

# **Title: Palaeoproteomic evidence identifies archaic hominins associated with the Châtelperronian at the Grotte du Renne.**

## **Authors:**

F. Welker<sup>a,b,1</sup>, M. Hajdinjak<sup>c</sup>, S. Talamo<sup>a</sup>, K. Jaouen<sup>a</sup>, M. Dannemann<sup>c,d</sup>, F. David<sup>e</sup>, M. Julien<sup>e</sup>, M. Meyer<sup>c</sup>, J. Kelso<sup>c</sup>, I. Barnes<sup>f</sup>, S. Brace<sup>f</sup>, P. Kamminga<sup>g</sup>, R. Fischer<sup>h</sup>, B. Kessler<sup>h</sup>, J.R. Stewart<sup>i</sup>, S. Pääbo<sup>c</sup>, M.J. Collins<sup>b</sup>, J.-J. Hublin<sup>a</sup>

<sup>a</sup>Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany.

<sup>b</sup>BioArCh, University of York, York, YO10 5DD, United Kingdom.

<sup>c</sup>Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany.

<sup>d</sup>Medical Faculty, University of Leipzig, 04103 Leipzig, Germany.

<sup>e</sup>Paris Unité Mixte de Recherche 7041, Archéologies et Sciences de l'Antiquité, Centre National de la Recherche Scientifique, 92023 Nanterre, France

<sup>f</sup>Department of Earth Sciences, Natural History Museum, London SW7 5BD, United Kingdom.

<sup>g</sup>Naturalis Biodiversity Center, P.O. Box 9517, 2300 RA Leiden, The Netherlands.

<sup>h</sup>Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, United Kingdom.

<sup>i</sup>Faculty of Science and Technology, Bournemouth University, Talbot Campus, Fern Barrow, Poole, Dorset, BH12 5BB, United Kingdom.

<sup>1</sup>To whom correspondence should be addressed. Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany. Telephone: +49(0)3413550749. Email: frido\_welker@eva.mpg.de/frido.welker@palaeo.eu

## **Author contributions:**

F.W., J.-J.H., J.R.S. and M.J.C. designed the research. I.B., S.B., P.K., B.K., and F.D. provided reagents, samples and laboratory equipment. F.W., M.H., S.T., K.J., M.D. performed experiments. F.W., M.H., S.T., K.J., M.D., M.M., J.K., R.F., S.P., M.J.C., J.-J.H. analysed and interpreted results. F.W., J.-J.H. and M.J.C. wrote the manuscript with contributions of all other authors.

## **Classification:**

Biological/Social Sciences: Anthropology

## **Keywords:**

Palaeoproteomics, Châtelperronian, ZooMS

**Abstract:**

In western Europe, the Middle to Upper Palaeolithic transition is associated with the disappearance of Neandertals and the spread of Anatomically Modern Humans (AMHs). Current chronological, behavioural and biological models of this transitional period hinge on the Châtelperronian technocomplex. At the site of the Grotte du Renne, Arcy-sur-Cure, morphological Neandertal specimens are not directly dated but contextually associated to the Châtelperronian, which contains bone points and beads. The association between Neandertals and this 'transitional' assemblage has been controversial, because of the lack of either a direct hominin radiocarbon date or molecular confirmation of the Neandertal affiliation. Here we provide further evidence for a Neandertal-Châtelperronian association at the Grotte du Renne through biomolecular and chronological analysis. We identified 28 additional hominin specimens through ZooMS (Zooarchaeology by Mass Spectrometry) screening of morphologically uninformative bone specimens from Châtelperronian layers at the Grotte du Renne. Next, we obtain an ancient hominin bone proteome through LC-MS/MS analysis and error-tolerant amino acid sequence analysis. Analysis of this palaeoproteome allows us to provide phylogenetic and physiological information on these ancient hominin specimens. We distinguish, for the first time, Late Pleistocene clades within the genus *Homo* based on ancient protein evidence through the identification of an archaic-derived amino acid sequence for the *COL10a1* protein. We support this by obtaining ancient mtDNA sequences, which indicate a Neandertal ancestry for these specimens. Direct AMS radiocarbon dating and Bayesian modeling confirm that the hominin specimens date to the Châtelperronian at the Grotte du Renne.

**Significance statement:**

The displacement of Neandertals by Anatomically Modern Humans (AMHs) 50-40,000 years ago in Europe has considerable biological and behavioural implications. The Châtelperronian at the Grotte du Renne (France) takes a central role in models explaining the transition but the association of hominin fossils at this site with the Châtelperronian is debated. We identify additional hominin specimens at the site through proteomic ZooMS screening and obtain molecular (ancient DNA, ancient proteins) and chronometric data to demonstrate that these represent Neandertals that date to the Châtelperronian. The identification of an amino acid sequence specific to a clade within the genus *Homo* demonstrates the potential of palaeoproteomic analysis in the study of hominin taxonomy in the Late Pleistocene, and warrants further exploration.

\body

## **Introduction:**

In order to understand the cultural and genetic interaction between the last Neandertals and some of the earliest Anatomically Modern Humans (AMHs) in Europe we need to resolve the taxonomic affiliation of the hominins associated with the “transitional” industries characterizing the replacement period, such as the Châtelperronian (1, 2). The well-characterized Châtelperronian lithic technology has recently been re-classified as fully Upper Palaeolithic (3), and is associated at several sites with bone awls, bone pendants and colorants (4, 5). The Grotte du Renne at Arcy-sur-Cure, France, is critical to competing behavioral and chronological models for the Châtelperronian, as at this site the Châtelperronian is stratigraphically associated with hominin remains that are morphologically identified as Neandertals (6–8). Hypotheses explaining this association range from **a**) “acculturation” by AMHs (9), **b**) independent development of such artefacts by Neandertals (5), **c**) movement of pendants and bone artefacts from the overlying Aurignacian into the Châtelperronian layers (10, 11), or **d**) movement of the hominins specimens from the underlying Mousterian into the Châtelperronian layers (10, 12). The first two hypotheses assume that the stratigraphic association of the hominins and the Châtelperronian assemblage is genuine, while the latter two hypotheses counter that the association is due to large-scale, taphonomic, movement of material. In all scenarios, the morphological identification of these hominins as Neandertals is accepted but unsupported by molecular evidence.

To test the chronostratigraphic coherence of the site, Bayesian models of radiocarbon dates for the site have been constructed (10, 13). The results of these two models contradict each other in the extent to which archaeological material moved between the Châtelperronian and non-Châtelperronian archaeological layers. Furthermore, they have been criticized on various methodological aspects (13, 14), and the first (10) is at odds with some archaeological evidence that suggests that large-scale displacement of material into the Châtelperronian from either the overlying or underlying layers is unlikely (14). Both Bayesian models are only indirect tests of the hominin-Châtelperronian association as no direct radiocarbon dates of the hominins are available.

Pending the discovery of further hominin specimens at other Châtelperronian sites, the Châtelperronian at the Grotte du Renne remains crucial in order to obtain a coherent biological and chronological view of the transitional period in Europe. It has been demonstrated previously that palaeoproteomics allows the identification of additional hominin specimens among unidentified Pleistocene faunal remains (ZooMS: Zooarchaeology by Mass Spectrometry; (15–17)) although here the value of doing so has for the first time enabled the direct dating and unambiguous identification of the Neandertal association with the Châtelperronian. We successfully apply this to the Grotte du Renne Châtelperronian. We obtain a direct hominin radiocarbon date at the site thereby directly addressing the chronostratigraphic context of this specimen in relation to hypothesis **d** and provide biomolecular data (palaeoproteomics, aDNA) on the genetic ancestry of the Grotte du Renne Châtelperronian hominins. Although proteomic data on Pleistocene hominin bone specimens has been presented before (17, 18), the phylogenetic and physiological implications of such datasets has, so far, not been fully explored. Here we utilize the potential of error-tolerant MS/MS database searches in relation to the biological questions associated with the Châtelperronian. Such technical advances have not been applied to entire palaeoproteomes from Late Pleistocene hominins before (19). In addition, we

demonstrate that the bone proteome reflects the developmental state of an ancient hominin individual. Throughout our study, we develop and employ tools designed to minimize, identify and exclude protein and DNA contamination (Fig. S1).

## Results:

### *ZooMS screening:*

We screened 196 taxonomically unidentifiable or morphologically dubious bone specimens (commonly <20 mg of bone) using ZooMS (Zooarchaeology by Mass Spectrometry) from the areas of the Grotte du Renne that had previously yielded hominin remains (20). This required us to construct a collagen type I (COL1) sequence database including at least one species of each medium or larger-sized genus in existence in western Europe during the Late Pleistocene (19) (Dataset S1) and from this derived a ZooMS peptide marker library (Dataset S2). ZooMS utilizes differences in tryptic peptide masses from *COL1 $\alpha$ 1* and *COL1 $\alpha$ 2* amino acid chains to taxonomically identify bone and tooth specimens (15). The peptide marker library combines newly obtained and published COL1 sequences with published ZooMS peptide markers (15, 21), and enabled us to confidently identify 28 bone fragments within the extant *Pan-Homo* clade, to the exclusion of other Hominidae (Figs. S1, S2; Tab. S1). Together with other studies, this confirms the suitability of ZooMS as a screening technique to identify hominin specimens among unidentified fragmentary bone specimens (17, 22). We confirmed these identifications for samples AR-7, AR-16 and AR-30 by analyzing the same extracts using shotgun proteomics and spectra assignment against our COL1 sequence database (Fig. S3). In each case, taxonomic assignment to the genus *Homo* had the highest score (see SI Appendix Section 1).

Molecular contamination is an important issue when studying ancient biomolecules, especially when it concerns ancient hominins. Extraction blanks were included throughout all analysis stages to monitor the introduction of potential contamination, although such controls only provide insight into contamination introduced during the laboratory analysis. MALDI-TOF-MS analysis of these blanks showed no presence of COL1 peptides (Fig. S2b). Furthermore, as a marker of diagenetic alteration of amino acids (23), glutamine deamidation values based on ammonium-bicarbonate ZooMS hominin spectra indicated that the analyzed collagen has glutamine deamidation values significantly different from modern bone specimens (t-test:  $p=2.31\text{E-}11$ ; Fig. 1A), but similar to deamidation values obtained for faunal specimens analyzed from the Grotte du Renne (t-tests; peptide P1105:  $p=0.85$ ; peptide P1706:  $p=0.55$ ; (24)). We interpret this to support the identification of endogenous, non-contaminated hominin COL1.

### *Palaeoproteomic (LC-MS/MS) analysis:*

After identifying additional hominin specimens at the Grotte du Renne by ZooMS, we undertook palaeoproteomic and genetic analyses to establish whether these newly identified hominins represent AMHs or Neandertals. Error-tolerant LC-MS/MS analysis of the protein content of the ZooMS extracts of AR-7, AR-16 and AR-30 and two additional palaeoproteomic extractions performed on AR-30 resulted in the identification of 73 proteins (Tabs. 1, S2). We base our assessment of endogenous and possibly exogenous proteins on four lines of evidence:

First, we analyzed our extraction blanks by LC-MS/MS analysis. This allowed us to identify several proteins introduced as contaminants during the analysis (human keratins, histones and HBB; Tabs. 1, S2), and matches to these proteins for AR-7, AR-16 and AR-30 were excluded from further in-depth analysis.

Second, we searched our data against the complete UniProt database, which contains additional non-human and non-vertebrate proteins from various sources that could have contaminated our extracts. Spectral matches to non-vertebrate proteins comprise <1.0% of the total number of matched spectra, indicating a minimal presence of non-vertebrate contamination (Fig. S4). These spectral matches were subsequently excluded from analysis.

Third, unsupervised cluster analysis based on glutamine and asparagine deamidation frequencies observed for all identified vertebrate proteins revealed a clear separation in three clusters (Figs. 1B, S5). The first group of 14 proteins display almost no deamidated asparagine and glutamine positions. Many of these (keratins, trypsin, bovine *CSN2*), but not all (*COL4a6*, *UBB*, *DCD*), have previously been reported as contaminants (Fig. 1B, filled triangles; (25)). All proteins identified in our extraction blanks have deamidation frequencies that fall into this group, further suggesting that these 14 proteins are contaminants. The second and third groups comprise a total of 35 proteins with elevated levels of deamidated asparagine and glutamine residues (Fig. 1B, filled circles and squares). These include various collagens and non-collagenous proteins (NCPs) previously reported in (ancient) bone proteomes, and we interpret that these proteins are endogenous to the analyzed bone specimens. We were unable to obtain sufficient deamidation spectral frequency data for 24 additional proteins, as insufficient numbers of asparagine or glutamine-containing peptides were present (Fig. 1B, open circles). Some of these 24 proteins have previously been identified in non-hominin bone proteomes or are involved in bone formation and maintenance (*POSTN*, *THBS1*, *ACTB*, *C3*, *IGHG1*, *DLL*), and are therefore likely endogenous to the bone specimens as well.

Fourth, after exclusion of contaminants, protein composition was similar to other non-hominin bone palaeoproteomes (25, 26). For these proteins we observed the presence of additional diagenetic and *in vivo* post-translational modifications (Fig. S6, Tab. S3, S4), likewise suggesting the retrieval of an endogenous hominin palaeoproteome. Based on these results, we suggest that matches to human proteins in palaeoproteomic analysis should be supported by additional lines of evidence to substantiate the claim that these represent proteins endogenous to the analyzed tissue and not derive from exogenous contamination derived from either handling of the bone specimen or contamination introduced during the analytical procedure.

Among the non-contaminant proteins there are several proteins that are (specifically) expressed by (pre)hypertrophic chondrocytes (*COL10a1* and *COL27a1* (27)) and osteoblasts (*DLL3* (28) and *COL24a1* (29)) during bone formation. The presence of *COL10a1* is particularly noteworthy. It is preferentially secreted by (pre)hypertrophic chondrocytes during initial bone ossification in bone formation, including cranial sutures (30, 31), and would therefore be removed from the bone matrix dependent on the rate of bone remodelling of a given mineralized tissue. In line with this, Gene Ontology (GO) annotation analysis indicates enrichment for GO biological processes related to cartilage development and bone ossification (Tab. S5). These observations are consistent with osteological and isotopic observations, which suggest the identified bone specimens belong to a breastfeeding infant (see below). GO annotation analysis further identified a significant group of blood microparticles (GO:00725262) such as albumin (*ALB*). These proteins have been identified in non-hominin palaeoproteomes as well and are consistently incorporated into the mineralized bone matrix (25, 26).

Error-tolerant analysis of MS/MS spectra has the ability to identify amino acid variants not present in protein databases or reference genomes, potentially revealing relevant phylogenetic information. We used an error-tolerant search engine (PEAKS) against the human reference proteome, and compared our protein sequence data against available amino acid

sequence variation known through genomic research for modern humans (32), a Denisovan genome (33), and the coding regions of three Neandertals (34). Using this approach, we confidently identify five proteins that contain a total of seven amino acid positions with non-synonymous SNPs with both alleles at frequencies  $\geq 1.0\%$  in present-day humans (Tab. S6). In six cases we observed the ancestral Hominidae state in the proteome data, which is also present in Denisovan and Neandertal protein sequences (34). These include one position for which a majority of AMHs (93.5%) carry a derived substitution (*COL28a1*; dbSNP rs17177927) and where our data contains the ancestral position (amino acid P). For the seventh case, *COL10a1*, we observed an amino acid state present in Denisovans, Neandertals and 0.9% of modern humans haplotypes (46/5008 1000G haplotypes; Tabs. S6, S9), but not in any other Hominidae (*Pongo abelii*, *Gorilla gorilla* or *Pan troglodytes*; Tab. S7).

We identified *COL10a1* in all three analyzed bone specimens (Tab. S2), but not in our extraction blank. The deamidation frequency observed for *COL10a1* (Fig. 1B) and the excretion of *COL10a1* by (pre)hypertrophic chondrocytes during ossification (30) indicates an endogenous origin of the identified *COL10a1* peptides. We identify one peptide for *COL10a1* (Tab. S8) that contains an amino acid position indicative of an archaic sequence (Neandertal or Denisovan). The peptide of interest is represented in two palaeoproteomic analyses performed on AR-30, with three spectrum-peptide matches in total (Fig. S7). Correct precursor mass and fragment ion assignment were validated manually to exclude false assignment of  $^{13}\text{C}$  derived isotopic peaks as deamidated variants of the peptide, which led to the exclusion of a fourth spectrum (see SI Appendix Section 3). All three spectra represent semi-tryptic peptides (Fig. S7). As in other palaeoproteomes, the presence of such semi-tryptic peptides is not uncommon and likely the result of protein diagenesis (23, 25, 35–37). In addition, all three spectra contain a hydroxylated proline on the same position (*COL10a1* position 135), further demonstrating consistency among our peptide-spectrum matches. The replication of our results in two independent analyses and the inferred presence of post-translational modifications in all three spectra, one of which is identically placed, further support the notion that these are endogenous to the analyzed bone specimen.

The nucleotide position of interest is located at chr6:116442897 (hg19, dbSNP rs142463796), which corresponds to amino acid position 128 in *COL10a1* (UniProt Q03692). For all three available Neandertal sequences this position carries the nucleotide T (34), which translates into the amino acid N (codon “Aat”, 3’ to 5’). The position is heterozygous N/D in the Denisovan genome (33), a D in the Ust’-Ishim  $\approx 45,000$  BP AMH genome (38), and D in 99.1% of modern humans (32) (codon “Gat”, 3’ to 5’). The remaining 0.9% of modern human individuals analyzed match the Neandertal sequence. All these individuals are outside sub-Saharan Africa (Fig. 2). For amino acid position *COL10a1* 128, the amino acid N represents the derived state, and the amino acid D is the ancestral state (Tab. S7).

#### *COL10a1 introgression into modern humans:*

When present in modern humans, the archaic-like allele is found in populations known to have archaic introgression (39): South-East Asia (1-6%), Oceania (33-47%) and with high frequencies in Papua New Guinea (47%; Fig. 2; Tab. S9). This archaic-like allele is found on extended archaic-like haplotypes that have a minimum length of 146kb (Fig. S8). Given the recombination rate of 0.4 cM/Mb in this region (40) and the age of the Neandertal and Denisovan samples (39), we compute that this haplotype length is more consistent with archaic introgression than with incomplete lineage sorting (see SI Appendix Section 4). Since *COL10a1* 128N is present in  $<1\%$

of present-day humans as a consequence of archaic introgression, its presence in sample AR-30 suggests archaic (Neandertal+Denisovans) ancestry for at least part of its nuclear genome.

#### *mtDNA analysis:*

To support the palaeoproteomic evidence, we extracted mtDNA from AR-14 and AR-30 (SI Appendix Section 5, Tab S10). Elevated C to T substitution frequencies at terminal sequence ends (up to 12.1% for AR-14 and 28.1% for AR-30) suggest that at least some of the recovered sequences for both specimens are of ancient origin (41) (Tab. S11). When restricting the analysis to these deaminated mtDNA fragments, support for the Neandertal branch in a panel of diagnostic mtDNA positions is above 70% (Fig. S9), but is without support for the Denisovan branch. This is confirmed when only diagnostic positions differing between Neandertals and present-day humans are included (Tab. S12). The uniparental mode of maternal inheritance for mtDNA, the absence of notable mtDNA contamination from Neandertal mtDNA in the extraction blanks, and the dominance of deaminated mtDNA sequences aligning to the Neandertal mtDNA branch all demonstrate that AR-14 and AR-30 are mitochondrial Neandertals. Residual modern human contamination in the fraction of deaminated sequences makes reconstruction of Neandertal mtDNA consensus sequences impossible for both bone specimens. We were therefore unable to test whether AR-14 and AR-30 are maternally related. Nevertheless, the above analyses allow us to conclude that both specimens carry mtDNA of the type seen in Late Pleistocene Neandertals.

#### *Isotopic analysis and radiocarbon dating:*

We extracted collagen from specimen AR-14 (MAMS-25149) to provide a direct AMS date from a Grotte du Renne hominin and thereby address the possibility that these hominins derive from the underlying Middle Palaeolithic Mousterian (hypothesis **d**; (10)). In addition, we extracted collagen from 21 additional faunal bone specimens, identified by ZooMS and from the Châtelperronian Layers IX and X at the Grotte du Renne, to provide an isotopic context for the  $\delta^{15}\text{N}$  and  $\delta^{13}\text{C}$  stable isotopes obtained for AR-14. The collagen quality criteria are within accepted ranges (Tab. S14). The stable isotopes indicate the  $\delta^{15}\text{N}$  is 5.4‰ higher for AR-14 compared to associated carnivores (Tabs. S14, S15; Fig. S11). This suggests breastfeeding as a major dietary protein source (42) (see SI Appendix Section 7) which is in agreement with the presence of *COL10a1* and the presence of an unfused vertebral hemiarch, and in support of the interpretation that these remains represent a breastfeeding infant (Fig. S10; Tab. S13).

To test different chronostratigraphic scenarios we constructed four different Bayesian models to address various criticisms raised against previous models: 1) the Hublin *et al.* model which includes all ages and priors following (13), 2) the hominin-modified model, which includes only ages from hominin-modified bone specimens (priors and dates from (13)), 3) the Discamps *et al.* model, which excludes radiocarbon ages obtained from an area of the site considered reworked by some (43) (priors and dates from (13)), and 4) the Higham *et al.* model, which includes all ages and priors following (44) (a model which also includes hominin-modified bones only). The Bayesian CQL code for each of the four models is included in the SI (Appendix Section 6). All four models treat the Châtelperronian Layers IX and X as a single chronological phase based on lithic refits between these two layers (13, 45). The measured age for AR-14,  $36,840 \pm 660$   $^{14}\text{C}$  BP, fits within the Châtelperronian chronological boundaries calculated in all four Bayesian models (Fig. 3, Tab. S16; (10, 13)). With a posterior outlier probability of 4 to 8%, AR-14 is unlikely to derive from the underlying Mousterian or the

overlying Aurignacian at the Grotte du Renne (Tab. S16) even when only hominin-modified bone specimens are included or bone specimens from potentially reworked areas are excluded (following (43)).

### **Discussion and conclusion:**

Despite their spatial proximity within Layer X and squares C7 and C8, none of the newly described specimens could be fitted together to form larger fragments. The morphologically informative specimens, however, seem to represent small fragments of an immature cranium and an unfused vertebral hemiarch of neonatal age (Fig. S10; Tab. S13). This is supported by isotopic evidence suggesting that these fragments belonged to a non-weaned infant and by proteomic evidence in the form of proteins present in bone before bone remodelling has started. All the newly identified specimens were found in close spatial association with a previously described hominin temporal bone from square C7, Layer Xb, assigned to an infant around 1 year old, as well as 10 dental specimens from squares C7 and C8 (6, 7). These dental specimens overlap in developmental age, suggesting that they represent one or possibly two individuals between 6 and 18 months old (7). The 28 newly identified specimens together with already described specimens may therefore represent the skeletal remains of a single infant.

Our study specifically aimed to provide molecular support that the hominin remains present in the Châtelperronian layers at the Grotte du Renne are Neandertals. Moreover, we test the hypothesis that these hominins derive from the underlying Mousterian (hypothesis **d**). This hypothesis must be rejected according to the chronological data presented here, and is in accordance with the spatial positioning of both the newly identified hominin specimens and the previously identified specimens, as these are located outside the sloping area of the Grotte du Renne (rows 1-6) or the areas that might have been affected by digging and levelling activities (12, 13, 46). We cannot contribute more substantially on the hypothesis that some cultural artefacts, in particular those interpreted as body ornaments and bone awls, in the Châtelperronian layers derive from the overlying Aurignacian (hypothesis **c**; (10, 11)). This has been contested elsewhere (14, 47). There is chronological evidence that at least some of these bone artefacts are relatively *in situ* (10), and archaeological arguments that large-scale displacement of lithic artefacts is unlikely (14). Similar cultural artefacts are also present in the Châtelperronian at Quinçay (4), although no hominin remains are present there. At Quinçay there are no later Upper Paleolithic levels overlying the Châtelperronian, which is partly sealed by roof collapse, and so artefact intrusion from overlying levels cannot account for the bone artefact association with a Châtelperronian lithic assemblage characterized technologically as Upper Palaeolithic (3). So, if these cultural artefacts are considered relatively *in situ* at either site then these belong within the Châtelperronian Neandertal cultural repertoire.

Our biomolecular data provides evidence that hominins contemporaneous with the Châtelperronian layers have archaic nuclear and Neandertal mitochondrial ancestry, supporting previous morphological studies (6, 7). They are therefore among some of the latest Neandertals in western Eurasia, and possible candidates to be involved in gene flow from Neandertals into AMHs (or *vice versa*; (48)). Future analysis of the nuclear genome of these or other Châtelperronian specimens might be able to provide further insights into the direction, extent and age of gene flow between Late Pleistocene western European Neandertals and “incoming” AMHs (49, 50).

Our results reveal that the bone proteome is a dynamic tissue, reflecting ontogeny during the early stages of bone development. This realization requires further investigation as this has



not been explored in-depth previously in the palaeoproteomic or the clinical literature (25, 26). We hypothesize that *COL10a1* or other cartilage-associated proteins could be retained in mineralized tissues that have reduced rates of remodelling (cranial sutures, epiphyseal plates), mineralized tissues that do not remodel (dentine), or bone regions that are in intimate contact with other tissue types (trabecular bone, articular surfaces). The excretion of *COL10a1* by (pre)hypertrophic chondrocytes implies this protein will be lost from the bone proteome during bone development and maintenance and may not be present routinely in adult fossil hominin specimens (51). The presence of such proteins in this study is consistent with the level of ontogenetic development of the bone specimen analysed, as observed through isotopic and morphological analyses. As a result of these observations, future quantitative (palaeo)proteomics might explain phenotypic differences observed between members of our genus and other hominids by studying protein composition and *in vivo* protein modification of ancient hominin specimens.

The identification of *COL10a1* and additional proteins predicted to carry derived amino acid substitutions specific for Late Pleistocene clades within the genus *Homo* (34) in the data presented here and elsewhere for non-hominin palaeoproteomes (25, 26) suggests that ancient proteins are a viable approach to study the taxonomic affiliation of Pleistocene fossil hominins, in particular when ancient DNA is poorly or not preserved. So, the analysis of ancient proteins provides a second biomolecular method capable of differentiating between Late Pleistocene clades within our genus when adequate error-tolerant search algorithms are used. Furthermore, the contextual analysis of these ancient proteins has the potential to provide *in vivo* details on ancient hominin ontogeny, physiology and phenotype.

## Methods:

We screened 196 bone specimens using ZooMS (Fig. S1) and taxonomically identified these using previously published and newly obtained ZooMS COL1 peptide marker masses (15, 21). Osteological analysis of bone specimens identified as homininae suggest that these specimens possibly represent an immature infant. LC-MS/MS analysis was conducted on three (Fig. S1) of the hominin bone specimens as published previously for non-hominin bone specimens (19), as well as two additional analyses of one palaeoproteomic extract generated following a modified protein extraction protocol (52). We took measures to avoid and detect protein contamination throughout our palaeoproteomic workflow. Ancient DNA analysis followed protocols outlined elsewhere (53, 54), as did stable isotope and radiocarbon analysis of collagen extracts (55). All biomolecular extractions were performed in dedicated facilities at the Max-Planck-Institute for Evolutionary Anthropology (Leipzig, Germany). Extended methods can be found online in the supplementary material. COL1 database samples can be accessed via ProteomeXchange with identifier PXD003190 and LC-MS/MS data for AR-7, AR-16 and AR-30 via PXD003208. The DNA sequences analyzed in this study were deposited in the European Nucleotide Archive under accession number PRJEB14504.

## Acknowledgements:

The authors thank S. van der Mije (Naturalis, NL), F. Boschini (University of Siena, Italy), G. Perry (State University, USA), P. Kosintsev (Russian Academy of Sciences, Ekaterinburg), J.-J. Cleyet-Merle (Musée National de Préhistoire, Les Eyzies, France) and A. Lister (NHM, London) for providing COL1 reference samples. A. Reiner, L. Klausnitzer, S. Steinbrenner, L. Westphal, B. Höber, B. Nickel, U. Stenzel and A. Weihmann are thanked for technical support. S.

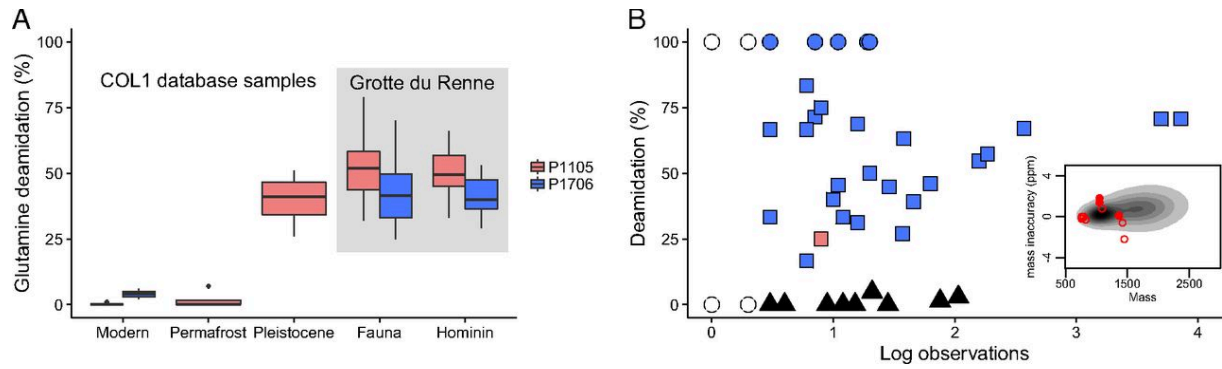
McPherron, M.P. Richards and our reviewers are thanked for their constructive comments, which greatly improved the paper. This research was supported by funding through the Max-Planck-Gesellschaft, European Research Council (ERC) Advanced Award CodeX, Deutsche Forschungsgemeinschaft (SFB1052, project A02) and Engineering and Physical Sciences Research Council NE/G012237/1.

## References:

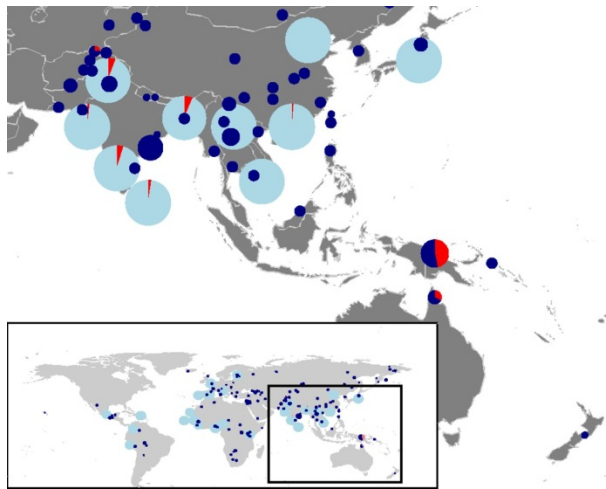
1. Green RE, et al. (2010) A draft sequence of the Neandertal genome. *Science* 328(5979):710–722.
2. Churchill SE, Smith FH (2000) Makers of the early Aurignacian of Europe. *Yearbook Phys Anthropol* 31:61–115.
3. Roussel M, Soressi M, Hublin J-J (2016) The Châtelperronian conundrum: Blade and bladelet lithic technologies from Quinçay, France. *J Hum Evol* 95:13–32.
4. Granger J-M, Lévêque F (1997) Parure castelperronienne et aurignacienne: étude de trois séries inédites de dents percées et comparaisons. *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science* 325(7):537–543.
5. d'Errico F, Julien M, Liolios D, Vanhaeren M, Baffier D (2003) Many awls in our argument. Bone tool manufacture and use in the Châtelperronian and Aurignacian levels of the Grotte du Renne at Arcy-sur-Cure. *The Chronology of the Aurignacian and of the Transitional Technocomplexes. Dating, Stratigraphies, Cultural Implications*, Trabalhos de arqueologia., eds Zilhão J, d'Errico F (Instituto Português de Arqueologia, Lisboa), pp 247–270.
6. Hublin J-J, Spoor F, Braun M, Zonneveld F, Condemi S (1996) A late Neanderthal associated with Upper Palaeolithic artefacts. *Nature* 381(6579):224–226.
7. Bailey SE, Hublin J-J (2006) Dental remains from the Grotte du Renne at Arcy-sur-Cure (Yonne). *J Hum Evol* 50(5):485–508.
8. Bailey SE, Weaver TD, Hublin J-J (2009) Who made the Aurignacian and other early Upper Paleolithic industries? *J Hum Evol* 57(1):11–26.
9. Hublin J-J (2015) The modern human colonization of western Eurasia: when and where? *Quat Sci Rev* 118:194–210.
10. Higham T, et al. (2010) Chronology of the Grotte du Renne (France) and implications for the context of ornaments and human remains within the Châtelperronian. *Proc Natl Acad Sci USA* 107(47):20234–20239.
11. White R (2001) Personal ornaments from the Grotte du Renne at Arcy-sur-Cure. *Athena Review* 2(4):41–46.
12. Bar-Yosef O, Bordes J-G (2010) Who were the makers of the Châtelperronian culture? *J Hum Evol* 59(5):586–593.
13. Hublin J-J, et al. (2012) Radiocarbon dates from the Grotte du Renne and Saint-Césaire support a Neandertal origin for the Châtelperronian. *Proc Natl Acad Sci USA* 109(46):18743–18748.
14. Caron F, d'Errico F, Del Moral P, Santos F, Zilhão J (2011) The reality of Neandertal symbolic behavior at the Grotte du Renne, Arcy-sur-Cure, France. *PLoS One* 6(6):e21545.
15. Buckley M, Collins MJ, Thomas-Oates J, Wilson JC (2009) Species identification by analysis of bone collagen using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 23(23):3843–3854.
16. Welker F, Soressi M, Rendu W, Hublin J-J, Collins MJ (2015) Using ZooMS to identify fragmentary bone from the Late Middle/Early Upper Palaeolithic sequence of Les Cottés, France. *J Archaeol Sci* 54:279–286.
17. Brown S, et al. (2016) Identification of a new hominin bone from Denisova Cave, Siberia using collagen fingerprinting and mitochondrial DNA analysis. *Sci Rep* 6:23559.
18. Nielsen-Marsh CM, et al. (2005) Osteocalcin protein sequences of Neanderthals and modern primates. *Proc Natl Acad Sci USA* 102(12):4409–4413.

19. Welker F, et al. (2015) Ancient proteins resolve the evolutionary history of Darwin's South American ungulates. *Nature* 522(7554):81–84.
20. van Doorn NL, Hollund H, Collins MJ (2011) A novel and non-destructive approach for ZooMS analysis: ammonium bicarbonate buffer extraction. *Archaeol Anthropol Sci* 3(3):281–289.
21. Buckley M, Kansa SW (2011) Collagen fingerprinting of archaeological bone and teeth remains from Domuztepe, South Eastern Turkey. *Archaeol Anthropol Sci* 3(3):271–280.
22. Evans S, et al. (2015) Using combined biomolecular methods to explore whale exploitation and social aggregation in hunter–gatherer–fisher society in Tierra del Fuego. *J Archaeol Sci: Reports*. 6: 757–767.
23. Robinson NE, et al. (2004) Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides. *J Pept Res* 63(5):426–436.
24. van Doorn NL, Wilson J, Hollund H, Soressi M, Collins MJ (2012) Site-specific deamidation of glutamine: a new marker of bone collagen deterioration. *Rapid Commun Mass Spectrom* 26(19):2319–2327.
25. Cappellini E, et al. (2012) Proteomic analysis of a pleistocene mammoth femur reveals more than one hundred ancient bone proteins. *J Proteome Res* 11(2):917–926.
26. Wadsworth C, Buckley M (2014) Proteome degradation in fossils: investigating the longevity of protein survival in ancient bone. *Rapid Commun Mass Spectrom* 28(6):605–615.
27. Hjørten R, et al. (2007) Type XXVII collagen at the transition of cartilage to bone during skeletogenesis. *Bone* 41(4):535–542.
28. Long F (2012) Building strong bones: molecular regulation of the osteoblast lineage. *Nat Rev Mol Cell Biol* 13(1):27–38.
29. Matsuo N, et al. (2008) Collagen XXIV (Col24a1) gene expression is a specific marker of osteoblast differentiation and bone formation. *Connect Tissue Res* 49(2):68–75.
30. Linsenmayer TF, Eavey RD, Schmid TM (1988) Type X collagen: a hypertrophic cartilage-specific molecule. *Pathol Immunopathol Res* 7(1-2):14–19.
31. Coussens AK, et al. (2007) Unravelling the molecular control of calvarial suture fusion in children with craniosynostosis. *BMC Genomics* 8:458.
32. 1000 Genomes Project Consortium, et al. (2015) A global reference for human genetic variation. *Nature* 526(7571):68–74.
33. Meyer M, et al. (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338(6104):222–226.
34. Castellano S, et al. (2014) Patterns of coding variation in the complete exomes of three Neandertals. *Proc Natl Acad Sci USA* 111(18):6666–6671.
35. Warinner C, et al. (2014) Pathogens and host immunity in the ancient human oral cavity. *Nat Genet* 46(4):336–344.
36. Kendall R, Hendy J, Collins MJ, Millard AR, Gowland RL (2016) Poor preservation of antibodies in archaeological human bone and dentine. *STAR: Science & Technology of Archaeological Research* 2(1):15–24.
37. Robinson NE, Robinson A (2004) *Molecular clocks: deamidation of asparaginyl and glutaminyl residues in peptides and proteins* (Althouse Press).
38. Fu Q, et al. (2014) Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* 514(7523):445–449.
39. Prüfer K, et al. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505(7481):43–49.
40. Kong A, et al. (2002) A high-resolution recombination map of the human genome. *Nat Genet* 31(3):241–247.
41. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7(3). doi:10.1371/journal.pone.0034131.

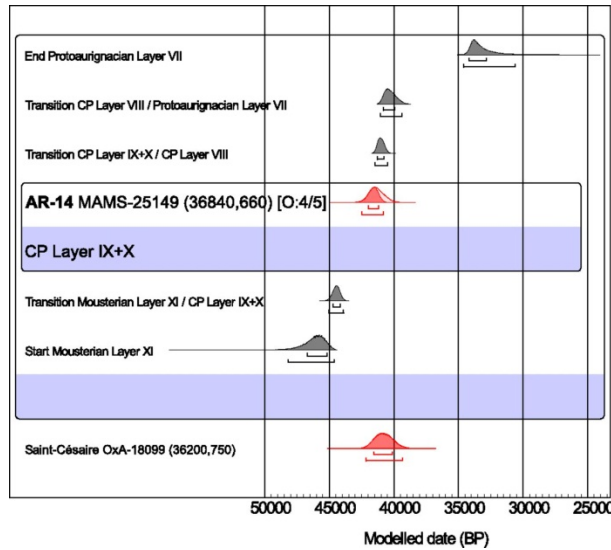
42. Richards MP, Mays S, Fuller BT (2002) Stable carbon and nitrogen isotope values of bone and teeth reflect weaning age at the Medieval Wharram Percy site, Yorkshire, UK. *Am J Phys Anthropol* 119(3):205–210.
43. Discamps E, Gravina B, Teyssandier N (2015) In the eye of the beholder: contextual issues for Bayesian modelling at the Middle-to-Upper Palaeolithic transition. *World Archaeol* 47(4):601–621.
44. Higham T, et al. (2014) The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* 512(7514):306–309.
45. Bodu P (1990) L'application de la méthode des remontages à l'étude du matériel lithique des premiers niveaux châtelperroniens d'Arcy-sur-Cure. *Paléolithique Moyen Récent et Paléolithique Supérieur Ancien En Europe. Ruptures et Transitions : Examen Critique Des Documents Archéologiques.*, ed Farzy C (Mémoires du Musée de Préhistoire d'Île de France), pp 309–312.
46. Movius HL (1969) The Châtelperronian in French Archaeology: the Evidence of Arcy-sur-Cure. *Antiquity* 43(170):111–123.
47. Zilhão J, d'Errico F, Julien M, David F (2011) Chronology of the site of Grotte du Renne, Arcy-sur-Cure, France: Implications for radiocarbon dating. *Before Farming* 2011(3):1–14.
48. Ruebens K, McPherron SJP, Hublin J-J (2015) On the local Mousterian origin of the Châtelperronian: Integrating typo-technological, chronostratigraphic and contextual data. *J Hum Evol* 86:55–91.
49. Fu Q, et al. (2015) An early modern human from Romania with a recent Neanderthal ancestor. *Nature* 524(7564):216–219.
50. Kuhlwilm M, et al. (2016) Ancient gene flow from early modern humans into Eastern Neanderthals. *Nature* 530(7591):429–433.
51. Hedges REM, Clement JG, Thomas CDL, O'Connell TC (2007) Collagen turnover in the adult femoral mid-shaft: Modeled from anthropogenic radiocarbon tracer measurements. *Am J Phys Anthropol* 133(2):808–816.
52. Orlando L, et al. (2013) Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499(7456):74–78.
53. Gansauge M-T, Meyer M (2014) Selective enrichment of damaged DNA molecules for ancient genome sequencing. *Genome Res* 24(9):1543–1549.
54. Korlević P, et al. (2015) Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *Biotechniques* 59(2):87–93.
55. Talamo S, Richards M (2011) A comparison of bone pretreatment methods for AMS dating of samples > 30,000 BP. *Radiocarbon* 53(3):443–449.



**Figure 1. Identification of damaged ancient proteins.** a) Deamidation values of bone ammonium-bicarbonate extracts for database samples and Arcy bone specimens analyzed using MALDI-TOF-MS. b) Deamidation frequency based on spectral counts obtained through LC-MS/MS analysis. Cluster analysis provides 3 clusters (filled triangles, squares and circles). Open circles represent proteins that have 2 or less spectral matches and were not included in cluster analysis. *COL10a1* is indicated in pink. 0% indicates no deamidation while 100% indicates complete deamidation. Inset: AR30 ppm error distribution by peptide mass for assigned spectra, with spectra matching to *COL10a1* in red open circles and those to *COL10a1* 128N in red filled circles.



**Figure 2. South-East Asian frequency of the archaic-like allele of rs142463796 in modern human populations.** The frequency of the archaic-like allele in modern human populations is displayed in red (Yoruba allele frequency in populations shown in light blue for 1000 Genomes data set and dark blue for SGDP data set). The diameter is proportional to the number of individuals in a given population which ranges between one and 16 individuals in the SGDP and between 66 and 113 in the 1000 Genomes phase 3 data.



**Figure 3. Bayesian model of radiocarbon dates for the Grotte du Renne.** Model constructed including all radiocarbon dates reported in (13). Archaeological layer boundaries are shown, with the direct dates on the Grotte du Renne (AR-14) and the Saint-Césaire hominins highlighted in red. Posterior/prior outlier probability for AR-14 is shown in square parenthesis (more details in the SI Appendix Section 6).

**Table 1. Palaeoproteomic summary statistics per analyzed bone specimen.**

Variable	Blank	AR-7	AR-16	AR-30 (ZooMS)	AR-30A	AR-30B
# of MS/MS scans	732	43,073	20,564	11,294	16,923	20,634
# of matching MS/MS scans	19 (2.6%)	7,035 (16.3%)	4,053 (19.7%)	2,232 (19.8%)	3,416 (20.2%)	3,769 (18.3%)
# of protein groups	7	42	30	24	39	40
of which endogenous	0	20	21	11	23	26
of which status unknown	0	10	6	8	12	8
of which exogenous	7	12	3	5	4	6
# of NCPs	-	7	7	7	13	12
# of unique protein groups	-	4	2	1	4	2

NCP= non-collagenous protein.