80. Welker, F. et al. The dental proteome of Homo antecessor. *Nature* **580**, 235–238 (2020).
81. Kontopoulos, I. et al. Screening archaeological bone for palaeogenetic and palaeoproteomic studies. *PLoS ONE* **15**, 1–17 (2020).
82. Henty, J. Ancient protein analysis in archaeology. *Sci. Adv.* **7**, 1–11 (2021).
83. Warinner, C., Korzow Richter, K. & Collins, M. J. Paleoproteomics. *Chem. Rev.* **122**, 13401–13446 (2022).
84. Scaggion, C. et al. A fresh perspective on infrared spectroscopy as a prescreening method for molecular and stable isotopes analyses on ancient human bones. *Sci. Rep.* **14**, 1–14 (2024).
85. Pérez-González, A. et al. Geología y estratigrafía del relleno de Galería de la Sierra de Atapuerca (Burgos). In *Atapuerca: Ocupaciones Humanas y Paleoecología del Yacimiento de Galería*, vol. 7 (eds Carbonell, E. et al.) 31–42 (Memorias de la Junta de Castilla y León, 1999).
86. Rey, C., Collins, B., Goehl, T., Dickson, I. R. & Glimcher, M. J. The carbonate environment in bone mineral: A resolution-enhanced fourier transform infrared spectroscopy study. *Calcif. Tissue Int.* **45**, 157–164 (1989).
87. Kontopoulos, I., Penkman, K., Liritzis, I. & Collins, M. J. Bone diagenesis in a Mycenaean secondary burial (Kastrouli, Greece). *Archaeol. Anthropol. Sci.* **11**, 5213–5230 (2019).
88. Fleet, M. E. Infrared spectra of carbonate apatites: ν₂-Region bands. *Biomaterials* **30**, 1473–1481 (2009).
89. Ruiz-Agudo, E., Putnis, C. V. & Putnis, A. Coupled dissolution and precipitation at mineral-fluid interfaces. *Chem. Geol.* **383**, 132–146 (2014).
90. France, C. A. M., Sugiyama, N. & Aguayo, E. Establishing a preservation index for bone, dentin, and enamel bioapatite mineral using ATR-FTIR. *J. Archaeol. Sci. Rep.* **33**, 102551 (2020).

Acknowledgements

The authors are thankful to the SRCiT of the URV and the Department of Mineralogy and Petrology of the UGR for their technical support and advice. We would also like to thank the Atapuerca team for their support, collaboration, and availability.

Bone samples collection

All bone samples included in the study were obtained with the appropriate permits. All samples were collected in accordance with current ethical and legal regulations.

Author contributions

HDV wrote the original draft and performed the investigation and data analysis. ARN and AB reviewed and edited data and text. PGM and IC reviewed and supervised text and lead the fieldwork. HDV and IC designed research and conceptualization. All author contributed to the main text and supplementary texts.

Funding

This research is part of the I+D+i project PID2021-122355NB-C32, funded by MCIN/AEI/<https://doi.org/10.13039/501100011033> and by ERDF A way of making Europe, and partly supported by: MICINN- PID2020-116660GB-I00, 2021 SGR 01238 (AGAUR), 2023 PFR-URV-01238 (URV), LATEUROPE ERC project (Grant agreement ID: 101052653) at CNRS. HDV research is funded by the Spanish Ministry of Science and Innovation through FPI Fellowship (PRE2021-100275), and PG-M was supported by a Beatriz de Pinós MSCA-COFUND (AGAUR) program at IPHES-CERCA. IPHES-CERCA has received financial support from the Spanish Ministry of Science and Innovation through the “María de Maeztu” program for Units of Excellence (CEX2019-000945-M). AM is funded by a postdoctoral contract grant from the Fyssen Foundation.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-88968-4>.

Correspondence and requests for materials should be addressed to H.V.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025