nature portfolio

https://doi.org/10.1038/s41586-026-10976-9

Supplementary information

Ancient proteins identify various Denisovan remains from Southwest China

In the format provided by the authors and unedited

Supplementary Information

Supplementary Text 1. Uranium-series (U-series) dating results of 3 hominin bone remains

To constrain the minimum ages on the three Bianfu Cave hominin bone samples, we performed U-series dating on both the bone samples and their attached carbonates using LA-MC-ICPMS (Supplementary Table 1). The carbonate samples from the parietal BFD769 and radius BFD771 were analyzed with an effort to calculate U-Th isochron ages. However, the measured $^{238}U/^{232}Th$ and $^{230}Th/^{232}Th$ activity ratios were clustered in narrow ranges ($^{238}U/^{232}Th$ ratios of ~1.06 – 1.45 for BFD769 and ~1.68 – 2.76 for BFD771, and $^{230}Th/^{232}Th$ ratios of ~1.12 – 1.38 for BFD769 and ~1.78 – 3.07 for BFD771), which prevented establishing precise $^{230}Th/U$ isochrons. For the hominin fossil samples, the parietal BFD767 yielded a relatively young age estimate of 23.9-31.3 ka. An older age of ~65.8 ka was detected in this sample, but it was strongly contaminated by sedimentary Th, as indicated by low $^{238}U/^{232}Th$ (~36) and $^{230}Th/^{232}Th$ (~19) activity ratios. The parietal BFD769 and the radius BFD771 showed apparent U-series ages of 133.9-148.9 ka and 92.8-109.6 ka, respectively. Both samples displayed uniform distributions of the $^{234}U/^{238}U$ and $^{230}Th/^{238}U$ activity ratios, as well as of the initial $^{234}U/^{238}U$ activity ratios. Furthermore, their U-series ages were consistent with their stratigraphic provenance (the radius BFD771 from Layer 7 and the parietal BFD769 from Layer 9), which indicates that the two samples experienced early U-uptake histories. Therefore, the averaged U-series ages of 143.4 ± 6.6 ka and 102.0 ± 5.7 ka are the reasonable minimum age estimates for the parietal BFD769 and the radius BFD771, respectively. Based on the U-uptake and decay pattern, the U-Th age can be influenced by delayed U-uptake or a secondary overprint, thus yielding a relatively younger age. Because the parietal BFD767 and radius BFD771 were from the same Layer 7, the young age of the parietal BFD767 was considered an underestimate. Here, we used the age of the radius BFD771 to represent the age of both hominin fossils.

Supplementary Text 2. Endogenous proteome for 5 hominin remains

To determine the endogenous proteome of the five hominins, we calculated the extent of deamidation of glutamine (Q) and asparagine (N) for each protein. Higher deamination levels of glutamine or asparagine indicate more extensive protein degradation and can serve as a criterion for validating the recovered ancient proteome[2,41,62]. Because Q is much more stable against deamidation than N, it is a more suitable indicator of archaeological degradation[89,90]. If at least two Q/N-containing peptides with valid intensity were identified in a single protein for deamidation calculation, we considered the values reliable.

Considering the high similarity between different collagen chains, some low-quality spectra may be misassigned to low-abundance collagens. To address this issue, we added additional "permutation and unipept confirmation" steps to validate their identifications (Supplementary Fig. 18). Except for the five most common collagens in bone or dentine (COL1A1, COL1A2, COL2A1, COL5A1, COL5A2), all the other collagens were considered low-abundant. For each low-abundance collagen, all its PSMs must be validated through the following steps.

(1) For each PSM, all the possible permutations (potential peptide candidates) that are isobaric to the annotated precursor are generated. These isobaric changes include positional isomer (when accounting for poorly-fragmented ion series, if the b- or y-ions do not split two amino acid residues, then those amino acids could be in either order) and single residue isobaric switch (if the peptide contains one of the paired isobaric residues, they could also be another). For example, if the original peptide contains isoleucine (I), it could be replaced with leucine (L). Besides the I/L isobaric pairs, we also considered another three common pairs, i.e. deamidated Q [Q(+0.98)] vs E, deamidated N [N(+0.98)] vs D, and py-Glu from E[E(-18.01)] vs py-Glu from Q[Q(-17.03)].

(2) All permutations are subjected to Unipept to retrieve their possible matches in the UniProt database. Note that due to the low coverage of y- and b-ions or the presence

of many isobaric amino acids in some PSMs, the number of permutations can be very large. To better balance computational resources, PSMs with more than 10,000 permutations are considered too ambiguous and excluded, while those with fewer than 10,000 permutations are retained for further validation. For each permutation, pep2prot (https://api.unipept.ugent.be/api/v2/pept2prot) is used to obtain all the UniProt entries in the entire UniProt database (including SwissProt and TrEMBL), which could match the given sequence.

(3) For each PSM, the matching results of all the permutations are summarized. If one of the UniProt matches belongs to another protein (with a different gene name from the low-abundance collagen, which may indicate incorrect identification) or to a domain other than Eukaryote, such as Bacteria (possible contamination sources), this PSM is considered a possible misassignment and is thus excluded from further calculations. PSMs that match a unique gene name and species limited to Eukaryotes are considered reliable matches to this low-abundance collagen.

After these “permutation and unipept confirmation” processes, the reliable PSMs for each low-abundance collagen were retained to calculate the numbers of Q/N-containing peptides with valid intensities and corresponding deamidation values. Combined with the results of other proteins, the endogenous proteins were determined based on the reliable deamidation values supported by at least two Q/N-containing peptides.

In the parietal BFD767, we identified 13 proteins with both reliable Q and N deamidation values, and 5 with only reliable Q deamidation values. The 13 proteins with both reliable Q and N deamidation values were subjected to cluster analysis to differentiate endogenous from contaminating proteins (Extended Data Fig. 3). Eight proteins, including trypsin and elastase introduced during the experimental procedures, exhibited low Q and N deamination values (average Q deamidation: 0.47 ± 1.28%; average N deamidation: 2.37 ± 4.97%). These proteins were deemed contaminants and excluded (group 1 in Extended Data Fig. 3), consistent with previously reported common contaminants[57,62]. The remaining 5 proteins exhibited elevated deamidation values (group 2 in Extended Data Fig. 3), indicating their authenticity as ancient

proteins. In addition, one protein exhibited a reliably elevated Q deamidation of 100% and was incorporated into the endogenous proteome (Supplementary Table 2). In total, 6 endogenous proteins were detected in the BFD767 specimen, all of which were collagens. In addition to the five most common collagens in bone or dentine, one low-abundance collagen (COL23A1) was also considered endogenous based on the elevated deamidation values, and this protein has been reported in other archaic hominins as well[3,5].

In the parietal BFD769, we obtained 4 proteins with both reliable Q and N deamidation values, and five proteins with only reliable Q deamidation values. The 4 proteins with both reliable Q and N deamidation values were subjected to cluster analysis to differentiate endogenous from contaminating proteins (Supplementary Fig. 1). Trypsin, introduced during the experimental procedures, exhibited low Q and N deamination (Q deamidation: 0.78%; average N deamidation: 0.33%), and was excluded as a contaminant (group 1 in Supplementary Fig. 1). The remaining 3 proteins demonstrated elevated deamidation (group 2 in Supplementary Fig. 1), indicating their authenticity as ancient proteins. Three additional proteins displayed reliably elevated Q deamidation, ranging from 83.33% to 100%, and were also incorporated into the endogenous proteome (Supplementary Table 2). In total, 6 endogenous proteins were detected in the BFD769 specimen, all of which were collagens. In addition to the five most common collagens in bone or dentine, one low-abundance collagen (COL15A1) was also considered endogenous based on the elevated deamidation values, and this protein has been reported in other archaic hominins as well[5].

In the radius BFD771, we obtained seven proteins with both reliable Q and N deamidation values, and no protein possessed only a reliable Q deamidation value. The seven proteins with both reliable Q and N deamidation values were subjected to cluster analysis to differentiate endogenous from contaminating proteins (Supplementary Fig. 2). Two proteins, including trypsin and non-POU domain-containing octamer-binding protein (NONO, a DNA- and RNA-binding protein), exhibited low N deamination values and exceptionally low values for the

slower deamidation of Q (average Q deamidation: 0%; average N deamidation: 16.74 ± 23.32%). These proteins were excluded as contaminants (group 1 in Supplementary Fig. 2). The remaining five proteins demonstrated elevated deamidation values (group 2 in Supplementary Fig. 2), indicating their authenticity as ancient proteins. In total, five endogenous proteins were detected in the BFD771 specimen, all of which were among the most common collagens in bone or dentine (Supplementary Table 2).

In the dentine extract of YHB3518, we identified 54 proteins with both reliable Q and N deamidation values, and 20 with only reliable Q deamidation values. The 54 proteins with both reliable Q and N deamidation values were subjected to cluster analysis to differentiate endogenous from contaminating proteins (Supplementary Fig. 3). 49 proteins, including trypsin introduced during the experimental procedures and keratins, exhibited low Q and N deamination values (average Q deamidation: 0.89 ± 3.25%; average N deamidation: 4.36 ± 10.55%). These proteins were excluded as contaminants (group 1 in Supplementary Fig. 3), consistent with previously reported common contaminants[57,62]. The remaining 5 proteins exhibited elevated deamidation values (group 2 in Supplementary Fig. 3), indicating their authenticity as ancient proteins. In addition, one protein displaying a reliably elevated Q deamidation rate of 100% was incorporated into the endogenous proteome (Supplementary Table 2). In total, 6 endogenous proteins were detected in the dentine extract of YHB3518, all of which were collagens. In addition to the five most common collagens in bone or dentine, one low-abundance collagen (COL11A2) was also considered endogenous based on the elevated deamidation values, and this protein has been reported in other archaic hominins as well[3-5,62]. In the enamel extract of YHB3518, we identified 6 proteins with both reliable Q and N deamidation values, and 3 with either reliable Q or N deamidation values. All 9 proteins exhibited elevated deamidation levels and were incorporated into the endogenous proteome (Supplementary Table 2). The deamidation rates of asparagine (N) and glutamine (Q) for the whole enamel proteome were also calculated, as well as peptide length distribution. The endogeneity of this enamel proteome was confirmed by high deamidation values and relatively

short peptide length caused by degradation (Supplementary Fig. 19). As two proteins (COL1A1 and COL1A2) were identified as endogenous proteins in both dentine and enamel of the tooth YHB3518, a total of 13 endogenous proteins were detected in this sample.

In the dentine extract of YHB3075, we identified 42 proteins with both reliable Q and N deamidation values, and 17 with only reliable Q deamidation values. The 42 proteins with both reliable Q and N deamidation values were subjected to cluster analysis to differentiate endogenous from contaminating proteins (Supplementary Fig. 4). 37 proteins, including trypsin introduced during the experimental procedures and keratins, exhibited low Q and N deamination values (average Q deamidation: 0.77 ± 2.64%; average N deamidation: 6.14 ± 11.04%). These proteins were excluded as contaminants (group 1 in Supplementary Fig. 4), consistent with previously reported common contaminants[57,62]. The remaining 5 proteins exhibited elevated deamidation values (group 2 in Supplementary Fig. 4), indicating their authenticity as ancient proteins. The deamidation values for the 17 proteins with only reliable Q deamidation ranged from 0% to 0.79%, and were therefore excluded from the endogenous proteome (Supplementary Table 2). In total, 5 endogenous proteins were detected in the dentine extract of YHB3075, all of which were among the most common collagens in bone or dentine. In the enamel extract of YHB3075, we identified 7 proteins with both reliable Q and N deamidation values, and 4 with either reliable Q or N deamidation values. All these 11 proteins exhibited elevated deamidation values and were incorporated into the endogenous proteome (Supplementary Table 2). The deamidation rates of asparagine (N) and glutamine (Q) for the whole enamel proteome were also calculated, as well as peptide length distribution. The endogeneity of this enamel proteome was confirmed by high deamidation values and relatively short peptide length caused by degradation (Supplementary Fig. 19). Thus, a total of 16 endogenous proteins were detected in the tooth YHB3075.

Supplementary Text 3. Evaluation of the informative variants within the *Homo* genus

The parietal BFD767

After screening for branch-specific single amino polymorphisms (SAPs) in the endogenous proteome and filtering with initial thresholds (PSM count ≥ 2, PSM ratio ≥ 10%, and intensity ratio ≥ 10%), 6 SAPs or amino acid positions were retained for further evaluation using multiple steps. First, the supporting peptides for each variant were BLASTed against the NCBInr database. Based on the BLAST results, we filtered out peptides with possible contaminant origins or that did not map uniquely, retaining 7 variants across 5 positions (2 possible heterozygous positions) that were covered by at least two peptides for further evaluation (Supplementary Data 4). The Xiahe 1-specific variant COL2A1 E583G[2] was excluded in this step because it lacked sufficient peptide support. The fragmentary y- and b-ions surrounding the SAP were checked for each PSM of the remaining peptides. Among the 7 variants, all contained at least one peptide with "MS2 support", but only 5 variants contained at least one peptide with "strict MS2 support". 3 variants had N-containing peptides for the calculation of N deamidation values, among which 2 exhibited elevated deamidation values. The Q deamidation values could be calculated for 4 variants, and all had elevated deamidation values. 2 variants were detected in both MaxQuant and pFind, with at least one peptide retained after BLAST filtering. The codon degeneracy of the 5 derived variants was checked, and all these variants could derive from a single base mutation in the DNA sequence.

Because all the variants originated from collagen, we also examined collagen features. Firstly, we assessed whether the SAP is in the triple helix and consists of complete GXY repeats. If so, we further evaluated whether some proline residues at the Y position are hydroxylated. Among the 7 variants, 6 were in the triple-helix region containing GXY repeats, and some of the prolines in the Y position were hydroxylated, consistent with collagen features. However, one variant, COL2A1 G1170S, was in the triple helix region of collagen type II alpha 1 chain, but the GXY repeat was hindered

by the mutation from G to S. One variant, COL1A2 R996K, was in the triple helix region of collagen type I alpha 2, and the sequences were complete GXY repeats. But all the related prolines were in the X position, and no Y position hydroxyproline was present.

In all, 2 variants were confidently identified based on the following criteria: (1) Based on the PEAKS result, at least two peptides were retained after a BLAST search; at least one peptide had "strict MS2 support" PSMs; the variant presented elevated N/Q deamidation rates (if calculated). (2) Based on MaxQuant and pFind results, at least one peptide was retained after a BLAST search. (3) The amino acid substitution can be caused by a single base mutation, and the peptide pattern had common collagen features if the SAP came from collagens. Among the confidently identified variants, one variant is phylogenetically informative, the diagnostic Denisovan marker (COL1A2 R996K, rs1792270869), which is present in all currently identified Denisovan individuals, including Harbin[5], Penghu[4], Xiahe 1[2], Xiahe 2[29], Denisova 25[30], and Denisova 3[31], suggesting that the parietal BFD767 also belongs to Denisovans (Table 1).

The parietal BFD769

After screening for branch-specific SAPs in the endogenous proteome, and conducting the initial threshold and BLAST filtering, 3 variants in 2 positions (with one possible heterozygous position) were retained with at least 2 peptides (Supplementary Data 4). All 3 variants contained at least one peptide with "MS2 support", but only 2 variants contained at least one peptide with "strict MS2 support". Q-deamidation values were calculated for the two variants, and both exhibited elevated deamidation. Two variants were also detected in both MaxQuant and pFind, with at least one peptide retained after the BLAST filtering. The codon degeneracy of the two derived variants was assessed, and both mutations could arise from a single base mutation in the DNA sequence. All 3 variants originated from collagens and were found in the triple-helix region containing the GXY repeats. Only 2 contained a Y position hydroxyproline, while the third variant, COL1A2 R996K, did not contain a proline in the Y position. Based on all the evaluations, a single variant was

confidently identified, i.e. the diagnostic Denisovan marker (COL1A2 R996K) (Table 1).

The radius BFD771

After screening for branch-specific SAPs in the endogenous proteome and conducting the initial threshold and BLAST filtering, 3 variants in 2 positions (one possibly heterozygous) were retained with at least two peptides (Supplementary Data 4). All 3 variants contained at least one peptide with "strict MS2 support". The N (one variant), and Q (two variants) deamidation values were calculated, and all the variants exhibited elevated deamidation. All 3 variants were detected in both MaxQuant and pFind, with at least one peptide retained after BLAST filtering. The codon degeneracy of the two derived variants was checked, and both mutations could occur with a single base mutation in the DNA sequence. All 3 variants came from collagens and were located in the triple helix region with the GXY repeats. Only 2 contained a Y position hydroxyproline, while the third, COL1A2 R996K, did not contain a proline in the Y position. Based on all evaluations, 3 variants were confidently identified, among which one is phylogenetically informative, i.e. the diagnostic Denisovan marker (COL1A2 R996K) (Table 1).

The tooth YHB3518

After screening for branch-specific SAPs in the endogenous proteome and applying filtering steps, 7 variants at 6 positions (one possibly heterozygous) were retained with at least 2 peptides (Supplementary Data 4). All 7 variants contained at least one peptide with "MS2 support", but only 6 contained at least one peptide with "strict MS2 support". The N (2 variants) and Q (5 variants) deamidation values were calculated, and all variants exhibited elevated deamidation. 6 variants were also detected in both MaxQuant and pFind, with at least one peptide retained after BLAST filtering. The codon degeneracy of the 4 derived variants was checked, and all the mutations could occur with a single base mutation in the DNA sequence. All 3 collagen variants were located in the triple helix region with the GXY repeats, and 2 variants contained a hydroxyproline at the Y position, whereas the third variant, COL1A2 R996K, did not contain a proline at the Y position. Based on all evaluations,

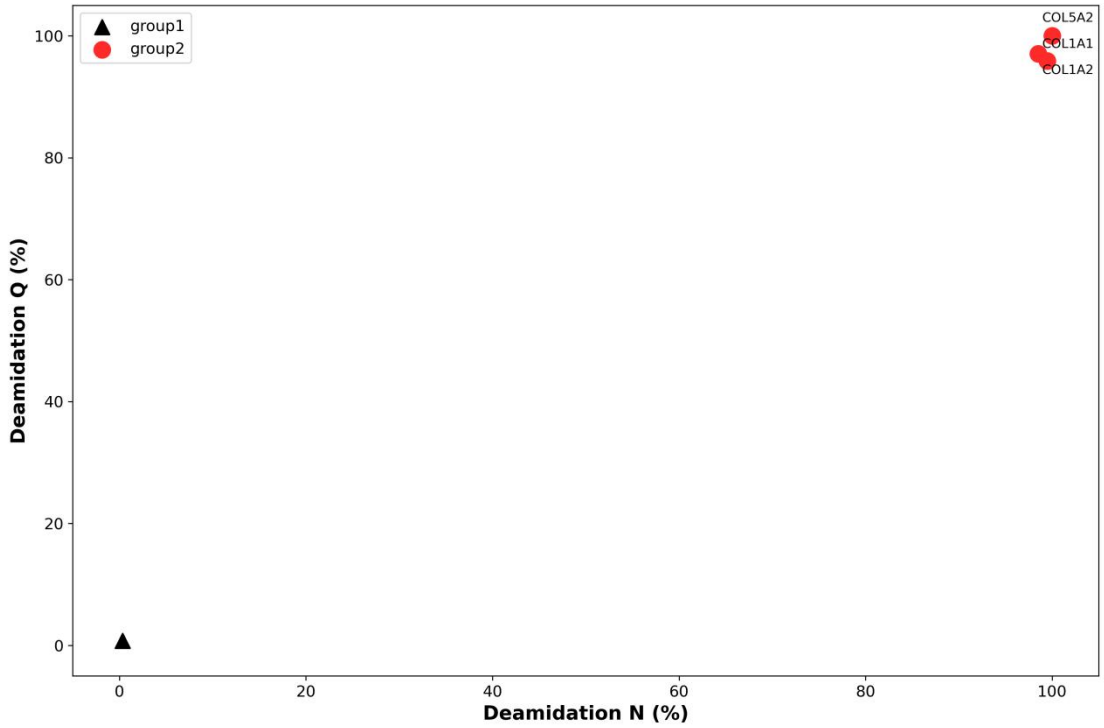
6 variants were confidently identified, among which one is sex determinative (AMELY 59M, assigning this individual as male) and 4 are phylogenetically informative (Table 1). Besides the diagnostic Denisovan marker (COL1A2 R996K), another informative variant AMBN M273V (rs564905233) is also present in East Asian *Homo erectus*[32], Harbin[32], Penghu[4], Denisova 3[31], Denisova 25[30], and a low frequency of modern humans with high Denisovan introgression[33]. Another variant (AMELY 179L) is shared in Penghu[4], Harbin[32], Denisova 25[30], Neanderthals[34-39], Ust'-Ishim[40] and all the modern humans in the gnomAD database, which is different from the AMELY 179M variant present in *Homo antecessor*[41] and great apes. Thus, the identification of AMELY 179L attributes this individual to the *Homo sapiens*, Neanderthal, and Denisovan clade. The last variant, AMBN 253A is present in all the comparative samples except East Asian *H. erectus*[32]. Therefore, based on all the confidently identified variants, the tooth YHB3518 can be attributed to Denisovans.

The tooth YHB3075

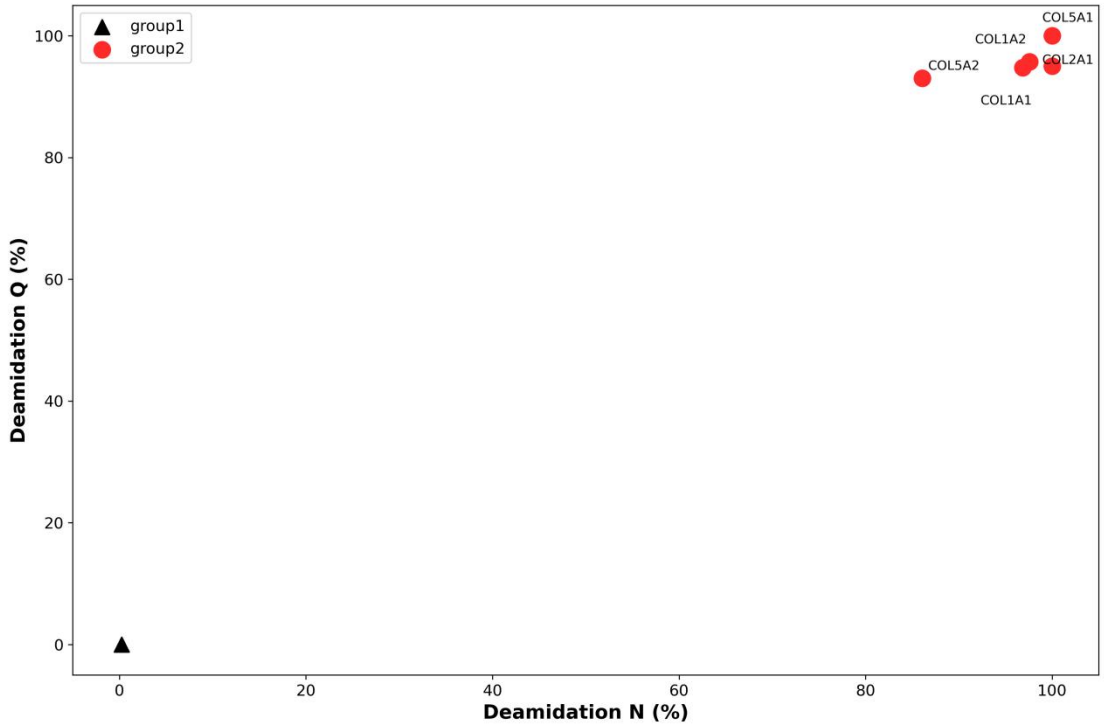
After screening for branch-specific SAPs in the endogenous proteome and conducting filtering steps, 8 variants in 7 positions (one possibly heterozygous) were retained with at least two peptides (Supplementary Data 4). Among the 8 variants, 7 contained at least one peptide with "MS2 support", and 6 variants contained at least one peptide with "strict MS2 support". The N (3 variants), and Q (6 variants) deamidation values were calculated, and all the variants exhibited elevated deamidation. 6 variants were also detected in both MaxQuant and pFind, with at least one peptide retained after the BLAST filtering step. The codon degeneracy of the 5 derived variants was checked, and all the mutations could be caused by a single base substitution in the DNA sequence. Of the 4 variants from collagens, 3 were within the GXY repeats of the triple helix region, whereas the GXY repeats of the fourth variant, COL2A1 G1170S, were hindered by the G to S substitution. Among the 3 variants with GXY repeats, 2 contained a hydroxyproline at the Y position, while the third variant, COL1A2 R996K, did not contain a proline at the Y position. Based on all evaluations, 6 variants were confidently identified, among which one is sex determinative (AMELY 59M, indicating this individual was male) and 4 are phylogenetically informative (Table 1).

The four informative variants were the same with those found in the tooth YHB3518, i.e., COL1A2 R996K (the diagnostic Denisovan marker), AMBN M273V (Denisovan related variant), AMELY 179L (attributed to the *H. sapiens*, Neanderthal, and Denisovan clade), and AMBN 253A (different from East Asian *H. erectus).* Combining all the confidently identified variants, the tooth YHB3075 was attributed to Denisovans.

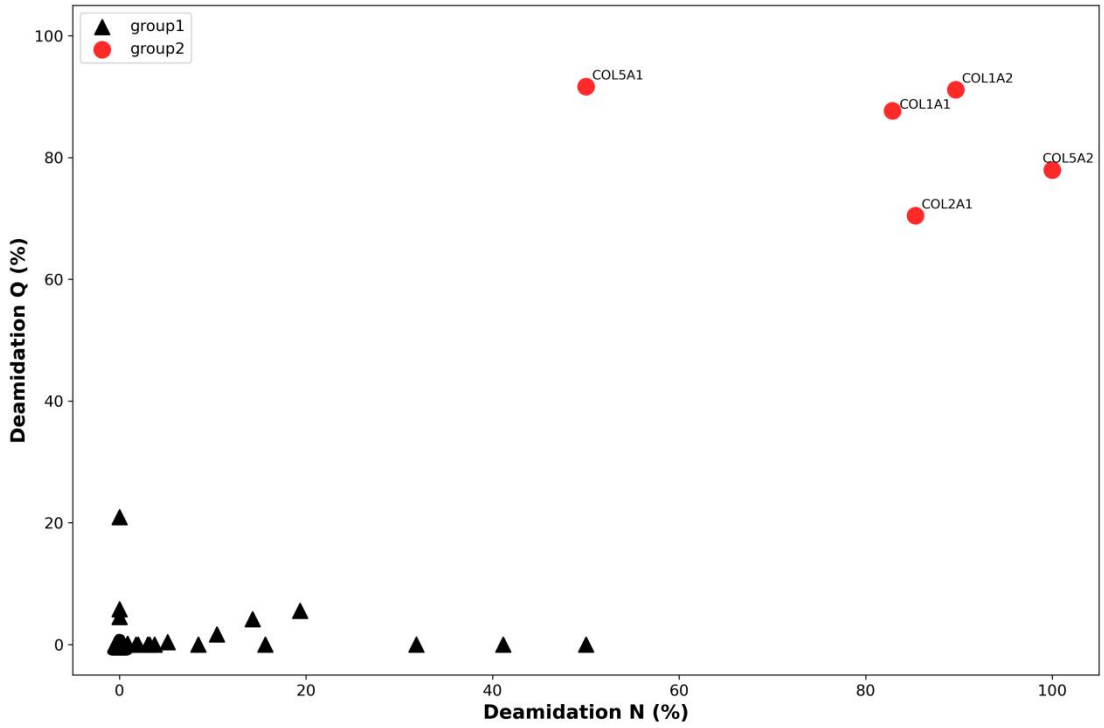
Supplementary Figures



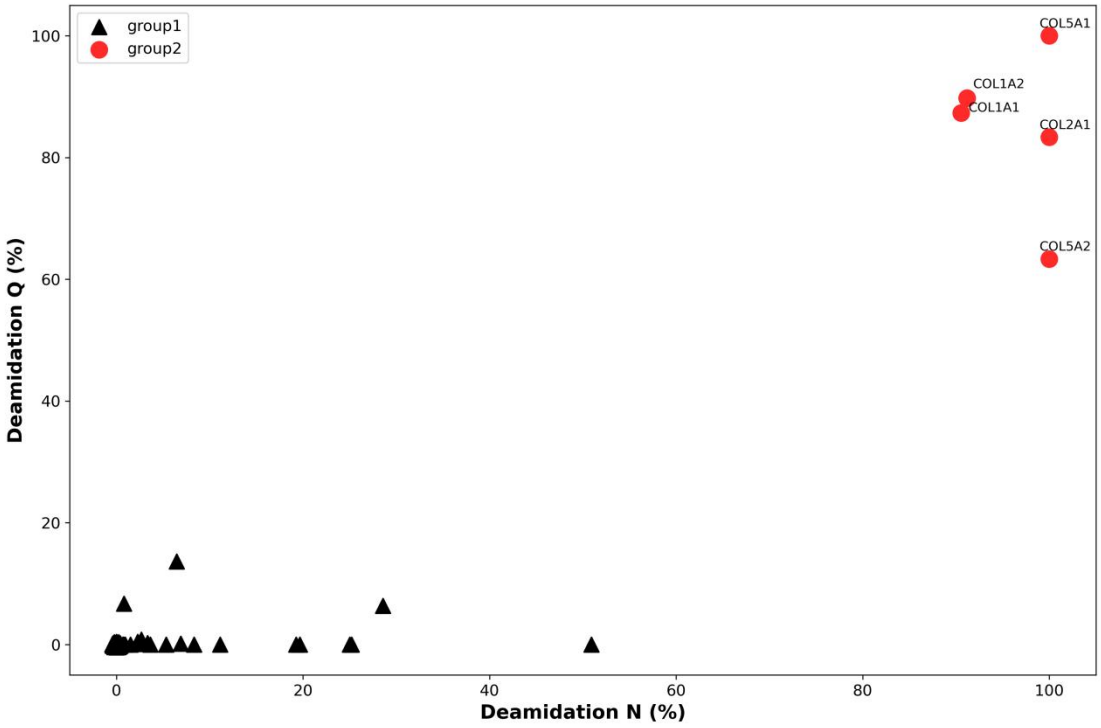
Supplementary Fig. 1. N/Q deamidation values of identified proteins with at least two N/Q-containing peptides in the parietal BFD769. Based on K-means clustering, a clear separation is observed between contaminating proteins (group 1: black) and endogenous proteins (group 2: red).



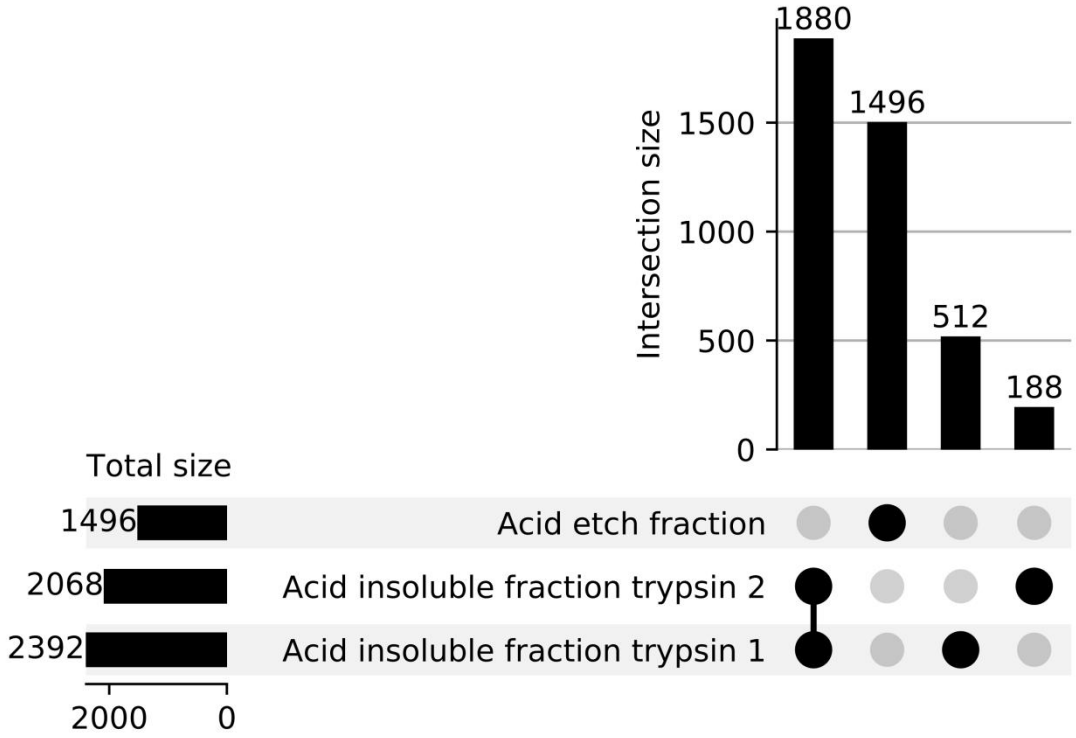
Supplementary Fig. 2. N/Q deamidation values of identified proteins with at least two N/Q-containing peptides in the radius BFD771. Based on K-means clustering, a clear separation is observed between contaminating proteins (group 1: black) and endogenous proteins (group 2: red).



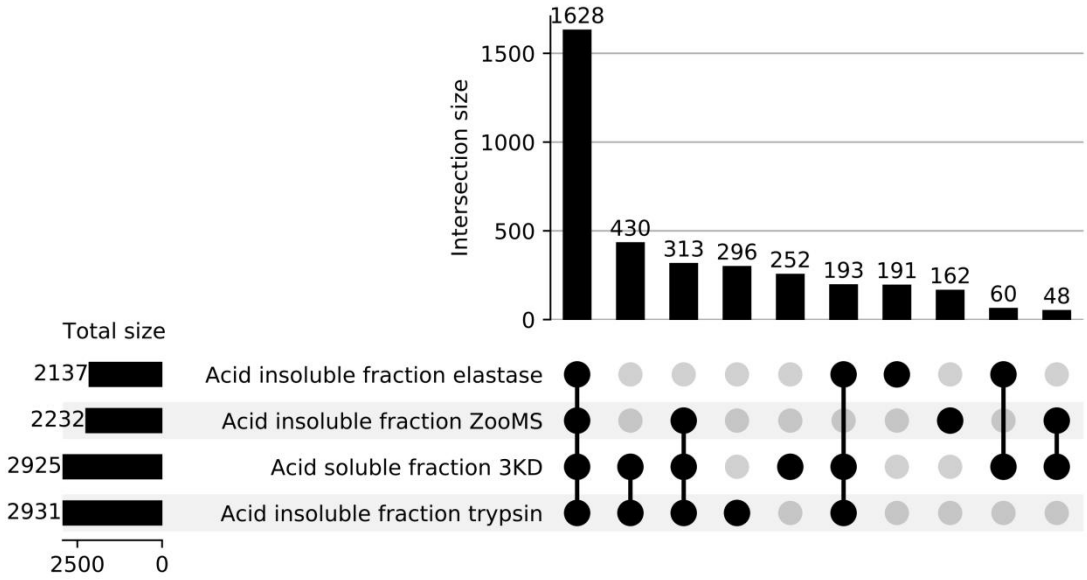
Supplementary Fig. 3. N/Q deamidation values of identified proteins with at least two N/Q-containing peptides in the YHB3518 dentine. Based on K-means clustering, a clear separation is observed between contaminating proteins (group 1: black) and endogenous proteins (group 2: red).



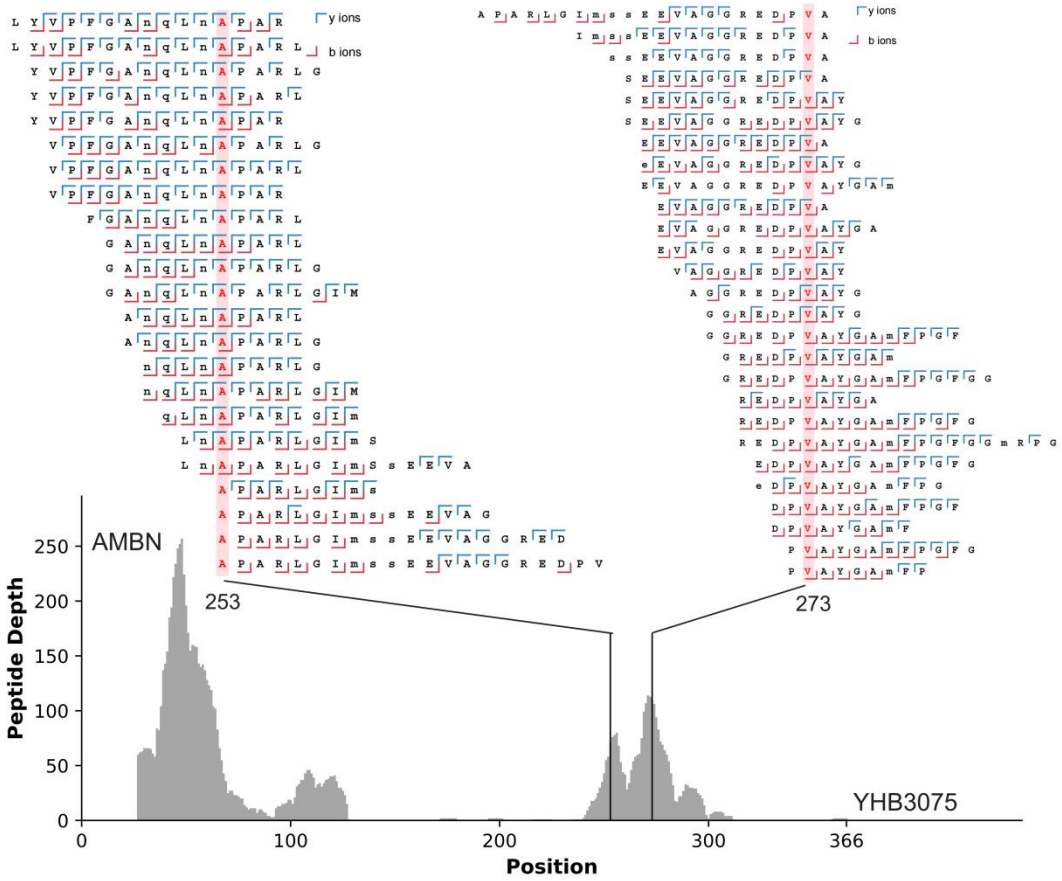
Supplementary Fig. 4. N/Q deamidation values of identified proteins with at least two N/Q-containing peptides in the YHB3075 dentine. Based on K-means clustering, a clear separation is observed between contaminating proteins (group 1: black) and endogenous proteins (group 2: red).



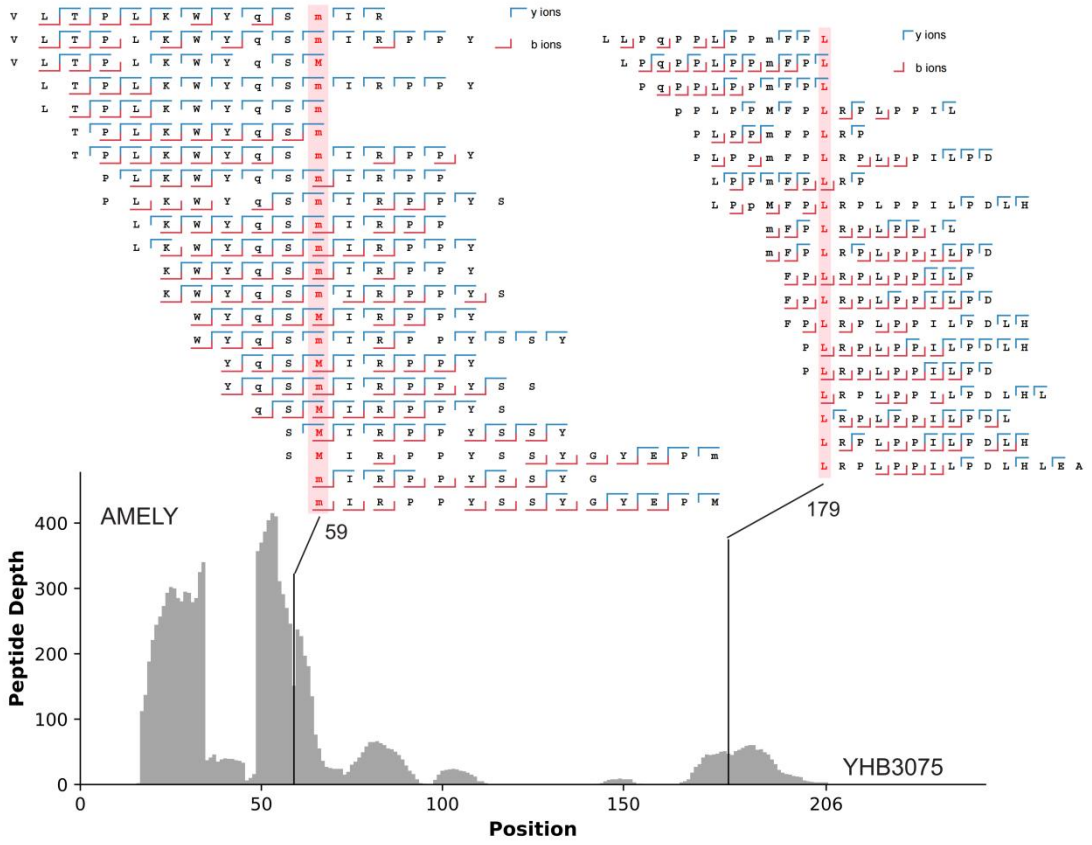
Supplementary Fig. 5. Upset plot of the identified amino acid residues from different extraction fractions of the tooth YHB3075. Note that the data from "Acid insoluble fraction trypsin 1" actually contains the peptides from two fractions, as "Acid soluble fraction 3KD" and "Acid insoluble fraction trypsin 1" of this tooth were combined for further LC-MS/MS analysis (see Materials and Methods).



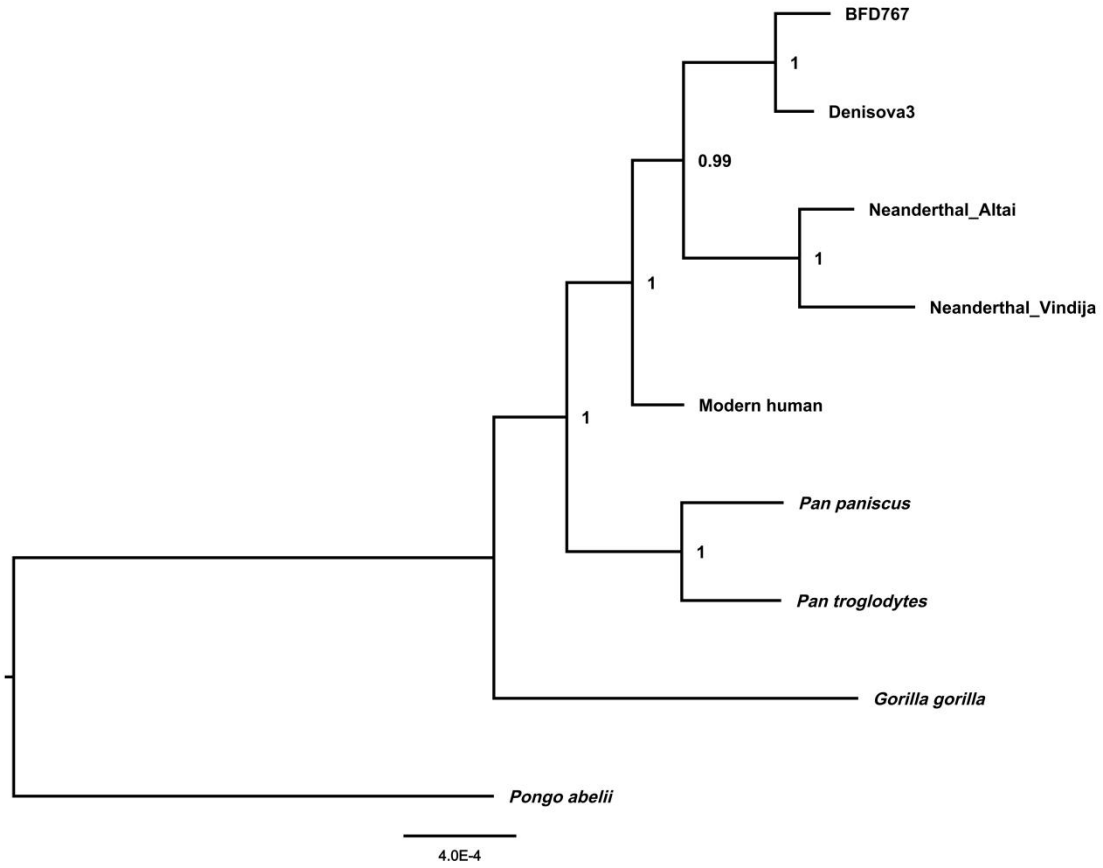
Supplementary Fig. 6. Upset plot of the identified amino acid residues from different extraction fractions of the parietal BFD767.



Supplementary Fig. 7. Details of the variants in AMBN from the tooth YHB3075. The amino acid coverage depth of AMBN, the sequence alignments covering the phylogenetically informative variants AMBN 253A and AMBN 273V (highlighted in red) are shown. Modified amino acids are indicated as lowercase letters: Deamidation (n), Deamidation (q), Pyro-glu from E (e), Oxidation (m), and Phosphorylation (s).



Supplementary Fig. 8. Details of the variants in AMELY from the tooth YHB3075. The amino acid coverage depth of AMELY, the sequence alignments covering the sex determinative variant AMELY 59M, and the phylogenetically informative variant AMELY 179L (highlighted in red) are shown. Modified amino acids are indicated as lowercase letters: Deamidation (q), Hydroxylation (p), and Oxidation (m).

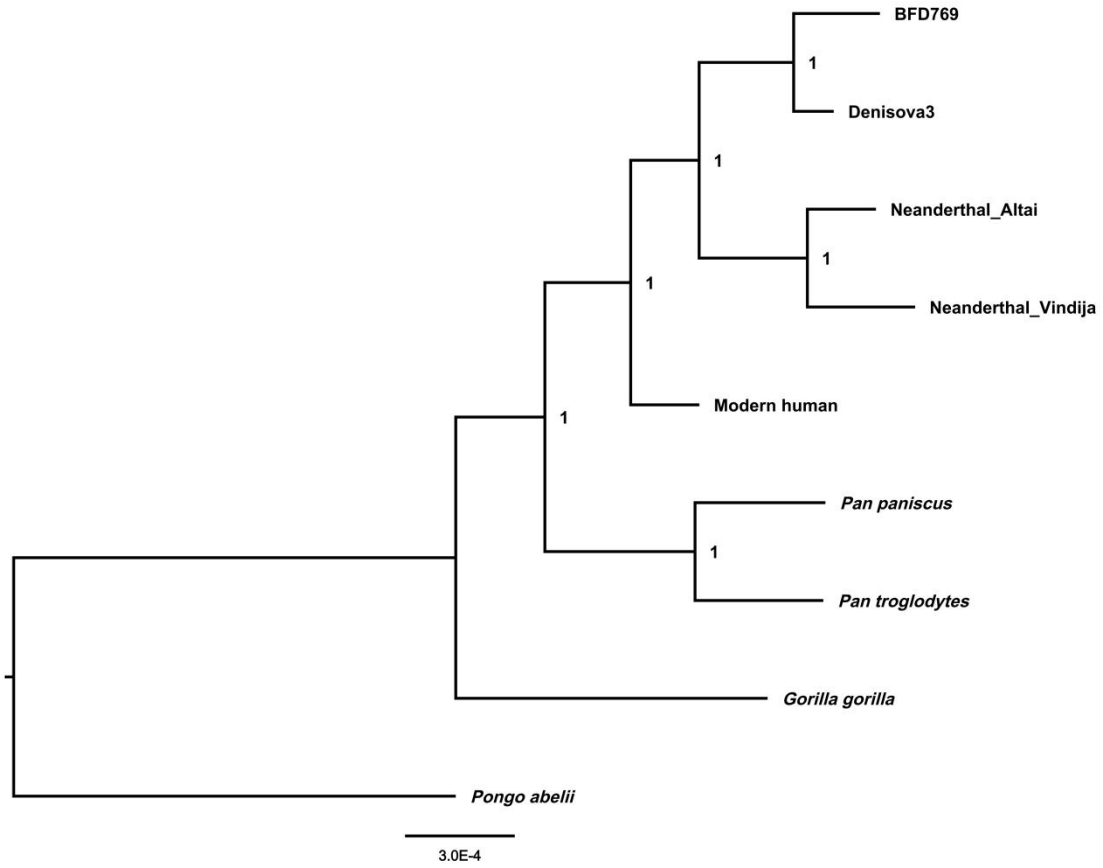


353

354 **Supplementary Fig. 9. Bayesian phylogenetic tree to determine the placement of**

355 **the parietal BFD767 using *Pongo abelii* as an outgroup.**

356

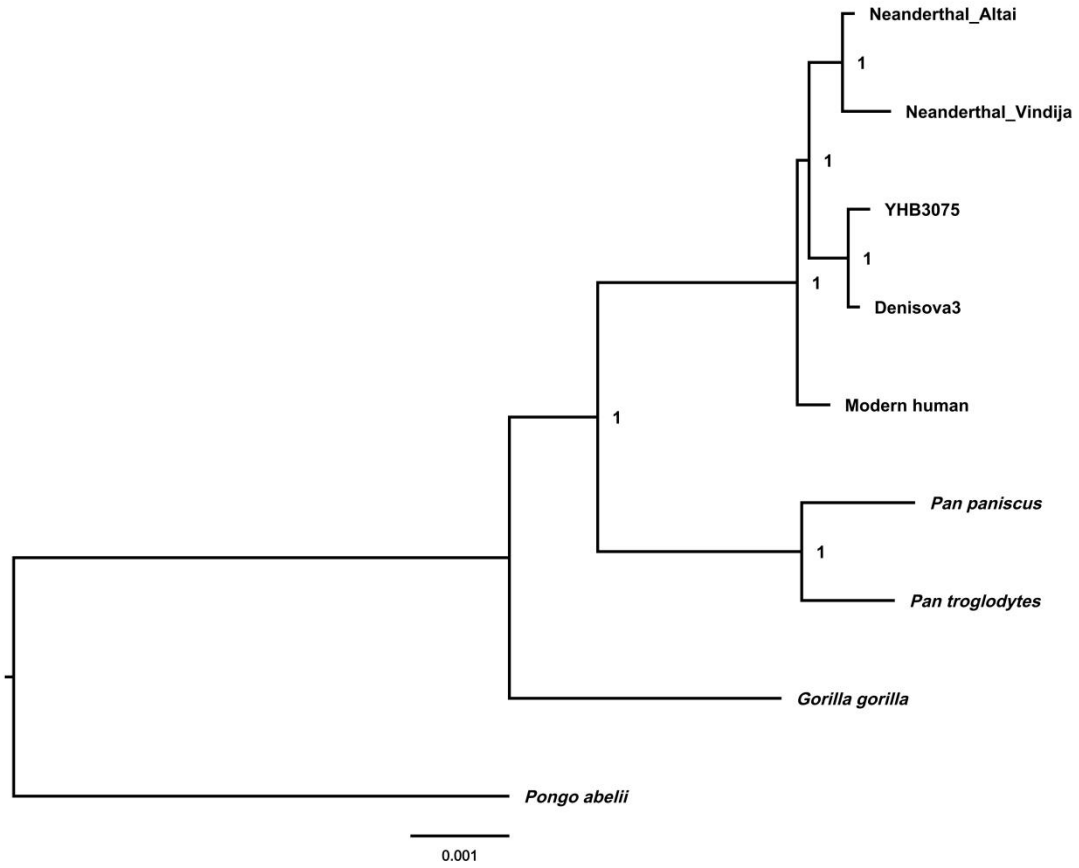


357

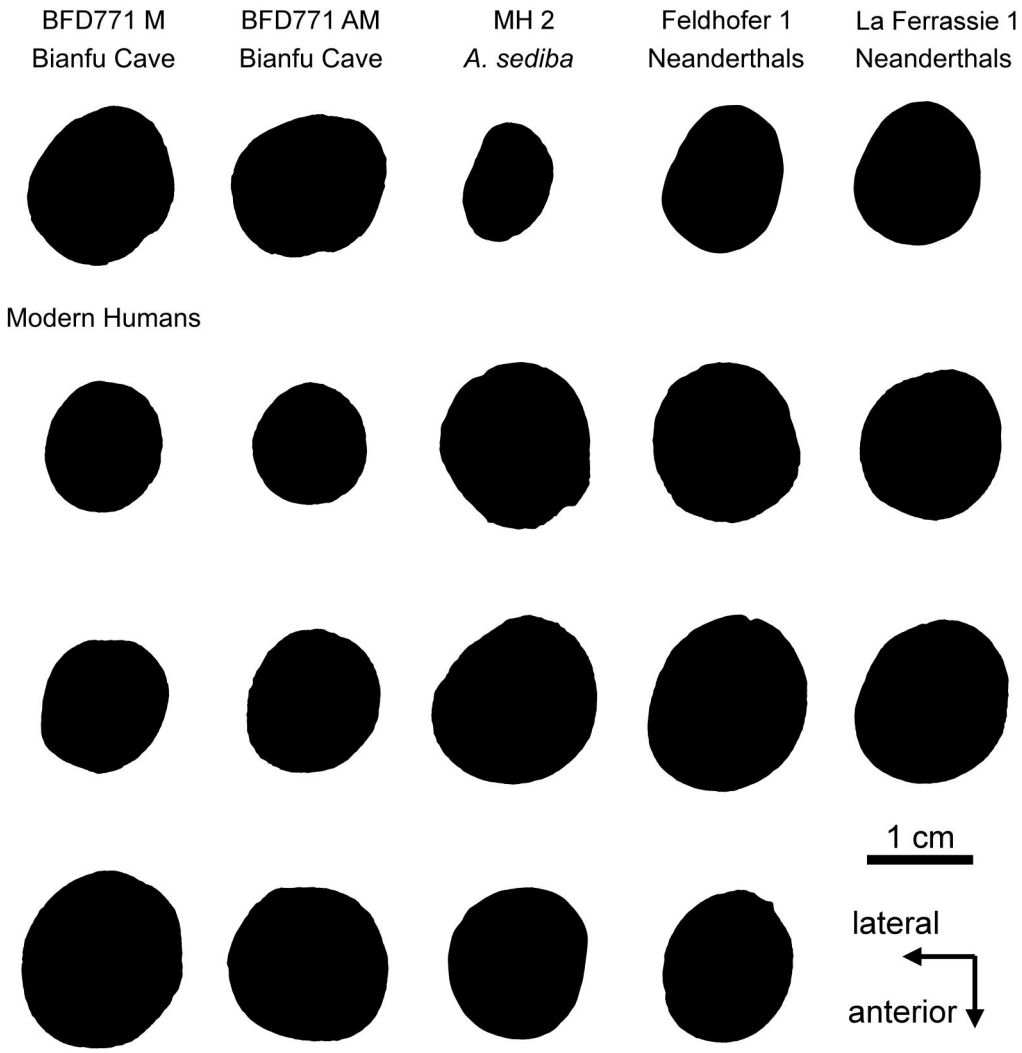
358 **Supplementary Fig. 10. Bayesian phylogenetic tree to determine the placement of**

359 **the parietal BFD769 using *P. abelii* as an outgroup.**

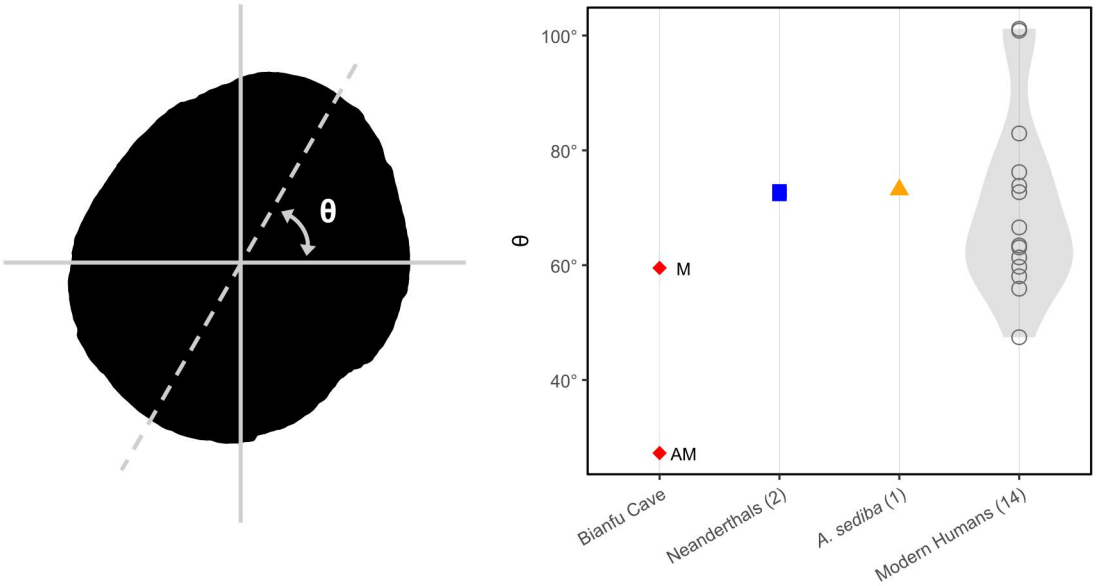
360



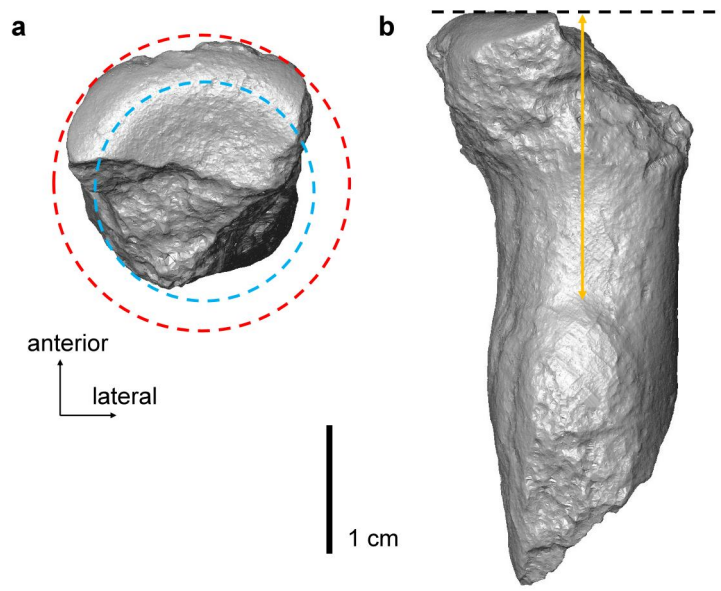
Supplementary Fig. 11. Bayesian phylogenetic tree to determine the placement of the tooth YHB3075 using *P. abelii* as an outgroup.



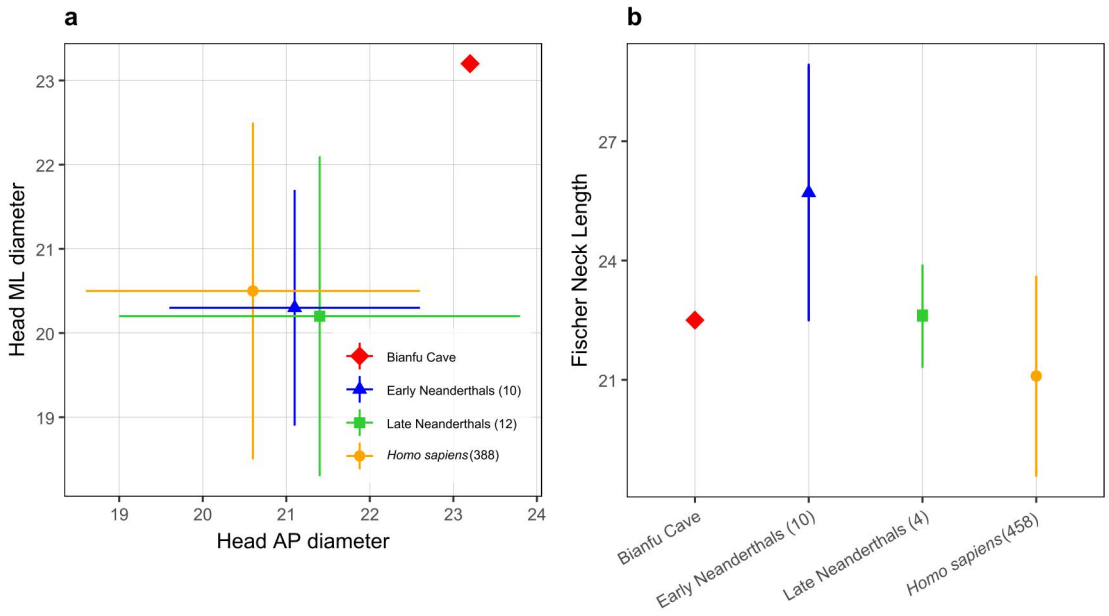
Supplementary Fig. 12. Comparative cross-sectional analysis of the radial mid-neck. The proximal radius fossil from Bianfu Cave (BFD771) is compared against the right radii of *Australopithecus sediba* (MH 2), Neanderthals (Feldhofer 1 and La Ferrassie 1), and 14 modern humans. The intact preservation of comparative specimens enables precise alignment following standard anatomical orientation, with the interosseous crest defining the medial aspect[44]. Cross-sections were extracted at the midpoint between the distal margin of the radial head and the proximal limit of the radial tuberosity. In the majority of comparative specimens, the major axis of the cross-section is oriented anterolaterally-posteromedially. BFD771 aligns with this morphology when oriented as a right radius. Two hypothetical orientations are modeled for BFD771: medial (M) and anteromedial (AM) positioning of the radial tuberosity.



Supplementary Fig. 13. Comparison of Theta (θ) in the radial mid-neck cross-section. The proximal radius fossil from Bianfu Cave (BFD771) is compared against the right radii of *Australopithecus sediba* (MH 2), Neanderthals (Feldhofer 1 and La Ferrassie 1), and 14 modern humans. Sample sizes for comparative taxa are indicated in parentheses. Theta represents the orientation of maximum bending rigidity[10]. Two hypothetical orientations for the radial tuberosity of BFD771 are present: medial (M) and anteromedial (AM). The medial orientation is considered most probable, as the major axis inclination of the BFD771 (AM) reconstruction represents an extreme outlier relative to the comparative specimens.



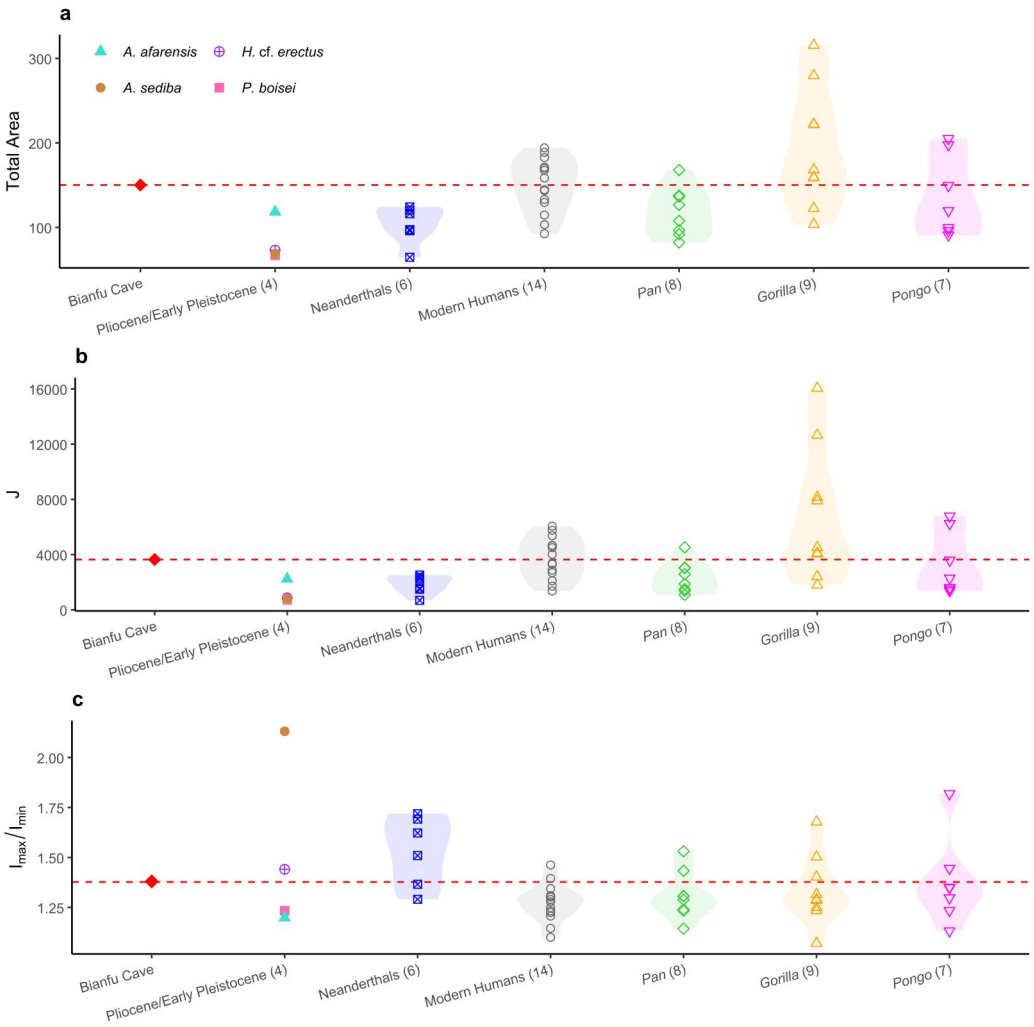
Supplementary Fig. 14. Osteometric measurement protocols for the proximal radius fossil from Bianfu Cave (BFD771). a, Estimation of radial head (red) and fovea (blue) diameters. The radial head diameter is estimated using a best-fit circle constrained by the preserved anteromedial and anterolateral margins. The radial fovea diameter is estimated through a best-fit circle aligned with the path of the preserved articular edge. The foveal circle serves as a reliable geometric constraint, providing a secondary dimensional reference for the estimation of the radial head. **b,** Radial neck length (Fischer neck length), defined as the distance between the superior medial margin of the radial head and the most proximal point of the radial tuberosity, following the definition of Rodríguez, et al.[91].



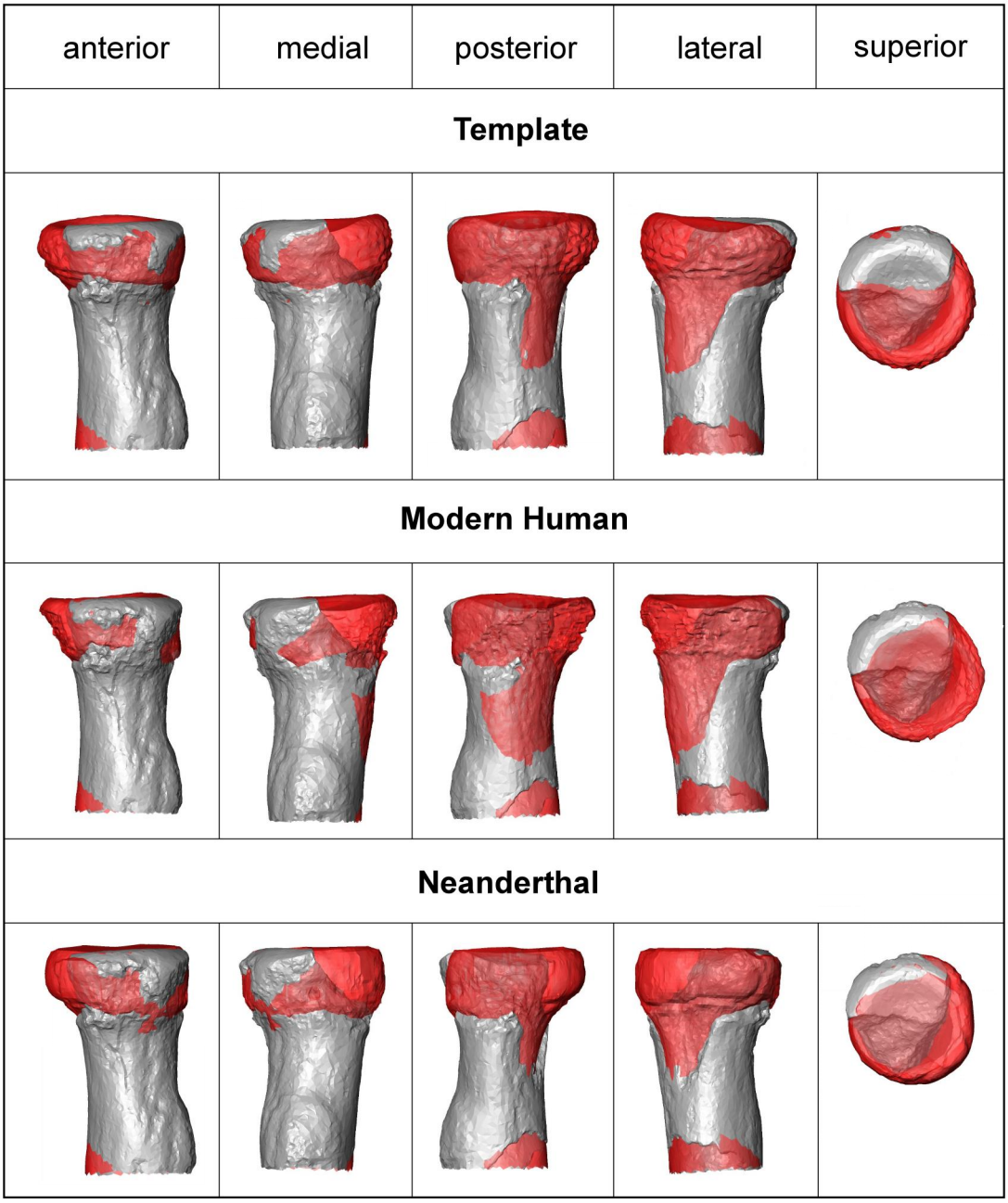
405

406 **Supplementary Fig. 15. Osteometric analysis of the proximal radius fossil from**
407 **Bianfu Cave (BFD771). a,** Bivariate plot of radial head dimensions (anteroposterior
408 versus mediolateral diameters). **b,** Comparative analysis of radial neck length (Fischer
409 neck length). Error bars indicate the mean ± one standard deviation. Comparative data
410 are sourced from Rodríguez, et al.[45]. The number in the parentheses indicates the
411 sample size.

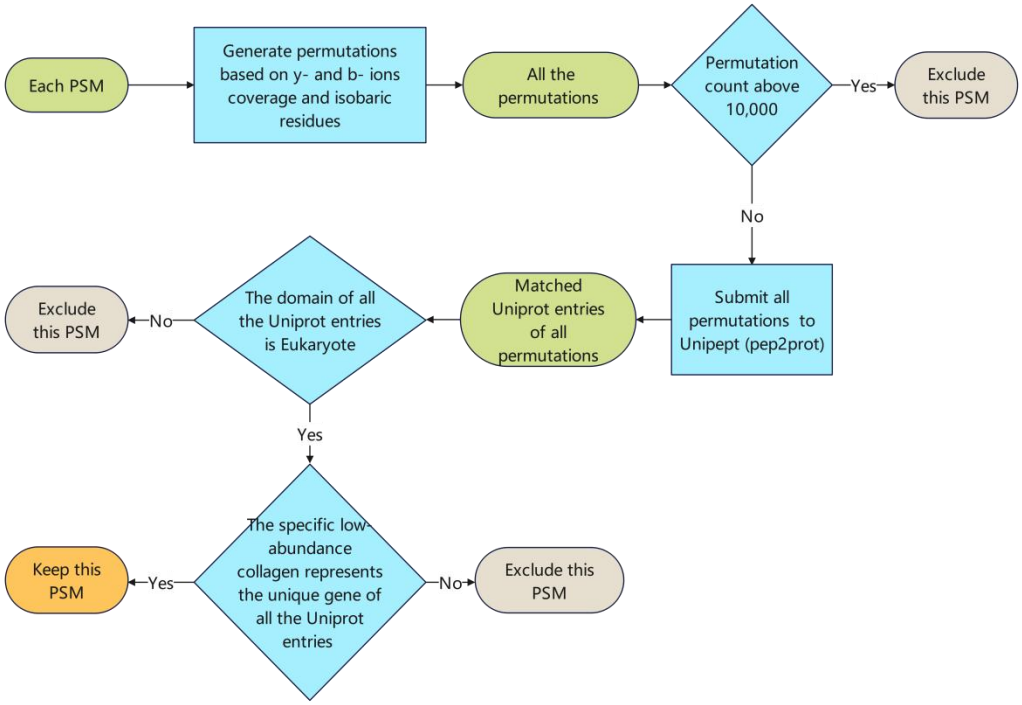
412



Supplementary Fig. 16. Cross-sectional geometric (CSG) analysis of the mid-neck region of the proximal radius fossil from Bianfu Cave (BFD771). Cross-sections were extracted at the midpoint between the distal margin of the radial head and the proximal limit of the radial tuberosity. **a,** Total area (TA), representing the area within the subperiosteal surface. **b,** Polar moment of area (J), reflecting torsional and average bending rigidity. **c,** The ratio of maximum to minimum second moments of area (I_{max}/I_{min}), representing the bending rigidity ratio and overall cross-sectional shape independent of orientation. Cross-sectional geometric parameters were calculated exclusively from the subperiosteal surface. Comparative Pliocene and Early Pleistocene hominin sample: *Homo* cf. *erectus* (SK 18b), *Australopithecus afarensis* (A.L. 288-1), *Australopithecus sediba* (UW88-85), *Paranthropus boisei* (KNM-ER 1500). Sample sizes for comparative taxa are indicated in parentheses.



Supplementary Fig. 17. Virtual restoration protocols for the proximal radius fossil from Bianfu Cave (BFD771) for diffeomorphic surface matching analysis. The template represents the mean shape model derived from a taxonomically diverse comparative sample, including modern humans, Neanderthals, *Australopithecus*, and extant great apes. The template, a representative modern human, and Neanderthal La Ferrassie 1 were truncated to match the preserved anatomy of BFD771. Meshes were subsequently subjected to rigid body transformation and uniform scaling, followed by translation and rotation to achieve a best-fit alignment with BFD771.

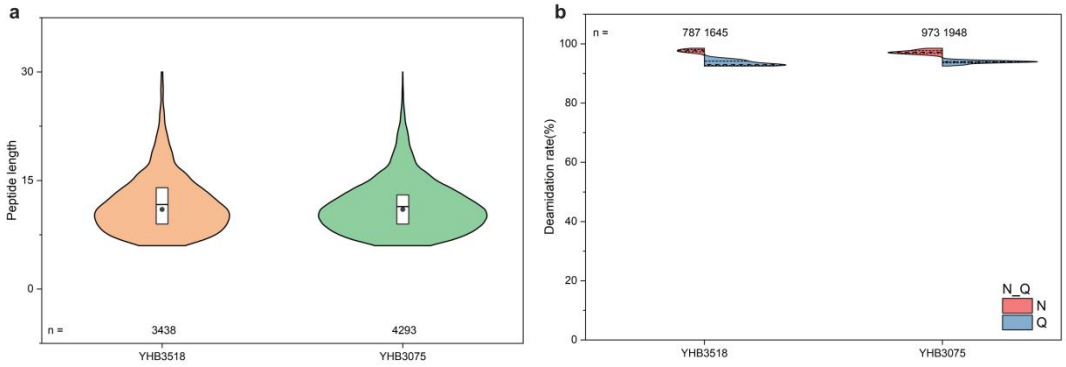


437

438 **Supplementary Fig. 18. The workflow for the "permutation and unipept**

439 **confirmation" steps to validate the identification of low-abundance collagens.**

440



Supplementary Fig. 19. Enamel proteome degradation of the teeth YHB3518 and YHB3075. a, Peptide length distribution. The number of peptides is shown for each tooth. Within the violin plots, the box plots represent the data range. The box edges show the lower and upper quartiles, while the center lines represent the means, and the black dots represent the medians. **b,** Asparagine (N) and glutamine (Q) deamidation profiles. Values are derived from 1,000 bootstrap replicates of sequence deamidation. The number of peptides containing N or Q with valid intensity is shown for each tooth.

Supplementary Tables

Supplementary Table 1. U-series dating results of three hominin bones and attached carbonates.

Sample	Spots	^{238}U [ppm]		($^{234}U/^{238}U$)		($^{238}U/^{232}Th$)		($^{230}Th/^{238}U$)		($^{230}Th/^{232}Th$)		Age [ka]		($^{234}U/^{238}U)_{initial}$	
BFD767 parietal	#1	7.6	±2.4	1.133	±0.008	2113	±492	0.225	±0.006	475	±111	23.9	±0.8	1.142	±0.009
	#2	7.8	±2.4	1.133	±0.009	17253	±2271	0.236	±0.007	4072	±548	25.3	±0.8	1.143	±0.010
	#3	7.5	±2.3	1.137	±0.009	12327	±1680	0.243	±0.006	2992	±414	26.0	±0.7	1.147	±0.009
	#4	6.4	±2.0	1.136	±0.010	6238	±1397	0.274	±0.008	1709	±386	29.8	±1.0	1.148	±0.010
	#5	6.1	±1.9	1.130	±0.010	12750	±1919	0.273	±0.006	3475	±529	29.9	±0.8	1.141	±0.010
	#6	6.5	±2.0	1.132	±0.008	3023	±764	0.251	±0.006	760	±193	27.2	±0.8	1.142	±0.009
	#7	4.8	±1.5	1.139	±0.014	5360	±898	0.286	±0.011	1532	±250	31.2	±1.4	1.152	±0.015
	#8	5.7	±1.8	1.136	±0.013	6219	±1397	0.286	±0.013	1778	±391	31.3	±1.7	1.149	±0.015
	#9	5.1	±1.6	1.124	±0.014	6644	±1323	0.277	±0.012	1838	±357	30.6	±1.6	1.135	±0.016
	#10	4.7	±1.5	1.147	±0.014	3925	±552	0.282	±0.013	1107	±148	30.5	±1.6	1.160	±0.015
	#11	1.7	±0.7	1.145	±0.015	36	±6	0.525	±0.038	19	±3	65.8	±6.4	1.175	±0.018
	Average*	6.2	±0.7	1.135	±0.004	7585	±3100	0.263	±0.014	1974	±745	28.6	±1.7	1.146	±0.004
BFD769 parietal	#1	5.5	±1.8	1.192	±0.013	4691	±585	0.914	±0.024	4288	±523	148.7	±8.6	1.292	±0.021
	#2	5.4	±1.7	1.199	±0.014	6492	±996	0.920	±0.024	5976	±904	148.9	±8.6	1.302	±0.022
	#3	5.3	±1.7	1.218	±0.013	5285	±1182	0.888	±0.022	4691	±1043	133.9	±6.8	1.318	±0.020
	#4	5.5	±1.8	1.226	±0.015	3235	±561	0.904	±0.029	2925	±499	136.8	±8.7	1.332	±0.023
	#5	5.9	±1.9	1.210	±0.013	6815	±980	0.930	±0.023	6339	±897	148.7	±8.2	1.320	±0.021
	Average	5.5	±0.2	1.209	±0.012	5303	±1292	0.911	±0.015	4844	±1227	143.4	±6.6	1.313	±0.014
BFD769 carbonates	#1	0.7	±0.1	0.912	±0.013	1.13	±0.04	1.027	±0.024	1.16	±0.05	-	-	-	-
	#2	0.9	±0.1	0.870	±0.012	1.45	±0.07	0.951	±0.020	1.38	±0.07	-	-	-	-
	#3	1.5	±0.1	0.942	±0.016	1.35	±0.04	0.973	±0.018	1.31	±0.04	-	-	-	-
	#4	0.7	±0.1	0.905	±0.016	1.42	±0.03	0.965	±0.024	1.37	±0.04	-	-	-	-
	#5	0.7	±0.0	0.915	±0.012	1.17	±0.04	1.009	±0.023	1.18	±0.05	-	-	-	-
	#6	1.2	±0.1	0.873	±0.011	1.06	±0.01	1.060	±0.018	1.12	±0.02	-	-	-	-

Sample	Spots	^{238}U [ppm]		($^{234}U/^{238}U$)		($^{238}U/^{232}Th$)		($^{230}Th/^{238}U$)		($^{230}Th/^{232}Th$)		Age [ka]		($^{234}U/^{238}U)_{initial}$	
	Average	0.9	±0.3	0.903	±0.022	1.26	±0.14	0.998	±0.034	1.26	±0.09				
BFD771 radius	#1	2.8	±0.9	1.161	±0.017	1707	±266	0.701	±0.022	1196	±183	98.0	±5.4	1.212	±0.022
	#2	3.1	±1.0	1.172	±0.016	1935	±316	0.705	±0.024	1364	±218	97.2	±5.6	1.226	±0.021
	#3	2.5	±0.8	1.154	±0.017	227	±36	0.746	±0.026	170	±26	109.6	±7.0	1.209	±0.023
	#4	3.1	±1.0	1.171	±0.018	1774	±327	0.742	±0.024	1316	±239	105.8	±6.2	1.230	±0.024
	#5	3.8	±1.2	1.166	±0.015	3511	±704	0.751	±0.021	2636	±524	108.7	±5.6	1.226	±0.021
	#6	4.1	±1.3	1.159	±0.015	2176	±506	0.676	±0.025	1470	±338	92.8	±5.5	1.207	±0.020
	Average	3.3	±0.5	1.164	±0.006	1888	±858	0.720	±0.025	1359	±642	102.0	±5.7	1.218	±0.008
BFD771 carbonates	#1	1.1	±0.1	1.095	±0.010	2.12	±0.06	1.039	±0.020	2.21	±0.08	-	-	-	-
	#2	1.5	±0.1	1.070	±0.008	1.87	±0.06	1.088	±0.017	2.03	±0.07	-	-	-	-
	#3	1.6	±0.1	1.047	±0.007	2.76	±0.06	1.112	±0.017	3.07	±0.08	-	-	-	-
	#4	0.9	±0.1	0.999	±0.011	1.68	±0.05	1.058	±0.021	1.78	±0.06	-	-	-	-
	#5	1.5	±0.1	1.106	±0.011	2.36	±0.04	1.119	±0.019	2.65	±0.07	-	-	-	-
	#6	0.8	±0.1	1.036	±0.009	2.36	±0.05	1.044	±0.022	2.46	±0.08	-	-	-	-
	Average	1.2	±0.3	1.059	±0.032	2.19	±0.32	1.077	±0.028	2.36	±0.38				

452 * The average value of the BFD767 parietal includes the measurements of #1 to 10. Spot #11 was excluded, because it was contaminated by
453 sedimentary thorium.
454

455 **Supplementary Table 2. The deamidation values and coverage of endogenous proteins identified in five hominin samples.***

Sample name	Protein	Accession	Deamidation N (%)	Deamidation Q (%)	Peptides (N)	Peptides (Q)	Peptides (Total)	#Spec	Amino acid coverage (%)
YHB3518 enamel	AHSG	C9JV77		100.00	0	3	6	9	7.07
	AMBN	Q9NP70	99.88	95.50	205	436	741	3679	46.53
	AMELX	Q99217-3	94.79	93.63	328	626	1445	11754	91.71
	AMELY	Q99218	91.95	94.96	257	382	999(286)	8278(1154)	72.33
	COL17A1	Q9UMD9		99.13	0	26	42	102	5.01
	COL1A1	P02452	100.00	100.00	12	9	74	114	20.83
	COL1A2	P08123	100.00	90.48	21	7	48	73	19.18
	ENAM	Q9NRM1	98.58	91.73	200	373	751	3606	18.74
	MMP20	O60882	100.00		2	0	7	8	7.87
YHB3518 dentine	COL11A2	P13942		100.00	0	2	32	78	10.2
	COL1A1	P02452	82.87	87.67	261	496	1383	4816	64.62
	COL1A2	P08123	89.64	91.13	299	150	887	3610	53.66
	COL2A1	P02458	85.33	70.46	12	37	78	222	33.69
	COL5A1	P20908	50.00	91.64	2	7	20	63	11.1
	COL5A2	P05997	100.00	77.96	26	25	84	276	25.75
YHB3075 enamel	AHSG	C9JV77		100.00	0	5	9	17	7.61
	ALB	P02768	99.64	97.73	26	22	78	273	41.38
	AMBN	Q9NP70	99.48	95.92	232	515	899	3832	47.65
	AMELX	Q99217-3	94.89	92.92	391	683	1650	12903	95.12
	AMELY	Q99218	92.06	94.74	305	405	1146(349)	9559(1492)	72.33
	AMTN	Q6UX39		100.00	0	6	10	15	13.88
	COL17A1	Q9UMD9		100.00	0	13	29	62	5.48
	ENAM	Q9NRM1	98.12	92.64	323	528	1259	6157	22.33
	MMP20	O60882	100.00	100.00	2	1	20	27	15.53

Sample name	Protein	Accession	Deamidation N (%)	Deamidation Q (%)	Peptides (N)	Peptides (Q)	Peptides (Total)	#Spec	Amino acid coverage (%)
	SERPINA1	P01009	100.00	87.50	3	8	30	68	28.23
	SERPINC1	P01008	81.50	69.68	5	14	55	305	21.55
YHB3075 dentine	COL1A1	P02452	90.55	87.33	192	387	1129	3150	62.16
	COL1A2	P08123	91.19	89.76	272	146	797	2772	58.93
	COL2A1	P02458	100.00	83.33	5	10	26	41	23.2
	COL5A1	P20908	100.00	100.00	5	10	25	73	9.9
	COL5A2	P05997	100.00	63.33	10	11	54	172	22.55
BFD767 bone	COL1A1	P02452	87.16	89.19	883	1473	5254	24382	66.19
	COL1A2	P08123	89.12	88.37	1018	666	3232	18473	70.35
	COL23A1	Q86Y22		100.00	0	2	7	8	12.78
	COL2A1	P02458	57.56	73.67	38	91	209	441	44.25
	COL5A1	P20908	62.38	79.13	13	31	66	205	25.57
	COL5A2	P05997	90.48	80.17	34	44	143	536	35.96
BFD769 bone	COL15A1	P39059	0.00	100.00	1	2	3	3	1.95
	COL1A1	P02452	98.51	97.05	156	296	795	1464	51.23
	COL1A2	P08123	99.48	95.89	177	122	619	1192	56.73
	COL2A1	P02458	100.00	100.00	1	2	8	10	7.06
	COL5A1	P20908	100.00	83.33	1	3	11	11	7.13
	COL5A2	P05997	100.00	100.00	3	5	18	26	12.27
BFD771 bone	COL1A1	P02452	96.86	94.77	177	337	1136	2104	56.63
	COL1A2	P08123	97.58	95.73	232	150	792	1577	55.56
	COL2A1	P02458	100.00	95.00	4	11	49	62	25.62
	COL5A1	P20908	100.00	100.00	3	9	23	32	10.99
	COL5A2	P05997	86.08	93.04	5	4	27	42	16.74

*In the “Peptides(N)” and “Peptides(Q)” columns, the numbers are the counts of N/Q-containing peptides with valid intensities, which are used

457 in the deamidation calculation. The “Amino acid coverage (%)” was calculated based on the sequence length of each protein from the human
458 reference accession listed.
459

Supplementary Table 3. The bone thickness of the parietal at the eminence for different comparative specimens.

Taxonomy	Bone thickness range at the parietal eminence (in mm)	Data source
Zhoukoudian *H. erectus* (n = 8)	5–16	References[92]
East Asian late Middle Pleistocene archaic *Homo* (n = 6)*	6–12.6	References[93-95]
Neanderthals (n = 32)	5–17	References[96]
Homo heidelbergensis (n = 9)	6.5–11.5	References[96,97]
Early modern human (n = 33)	3.5–12	References[96,98-101]

* Specimens of East Asian late Middle Pleistocene archaic *Homo* include the following: Jinniushan (6 mm), Maba (7 mm), Xuchang (7.9 mm), Dali (11.2 mm), Xujiayao (7-12.6 mm).

Supplementary Table 4. Virtual datasets used for radial morphological analysis obtained from MorphoSource.

Taxa	Media ID	Manager	Organization	ARK	DOI
A. sediba	000050419	Lee Berger	Evolutionary Studies Institute	ark:/87602/m4/M50419	
Pan	000091844	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M91844	10.17602/M2/M91844
Pan	000091928	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M91928	10.17602/M2/M91928
Pan	000095008	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M95008	10.17602/M2/M95008
Pan	000167806	Sergio Almecija	AMNH Mammal Collections	ark:/87602/m4/M167806	10.17602/M2/M167806
Pan	000725317	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725317	
Pan	000725398	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725398	
Pan	000725430	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725430	
Pan	000725455	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725455	
Gorilla	000085875	Alisha Anaya	AMNH Mammal Collections	ark:/87602/m4/M85875	10.17602/M2/M85875
Gorilla	000086646	Alisha Anaya	AMNH Mammal Collections	ark:/87602/m4/M86646	10.17602/M2/M86646
Gorilla	000090825	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M90825	10.17602/M2/M90825
Gorilla	000091645	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M91645	10.17602/M2/M91645
Gorilla	000468985	Sergio Almecija	Hamann-Todd Non-Human Primate Collection	ark:/87602/m4/468985	10.17602/M2/M468985
Gorilla	000725268	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725268	
Gorilla	000725343	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725343	
Gorilla	000725374	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725374	
Gorilla	000725558	Jeffrey Spear	RMCA - Mammal Collections	ark:/87602/m4/725558	
Pongo	000088099	Alisha Anaya	AMNH Mammal Collections	ark:/87602/m4/M88099	10.17602/M2/M88099
Pongo	000093969	Maryse Biernat	NMNH - Division of Mammals	ark:/87602/m4/M93969	10.17602/M2/M93969

Taxa	Media ID	Manager	Organization	ARK	DOI
Pongo	000095247	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M95247	10.17602/M2/M95247
Pongo	000095255	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M95255	10.17602/M2/M95255
Pongo	000095420	Eleanor Cordiner	NMNH - Division of Mammals	ark:/87602/m4/M95420	10.17602/M2/M95420
Pongo	000160955	Roger Benson	NHMUK Life Sciences Mammal Section	ark:/87602/m4/M160955	
Pongo	000735170	Jeffrey Spear	Zoological Department (Vertebrates)	ark:/87602/m4/735170	

Supplementary References

89 Schroeter, E. R. & Cleland, T. P. Glutamine deamidation: an indicator of antiquity, or preservational quality? *Rapid Communications in Mass Spectrometry* **30**, 251-255 (2016).

90 Welker, F. *et al.* Variations in glutamine deamidation for a Châtelperronian bone assemblage as measured by peptide mass fingerprinting of collagen. *STAR: Science & Technology of Archaeological Research* **3**, 15-27, doi:10.1080/20548923.2016.1258825 (2017).

91 Rodríguez, L. *et al.* Fossil hominin radii from the Sima de los Huesos Middle Pleistocene site (Sierra de Atapuerca, Spain). *Journal of Human Evolution* **90**, 55-73 (2016).

92 Weidenreich, F. The skull of Sinanthropus pekinensis: a comparative study on a primitive hominid skull. *Paleontologia Sinica, Series D* **10**, 96-157 (1943).

93 Wu, X. in *Ancient humans in China* (eds Rukang Wu, Xinzhi Wu, & Senshui Zhang) 24-41 (Science Press, 1989).

94 Wu, X. *Middle Pleistocene human skull from Dali, China*. Vol. 13 1-205 (Science Press, 2020).

95 Li, Z. *et al.* Late Pleistocene archaic human crania from Xuchang, China. **355**, 969-972, doi:doi:10.1126/science.aal2482 (2017).

96 Copes, L. E. & Kimbel, W. H. Cranial vault thickness in primates: Homo erectus does not have uniquely thick vault bones. *Journal of Human Evolution* **90**, 120-134, doi:https://doi.org/10.1016/j.jhevol.2015.08.008 (2016).

97 Kennedy, K. A. R., Sonakia, A., Chiment, J. & Verma, K. K. Is the Narmada hominid an Indian Homo erectus? *American journal of physical anthropology* **86**, 475-496, doi:https://doi.org/10.1002/ajpa.1330860404 (1991).

98 Curnoe, D. *et al.* Human remains from the Pleistocene-Holocene transition of southwest China suggest a complex evolutionary history for East Asians. *PloS one* **7**, e31918 (2012).

99 Fu, R. *et al.* Modern Homo sapiens skeleton from Qianyang Cave in Liaoning, northeastern China and its U-series dating. *Journal of human evolution* **55**, 349-352 (2008).

100 Curnoe, D., Datan, I., Taçon, P. S., Leh Moi Ung, C. & Sauffi, M. S. Deep skull from Niah Cave and the Pleistocene peopling of Southeast Asia. *Frontiers in Ecology Evolution* **4**, 75 (2016).

101 Suzuki, H. in *The Minatogawa Man: The Upper Pleistocene Man from the Island of Okinawa* (eds H. Suzuki & K Hanihara) 7-49 (University of Tokyo Press, 1982).