nature

https://doi.org/10.1038/s41586-020-2509-0

Supplementary information

Evidence of human occupation in Mexico around the Last Glacial Maximum

In the format provided by the authors and unedited

Ciprian F. Ardelean✉, Lorena Becerra-Valdivia, Mikkel Winther Pedersen, Jean-Luc Schwenninger, Charles G. Oviatt, Juan I. Macías-Quintero, Joaquin Arroyo-Cabrales, Martin Sikora, Yam Zul E. Ocampo-Díaz, Igor I. Rubio-Cisneros, Jennifer G. Watling, Vanda B. de Medeiros, Paulo E. De Oliveira, Luis Barba-Pingarón, Agustín Ortiz-Butrón, Jorge Blancas-Vázquez, Irán Rivera-González, Corina Solís-Rosales, María Rodríguez-Ceja, Devlin A. Gandy, Zamara Navarro-Gutierrez, Jesús J. De La Rosa-Díaz, Vladimir Huerta-Arellano, Marco B. Marroquín-Fernández, L. Martin Martínez-Riojas, Alejandro López-Jiménez, Thomas Higham & Eske Willerslev✉

SUPPLEMENTARY INFORMATION FILE SI1

Evidence of human occupation in Mexico around the Last Glacial Maximum

Ciprian F. Ardelean[1,2*], Lorena Becerra-Valdivia[3,4], Mikkel Winther Pedersen[5], Jean-Luc Schwenninger[3], Charles G. Oviatt[6], Juan I. Macías-Quintero[7], Joaquin Arroyo-Cabrales[8], Martin Sikora[5], Yam Zul E. Ocampo-Díaz[9,10], Igor I. Rubio-Cisneros[10], Jennifer G. Watling[11], Vanda B. de Medeiros[12], Paulo E. De Oliveira[12,13], Luis Barba-Pingarón[14], Agustín Ortiz-Butrón[14], Jorge Blancas-Vázquez[14], Irán Rivera-González[15], Corina Solís-Rosales[16], María Rodríguez-Ceja[16], Devlin A. Gandy[17], Zamara Navarro-Gutierrez[1], Jesús J. De La Rosa-Díaz[1], Vladimir Huerta-Arellano[1], Marco B. Marroquín-Fernández[7], L. Martin Martínez-Riojas[1], Alejandro López-Jiménez[8], Thomas Higham[3], and Eske Willerslev[5,18,19, 20**].

[1]Unidad Académica de Antropología, Universidad Autónoma de Zacatecas.

[2]Department of Archaeology, University of Exeter.

[3]Research Laboratory for Archaeology and History of Art, University of Oxford

[4]Chronos 14C-Cycle Facility, University of New South Wales

[5]Lundbeck Foundation GeoGenetics Centre, Globe Institute, University of Copenhagen.

[6]Department of Geology, Kansas State University.

[7]Escuela de Arqueología, Universidad de Ciencias y Artes de Chiapas.

[8]Laboratorio de Arqueozoología, Subdirección de Laboratorios y Apoyo Académico, Instituto Nacional de Antropología e Historia.

[9]Facultad de Ingeniería, Universidad Autónoma de San Luís Potosí.

[10]Grupo de Geología Exógena y del Sedimentario, México, freelance.

[11]Laboratório de Arqueologia dos Trópicos, Museu de Arqueologia e Etnologia, Universidade de São Paulo.

[12]Laboratório de Micropaleontologia, Instituto de Geociências, Universidade de São Paulo.

[13]Botany Department, The Field Museum of Natural History.

[14]Laboratorio de Prospección Arqueológica, Instituto de Investigaciones Antropológicas (IIA), Universidad Nacional Autónoma de México (UNAM).

[15]Laboratorio de Palinología, Escuela Nacional de Antropología e Historia (ENAH).

[16]Laboratorio de Espectrometría de Masas con Aceleradores, Instituto de Física, Universidad Nacional Autónoma de México.

[17]Department of Archaeology, University of Cambridge.

[18]Welcome Trust, Sanger Institute.

[19]The Danish Institute for Advanced Study at The University of Southern Denmark

[20]Lundbeck Foundation GeoGenetics Centre, Department of Zoology, University of Cambridge.

*e-mail: aeci000206@uaz.edu.mx

**e-mail: ewillerslev@sund.ku.dk

Table of Contents

1. Geology

1.1. Field Methods

The bedrock geology and geomorphology of the landscape surrounding Cueva del Chiquihuite, and the cave-floor deposits, which had been previously partially excavated, were examined and interpreted in November, 2017, by C.G. Oviatt. Observations of the bedrock at the cave were quantified by measurements of bedding attitude using a Brunton compass, and major faults, masses of brecciated bedrock and associated structural features, were recorded in field notes. Within the cave, the bedrock bedding attitude and likely presence of faults were noted, as well as the occurrence of roof-fall blocks and accumulations of finer-grained gravel and dust, speleothems, and drip- or flow-stone. Dissolution cavities and tunnels were noted. Cave-floor deposits were primarily examined where they were exposed in the walls of the archaeological excavations. Macroscopic visual observations were supplemented by close-up examinations of small samples plucked from the excavation walls using a hand lens and a Dino-Lite digital microscope. Alternating beds of "muddy" and "gravelly" sediments were apparent in the exposures. All units are dominated in grain size by particles of sand size or larger (mostly gravel), but the fine-grained or "muddy" character of some units was visually striking in exposed profiles. In this context, the word "mud" is used in a relative sense, in contrast to the coarse gravels, and refers to a mixture of sand, silt, and clay (see below the results of grain size analyses). All surfaces within the cave, including the surfaces of roof-fall blocks and all other debris that has been washed or carried into the cave, are covered with dust (see below the results of analyses of this dust).

1.2. Laboratory Methods

Samples were collected at Cueva del Chiquihuite, transported to the United States, and analyzed in various ways (see below, 1.3).

Table S1. Samples collected at Cueva del Chiquihuite by C.G. Oviatt in November, 2017.

sample	material	amount	location	date of collection
1209	sediment (gravel)	small zip-lock bag	excavation wall I 5/6	11/2/2017
1210	sediment (very top of 1212)	small zip-lock bag	excavation wall I 5/6	11/2/2017
1212	sediment (muddy gravel)	small zip-lock bag	excavation wall I 5/6	11/2/2017
1217	sediment (muddy gravel)	small zip-lock bag	excavation wall I 5/6	11/2/2017

sample	material	amount	location	date of collection
1219	sediment (gravel)	small zip-lock bag	excavation wall I 5/6	11/2/2017
unlabeled	limestone bedrock	large pebble	outside mouth of cave	11/2/2017
dust	sediment (dust)	about 5 g in a small zip-lock bag	brushed from the surface of a roof-fall boulder inside the cave	11/2/2017

1.3. Analyses

1.3.1. Acid-insoluble fraction

Small subsamples were passed through a 180-micron sieve to remove larger particles. The sediment that passed the 180-micron sieve was then dissolved in dilute (10%) hydrochloric acid, passed through a 63-micron sieve, and the sediment caught on the sieve (mostly fine-sand size) was examined under a binocular microscope at 40 power. For the limestone sample, about 7 g of the limestone were dissolved in kitchen vinegar, then passed through the 63-micron sieve; the particles remaining on the 63-micron sieve were examined under the binocular microscope. This was done in November and December, 2017.

Table S2. The acid-insoluble fraction from samples collected.

sample	material dissolved	amount dissolved	visibly identifiable materials in the acid-insoluble residue
1209	fine-grained sediment adhering to the surfaces of gravel particles	about 1 g	quartz, muscovite, minor dark mineral grains
1210	fine-grained sediment	about 1 g	quartz, muscovite, minor dark mineral grains
1212	fine-grained sediment	about 1 g	quartz, muscovite, minor dark mineral grains
1217	fine-grained sediment	about 1 g	quartz, muscovite, minor dark mineral grains
1219	fine-grained sediment adhering to the surfaces of gravel particles	about 1 g	quartz, muscovite, minor dark mineral grains

sample	material dissolved	amount dissolved	visibly identifiable materials in the acid-insoluble residue
unlabeled	limestone bedrock	about 7 g	quartz, muscovite, minor dark mineral grains, dark amorphous material (looks organic)
dust	dust collected from the surface of a roof-fall block within the cave	about 1 g	quartz, muscovite, minor dark mineral grains, fragments of organic materials such as tiny twigs and insect parts

1.3.2. Coatings on Non-calcite grains

Two samples were examined to see if some calcium carbonate (calcite) was stuck to the grains of quartz and other mineral grains in the fine fraction of "muddy" units and in broken limestone. The samples were prepared by dispersing the mud fraction (silt and clay) in hot water and baking soda (no acid was used), then removing the mud fraction from the sample by allowing the mixture to settle for at least 30 seconds in quiet water, then pouring off the suspended grains (the silt and clay) with the excess water. The sample of sand-size and larger grains was then examined under a binocular microscope under 40x. The samples consisted of muddy sediment in stratum 1210, and limestone (bedrock) were broken with a hammer on an anvil. In both cases, the non-calcite grains were clean, and no carbonate or other residue was seen adhering to the grains. In other words, sand-size quartz and other mineral grains within the limestone are released on impact without a coating of carbonate. The muddy sediment in the stratigraphic sections in the excavations contains non-calcite mineral grains that do not have carbonate coatings.

1.3.3. Grain-size analysis

The percentage of sand, silt, and clay in sediment samples finer than 2 mm in stratigraphic units (UE) 1212 and 1217 (that is, the "muddy" fractions of these samples) was determined using pipette-analysis techniques, except that carbonates were not removed prior to analysis, supervised by James (Bruce) Harrison, soil scientist at New Mexico Institute of Mining and Technology, Socorro, NM, USA. Both samples (1212 and 1217) contain many fragments larger than 2 mm in diameter but these were not measured in the grain-size analyses.

Table S3. Grain-size results.

sample	% clay	% silt	%sand
1212	0.94	1.69	97.37
1217	0.76	0.28	98.96

The averages for these two samples are ~1% clay, ~1% silt, and ~98% sand. Note that because carbonates were not removed prior to the pipette analyses, and because of the possibility of flocculation of clay-size carbonate grains, the percentage of sand in the samples may be higher and the percentage of clay may be lower than if the carbonates had first been removed from the samples.

The sample of dust was too small to analyze for grain size; it contained abundant organic materials.

1.3.4. X-ray diffraction

Three sediment samples and one limestone sample were submitted to Kelsey McNamara at the X-ray diffraction laboratory, New Mexico Bureau of Geology and Mineral Resources in Socorro, NM, USA. Sediment samples for X-ray diffraction were prepared by first passing them through a 180-micron sieve to remove larger particles, then the fine-grained material was ground to a fine powder using a hand-held mortar and pestle. About 1 g of limestone was crushed and powdered using a mortar and pestle. The diffraction patterns for illite (a clay mineral) and muscovite (a mica mineral) are essentially the same and the two minerals cannot be distinguished by X-ray diffraction. Although it is possible that illite is present in some or all of the samples, muscovite is definitely present − it was observed in the insoluble fractions of some samples (see above).

Table S4. Samples submitted for X-ray diffraction

sample	**material X-rayed**	**results with or without illite (muscovite)**	**minerals listed in order of relative abundance; mineral percentages in parentheses**
1209	"muddy" sediment adhering to the surfaces of gravel particles	without	calcite (93) quartz (7)
1209	"muddy" sediment adhering to the surfaces of gravel particles	with	calcite (73.7) illite (muscovite) (20.2) quartz (6.1)
1212	"muddy" sediment	without	calcite (93) quartz (7)
1212	"muddy" sediment	with	calcite (76.2) illite (muscovite) (17.9) quartz (5.9)
dust	fine-grained sediment	without	calcite (67.3) quartz (19.8) gypsum (12.9)

sample	material X-rayed	results with or without illite (muscovite)	minerals listed in order of relative abundance; mineral percentages in parentheses
dust	fine-grained sediment	with	illite (muscovite) (54) calcite (31) quartz (10) gypsum (5)
limestone	powdered limestone	N/A	magnesium calcite (83) calcite (15) quartz (2)

2. Bayesian age modeling

2.1. Sensitivity testing

To determine the reproducibility of the Bayesian model in the main text ('Model A', Fig. 2 and Fig. S1), we ran the model multiple times and applied sensitivity testing. This Bayesian age model was compared against a series of different models with slightly different data included. For the first model test, we ran a Bayesian model with the excluded bulk sediment (Beta 436709, LEMA 575.1.2. and ICA-16OS/0510; which were deemed unreliable, minimum-age estimates) and bone collagen dates (LEMA-640.1.1 fails collagen quality control values set by the ORAU[1]; 'Model B'; Fig. S1b), including strata within SC-B ('Model C'; Fig. S2), and with the application of a 'Charcoal' instead of 'General' outlier model[2] for all charcoal samples ('Model D'; Fig. S3). Model D was run with SC-C and -B as separate sequences.

We note low convergence values (C) are present in some parts of the model and these are usually associated with parameters that have high outlier probabilities. High convergence values are an indication that the MCMC sampling is able to find a solution and obtain values that are stable or converge. Usually C values should be >95. The solution for poor convergence is for OxCal to continue to run the models until convergence is satisfactory. We ran our models for 4,224,000 iterations for this reason. Some parts of the model in particular disclose poor convergence; we note the section between the start of stratum 1217 and the beginning of SC-B where values are low due to, principally, few determinations over a wide period coupled with some variability in the results. To test whether this led to significant variations between the favoured model and the different models outlined above, we used the 'Difference' function within OxCal[3,4] and focused on the start boundaries for SC-C and SC-B (Table S5). Given that these overlap zero at 95.4% probability (Fig. S4), the exclusion of strata within SC-B, the manual removal of sediment dates and LEMA-640.1.1, and the consistent use of the 'General' outlier analysis have no significant impact on the modelled output. CQL code for the four models can be found below.

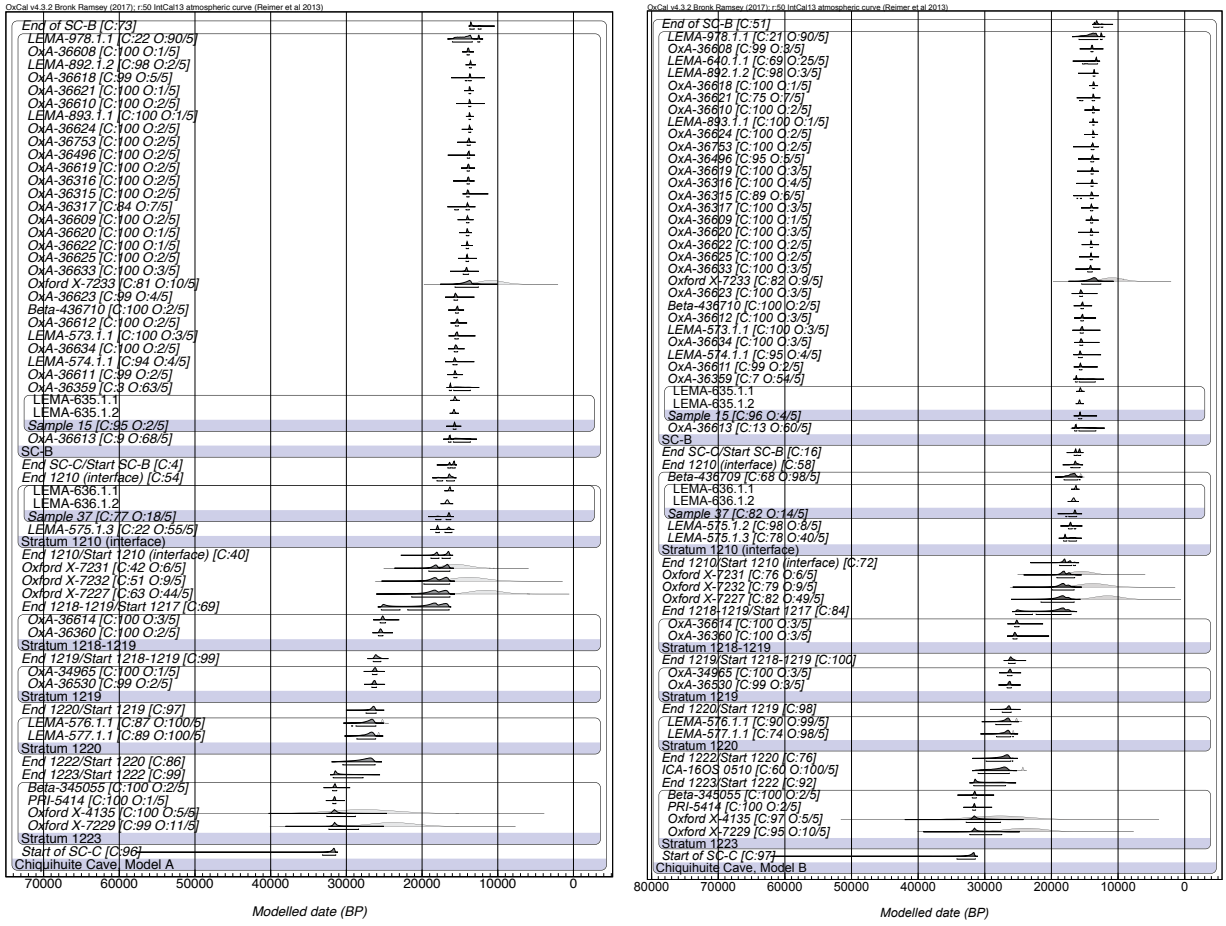


Fig. S1. Bayesian age 'Model A' (left) and 'B' (right). The latter is identical to 'Model A', but with LEMA 640.1.1 and the three sediment dates (Beta 436709, LEMA 575.1.2., and ICA-160S/0510). Brackets beneath each age estimate show 95.4% confidence interval. 'C' denotes convergence values, whilst 'O:prior/posterior probability' reflects the outlier analysis. Start boundary estimates for SC-C and -B are in Table S5.

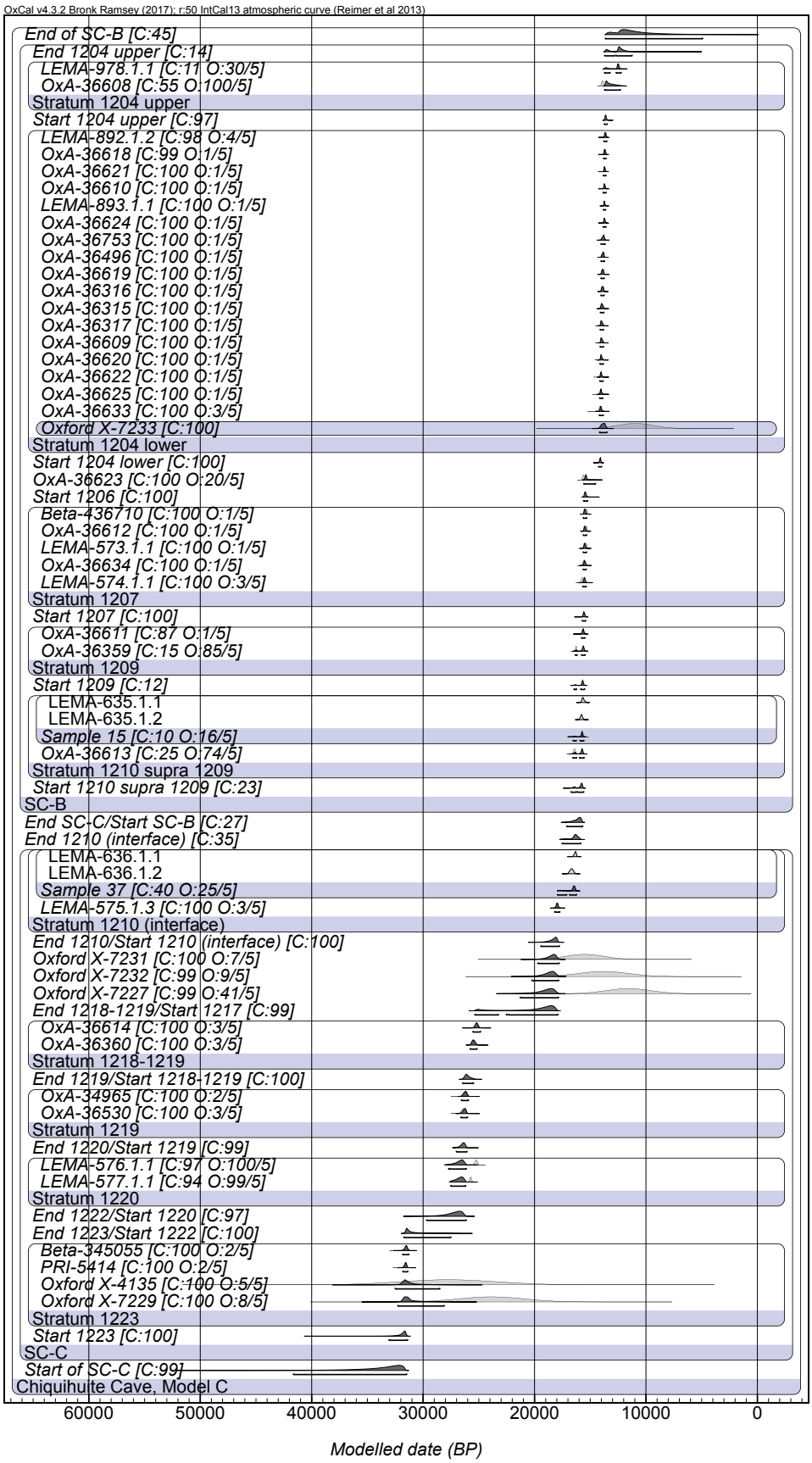


Fig. S2. Bayesian age 'Model C'. This is identical to 'Model A' (see main Fig. 2 and Fig. S1), but with strata in SC-B. Brackets beneath each age estimate show 95.4% confidence interval. 'C' denotes convergence values, whilst 'O:prior/posterior probability' reflects the outlier analysis. Start boundary estimates for SC-C and -B are in Table S5.

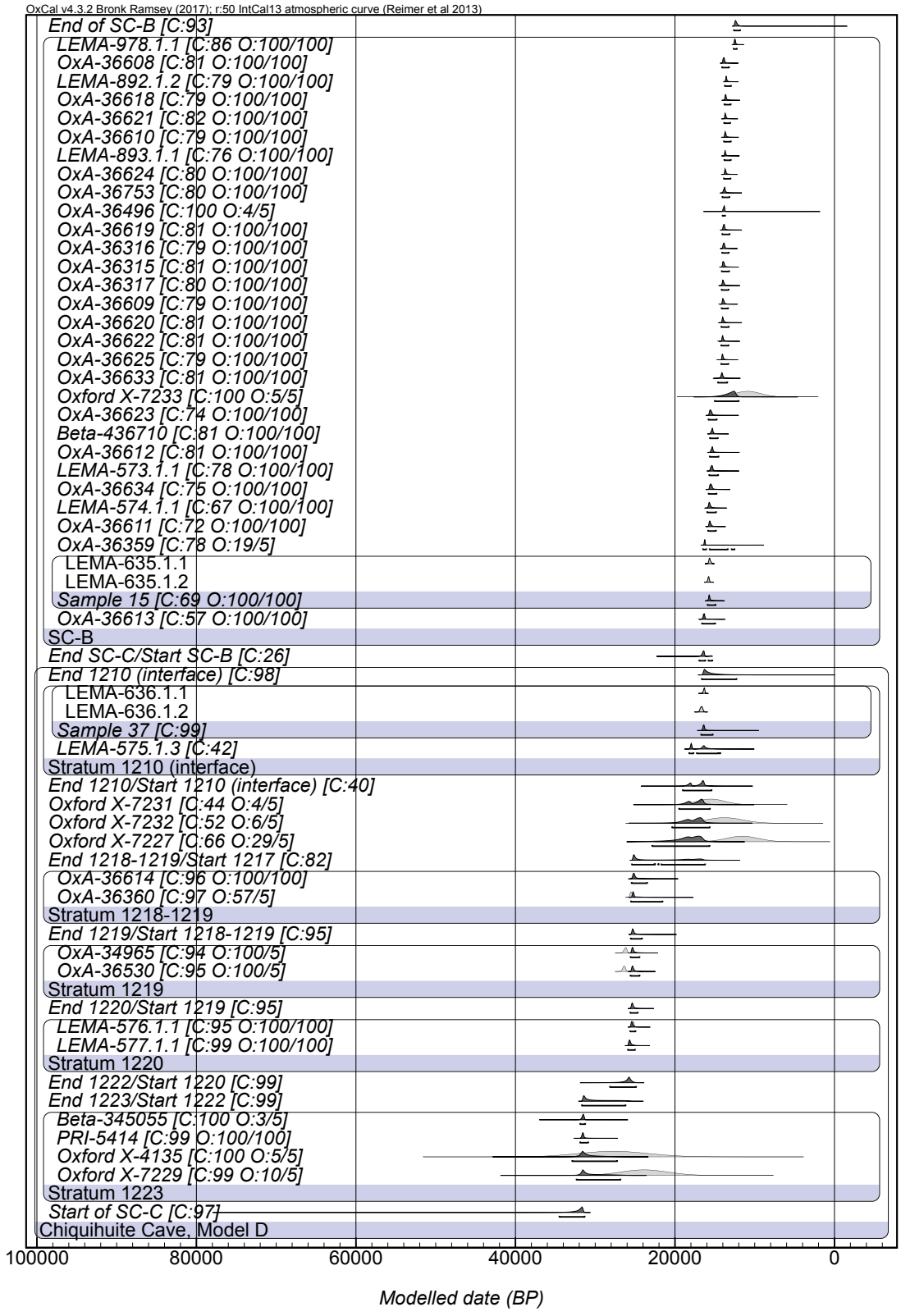


Fig. S3. Bayesian age 'Model D'. This model is identical to 'Model A' (see main Fig. 2 and Fig. S1), but with 'Charcoal' instead of 'General' outlier analysis[2] for all charcoal samples. Brackets beneath each age estimate show 95.4% confidence interval. 'C' denotes convergence values, whilst 'O:prior/posterior probability' reflects the outlier analysis. Start boundary estimates for SC-C and -B are in Table S5.

Table S5 (below).Bayesian age model output (start of SC-C/-B and end of SC-B) for models A (Fig. 2 and Fig. S1a), B (Model A with four excluded dates), C (Model A with strata in SC-B), and D (Model A with 'Charcoal' instead of 'General' outlier analysis for all charcoal samples).

Output	**Model**	**years cal. BP (95.4% confidence)**
Start of SC-C	A	33,150-31,405
	B	34,190-31,405
	C	41,635-31,470
	D	34,495-31,315
Start of SC-B	A	16,605-15,615
	B	16,605-15,615
	C	17,110-15,675
	D	16,955-15,310
End of SC-B	A	13,705-12,200
	B	13,675-12,280
	C	13,635-4,915
	D	12,615-11,825

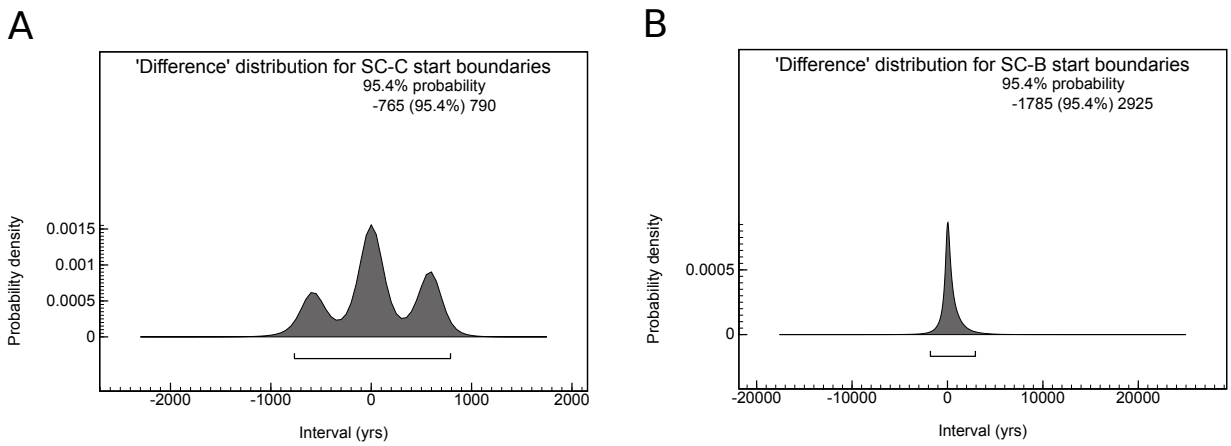


Fig. S4. **a**. Probability density function (PDF) for the difference between the start of SC-C in Model A, compared with that of Models B-D. **b**. PDF for the difference between the start of SC-B in Model A, compared with that of Models B-D. These results suggest that there is no significant difference between the modelled outputs, as the distributions overlap zero at 95.4% probability.

2.2. CQL code

Model A (main text, Fig. 3)

```
Options()
{
 Resolution=50;
};
Plot()
{
 Outlier_Model("General",T(5),U(0,4),"t");
 Outlier_Model("SSimple",N(0,2),0,"s");
 Sequence("Chiquihuite Cave, Model A")
 {
  Boundary("Start of SC-C");
  Phase("Stratum 1223")
```

```
{
 Date("Oxford X-7229", N(2017-23940, 2950))
 {
  Outlier("General", 0.05);
 };
 Date("Oxford X-4135", N(2017-27790, 4340))
 {
  Outlier("General", 0.05);
 };
 R_Date("PRI-5414", 27929, 82)
 {
  Outlier("General", 0.05);
 };
 R_Date("Beta-345055", 27830, 150)
 {
  Outlier("General", 0.05);
 };
};
Boundary("End 1223/Start 1222");
Boundary("End 1222/Start 1220");
Phase("Stratum 1220")
{
 R_Date("LEMA-577.1.1", 21401, 95)
 {
  Outlier("General", 0.05);
 };
 R_Date("LEMA-576.1.1", 20896, 80)
 {
  Outlier("General", 0.05);
 };
};
Boundary("End 1220/Start 1219");
Phase("Stratum 1219")
{
 R_Date("OxA-36530", 22170, 140)
 {
  Outlier("General", 0.05);
 };
 R_Date("OxA-34965", 21990, 170)
 {
  Outlier("General", 0.05);
 };
};
Boundary("End 1219/Start 1218-1219");
Phase("Stratum 1218-1219")
{
 R_Date("OxA-36360", 21140, 130)
 {
  Outlier("General", 0.05);
 };
 R_Date("OxA-36614", 20860, 100)
 {
  Outlier("General", 0.05);
 };
};
Boundary("End 1218-1219/Start 1217");
Date("Oxford X-7227", N(2017-11620, 2000))
{
 Outlier("General", 0.05);
};
Date("Oxford X-7232", N(2017-13870, 2250))
{
 Outlier("General", 0.05);
};
Date("Oxford X-7231", N(2017-15560, 1740))
{
 Outlier("General", 0.05);
};
Boundary("End 1210/Start 1210 (interface)");
Phase("Stratum 1210 (interface)")
{
 R_Date("LEMA-575.1.3", 14778, 77)
 {
  Outlier("General", 0.05);
 };
 R_Combine("Sample 37")
 {
  Outlier("General", 0.05);
  R_Date("LEMA-636.1.2", 13788, 90)
  {
   Outlier("SSimple", 0.05);
  };
  R_Date("LEMA-636.1.1", 13569, 60)
  {
   Outlier("SSimple", 0.05);
  };
 };
};
Boundary("End 1210 (interface)");
Boundary("End SC-C/Start SC-B");
Phase("SC-B")
{
 R_Date("OxA-36613", 13630, 55)
 {
  Outlier("General", 0.05);
 };
 R_Combine("Sample 15")
 {
  Outlier("General", 0.05);
  R_Date("LEMA-635.1.2", 13142, 60)
  {
   Outlier("SSimple", 0.05);
  };
  R_Date("LEMA-635.1.1", 13054, 60)
  {
   Outlier("SSimple", 0.05);
  };
 };
 R_Date("OxA-36359", 13525, 35)
 {
  Outlier("General", 0.05);
 };
 R_Date("OxA-36611", 13050, 50)
 {
  Outlier("General", 0.05);
 };
 R_Date("LEMA-574.1.1", 13092, 63)
 {
  Outlier("General", 0.05);
 };
```

```
 R_Date("OxA-36634", 12990, 55)
 {
 Outlier("General", 0.05);
 };
 R_Date("LEMA-573.1.1", 12916, 58)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36612", 12885, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("Beta-436710", 12880, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36623", 13010, 55)
 {
 Outlier("General", 0.05);
 };
 Date("Oxford X-7233", N(2017-10960, 1610))
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36633", 12235, 75)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36625", 12170, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36622", 12155, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36620", 12140, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36609", 12120, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36317", 12120, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36315", 12095, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36316", 12050, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36619", 12040, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36496", 12005, 55)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36753", 11975, 70)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36624", 11900, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("LEMA-893.1.1", 11897, 35)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36610", 11895, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36621", 11890, 45)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36618", 11855, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("LEMA-892.1.2", 11770, 35)
 {
 Outlier("General", 0.05);
 };
 R_Date("OxA-36608", 12060, 50)
 {
 Outlier("General", 0.05);
 };
 R_Date("LEMA-978.1.1", 10513, 50)
 {
 Outlier("General", 0.05);
 };
 };
 Boundary("End of SC-B");
 Span("Span of sequence");
 };
};
```

Model B (Fig. S1)

```
Options()
{
 Resolution=50;
};
Plot()
{
```

```
Outlier_Model("General",T(5),U(0,4),"t");
Outlier_Model("SSimple",N(0,2),0,"s");
Sequence("Chiquihuite Cave, Model B")
{
 Boundary("Start of SC-C");
 Phase("Stratum 1223")
 {
  Date("Oxford X-7229", N(2017-23940, 2950))
  {
   Outlier("General", 0.05);
  };
  Date("Oxford X-4135", N(2017-27790, 4340))
  {
   Outlier("General", 0.05);
  };
  R_Date("PRI-5414", 27929, 82)
  {
   Outlier("General", 0.05);
  };
  R_Date("Beta-345055", 27830, 150)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("End 1223/Start 1222");
 R_Date("ICA-16OS 0510", 20220, 80)
 {
  Outlier("General", 0.05);
 };
 Boundary("End 1222/Start 1220");
 Phase("Stratum 1220")
 {
  R_Date("LEMA-577.1.1", 21401, 95)
  {
   Outlier("General", 0.05);
  };
  R_Date("LEMA-576.1.1", 20896, 80)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("End 1220/Start 1219");
 Phase("Stratum 1219")
 {
  R_Date("OxA-36530", 22170, 140)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-34965", 21990, 170)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("End 1219/Start 1218-1219");
 Phase("Stratum 1218-1219")
 {
  R_Date("OxA-36360", 21140, 130)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36614", 20860, 100)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("End 1218-1219/Start 1217");
 Date("Oxford X-7227", N(2017-11620, 2000))
 {
  Outlier("General", 0.05);
 };
 Date("Oxford X-7232", N(2017-13870, 2250))
 {
  Outlier("General", 0.05);
 };
 Date("Oxford X-7231", N(2017-15560, 1740))
 {
  Outlier("General", 0.05);
 };
 Boundary("End 1210/Start 1210 (interface)");
 Phase("Stratum 1210 (interface)")
 {
  R_Date("LEMA-575.1.3", 14778, 77)
  {
   Outlier("General", 0.05);
  };
  R_Date("LEMA-575.1.2", 14107, 64)
  {
   Outlier("General", 0.05);
  };
  R_Combine("Sample 37")
  {
   Outlier("General", 0.05);
   R_Date("LEMA-636.1.2", 13788, 90)
   {
    Outlier("SSimple", 0.05);
   };
   R_Date("LEMA-636.1.1", 13569, 60)
   {
    Outlier("SSimple", 0.05);
   };
  };
  R_Date("Beta-436709", 13010, 50)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("End 1210 (interface)");
 Boundary("End SC-C/Start SC-B");
 Phase("SC-B")
 {
  R_Date("OxA-36613", 13630, 55)
  {
   Outlier("General", 0.05);
  };
  R_Combine("Sample 15")
  {
   Outlier("General", 0.05);
   R_Date("LEMA-635.1.2", 13142, 60)
   {
    Outlier("SSimple", 0.05);
```

```
};
R_Date("LEMA-635.1.1", 13054, 60)
{
 Outlier("SSimple", 0.05);
};
};
R_Date("OxA-36359", 13525, 35)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36611", 13050, 50)
{
 Outlier("General", 0.05);
};
R_Date("LEMA-574.1.1", 13092, 63)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36634", 12990, 55)
{
 Outlier("General", 0.05);
};
R_Date("LEMA-573.1.1", 12916, 58)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36612", 12885, 50)
{
 Outlier("General", 0.05);
};
R_Date("Beta-436710", 12880, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36623", 13010, 55)
{
 Outlier("General", 0.05);
};
Date("Oxford X-7233", N(2017-10960, 1610))
{
 Outlier("General", 0.05);
};
R_Date("OxA-36633", 12235, 75)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36625", 12170, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36622", 12155, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36620", 12140, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36609", 12120, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36317", 12120, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36315", 12095, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36316", 12050, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36619", 12040, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36496", 12005, 55)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36753", 11975, 70)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36624", 11900, 50)
{
 Outlier("General", 0.05);
};
R_Date("LEMA-893.1.1", 11897, 35)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36610", 11895, 50)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36621", 11890, 45)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36618", 11855, 50)
{
 Outlier("General", 0.05);
};
R_Date("LEMA-892.1.2", 11770, 35)
{
 Outlier("General", 0.05);
};
R_Date("LEMA-640.1.1", 11403, 60)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36608", 12060, 50)
{
 Outlier("General", 0.05);
};
R_Date("LEMA-978.1.1", 10513, 50)
{
```

```
  Outlier("General", 0.05);
 };
};
Boundary("End of SC-B");
Span("Span of sequence");
};
};
```

Model C (Fig. S2)

```
Options()
{
 Resolution=50;
};
Plot()
{
 Outlier_Model("General",T(5),U(0,4),"t");
 Outlier_Model("SSimple",N(0,2),0,"s");
 Sequence("Chiquihuite Cave, Model C")
 {
  Boundary("Start of SC-C");
  Sequence("SC-C")
  {
   Boundary("Start 1223");
   Phase("Stratum 1223")
   {
    Date("Oxford X-7229", N(2017-23940, 2950))
    {
     Outlier("General", 0.05);
    };
    Date("Oxford X-4135", N(2017-27790, 4340))
    {
     Outlier("General", 0.05);
    };
    R_Date("PRI-5414", 27929, 82)
    {
     Outlier("General", 0.05);
    };
    R_Date("Beta-345055", 27830, 150)
    {
     Outlier("General", 0.05);
    };
   };
   Boundary("End 1223/Start 1222");
   Boundary("End 1222/Start 1220");
   Phase("Stratum 1220")
   {
    R_Date("LEMA-577.1.1", 21401, 95)
    {
     Outlier("General", 0.05);
    };
    R_Date("LEMA-576.1.1", 20896, 80)
    {
     Outlier("General", 0.05);
    };
   };
   Boundary("End 1220/Start 1219");
   Phase("Stratum 1219")
   {
    R_Date("OxA-36530", 22170, 140)
    {
     Outlier("General", 0.05);
    };
    R_Date("OxA-34965", 21990, 170)
    {
     Outlier("General", 0.05);
    };
   };
   Boundary("End 1219/Start 1218-1219");
   Phase("Stratum 1218-1219")
   {
    R_Date("OxA-36360", 21140, 130)
    {
     Outlier("General", 0.05);
    };
    R_Date("OxA-36614", 20860, 100)
    {
     Outlier("General", 0.05);
    };
   };
   Boundary("End 1218-1219/Start 1217");
   Date("Oxford X-7227", N(2017-11620, 2000))
   {
    Outlier("General", 0.05);
   };
   Date("Oxford X-7232", N(2017-13870, 2250))
   {
    Outlier("General", 0.05);
   };
   Date("Oxford X-7231", N(2017-15560, 1740))
   {
    Outlier("General", 0.05);
   };
   Boundary("End 1210/Start 1210 (interface)");
   Phase("Stratum 1210 (interface)")
   {
    R_Date("LEMA-575.1.3", 14778, 77)
    {
     Outlier("General", 0.05);
    };
    R_Combine("Sample 37")
    {
     Outlier("General", 0.05);
     R_Date("LEMA-636.1.2", 13788, 90)
     {
      Outlier("SSimple", 0.05);
     };
     R_Date("LEMA-636.1.1", 13569, 60)
     {
      Outlier("SSimple", 0.05);
     };
    };
   };
  };
  Boundary("End 1210 (interface)");
  Boundary("End SC-C/Start SC-B");
```

```
Sequence("SC-B")
{
 Boundary("Start 1210 supra 1209");
 Phase("Stratum 1210 supra 1209")
 {
  R_Date("OxA-36613", 13630, 55)
  {
   Outlier("General", 0.05);
  };
  R_Combine("Sample 15")
  {
   R_Date("LEMA-635.1.2", 13142, 60)
   {
    Outlier("SSimple", 0.05);
   };
   R_Date("LEMA-635.1.1", 13054, 60)
   {
    Outlier("SSimple", 0.05);
   };
   Outlier("General", 0.05);
  };
 };
 Boundary("Start 1209");
 Phase("Stratum 1209")
 {
  R_Date("OxA-36359", 13525, 35)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36611", 13050, 50)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("Start 1207");
 Phase("Stratum 1207")
 {
  R_Date("LEMA-574.1.1", 13092, 63)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36634", 12990, 55)
  {
   Outlier("General", 0.05);
  };
  R_Date("LEMA-573.1.1", 12916, 58)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36612", 12885, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("Beta-436710", 12880, 50)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("Start 1206");
 R_Date("OxA-36623", 13010, 55)
 {
  Outlier("General", 0.05);
 };
 Boundary("Start 1204 lower");
 Phase("Stratum 1204 lower")
 {
  Age("Oxford X-7233", N(10960, 1610))
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36633", 12235, 75)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36625", 12170, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36622", 12155, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36620", 12140, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36609", 12120, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36317", 12120, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36315", 12095, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36316", 12050, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36619", 12040, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36496", 12005, 55)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36753", 11975, 70)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36624", 11900, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("LEMA-893.1.1", 11897, 35)
```

```
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36610", 11895, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36621", 11890, 45)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36618", 11855, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("LEMA-892.1.2", 11770, 35)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("Start 1204 upper");
 Phase("Stratum 1204 upper")
 {
  R_Date("OxA-36608", 12060, 50)
  {
   Outlier("General", 0.05);
  };
  R_Date("LEMA-978.1.1", 10513, 50)
  {
   Outlier("General", 0.05);
  };
 };
 Boundary("End 1204 upper");
};
Boundary("End of SC-B");
};
};
```

Model D (Fig. S3)

```
Options()
{
 Resolution=50;
};
Plot()
{
 Outlier_Model("General",T(5),U(0,4),"t");
 Outlier_Model("Charcoal",Exp(1,-10,0),U(0,3),"t");
 Outlier_Model("SSimple",N(0,2),0,"s");
 Sequence("Chiquihuite Cave, Model D")
 {
  Boundary("Start of SC-C");
  Phase("Stratum 1223")
  {
   Date("Oxford X-7229", N(2017-23940, 2950))
   {
    Outlier("General", 0.05);
   };
   Date("Oxford X-4135", N(2017-27790, 4340))
   {
    Outlier("General", 0.05);
   };
   R_Date("PRI-5414", 27929, 82)
   {
    Outlier("Charcoal", 1);
   };
   R_Date("Beta-345055", 27830, 150)
   {
    Outlier("General", 0.05);
   };
  };
  Boundary("End 1223/Start 1222");
  Boundary("End 1222/Start 1220");
  Phase("Stratum 1220")
  {
   R_Date("LEMA-577.1.1", 21401, 95)
   {
    Outlier("Charcoal", 1);
   };
   R_Date("LEMA-576.1.1", 20896, 80)
   {
    Outlier("Charcoal", 1);
   };
  };
  Boundary("End 1220/Start 1219");
  Phase("Stratum 1219")
  {
   R_Date("OxA-36530", 22170, 140)
   {
    Outlier("General", 0.05);
   };
   R_Date("OxA-34965", 21990, 170)
   {
    Outlier("General", 0.05);
   };
  };
  Boundary("End 1219/Start 1218-1219");
  Phase("Stratum 1218-1219")
  {
   R_Date("OxA-36360", 21140, 130)
   {
    Outlier("General", 0.05);
   };
   R_Date("OxA-36614", 20860, 100)
   {
    Outlier("Charcoal", 1);
   };
  };
  Boundary("End 1218-1219/Start 1217");
  Date("Oxford X-7227", N(2017-11620, 2000))
  {
   Outlier("General", 0.05);
  };
  Date("Oxford X-7232", N(2017-13870, 2250))
  {
```

```
 Outlier("General", 0.05);
};
Date("Oxford X-7231", N(2017-15560, 1740))
{
 Outlier("General", 0.05);
};
Boundary("End 1210/Start 1210 (interface)");
Phase("Stratum 1210 (interface)")
{
 R_Date("LEMA-575.1.3", 14778, 77)
 {
  Outlier("General", 0.05);
 };
 R_Combine("Sample 37")
 {
  R_Date("LEMA-636.1.2", 13788, 90)
  {
   Outlier("SSimple", 0.05);
  };
  R_Date("LEMA-636.1.1", 13569, 60)
  {
   Outlier("SSimple", 0.05);
  };
  Outlier("Charcoal", 1);
 };
};
Boundary("End 1210 (interface)");
Boundary("End SC-C/Start SC-B");
};
};
Options()
{
Resolution=50;
};
Plot()
{
Outlier_Model("General",T(5),U(0,4),"t");
Outlier_Model("Charcoal",Exp(1,-10,0),U(0,3),"t");
Outlier_Model("SSimple",N(0,2),0,"s");
Sequence("Chiquihuite Cave, Model D")
{
Boundary("End SC-C/Start SC-B");
Phase("SC-B")
{
R_Date("OxA-36613", 13630, 55)
{
 Outlier("Charcoal", 1);
};
R_Combine("Sample 15")
{
 R_Date("LEMA-635.1.2", 13142, 60)
 {
  Outlier("SSimple", 0.05);
 };
 R_Date("LEMA-635.1.1", 13054, 60)
 {
  Outlier("SSimple", 0.05);
 };
 Outlier("Charcoal", 1);
};
R_Date("OxA-36359", 13525, 35)
{
 Outlier("General", 0.05);
};
R_Date("OxA-36611", 13050, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("LEMA-574.1.1", 13092, 63)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36634", 12990, 55)
{
 Outlier("Charcoal", 1);
};
R_Date("LEMA-573.1.1", 12916, 58)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36612", 12885, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("Beta-436710", 12880, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36623", 13010, 55)
{
 Outlier("Charcoal", 1);
};
Date("Oxford X-7233", N(2017-10960, 1610))
{
 Outlier("General", 0.05);
};
R_Date("OxA-36633", 12235, 75)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36625", 12170, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36622", 12155, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36620", 12140, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36609", 12120, 50)
{
 Outlier("Charcoal", 1);
};
R_Date("OxA-36317", 12120, 50)
{
 Outlier("Charcoal", 1);
```

```
  };
  R_Date("OxA-36315", 12095, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36316", 12050, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36619", 12040, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36496", 12005, 55)
  {
   Outlier("General", 0.05);
  };
  R_Date("OxA-36753", 11975, 70)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36624", 11900, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("LEMA-893.1.1", 11897, 35)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36610", 11895, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36621", 11890, 45)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36618", 11855, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("LEMA-892.1.2", 11770, 35)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("OxA-36608", 12060, 50)
  {
   Outlier("Charcoal", 1);
  };
  R_Date("LEMA-978.1.1", 10513, 50)
  {
   Outlier("Charcoal", 1);
  };
 };
 Boundary("End of SC-B");
};
};
```

3. Lithic artefact metrics

Table S6. Contextual and metric values of the 91 artefacts depicted in illustrations (4.71% of the total).

Artefact's ID number (bag - inventory)	Figure no.	Square-subsquare	Depth range or depth from datum Z(D)	Stratigraphic component (SC)	Max. length (mm)	Max. width (mm)	Max. thickness (mm)	Weight (g)
1541-12309	3a	N6-NE	-1.90/ -2.00	B	42	36	20	28
1200-12510	3b	N4-SW	-1.70/ -1.80	B	26.2	19.3	6	3
1395-12532	3c	N6-NW	-1.80/ -1.90	B	21	23	4.5	2.5
1979-12734	3d	M7-SW	-2.50/ -2.60	**C**	23	21	6.4	2
1266-12519	3e	O4-NW	-1.70/ -1.80	B	38	25.7	7	7
391-10774	3f	i5-E/C	-1.10/ -1.25	B	40	12.1	4.7	2.5
444-9873	3g	i4-E	-	B	35.5	9.4	3.5	1
1624-12873	3h	N4-SE	-2.00/ -2.10	B	42	9.7	6.6	2.5
369-9733	3i	K3	-1.60/ -1.70	B	22.2	22	3.7	2
1836-12907	3j	N4-SW	-2.30/ -2.40	B	32	6.4	3.2	0.5
279-9539	3k	M4-W	-1.90	B	47	22.8	7	7
370-9734	3l	K3	-1.70/ -1.80	B	47.2	21.7	9.3	9.5
570-10056	3m	i5-E/C	-2.55	**C**	55.7	16.4	9.2	7
1925-13709	3n	M4-NW	-2.50/ -2.60	B	46.8	22.1	7.4	8.5
1554-13487	3o	O3-SW	-1.96	B	29.3	15	4.4	2
404-10915	Extended Data 5a	K3	-1.60/ -1.70	B	20.3	35.8	33	25.5
1209-12301	Extended Data 5b	M4-SW	-1.70/ -1.80	B	55.3	48.7	22.7	54
1608-12313	Extended Data 5c	N4-SW	-2.00/ -2.10	B	19.8	29.6	12.3	8
297-10638	Extended Data 5d	G3/H2/H3	-0.50/ -0.60	B	51.3	39.5	21.5	40
1074-12371	Extended Data 5e	N5-SE	-1.60/ -1.70	B	51.1	36.5	24.1	47
1081-12491	Extended Data 5f	N6-SW	-1.60/ -1.70	B	35.5	23	8.3	5
1779-12649	Extended Data 5g	N6-SW	-2.20/ -2.30	B	43.2	20.7	8.7	7
51-8916	Extended Data 5h	B1	+2.26/ 2.16	B	45	25.5	11.4	10.5
354-10664	Extended Data 5i	-	-	**C**	51.4	33.8	8.3	13.5
1475-12541	Extended Data 5j	M4-SE	-1.90/ -2.00	B	18	34.5	6.8	3.5

Artefact's ID number (bag - inventory)	Figure no.	Square-subsquare	Depth range or depth from datum Z(D)	Stratigraphic component (SC)	Max. length (mm)	Max. width (mm)	Max. thickness (mm)	Weight (g)
273-10598	Extended Data 5k	L3	-1.50/ -1.60	B	21.5	38	7.4	5
1530-12569	Extended Data 5l	N5-NW	-1.90/ -2.00	B	16.6	20.4	2.7	1
1631-12604	Extended Data 5m	N4-SE	-2.00/ -2.10	B	25.7	32.3	6	5
2132-13241	Extended Data 5n	L6-NE	-2.90/ -3.00	**C**	21.7	27.3	6.5	3
202-9335	Extended Data 5o	F2	+0.30/ 0	B	26.3	10.5	4.1	1.5
357-10674	Extended Data 5p	M6	-2.20/ -2.30	B	34	13.5	7.5	4
1044-12790	Extended Data 5q	N6-NW	-1.60/ -1.70	B	28	12.2	4.3	1
1046-12791	Extended Data 5r	N4-NE	-1.60/ -1.70	B	49.5	21	7.3	6.5
1763-12896	Extended Data 5s	M5-SE	-2.20/ -2.30	B	45.7	14.6	7.7	4
1884-13169	Extended Data 5t	M4-SW	-2.45/ -2.50	B	32.5	18.6	6.9	2.5
1519-12846	Extended Data 5u	N4-SW	-1.90/ -2.00	B	15.8	13	3.7	1
1839-12909	Extended Data 5v	O5-SW	-2.10/ -2.20	B	13.9	10.5	3.2	0.5
1229-13022	Extended Data 5x	N3-NW	-1.70/ -1.80	B	16.8	27	4.1	2
1888-13172	Extended Data 5y	M6-SE	-2.30/ -2.40	B	23.2	26	7.9	4.5
1202-13014	Extended Data 5w	N6-NW	-1.70/ -1.80	B	32.2	21	4	4
1926-13298	Extended Data 5z	M6-SE	-2.51	B	45	30.1	11	13
1289-13277	Extended Data 5a'	N4-SE	-1.70/ -1.80	B	103.8	29.6	15.5	43
589-10111	Extended Data 5b'	i6-NE	-2.53	**C**	36	22.6	9.3	7
643-10204	Extended Data 5c'	L6	-3.40/ -3.50	**C**	36	18.6	5.1	3.5
910-13315	Extended Data 5d'	-	-1.50/ -1.60	B	37	18.9	9.7	5.5
976-13332	Extended Data 5e'	O5-NE	-1.58	B	37.4	19.1	10.2	7
560-10012	Extended Data 5f'	i6	-2.36	**C**	45	16.8	6.2	5

Artefact's ID number (bag - inventory)	Figure no.	Square-subsquare	Depth range or depth from datum Z(D)	Stratigraphic component (SC)	Max. length (mm)	Max. width (mm)	Max. thickness (mm)	Weight (g)
1546-13483	Extended Data 5g'	N3-NE	-1.90/ -2.00	B	34.7	13.6	5.5	2
1660-13520	Extended Data 5h'	N4-NW	-2.18	B	35.3	12.6	6.1	3
1018-13639	Extended Data 5i'	N4-SE	-1.63	B	33.1	20.6	9.5	6
558-11720	Extended Data 5j'	i6	-2.20/ -2.30	B	31	14.1	6.6	3
1307-13671	Extended Data 5k'	O3-Nw	-1.82	B	31.7	21.8	10	6.5
956-13765	Extended Data 5l'	O4-SW	-1.50/ -1.60	B	20.5	17	4	2
1530-13770	Extended Data 5m'	N5-NW	-1.90/ -2.00	B	25.4	24.9	8.5	6
1845-13762	Extended Data 5n'	O5-SE	-2.10/ -2.20	B	18.1	16.4	5	2
2107-13778	Extended Data 5o'	N6-SW	-2.90/ -3.00	B	21	21.4	6.5	5
1719-13775	Extended Data 5p'	N4-SE	-2.10/ -2.20	B	19.5	18.8	4.6	2.5
1609-12314	Extended Data 6a	M4-NE	-2.00/ -2.10	B	28.1	27.6	14	10.5
1471-12540	Extended Data 6b	M5-SE	-1.90/ -2.00	B	20.6	20.3	8	3
2019-12744	Extended Data 6c	N4-SW	-2.60/ -2.70	B	21	31.3	4.7	3
1486-12837	Extended Data 6d	M6-SE	-1.90/ -2.00	B	16	13.6	3	1
1741-13138	Extended Data 6e	N6-NE	-2.10/ -2.20	B	22.7	20.8	5.9	4
1430-12534	Extended Data 6f	N4-NE	-1.80/ -1.90	B	33	26	7	6
1628-12874	Extended Data 6g	N4-SE	-2.00/ -2.10	B	107	38	18	75.5
1779-13148	Extended Data 6h	N6-SW	-2.20/ -2.30	B	33.6	12	4.8	2.5
1919-13194	Extended Data 6i	M4-NW	-2.50/-2.60	B	40.9	14.2	7.2	4
167-9233	Extended Data 6j	L4	-1.40/ -1.50	B	48.3	48	10.3	28
573-10065	Extended Data 6k	i7	-2.45/ -2.50	B	67.3	48.3	21	60
406-9792	Extended Data 6l	i4-W	-0.98/ -1.10	B	46.1	20	5.9	6

Artefact's ID number (bag - inventory)	Figure no.	Square-subsquare	Depth range or depth from datum Z(D)	Stratigraphic component (SC)	Max. length (mm)	Max. width (mm)	Max. thickness (mm)	Weight (g)
1015-13637	Extended Data 6m	M3-SE	-1.64	B	37.4	15.5	7.7	5
2053-13588	Extended Data 6n	M5-SW	-2.70/ -2.80	B	29.5	14.2	4.5	2
885-13312	Extended Data 6o	N6-SE	-1.50/ -1.60	B	34.6	18.4	6.9	4
912-13316	Extended Data 6p	M4-NE	-1.50/ -1.60	B	26.1	16.3	7.4	3
979-13333	Extended Data 6q	O4-NE	-1.57	B	32.3	15.5	4.5	2
1034-13349	Extended Data 6r	N4-SE	-1.60/ -1.70	B	39.2	26.6	5.6	7
1045-13357	Extended Data 6s	N4-NE	-1.60/ -1.70	B	35.3	19.3	6.3	3.5
327-10660	Extended Data 6t	M5	-2.00/ -2.10	B	44	28.6	6.3	9
336-10662	Extended Data 6u	i4	-0.20/ -0.40	B	37.8	22.3	5	4
236-9460	Extended Data 6v	J3	-0.80/ -0.90	B	34.2	25.9	6.3	6
460-9924	Extended Data 6x	J4-SW	-1.77	**C**	32.3	17.1	4.9	3
484-9944	Extended Data 6y	i6	-1.60/ -1.70	**C**	29	17	6	2.5
487-11106	Extended Data 6w	L4	-2.33/ -2.43	B	33.1	20.7	5.3	3.5
540-9998	Extended Data 6z	K4	-2.73/ -2.85	**C**	33	21.7	7.5	4.5
1563-13494	Extended Data 6a'	M5-SE	-2.00/ -2.10	B	43.3	26.5	6.6	9
2124-13601	Extended Data 6b'	M5-SW	-3.08	**C**	35.6	18	7.1	4
907-13735	Extended Data 6c'	N4-NW	-1.50/ -1.60	B	30.8	21.1	3.8	3.5
1899-13576	Extended Data 6d'	N6-NW	-2.30/ -2.40	B	18.9	13.3	3.3	1
1474-13747	Extended Data 6e'	M6-SE	-1.90/ -2.00	B	36	19.8	6.1	5.5
1625-13513	Extended Data 6f'	N6-NW	-2.00/ -2.10	B	34.3	18.8	6.7	4.5
2104-13596	Extended Data 6g'	N5-SW	-2.90/ -3.00	**C**	35	19.9	7.9	5.5
2125-13764	Extended Data 6h'	M5-SW	-3.00/ -3.10	**C**	31.1	24.1	6.7	6

4. Chemical Residues

4.1. Methodology

Simple chemical tests have been successful in analyzing chemical residues, specially in soil and floor samples[5,6]. In this project, samples were obtained from an occupation surface (UE1210) by the excavation team following instructions from the laboratory (Fig. S6a,b). Samples bagged in polyethylene and tagged were sent to the lab. All floor samples were tested to detect phosphates, carbonates, protein, fatty acids and carbohydrates residues following procedures established in the laboratory[7]. Semi-quantitative results were mapped to produce distribution maps (where the colour saturation indicates higher values of each chemical indicator) (Figs. S5, S6). As a direct antecedent, we mention the study of organic chemical residues found in sediments surrounding mammoth bone remains dated to 18,000 years BP[8].

4.2. Results & discussion

4.2.1. Phosphate

There is a clear contrast among phosphate values. Low values are in the western part, while high values are in the northeast of the excavated area.

4.2.2. Carbonates

Carbonates are quite homogeneous. Taking into consideration that natural bedrock is a limestone , it was expected to have high carbonates values in almost all floor samples.

4.2.3. Protein residues

This chemical indicator is consequence of protein decay and has a rather similar distribution to that of phosphates, with the lowest values in the western part but the highest values towards the northeast. This overlapping pattern suggests that some cultural activity enriched the same areas with these residues.

4.2.4. Fatty acids

The distribution of this chemical residue is related with resins, oils and fats, and has a very similar pattern to the distribution of proteins and phosphates in the excavated area.
Carbohydrates are the consequence of ancient starch and sugars. They also have a similar pattern to the previous chemical indicators. In all cases, the western part has the lowest values. In

contrast, the northeastern part of the sampled area has the maximum values of chemical indicators.

We also plotted the spatial distribution of values for S, K and Zn made with XRF (Fig. S5). These chemical elements were selected because they are usually a product of chemical enrichment by human activities. The relationship between phosphates and Zn has been recognised as an indicator of refuse areas in previous works[6]; potassium concentrations are usually interpreted as indicators of cellulosic fuel ash accumulation, and sulphur is widely present in living cells and in this case. All three elements follow the same distribution patterns.

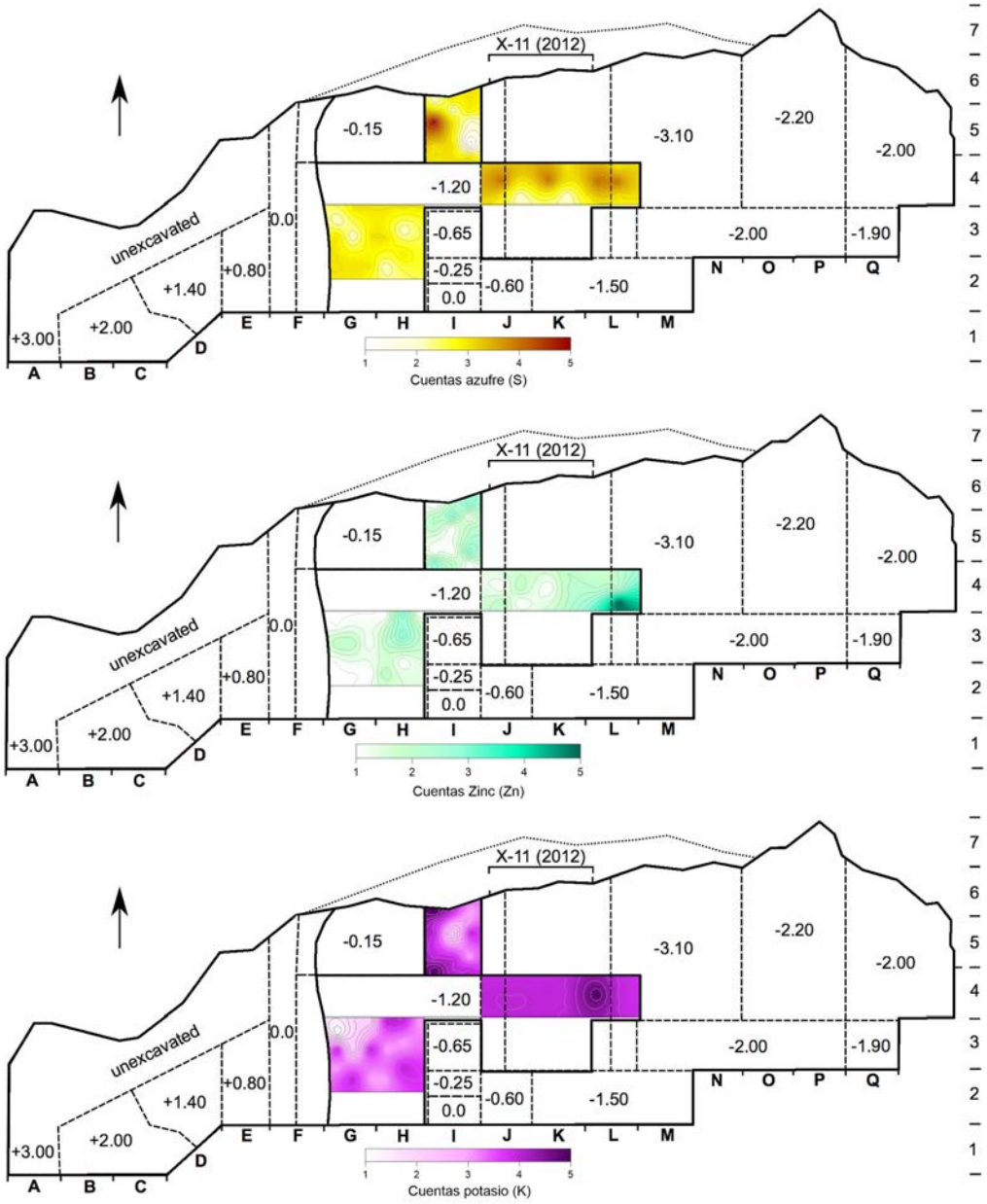


Fig. S5. Distribution maps of the XRF values for sulphur, zinc, and potassium, respectively (from top to bottom).

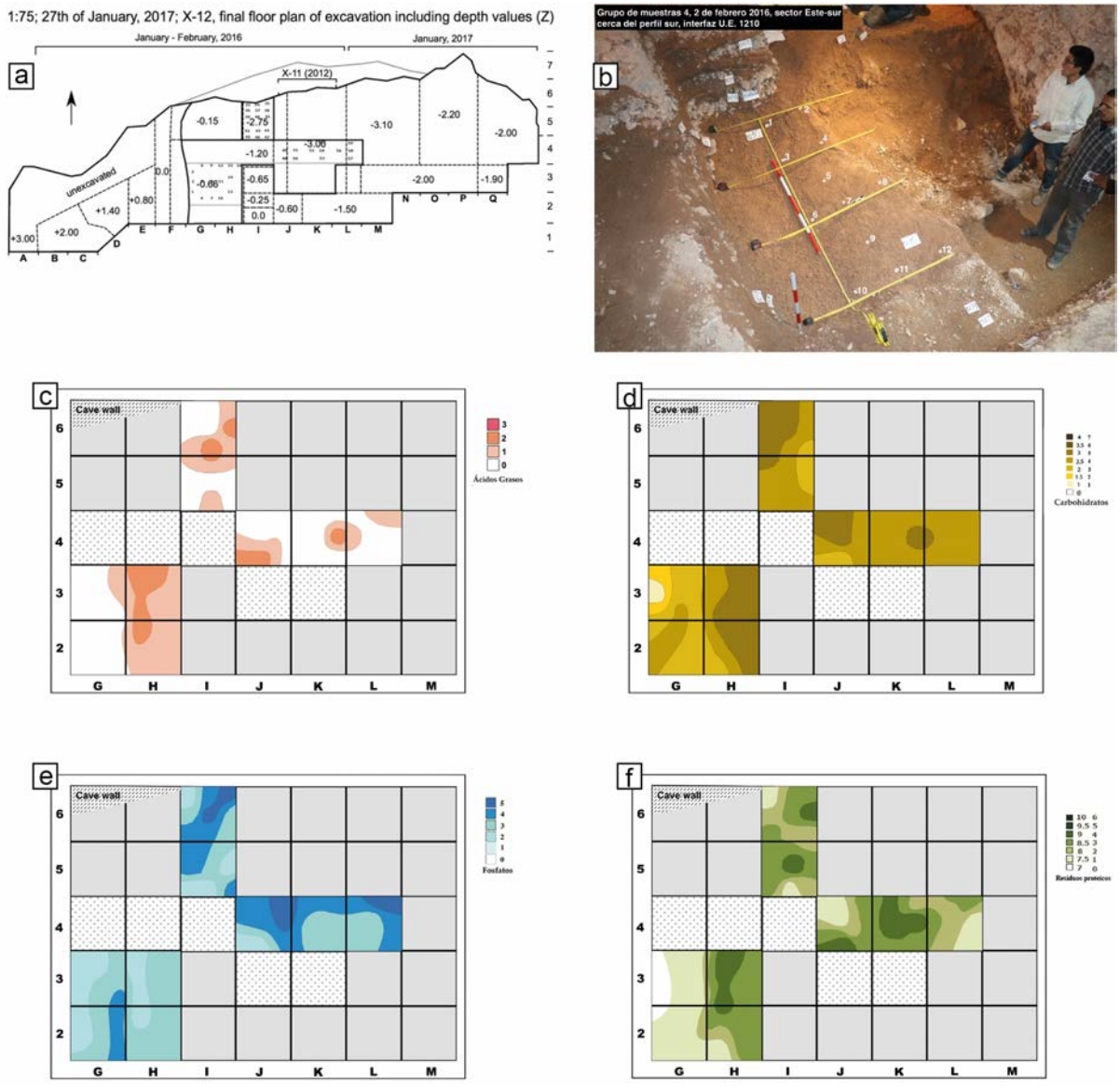


Fig. S6: Chemical residues spot test analyses on the interface of stratum 1210**.** **a**. Diagram of X-12, showing the location of samples. **b**. Sampling methodology, squares J4-L4. Results for: **c**. Fatty acids. **d**. Carbohydrates. **e**. Phosphates. **f**. Protein residues. **g**. Carbonates.

4.3. Final comments

The observed organic enrichment is not likely due to the natural decay of the limestone parent rock, but rather the product of human activities performed on top of this layer. To support the case for human presence in this floor, the same samples were analyzed by XRF and results displayed the same distribution patterns as the organic residues. The overlapping of these independent chemical indicators provides a higher degree of confidence in the interpretation of human activities producing chemical enrichment on this surface.

5. Faunal remains

Table S7. List of vertebrate species found in the excavation X-12.

Class	Order	Family	Taxon	Common name
Aves	Passeriformes	Picidae	*Melanerpes formicivorus*	woodpecker
		Emberizidae		sparrows; juncos
Mammalia	Eulipotyphla	Soricidae	*Notiosorex*	desert shrew
	Chiroptera	Phyllostomidae	*Leptonycteris*	long-tongued bat
		Vespertilionidae	*Antrozous pallidus*	pallid bat
			Myotis cf. planiceps	cave bat
			cf. Myotis	cave bat
	Lagomorpha	Leporidae	*Sylvilagus audubonii*	rabbit
			Sylvilagus floridanus	rabbit
			Sylvilagus sp.	rabbit
	Rodentia	Geomyidae	*Thomomys* sp.	gopher
		Muridae	*Microtus cf. mexicanus*	Mexican vole
			Neotoma cf. leucodon	woodrat
			Neotoma cf. goldmani	woodrat
			Onychomys cf. arenicola	grasshopper mouse
			Peromyscus melanophrys	plateau mouse
			Peromyscus sp.	white-footed mouse
			Reitrodontomys sp.	harvested mouse
	Carnivora	Ursidae	*Ursus cf. americanus*	black bear
	Artiodactyla	Cervidae	*Odocoileus virginianus*	white-tailed deer
		Antilocapridae	*Antilocapra*?	pronghorn

Fig. S7. Examples of faunal bone material. **a**, Articular condyle (jaw), probably Pleistocene condor, *Gymnogyps* sp. (SC-C, strata 1218-1219, dated to 21.1 ±130 ^{14}C kyr BP, OxA-36360). **b**, Canidae canine tooth, naturally split (SC-B, str. 1206, not dated, found with point shown in Fig. 3k). **c**, Medium-sized

mammal femur, probably otter, *Lontra* sp. (SC-C, str. 1223, not dated). **d**, Long bone fragment, with possible human modification (CC-C, str. 1219, dated to 22.1±140 ^{14}C kyr BP, OxA-36530). **e**, Black bear (*Ursus americanus*) penis bone, from trench X-11 (SC-C, str. 1223, dated to 27.8±150 ^{14}C kyr BP, Beta-345055). **f**, Mammal rib (SC-B, str. 1207D, dated to 12±55 ^{14}C kyr BP, OxA-36496). **g**, Passerine bird beaks (probably Turdidae), upper (i) and lower (ii, iii) parts, closely grouped in squares O-P (CC-B, str. 1204, not dated). **h**, Land snails taxa found in all strata, and living today near the cave: *Humboldtiana* sp. (i), and fam. Urocoptidae (mainly *Microceramus* sp. and *Urocoptis* sp.).

6. Phytolith and pollen

6.1. Sampling

Bulk sediment samples from Chiquihuite Cave, excavation X-12, were analyzed for pollen and phytoliths, to assess whether these proxies could provide paleoecological data and/or detect human influence at the site.

Nine samples were taken from the southern profile of unit M-N, where the natural stratigraphy slopes by up to 35-40 degrees from west to east. Samples were taken horizontally across the excavated sequence during the initial phases and prior to full understanding of the stratigraphy, thus resulting in a mixture of material from more than one natural stratum in each sample (Fig. S8). It is important to make clear that the intention of the excavators, during this sampling process, was merely *to evaluate the potential of the cave* site for such studies, not necessarily to carry a detailed micro-botanical analysis. Although this hinders a direct representation of the stratigraphical units, fluctuations in the pollen and phytolith data were detected along the sequence and used to find concordances with the ancient floristic eDNA detected.

6.2. Methods

Phytoliths were extracted from 100 ml of sediment following the wet oxidation method described elsewhere[9]. Sediment was sieved into silt (< 53 µm) and sand (53–250 µm) fractions to concentrate larger diagnostic morphotypes. Residue was mounted in Permount mounting medium, phytoliths counted under 400x (silt fraction) and 200x (sand fraction) magnification and photographs taken using Zen software. A phytolith count of 200 was sought in each sample and the graph made using C2 software[10]. Grass short cell phytoliths were identified according to published Poaceae reference collections from the Americas, Africa Asia and New Zealand[11-18].

At the São Paulo lab, samples were sieved for gravel removal (>250 µm) and 5 cm^3 of sediment processed for pollen grains following[19]. A final sieving stage (5 µm) was added at the end to remove clay and two Lycopodium sp. (exotic marker) tablets were used per sample. Samples were counted at 100x magnification using immersion oil and +150 palynomorphs counted. Photographs were taken using Zen software. Percentage and concentration values were calculated and plotted using TILIA, TILIAGRAPH software[20]. Identification was made with the IGC/USP reference collection and by comparison with pollen atlases from Colombia, Panama and Argentina[21-23]. Phytolith and pollen identifications were checked against lists of native species to improve taxonomic identifications[24-26].

6.3. Results

TILIA software identified four pollen "zones" using the sum of squares principle (CONISS), which we then superimposed onto the phytolith graph. We have maintained these zones in Figs. S9a-d to ease description of the results, but recognise that they do not represent ecological zones *stricto sensu* due to the admixture of microremains from different natural stratum in each sample. Micrographs of selected morphotypes are presented in Fig. S9. Abundant damaged pollen grains (>50% of counts) in all samples indicated transportation or pre-burial and/or long grain exposure time.

Zone 1 [which represents a mixture of terminal LGM (1212) and earlier LGM (1218) sediments (Fig. S8)] is characterized by the dominance of terrestrial herb pollen alongside phytoliths belonging to several grass subfamilies, suggesting the presence of an open, dry environment during this time. This type of vegetation is also reflected in the taxa identified by eDNA in the LGM strata (UE1212; 2017 season eDNA sample 3), as is the rapid replacement of warm-adapted PACMAD grasses to cold-adapted Pooideae grasses at the beginning of the LGM, recorded in the phytoliths. We suggest that the burnt globular echinate phytoliths recovered in this zone might represent material brought into the cave by humans, given the limited present distribution of palm species in the landscape (see main text).

In Zone 2, phytoliths from different grasses continue to be present, while the pollen records increasing levels of Agave pollen. The generalised low counts of phytoliths and pollen in this zone might be related to poor precipitation regimes that minimised debris-flows into the cave, however, the eDNA results from the bottom-most strata of this zone (UE1210; 2017 season eDNA sample 2) suggests a transition to forested vegetation (and wetter conditions), at least in the period immediately following the LGM, with Agave DNA only occurring in the YD-related upper strata (UE1204).

Zone 3 consists of samples from a mixture of four to five different sloping strata (1207-1204) and record a general increase in cold-adapted taxa (*Alnus*, *Pinus*) and a decrease in grasses. Bambusoideae phytoliths and wet/cold-adapted ferns also become more abundant and palm phytoliths peak in this zone.

Both phytoliths and pollen are most abundant in Zone 4, a mixture of sediments mainly from stratum 1204, representing the Younger Dryas. This would imply an increased influx of organics and possibly higher precipitation during this time. More humid conditions would also explain the decrease in Agave pollen and the increase in ferns, conifers, oak and wet/cold-adapted (Pooideae) grasses. These patterns strongly contrast with the DNA results from the same strata (1204 A/B/C; 2017 season eDNA sample 1), which show a general shift to dryer conditions as *Pinus* and algae decrease and Agave peaks for the first time.

On closer inspection, the phytolith and pollen results for the terminal LGM to YD (Zones 2-4) seem to record the opposite pattern to the DNA data, i.e. the establishment of a drier, more open environment in the terminal LGM, and a wetter, more forested environment in the YD. This discrepancy is likely related to the admixture of different strata during the microbotanical sampling, a fact that makes the eDNA results a more reliable proxy for this time period. Future pollen and phytolith analysis at equal resolution to the eDNA data would likely resolve this discrepancy.

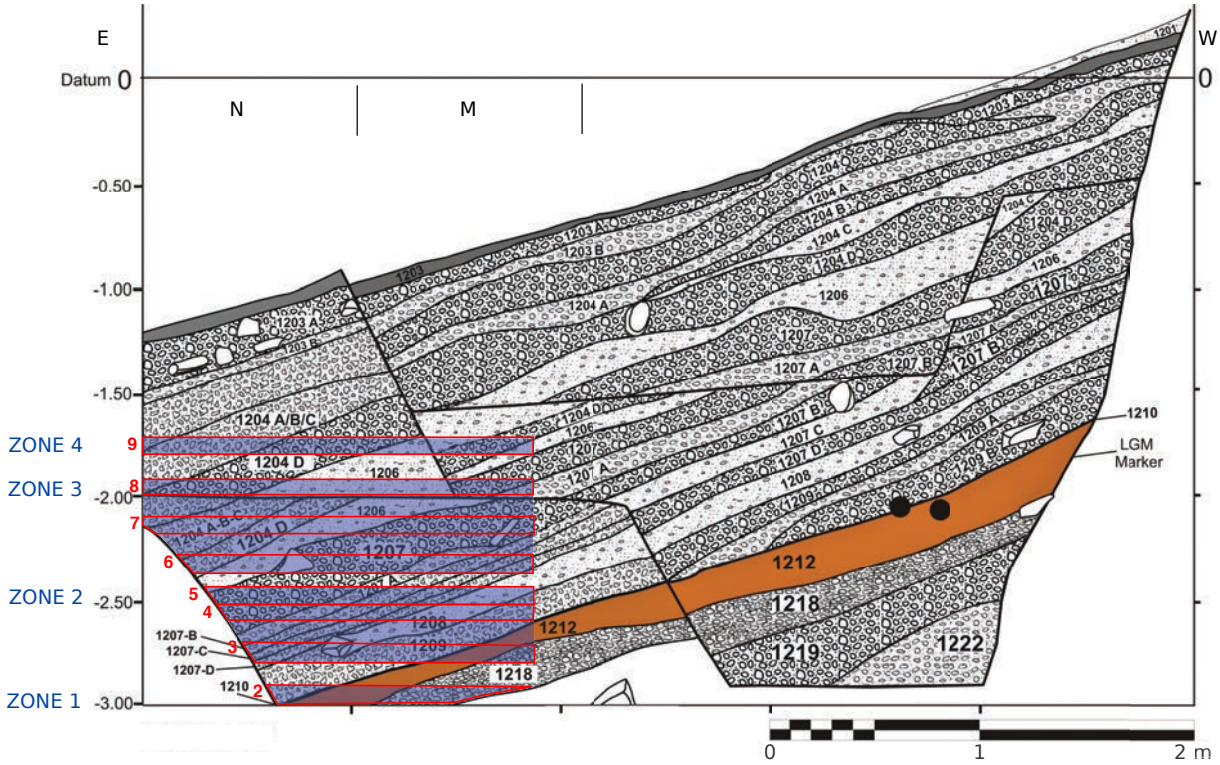


Fig. S8. Location of pollen and phytolith samples within the stratigraphy (in red). Sample 1, out of view, was extracted from SC-C only, immediately below sample 2. Pollen zones are noted in blue.

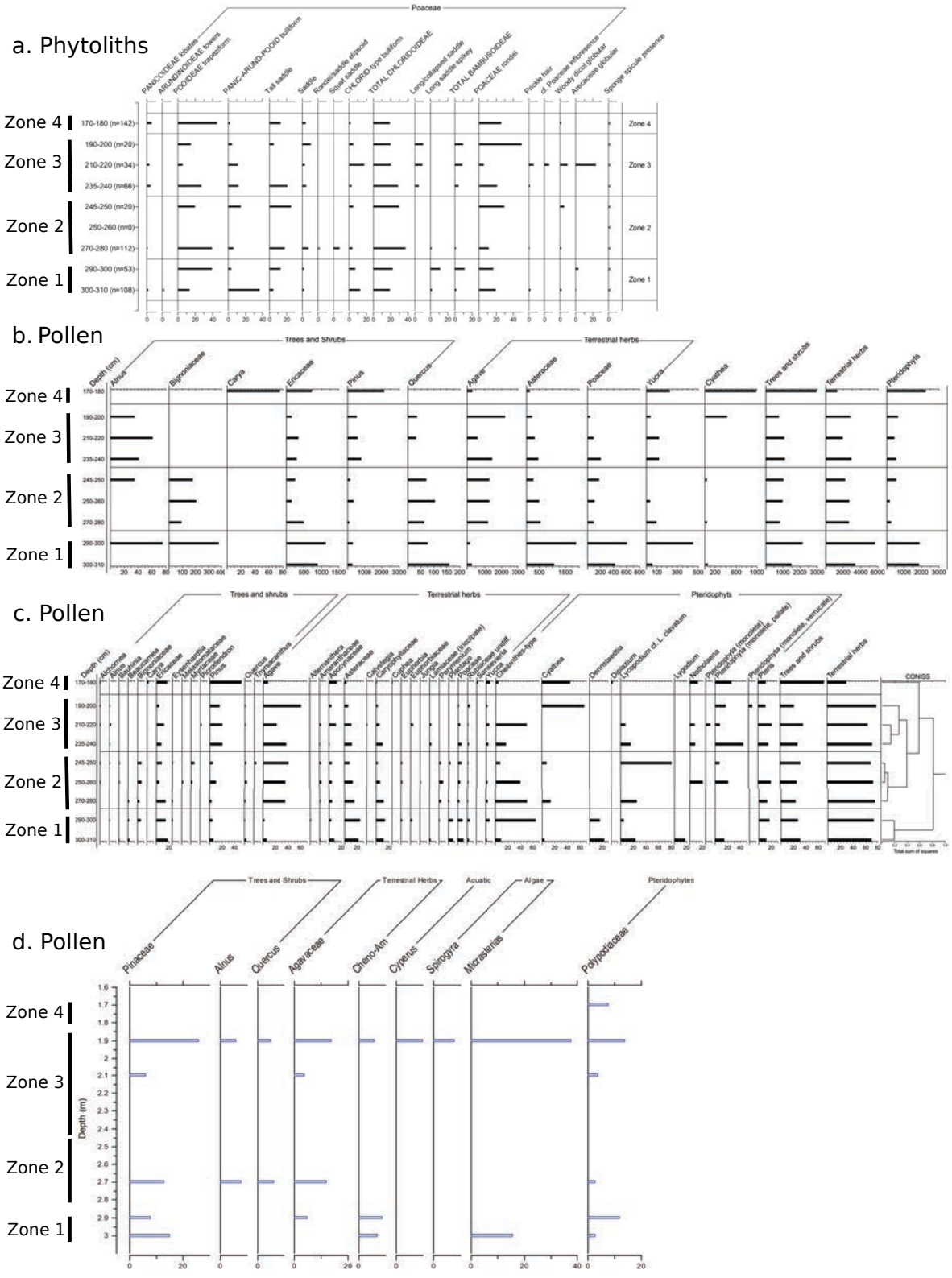


Fig. S9. **a**. Phytoliths results (in percentages). **b**. and **c**. Pollen results (pollen grain concentrations and percentage frequencies, respectively), São Paulo laboratory. **d**. Pollen results, Mexico City laboratory.

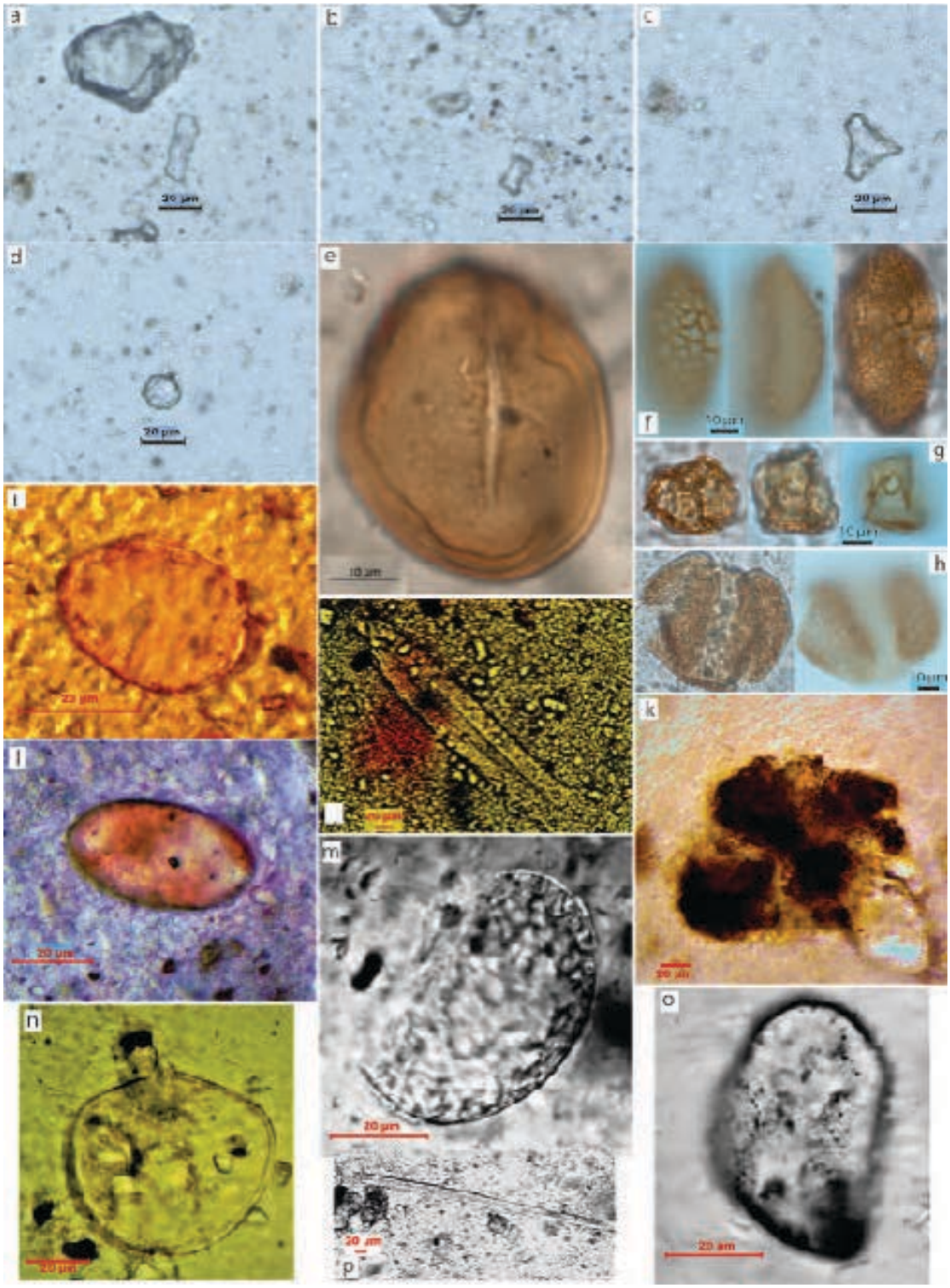


Fig. S10. Paleobotanical material. Phytoliths: **a.** Pooideae (trapeziform sinuate). **b.** Bambusoideae (long/ collapsed saddle). **c**. Chloridoideae (bulliform). **d**. Arecaceae (globular echinate). Pollen: **e**. *Yucca* sp. **f**. *Agave* sp. **g**. *Carya* sp. **h**. *Pinus* sp. Others: **i**. Alga spore, ovoidites of Spyrogira. **j**. Diatom. **k**. Alga spore, fam. Zignemataceae aff. Micrasterias. **l**. Fungus spore, unicelular. **m**. *Pinus* sp., alveolar sac. **n**. Fungus spore with hypha. **o**. Reed pollen, *Cyperus* sp. **p**. Sponge spicule.

7. Thin section and micro-morphology

The fifteen samples studied include sample M1 (cave floor, strata 1210-1212), M2-M5 (grey gravels from X-12), M6 (light-grey cobble from the slope), M7 (slope dark-grey gravel), M8 (slope greenish limestone), M9 (dark limestone artefact 1889-12698), M10-12 (greenish artefacts 1866-12685, 1899-12709, 1899-12710), M13 (grey artefact 2110-12949), M14 (rock fragment from the roof on the current entrance drip line), M15 (sample from the eastern wall, near the excavation), and M16 (from the ceiling above the western end of the dig) (Fig. S28).

7.1. Sample M-1 (cave floor, interface 1210 + upper cm of UE1212)

7.1.1. Macroscopic description

Sandy gravel with limestone fragments supported by silt and fine sand matrix. The sample displays both normal and inverse grading in different parts and incipient imbrication. The inversely graded sequence is 3 cm thick with a grain size that varies from very coarse sand (1 mm) to medium pebbles (10 mm). In contrast, the normally graded sequence is ~15 cm thick, and ranges in grain size from very large pebbles (50 mm) to granules (2 mm). The clasts are texturally subangular to subrounded, with sphericity ranging from low to moderate. The matrix makes up less than 20% of the total volume and consists of silt and fine sand. A brown color denotes a moderate state of oxidation.

The sediment in this sample is classified as Gmg, or Matrix Supported Gravel, with inverse to normal grading[27] that probably formed in pseudoplastic debris flows under low strength or viscosity conditions.

7.1.2. Microscopic description

Calcareous-sandy gravel composed of lithic fragments of limestone and metalimestone (Fig. S11a-f). Silt and fine sand matrix support the gravel. Carbonate fragments consist of mudstone of globigerinids and Radiolaria, but also wackestone primarily composed of globigerinids (Fig. S11a-c). The fragments are texturally subangular to subrounded with moderate sphericity. Metacarbonate fragments are classified as texture-types 1, 2, and 3, a numerical sequence that indicates increasing metamorphic grade[28]. Fragments classified as metacarbonate 1 have weakly developed cleavage, fragments of metacarbonate 2 have moderate cleavage, and fragments of metacarbonate 3 have well-formed cleavage and slight mica overgrowths similar to marble (Fig. S11e-g). The metacarbonate fragments are subangular with low to moderate sphericity. Subordinate fragments of phosphorite have parallel lamination and are subrounded (Fig. S11h).

Most fragments have a coating of iron oxide, 2 mm to 6 mm in thickness (Fig. S11a-h), and several mudstone and wackestone carbonate fragments have dissolution cavities, most likely caused by exposure to atmospheric CO_2 (Fig. S11b-c). In some fragments, microfractures in two orientations form conjugated joints that cut both the fragment and the oxide coating (Fig. S11c). This last feature is similar to triaxial deformation and fracture caused by exposure to confined pressure under compressional stress. In other words, the limestone in sample M-1 contains a pattern of microscopic surface cracks that suggest the rock was broken without complete separation of the parts.

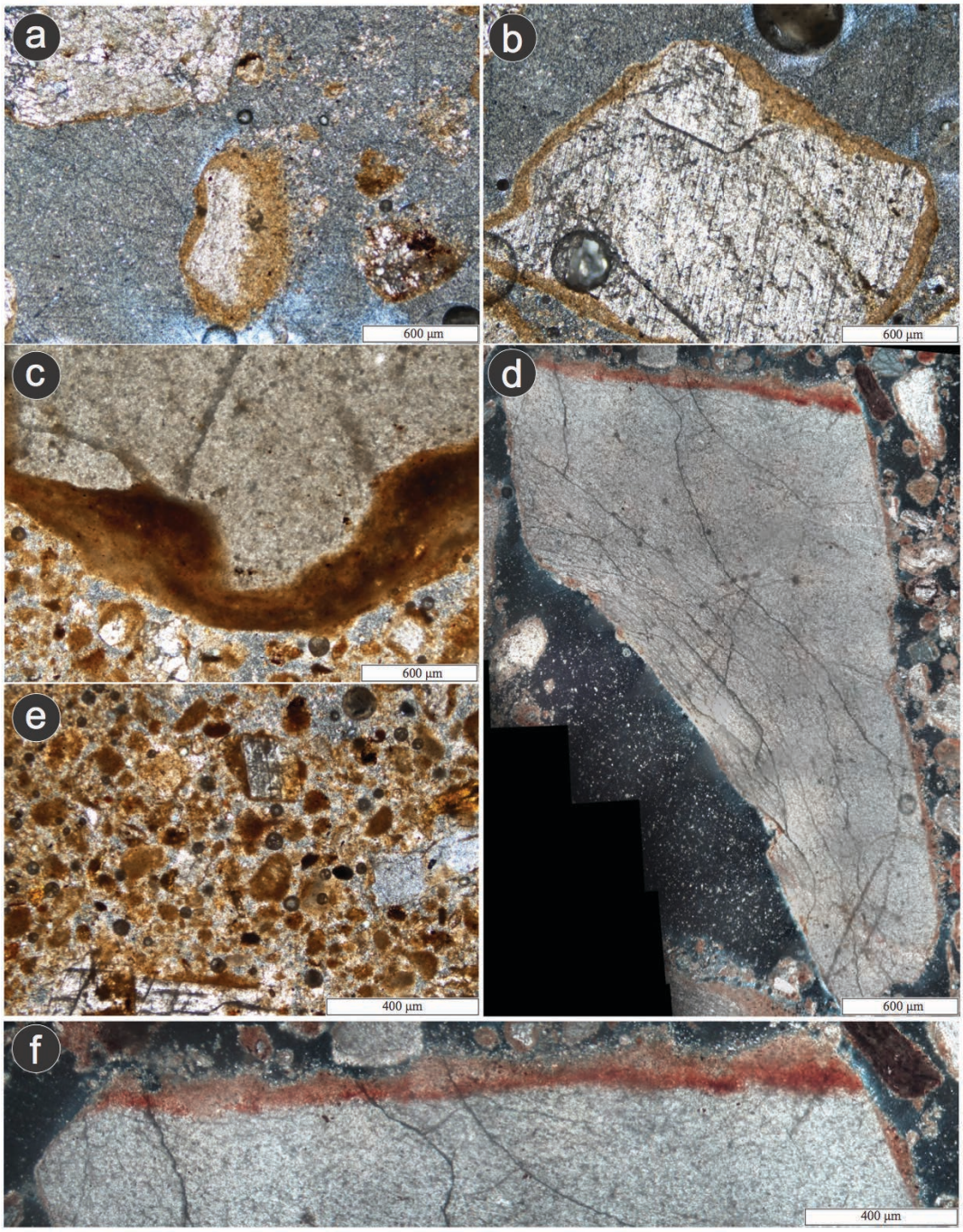


Fig. S11. Microphotographs from sample M-1. **a.** and **b.** carbonate grains from mudstone coated by iron oxide laminae; **c.** detail of the iron-oxide envelopes; note the dissolution at the edges of the grains probably associated with exposure to atmospheric CO_2; **d.** panoramic view of carbonate grain cut by triaxial deformation and fractures; note how fractures cut oxide laminae that cover grain surface. **e.** iron oxide grain coatings; **f.** detail at carbonate grain in Figure S8d, showing triaxial deformation and fractures cutting the oxide laminae that cover the surface of the grain.

7.2. Sample M-2

Mudstone–wackestone with pellets[29] or pelmicrite[30,31]; mid to high recrystallization (90% of the sample) (Fig. S9a-b). The orthochemical components are microcrystalline calcite (<60%) and calcite spar (<40%). The allochemical components are pellets (<65%) (Fig. S10c-d). Marine

protozoa (foraminifera) represent the skeletal components in the sample, and they include rotalinids, globigerinids, Globotruncana sp., and Radiolaria. Microstructures, such as walls in foraminifera, are difficult to discern in replacement calcite. Grains of various extrabasinal minerals, such as quartz and feldspar, are common. Some have remnants of crystal faces and are subangular to subrounded. Pellets are subrounded to rounded and have low sphericity and mean size of 60μm. These are dark brown (Fig. S12e-f). Cement is calcite spar partly dolomitized and ankeritized. Most dolomite is probably formed by the replacement of calcite and is distinguished from calcite in thin section by crystal habit. Replacement, in this case, consists of dissolution of the original calcite and precipitation of dolomite—mostly as the cement in voids. Most iron oxide (hematite) ranges in size from 10μm to 150μm. Microfractures or veins occur in three orientations filled with calcite, a trigonal polymorph, and an orthorhombic polymorph, such as aragonite. The thickness of fractures varies from 30μm to 250μm, with lengths that range from 15μm to 5cm. In microfractures, components have been cemented by calcite in grains that are much larger than the components themselves.

The microfacies characteristics of sample M-2 are similar to that of standard microfacies 3 and 4, which allow the interpretation of deep shelf margin deposits within facies belt 3[32].

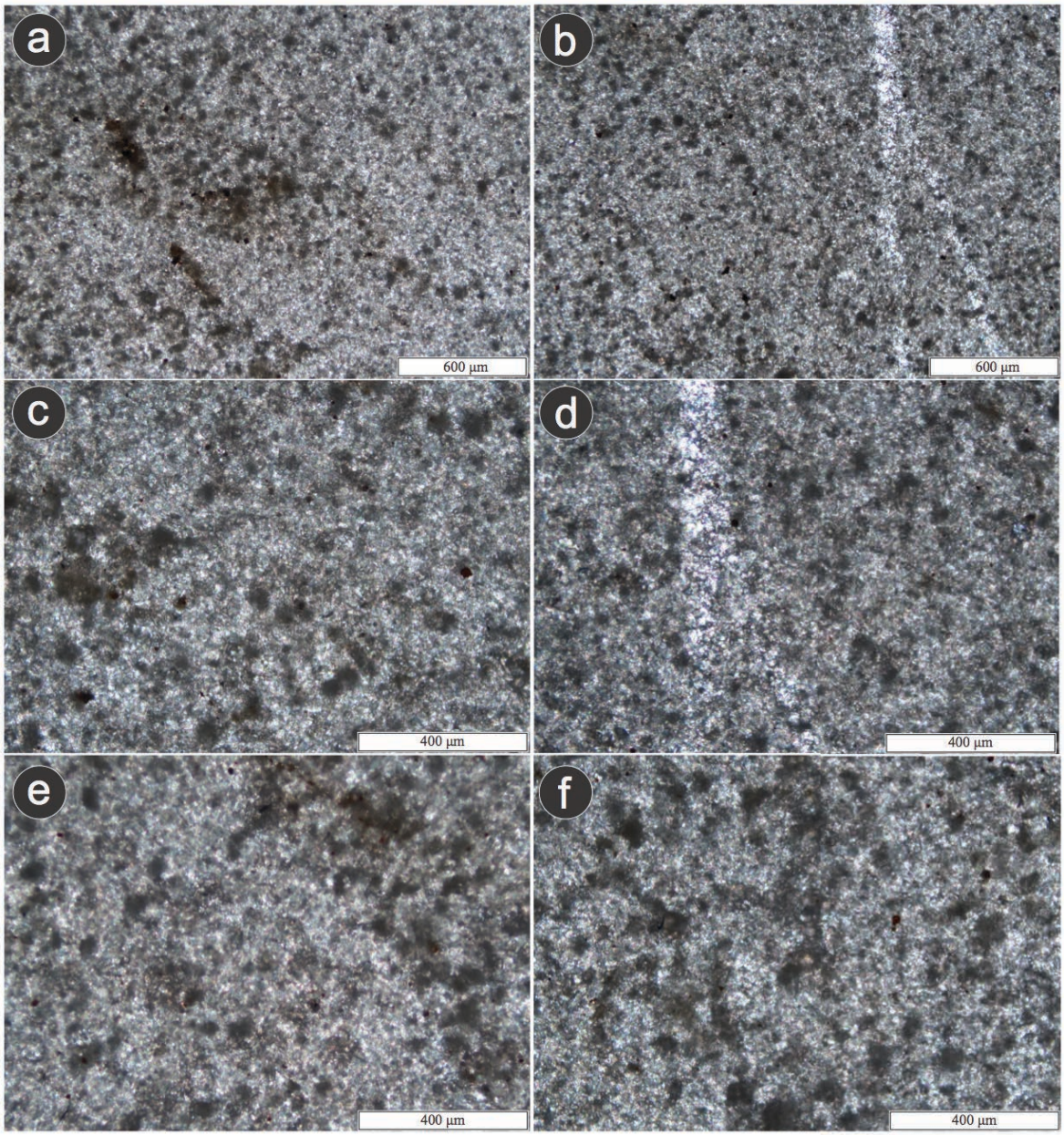


Fig. S12. Microphotograph in crossed polars for Sample M-2. **a.** and **b.** mudstone-wackestone with pellets, mid to high recrystallization; microfractures filled with calcite; **c.** and **d.** detail from a. and b. showing calcite recrystallization and calcite spar; **e.** and **f.** closeup of pellets.

7.3. Sample M-3

Mudstone–wackestone with pellets[29] or pelmicrite–pelsparite[30,31]; low-level recrystallization (20%). The orthochemical constituents consist of microcrystalline calcite (<60%) and calcite spar (<40%) (Fig. S13a-b). The allochemical constituents consist of pellets (<15%), foraminifera, such as rotalinids, globigerinids, and Radiolaria, as well as extrabasinal minerals such as quartz and feldspar. Pellets are subrounded to rounded and platy in shape, ranging in size from 20µm to 50 µm, and are dark brown. Cement is calcite spar that has been dolomitized to some extent. Mixed iron oxides, such as hematite, prevail, with sizes ranging from 10µm to 150µm. Fractures are present in three orientations filled with calcite and aragonite (Fig. S13c-e). The thickness of the microfractures varies from 30µm to 150µm, and 15µm to 3cm in length. At fractures, cement

crystals are large enough to show up in freshly broken hand specimens as “shiny cleavage” surfaces (“luster mottling”; Fig. S13f).

Microfacies characteristics are similar to those found in standard microfacies 3 and 4[32]. This indicates that the limestone in sample M-3 was deposited on a deep shelf margin within facies belt 3[32].

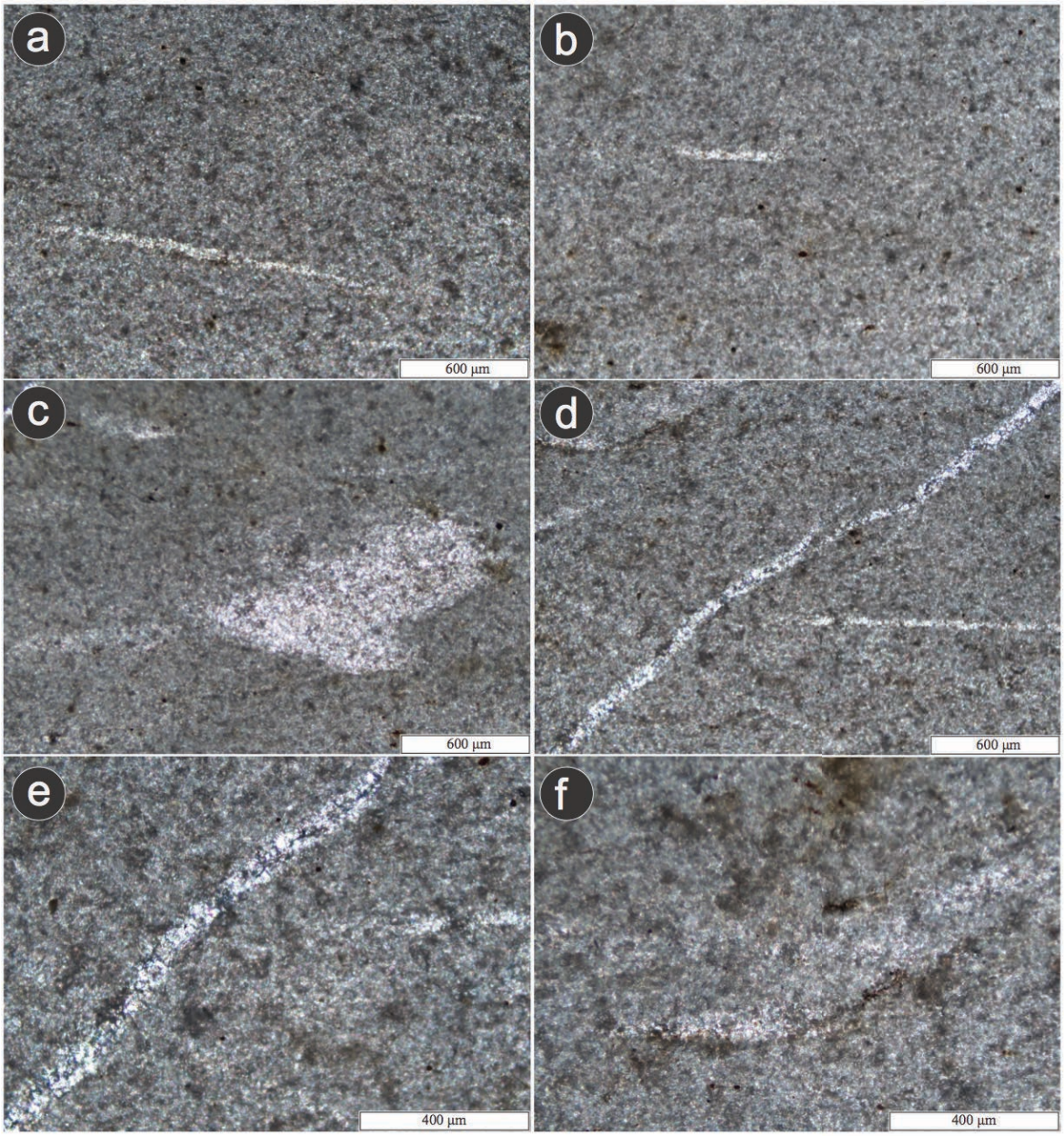


Fig. S13. Microphotograph in plane-polarized light of sample M-3. **a.** and **b.** mudstone-wackestone with pellets slightly recrystallized with microfractures distributed in three orientations and cavities filled with calcite; **c.** detail of pores and cavities filled by calcite spar; **d.** and **e.** fractures filled with calcite spar and aragonite; **f.** microfracture replaced by aragonite and opaque minerals, probably hematite.

7.4. Sample M-4

Mudstone–wackestone[29] or pelmicrite[30,31]; relatively highly recrystallized (70%) (Fig. S14a-b). The primary orthochemical constituents are microcrystalline calcite (<50%) and calcite spar (<50%). The allochemical components are pellets (<65%), foraminifera-like rotalinids,

globigerinids, Globotruncana sp., and Radiolaria. Extrabasinal minerals, such as quartz and feldspar, are common. Pellets are subrounded to rounded with elongated leaf-like forms that average 50μm in size and are of a dark brown color (Fig. S14b). Cement is calcite spar, moderately dolomitized, and ankeritized. Iron oxides are assumed to be hematite. Commonly, particles of hematite have diameters of 10μm to 150μm. Fractures have three orientations filled with spar (Fig. S14c). The width of the microfractures varies from 20μm to 350μm, and the length varies from 15 μm to 5cm (Fig. S14d-f). Calcite spar, which fills pores spaces, appears black in crossed polars. Pore space, also known as poikilitic spar microstructures or "luster mottling," extends 700μm-90mm, with a width of 300μm-700μm. Calcite-filled cracks, which are found together with sedimentary or tectonic microstylolite seems, are interpreted as shear structures. However, no distinction has been made to specify their genesis. The seams may be derived from mechanical compaction, possibly with burial and before cementation, and arrange as irregular surfaces within pseudo-bedding. They are characterized by mutual interpenetration of the two sides, with column pits and tooth-like projections on one side fitting into their counterparts on the other.

The microfacies characteristics for sample M-4 correspond to standard microfacies 3 and 4, interpreted as deep shelf-margin deposits within facies belt 3[32].

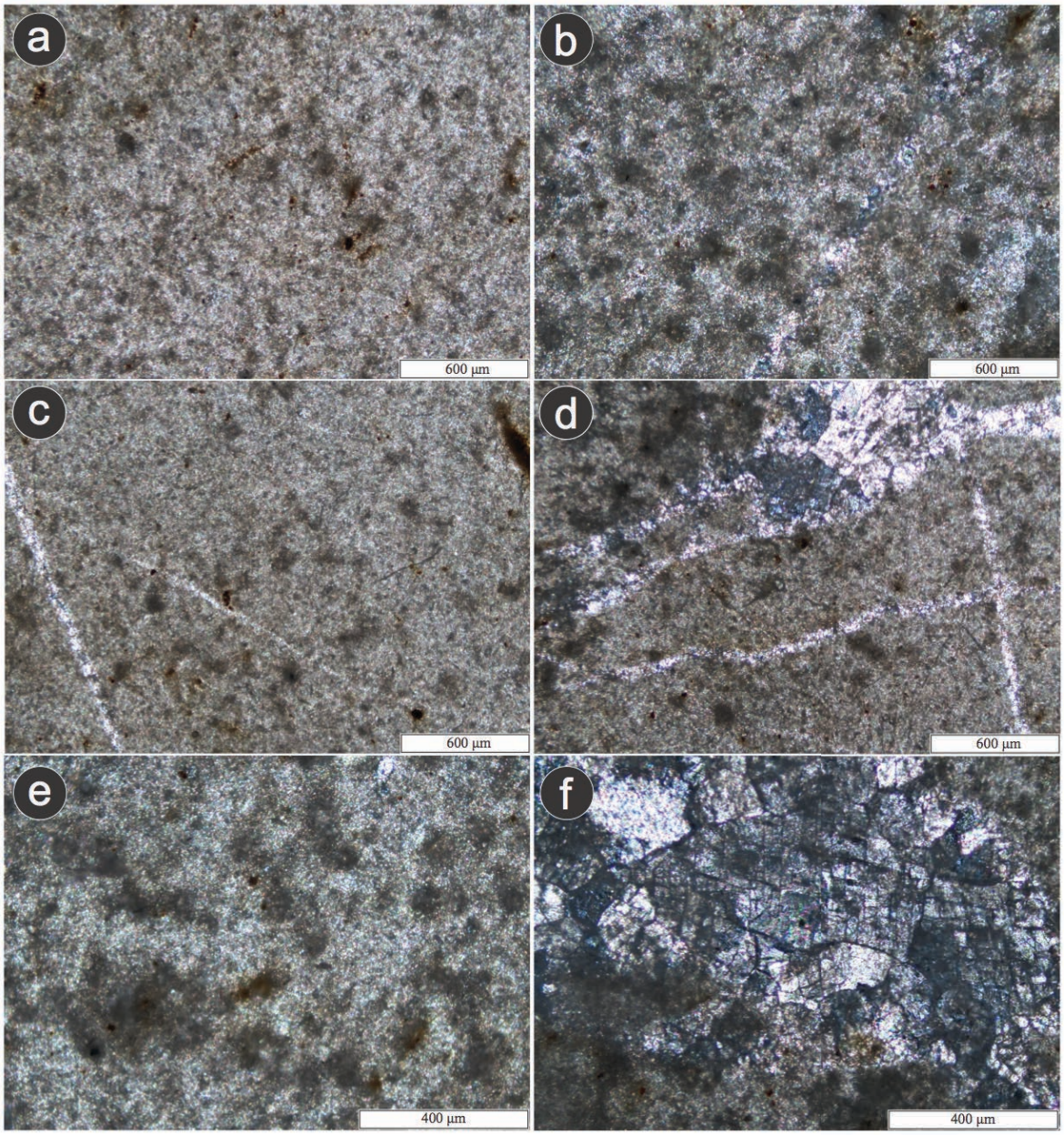


Fig. S14. Microphotograph in crossed polars of sample M-4. **a.** mudstone-wackestone with pellets, moderately to highly recrystallized with microfractures filled with calcite; **b.** detail of the gradation in crystallization between microcrystalline calcite and calcite spar, as well as within pellets, which are the allochemical components in the sample; **c.** and **d.** iron oxide filled fractures; **e.** and **f.** pores and cavities filled with calcite spar cement, partly dolomitized.

7.5. Sample M-5

Mudstone–wackstone of pellets[29] or pellsparite–pellmicrite[30-31]; mostly recrystallized (90%) (Fig. S15a). The orthochemical contents are microcrystalline calcite (<30%) and calcite spar (<70) (Fig. S15b). The allochemical components are pellets (<35%), foraminifera, such as rotalinids, Globigerina sp., Globotruncana sp., and Radiolaria, as well as various extrabasinal constituents, such as quartz and feldspar. Pellets are subrounded to rounded, about 50µm in diameter, and are dominantly dark-brown. The calcite spar cement is moderately dolomitized. Generally, some ankerite component is present in dolomite. Hematite particles are common with sizes that vary from 10µm to 200µm (Fig. S15c-d). Microfractures in three orientations contain calcite and aragonite (Fig. S15e). The thickness of the microfractures ranges from 30µm to 600µm, and

span, in length, between 15µm and 3cm (Fig. S15f). Several fractures are folded with wave heights (distance perpendicular to the axis of the waveform, from wave crest to adjacent wave trough) that measure 15µm to 3cm.

The microfacial characteristics in sample M-5 are similar to those for standard microfacies 3 and 4, suggesting deposition in deep shelf-margin of facies belt 3[32].

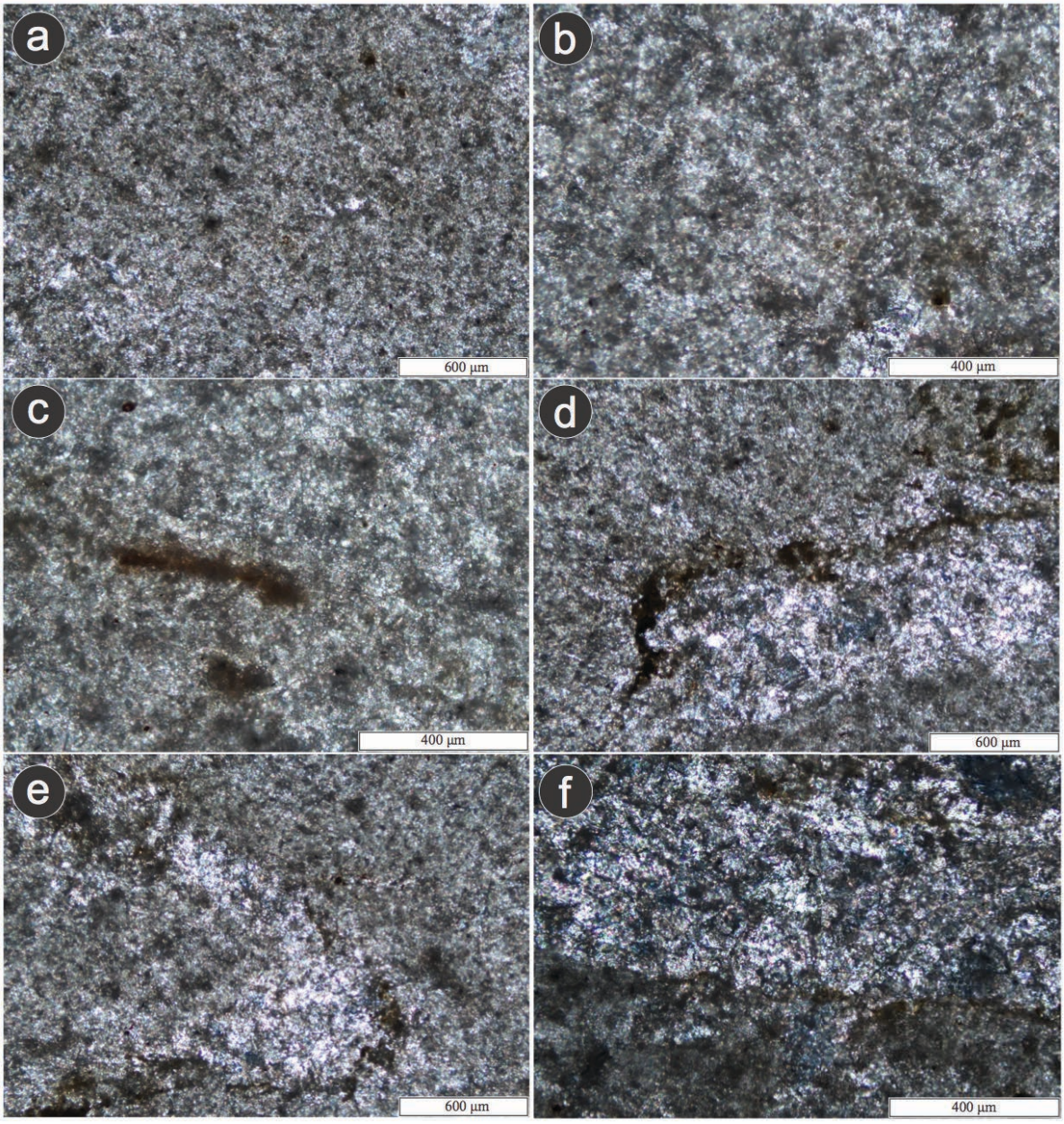


Fig. S15. Microphotograph in crossed polars for sample M-5. **a.** mudstone-wackestone with pellets highly recrystallized, containing microfractures filled with microcrystalline calcite; **b.** pellets and recrystallized grains; **c.** and **d.** calcite spar filling fractures in two directions; **e.** and **f.** poikilitic spar microstructures formed along with fractures.

7.6. Sample M-6

Packstone–grainstone of pellets and ooids[29] or pelsparite–pelmicrite[30-31]; highly recrystallized (95%) (Fig. S16a-b). The orthochemical constituents are microcrystalline calcite (<30%) and calcite spar (<70%). The allochemical components are pellets (<45%), ooids (<30), foraminifera

(globigerinids and Globotruncana sp.) and some extrabasinal minerals, such as quartz and feldspar. Pellets and ooids are subrounded to rounded, with a leaf-like or roller form, and sizes vary between 200µm and 450µm (Fig. S16c-d). The nucleus in ooids is mainly of extrabasinal quartz and lithic fragments, which are highly deformed. Flakes of clay minerals and/or mica align parallel to each other and create pseudo-bedding (Fig. S16c-d). The calcite spar cement is mildly dolomitized and ankeritized. Iron oxides (probable hematite) are present with dimensions near 10µm to 200µm. Microfractures come in three orientations and filled with calcite or aragonite (Fig. S16e-f). The thickness of the fractures varies from 30µm to 600µm. Veins are oriented in the same way as the deformation of pellets and ooids.

The microfacies characteristics in sample M-6 are similar to those in standard microfacies 11, 12, 13, and 15, suggesting deposition in deep shelf-margin of facies belt 3[32].

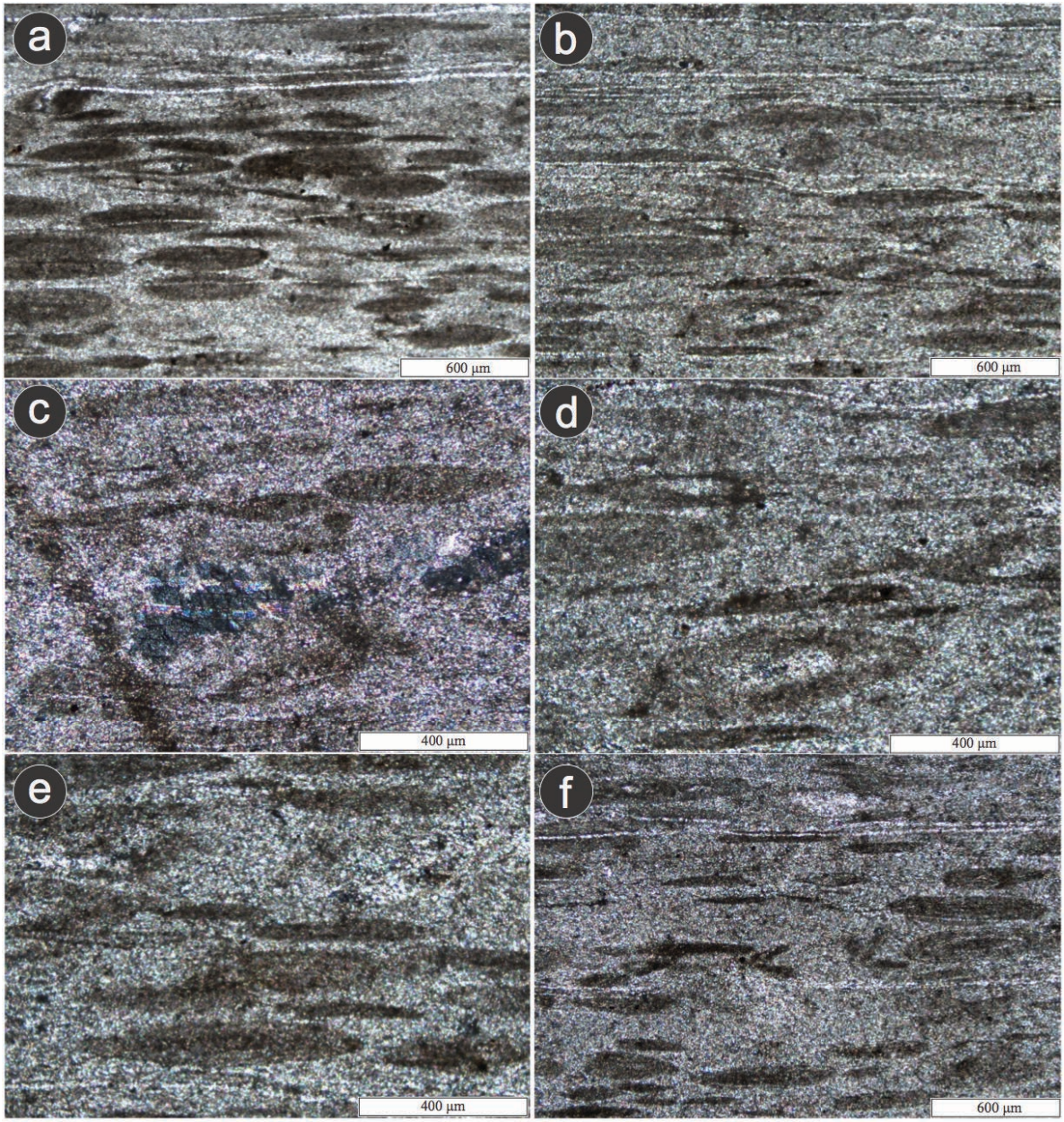


Fig. S16. Microphotograph in crossed polars for sample M-6. **a.** and **b.** packstone-grainstone with pellets and ooids, highly recrystallized with some shear planes; **c.**, **d.**, **e.** and **f.** allochemical constituents (pellets and ooids) highly deformed and recrystallized, notice the shear planes cutting the grains.

7.7. Sample M-7

Mudstone–wackestone with pellets[29] or pelsparite–pelmicrite[30-31]. The limestone in the sample is mildly recrystallized, either by diagenesis or tectonic deformation (30%) (Fig. S17a-b). Orthochemical components are microcrystalline calcite (<60%) and calcite spar (<40%). Cement is calcite spar, partly dolomitized, and ankeritized. The allochemical contents are pellets (<15%), foraminifera (rotalinids, globigerinids, Globotruncana sp., and Radiolaria) and various extrabasinal minerals, such as quartz and feldspar. Pellets are subrounded to rounded and platy in sizes near 30µm. These are of a dark brown color (Fig. S17b). Various iron oxides (probable hematite) are common in sizes that vary from 10µm to 150µm (Fig. S17b-c). Fractures are in three orientations and contain calcite or aragonite (Fig. S17d-f). The thickness of the microfractures varies from 20µm a 600µm, and their lengths range from approximately 15µm to 3cm. Some micro-fractures are folded, with wave heights (distance perpendicular to the axis of the waveform, from wave crest to adjacent wave trough) that measure 500µm to 3cm. Intercrystalline porosity shear planes are present. Pores and cavities are filled with aragonite and hematite, and measure from 200µm to 900µm in diameter (Fig. S17b-c). Some extraclasts exhibit cataclastic textures.

Microfacies characteristics in sample M-7 correspond to standard microfacies 2, 3, and 4, and are interpreted as deep shelf-margin deposits in facies belt 3[32].

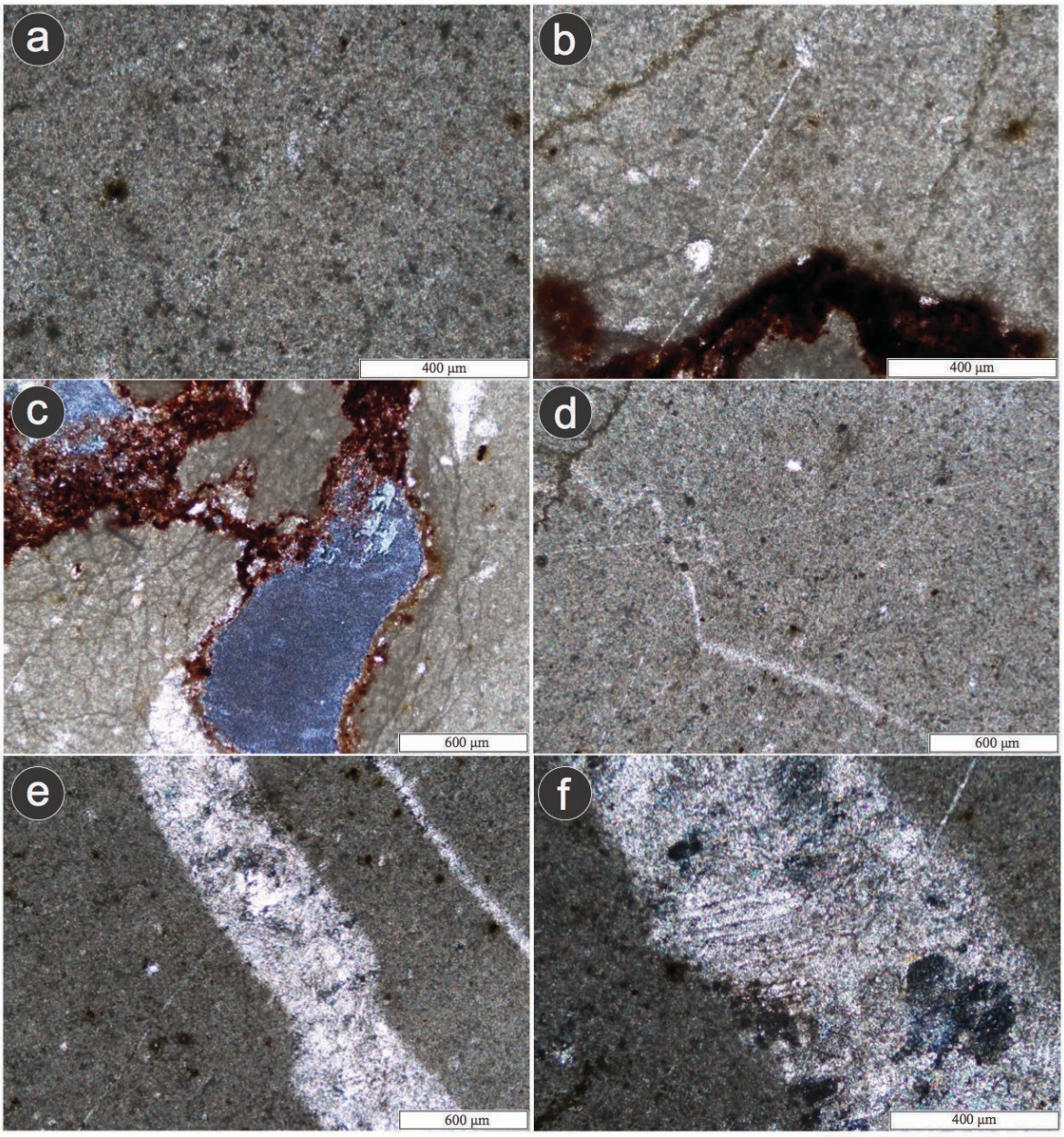


Fig. S17. Microphotographs in crossed polars of sample M-7. **a.** and **b.** mudstone-wackestone with pellets, mildly recrystallized, with cavities occupied by iron oxide; **c.** Iron oxides filling pores and cavities, and calcite spar partly dolomitized and ankeritized; **d.** calcite spar in fractures; **e.** and **f.** calcite spar partly dolomitized along with fractures.

7.8. Sample M-8

Mudstone–wackestone with pellets[29] or pelmicrite–pelsparite[30-31]. The sample is moderately recrystallized (50%). The orthochemical components are microcrystalline calcite (<60%) and calcite spar (<40%) (Fig. S18a-b). Cement is calcite spar, partly dolomitized, and ankeritized. The allochemical constituents are pellets (<35%), foraminifera (rotalinids, globigerinids, Globotruncana sp., and Radiolaria). Quartz and feldspar are the predominant extrabasinal minerals. Pellets are subrounded to rounded and leaf-like forms that average 30µm in diameter and are dark brown (Fig. S18b). Calcite spar cement is relatively dolomitized and ankeritized. Hematite is common in sizes that vary from 10µm to 150µm. Microfractures have three main orientations and are filled with calcite and aragonite. The width of fractures ranges from 20µm to 600µm, and their length is 15µm to 3cm (Fig. S18c-f).

The microfacial characteristics in sample M-8 are similar to the ones in standard microfacies 8, 9, 10, which suggests deep shelf margin deposits in facies belt 2[32].

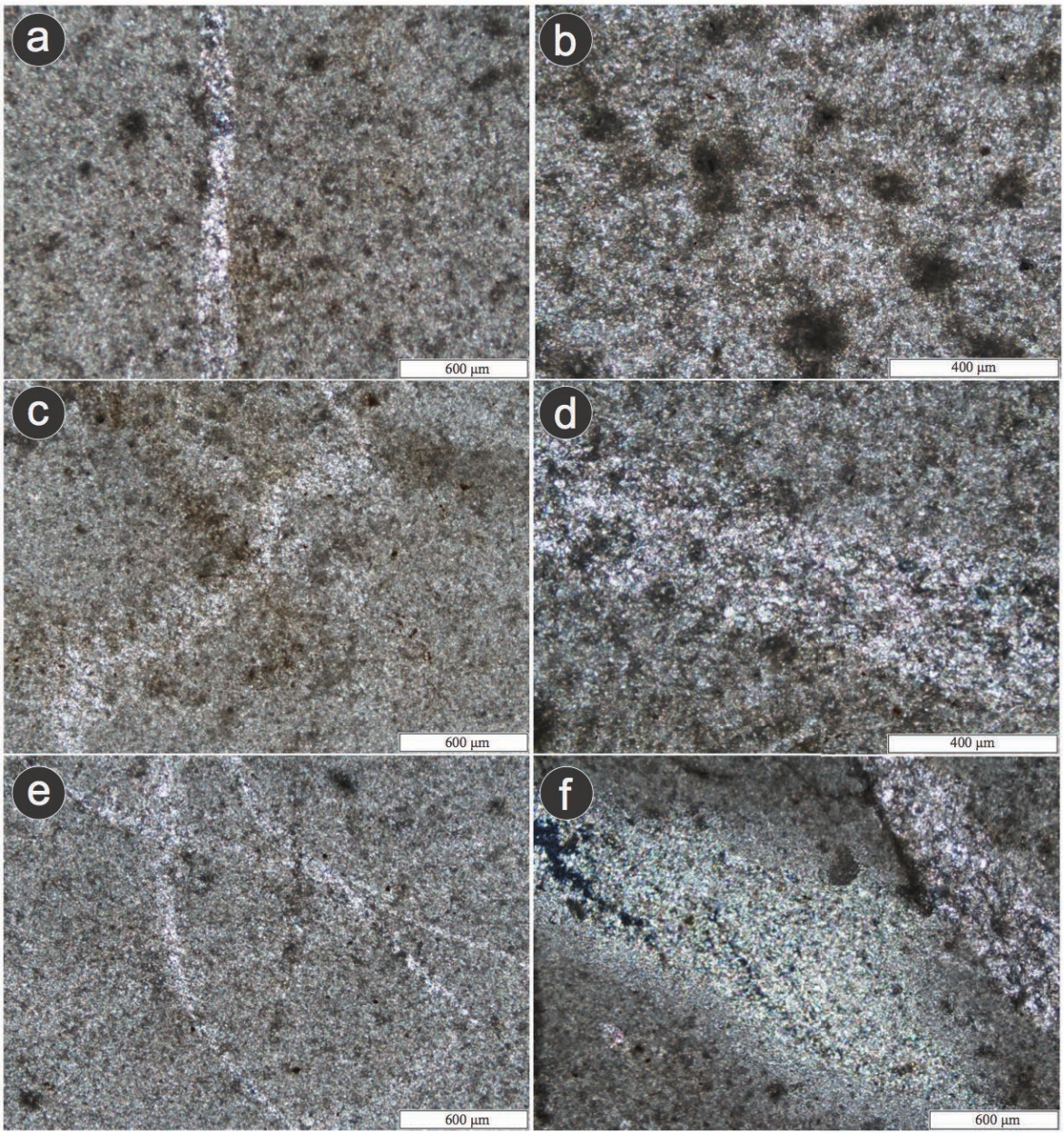


Fig. S18. Microphotograph in crossed polars for sample M-8. **a.** mudstone-wackestone with pellets, mild recrystallization with microfractures filled with microcrystalline calcite and aragonite **b.** allochemical and orthochemical components show a gradation in degree of recrystallization; **c.**, **d.**, **e.**, and **f.** detail of microfractures filled with aragonite and microcrystalline calcite.

7.9. Sample M-9

Packstone–grainstone with pellets and ooids[29] or pelmicrite–pelsparite[30,31]; highly recrystallized (90%) (Fig. S19a-b). The orthochemical components are microcrystalline calcite (<30%) and calcite spar (<70%). The allochemical components are pellets (<65%), ooids (<10), planktonic foraminifera (globigerinids and Globotruncana sp.), and extrabasinal minerals, such as quartz and feldspar. Pellets and scarce ooids are subrounded to rounded, platy, or with roller shapes. Their mean size is from 100µm to 450µm, and their color is dark brown. The nucleus of the ooids is commonly extrabasinal quartz and lithic fragments. Ooids are highly deformed. Significant

amounts of fecal material, i.e., pellets and peloids (any other micritic pellet-like forms), are found in the thin section. Pellets consist mainly of a petrographically distinct form called Fabreina, a type of nektic crustacean. The calcite spar cement is moderately dolomitized and ankeritized (Fig. S19c). Hematite oxides vary from 10µm to 150µm in diameter. Microfractures are in three orientations filled with calcite and aragonite (Fig. S19c-d). Fracture thickness is 30µm to 600µm, and the length is from 15µm to 3cm. Fractures are orientated in the same direction as the deformation in pellets and ooids (Fig. S19c-f).

The characteristics of the microfacies in sample M-9 are similar to the standard microfacies 11, 12, 13, and 15[32], which allow us to interpret the limestone as being deposited in a deep shelf-margin environment in facies belt 6[32].

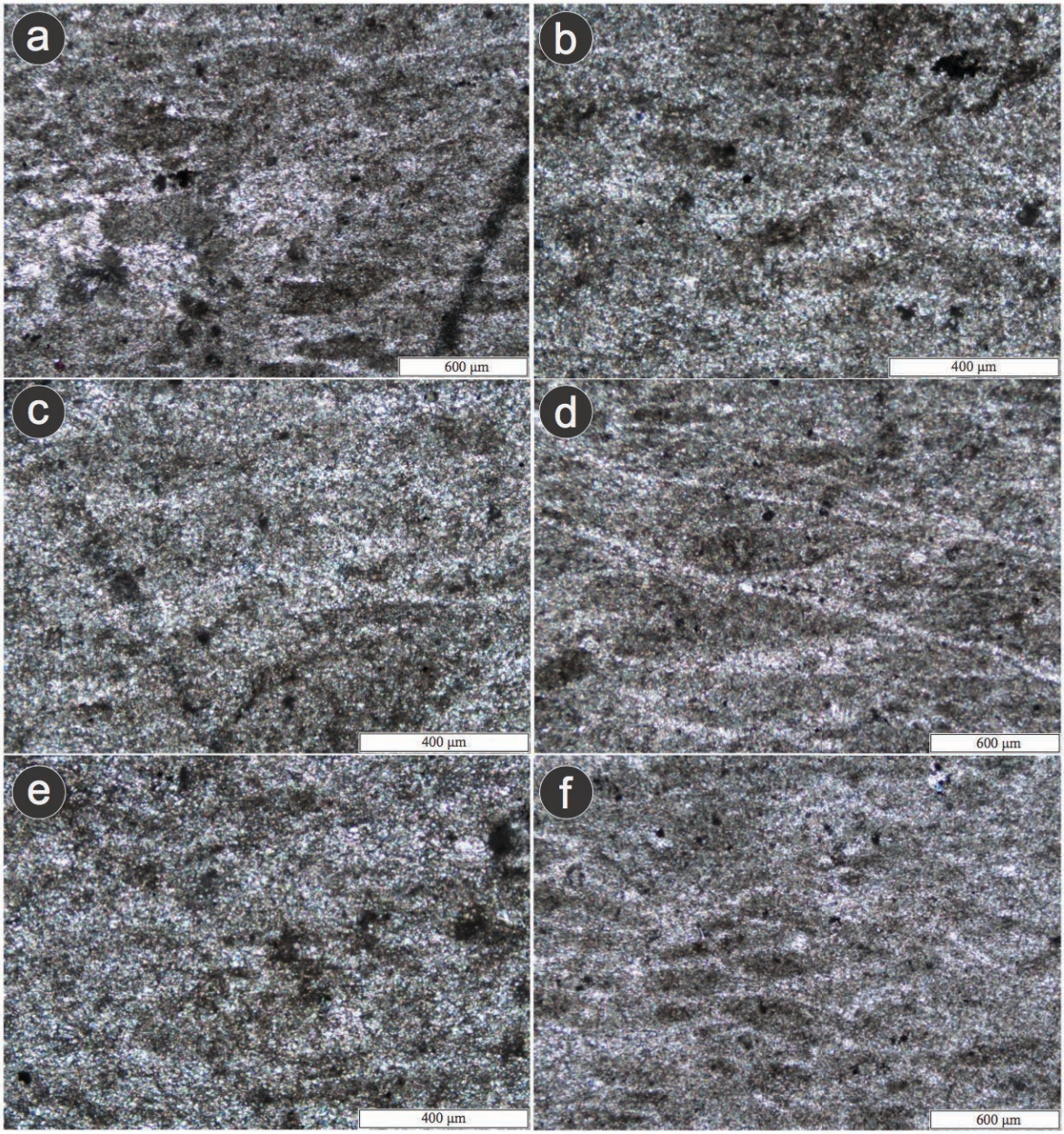


Fig. S19. Microphotograph in crossed polars for sample M-9. **a.** and **b.** packstone-grainstone with pellets and ooids, highly recrystallized with microfractures filled with microcrystalline calcite and aragonite; **c.** detail of a grade of crystallization in the components marked by sizable white aragonite and calcite crystals; **d.** allochemical "oolites" components are cut and reorganized by shear planes; **e.** and **f.** detail of microcrystalline calcite replacing calcite spar denoted by increasing size on white crystals.

7.10. Sample M-10

Mudstone–wackestone of pellets[29] or pelsparite–pelmicrite[30,31], with mild recrystallization (50%) (Fig. S20a-b). Recrystallization is either by diagenesis or deformation. The orthochemical components are microcrystalline calcite (<50%) and calcite spar (<50%). The calcite spar cement is partly dolomitized, and ankeritized. The allochemical components are pellets (<25%), foraminifera (rotalinids, globigerinids, Globotruncana sp., and Radiolaria), and extrabasinal minerals—mainly quartz and feldspar. Pellets are dark brown, subrounded to rounded, and have a leaf-like form with mean sizes of 30µm (Fig. S20b). Cement is calcite spar, moderately dolomitized and ankeritized. Several iron oxide minerals, such as hematite, are common with sizes that vary from 10µm to 130µm (Fig. S20c). Fractures have three orientations and are filled with calcite or aragonite. The thickness of fractures ranges from 20µm to 450µm, with lengths of 15µm to 2cm. Fractures are folded or bent, with wave heights (distance perpendicular to the axis of the waveform, from wave crest to adjacent wave trough) that measure 500µm to 2.5cm (Fig. S20d-f). Cavities range in size from 200µm to 900µm, and are filled with aragonite and hematite. Shear planes are common in multiple directions.

The microfacial characteristics in sample M-10 are similar to those reported for standard microfacies 2, 3, and 4, which relate to deep shelf-margin deposits within facies belt 3[32].

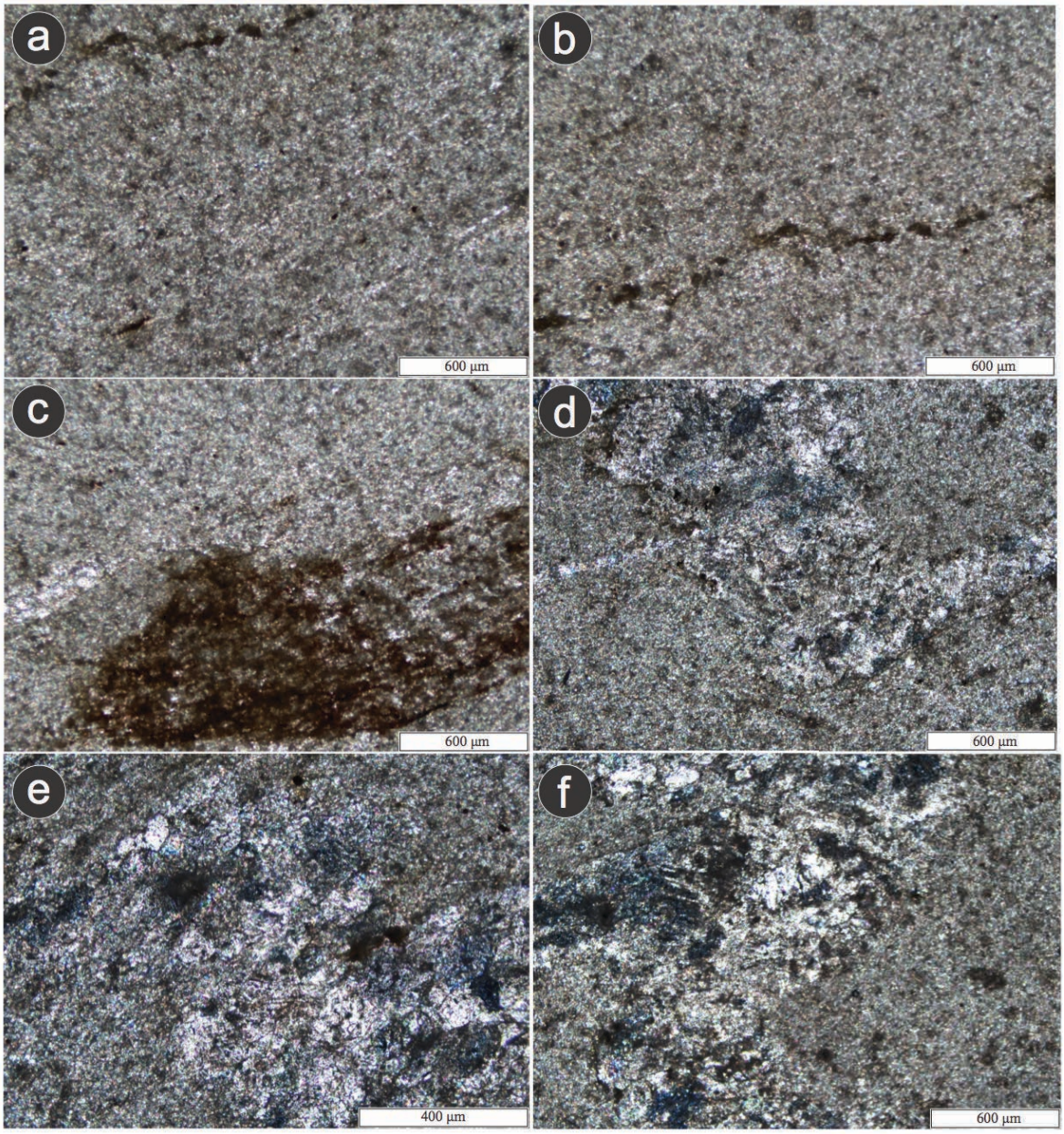


Fig. S20. Microphotograph in crossed polars for sample M-10. **a.** and **b.** mudstone-wackestone with pellets with mild recrystallization and calcite spar cement, partly dolomitized and ankeritized; **c.** iron oxides replace calcite cement; **d.** cavities filled with partly dolomitized cement; **e.** and **f.** folded fractures filled with calcite and aragonite.

7.11. Sample M-11

Mudstone–wackestone with peloids[29] or pelsparite–pelmicrite[30,31]; moderately recrystallized (50%) (Fig. S21a-b). Peloids are structureless or micritic intraclasts (Fig. S21b). The orthochemical components are microcrystalline calcite (<60%) and calcite spar (<40%). Cement is calcite spar, partly dolomitized, and ankeritized. The allochemical components are pellets (<45%), foraminifera (rotalinids, globigerinids, Globotruncana sp., and Radiolaria) and several extrabasinal minerals, such as quartz and feldspar. Pellets are dark brown, subrounded to rounded and platy, in sizes near 50µm (Fig. S21b). Various iron oxides such as hematite are common in sizes from 20µm to 100µm (Fig. S21c). Fractures are arranged in three orientations and filled with calcite and aragonite (Fig. S21d). The thickness of fractures varies from 20µm to 900µm, and lengths of 15µm to 2cm. Some fractures appear in fold crests and wave heights (distance

perpendicular to the axis of the waveform, from wave crest to adjacent wave trough) that measure 500µm to 2.5cm. Cavities are filled with aragonite and hematite, with sizes that range from 200µm to 900µm (Fig. S21e-f). Various shear planes are present.

The microfacial characteristics in sample M-11 correspond to those in standard microfacies 2, 3, and 4, occurring in deep shelf-margin deposits within facies belt 3[32].

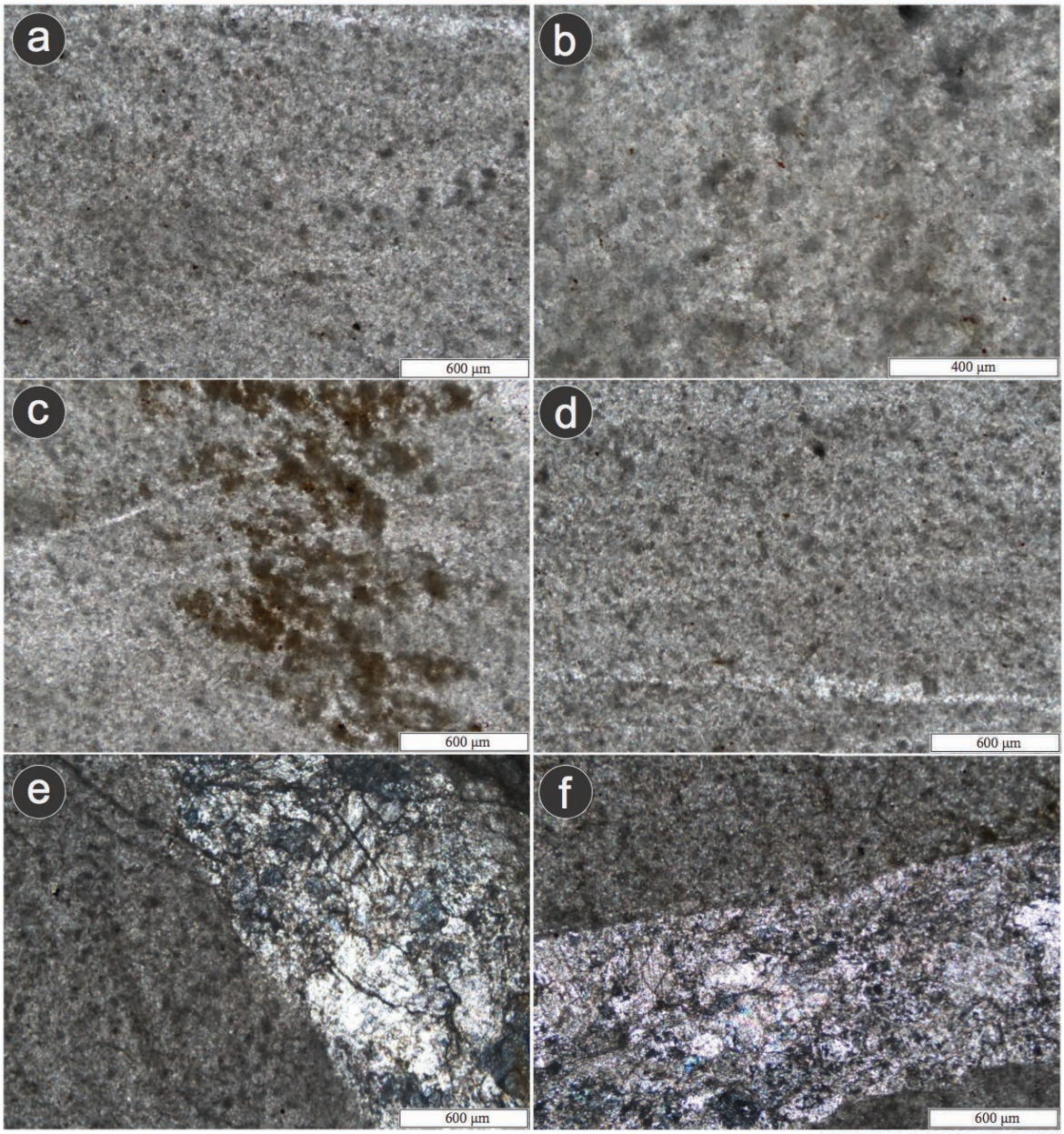


Fig. S21. Microphotograph in crossed polars for sample M-11. **a.** mudstone-wackestone with peloids, mild recrystallization, and calcite spar cement that is partly dolomitized and ankeritized; **b.** peloids and cement with mild recrystallization; **c.** iron oxides replacing calcite cement and fractures filled with calcite; **d.** fractures filled with calcite and aragonite; **e.** cavity filled with calcite and aragonite; **f.** detail of calcite and aragonite filling a fracture.

7.12. Sample M-12

Mudstone–wackestone of peloids[29] or pelsparite–pelmicrite[30,31]; moderately recrystallized (50%) (Fig. S22a-b). Peloids are silt- to sand-size aggregates of microcrystalline calcium carbonate that

lack internal structure. Orthochemical components are microcrystalline calcite (<60%) and calcite spar (<40%). Cement is calcite spar, partly dolomitized, and ankeritized. The allochemical components are pellets (<45%), foraminifera (rotalinids, globigerinids, Globotruncana sp., and Radiolaria), and extrabasinal minerals, such as quartz and feldspar. Pellets are subrounded to rounded, with leaf-like forms. These are approximately 50µm and of a dark-brown color (Fig. S22b). Some other peloids are ellipsoid to roughly spherical shape. These are rather uniform in size and built up by coarse silt to very fine sand size (Fig. S22c-d). Various iron oxide particles are present, e.g., hematite in sizes that go from 20µm to 100µm (Fig. S22e). Microfractures appear in three orientations filled with calcite and aragonite. The thickness of the fractures varies from 20µm to 900µm, and lengths range from 15µm to 2cm (Fig. S22e-f). Some fractures are folded or bent, with wave heights (distance perpendicular to the axis of the waveform, from wave crest to adjacent wave trough) that vary from 500µm to 2.5cm. Cavities contain aragonite and iron-rich clay in sizes that range from 200µm to 900µm. Several shear planes are also in the sample.

The microfacies characteristics in sample M-12 are similar to those from standard microfacies 2, 3, and 4, corresponding to deep shelf-margin deposits within facies belt 3[32].

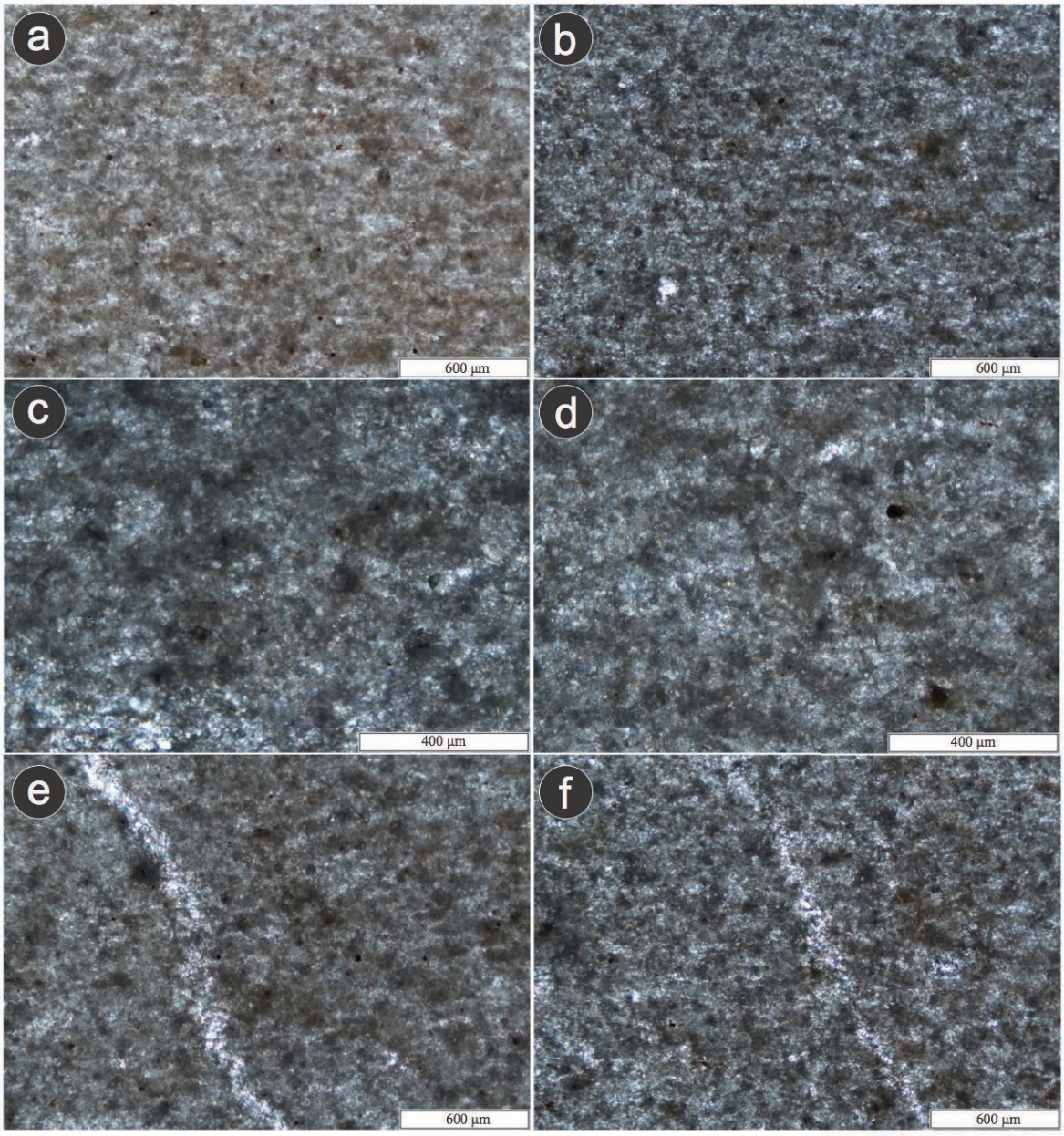


Fig. S22. Microphotograph in crossed polars for sample M-12. **a.** and **b.** mudstone-wackestone with peloids, showing mild recrystallization with calcite spar cement that is partly dolomitized and ankeritized; **c.** and **d.** detail on the grade of recrystallization in the cement and peloids denoted by the occurrence of calcite and aragonite large crystals; **e** and **f.** fractures filled with microcrystalline calcite and aragonite that cut the entire sample. Note the aggregates, some of which are iron oxides.

7.13. Sample M-13

Mudstone–wackestone of pellets[29] or pelsparite–pelmicrite[30,31]; with high recrystallization (90%) (Fig. S23a-b). The limestone is mildly recrystallized either by diagenesis or deformation. The orthochemical components are microcrystalline calcite (<30%) and highly recrystallized calcite spar (<70%). The cement is partly dolomitized and ankeritized calcite spar. The allochemical components are pellets (<25%), shale intraclasts, foraminifera (rotalinids, globigerinids, Globotruncana sp., and Radiolaria) and extrabasinal minerals, such as quartz and feldspar. Pellets are dark brown subrounded to rounded, with platy forms with mean sizes of 30µm. Shale intraclasts are highly deformed (porphyroclasts) with mean sizes of 300µm by 300µm, showing δ and θ deformation textures[33] (Fig. S23b, and S23c, respectively), with pressure shadows or

fringes that suggest sinistral displacement related to tectonic collision processes. The nuclei of porphyroclasts show blastesis denoted by the growth of dolomite (Fig. S23d-e). Cementing overgrowths are present in strain shadows. Flakes of clay minerals and/or mica are aligned in parallel. The calcite spar cement is moderately dolomitized and ankeritized. Several iron oxides, such as hematite, are common with sizes that vary from 10μm to 150μm. Microfractures are in three orientations filled with calcite and aragonite. The thickness of fractures varies in size from 20μm to 600μm, and lengths range from 15μm to 3cm. Most of the veins exhibit folding in the sense of strain on the deformation textures. The wave height (distance perpendicular to the axis of the waveform, from wave crest to adjacent wave trough) of the fold varies from 500μm to 3cm. Microcavities are filled with relatively coarse-grained mosaics of authigenic calcite or fined-grained carbonate spar, aragonite, and hematite, in sizes that vary from 200μm to 900μm. Several shear planes are also common. The edge of the sample is highly recrystallized and oxidized, probably a result of chemical alteration that allowed the precipitation of iron oxides (Fig. S23f).

The microfacial characteristics in sample M-13 are similar to the standard microfacies 2, 3, and 4, interpreted as deep shelf-margin deposition within facies belt 3[32].

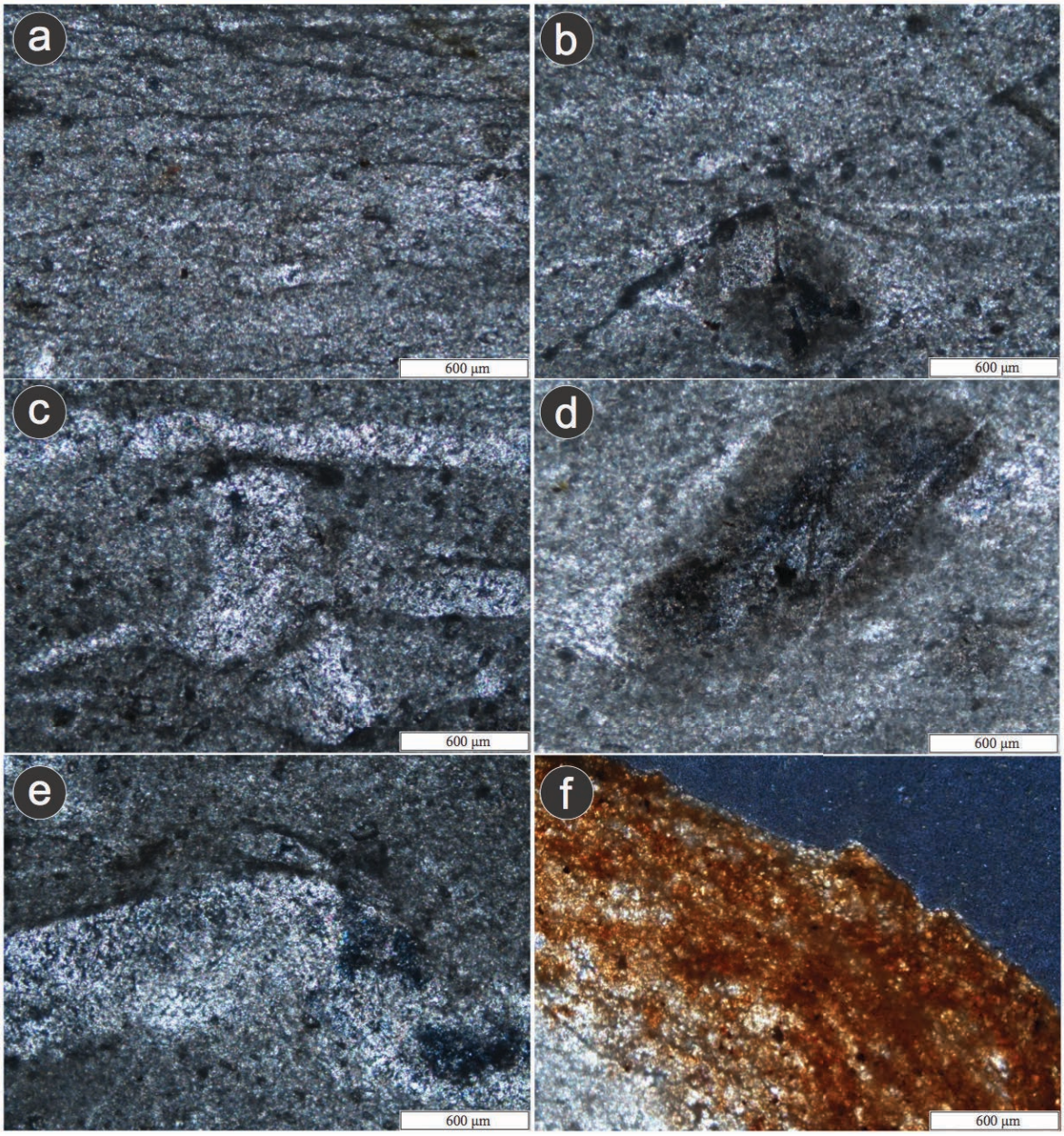


Fig. S23. Microphotograph in crossed polars for sample M-13. **a.** highly recrystallized and shearing mudstone-wackestone with pellets and calcite spar cement that is partly dolomitized and ankeritized in its shear planes; **b.** porphyroclasts type *δ* (Delta) with pressure shadows or fringes filled with calcite; **c.** porphyroclasts type θ (Theta) with pressure shadows filled with calcite; **d.** porphyroclasts showing blastesis; **e.** calcite spar filling cavities; **f.** clast highly oxidized.

7.14. Sample M-14

This sample is the first from a recent group of three samples extracted in 2019 in order to provide additional assessments for the aloctonous nature of the raw materials used for the manufacture of lithic artefacts. It comes from the limestone dintel above the current entrance to the cave (see Fig. S24a).

Mudstone–wackestone with pellets[29], or pelsparite-pelmicrite[30,31], highly recrystallized (Fig. S24b). The orthochemical content is microcrystalline calcite (<40%) and calcite spar (<60%). The allochemical components are pellets (<25%), ooids with quartz nuclei (<2%) (Fig. S24c-d), foraminifera (globigerinids, *Globotruncana* sp., and Radiolaria), and other extrabasinal minerals

such as quartz and feldspar are common. Pellets are subrounded to rounded with equant particles or as an elongated leaf-like shape with mean sizes of 40µm, light to dark brown colors prevail. Cement is calcite spar, moderately dolomitized, and ankeritized. Various iron oxides similar to hematite are typical filling in pores or as patches in sizes go from 10µm to 300µm (Fig. S24e). Fractures arrange in three orientations, filled by calcite and aragonite. Thicknesses of the fractures vary from 25µm to 650µm, and extend from 20µm to 3cm. Tectonic microstylolite seams cut some fracture systems. Iron oxides fill the shear planes that cut fractures (Fig. S24f). Some pores contain iron oxide and spar calcite, presenting deformation by the shear planes.

The microfacies characteristics are comparable to standard microfacies 3 and 4, interpreted as deep shelf-margin deposits within facies belt 3[32].

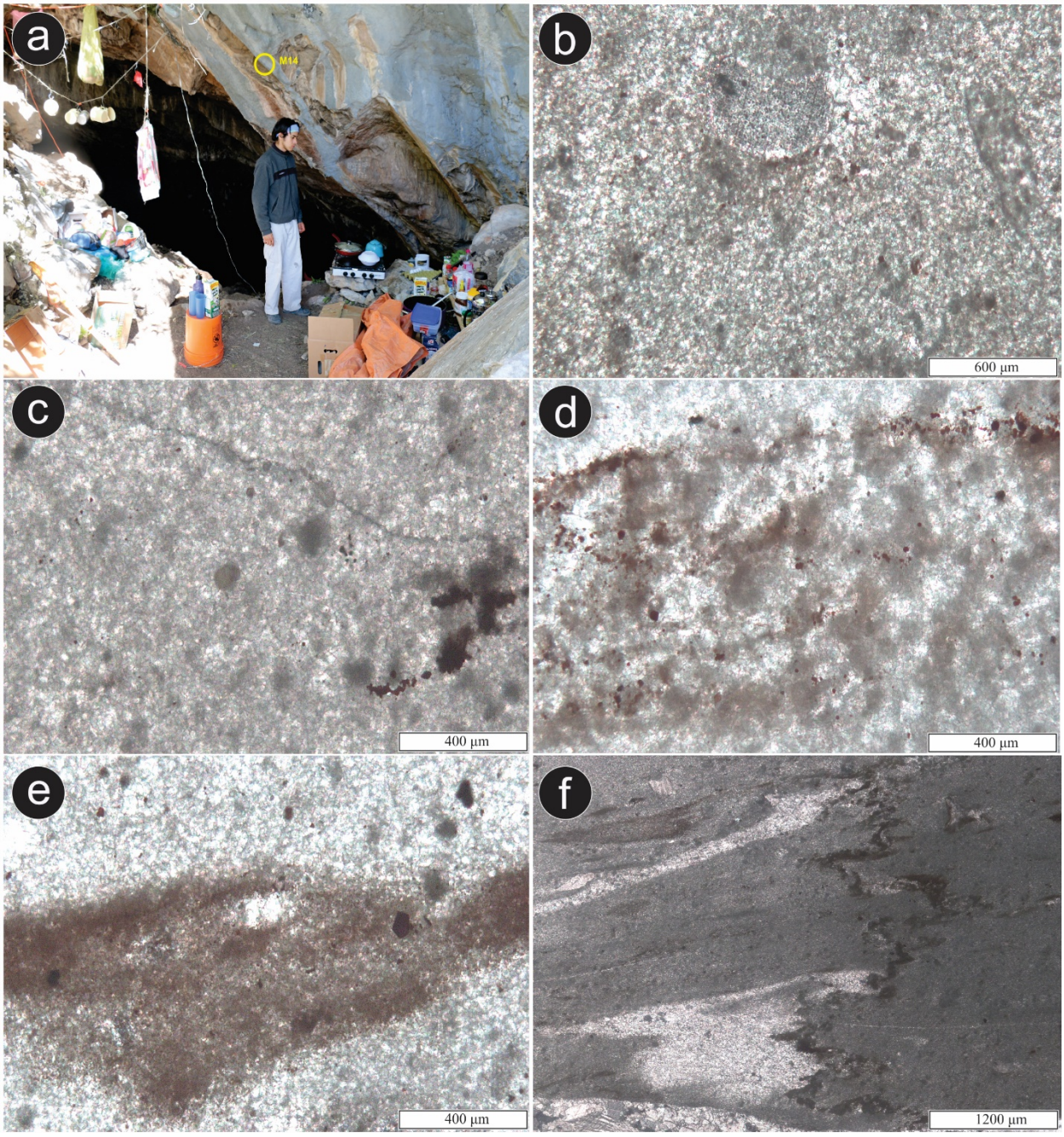


Fig. S24. **a**. the yellow circle indicates the location for sample M-14 at the entry of Chiquihuite Cave, in the state of Zacatecas, Mexico. **b**., **c**., **d**., **e**., and **f**., microphotographs in crossed polars for sample M-14.

b. and **c**. Mudstone-Wackstone with pellets, mild to a high grade of crystallization, and microfractures filled by calcite; **d**. and **e**. detail from recrystallized calcite, calcite spar, and iron oxides replacement; and **f**. detail of tectonic microstylolite filled by iron oxides that cut fractures.

7.15. Sample M-15

This sample has been extracted in 2019 from the eastern wall of the cave, close to the eastern end of the excavation X-12 (Fig. S25a).

Wackstone-packstone with pellets and ooids[29] or pelsparite-pelmicrite[30,31], highly recrystallized (Fig. S25b). Orthochemical constituents are microcrystalline calcite (<30%) and spar calcite highly recrystallized (<70%). Cement is spar calcite, partly dolomitized and ankeritized. The allochemical content is pellets (<25%), ooids with quartz nuclei strongly deformed, foraminifera (globigerinids, Globotruncana sp., and Radiolaria), different extrabasinal minerals such as quartz and feldspar (Figs. S25c-e). Pellets are subrounded to rounded, sometimes as leaf-like structures 35µm in size, displaying light to dark brown color. Various iron oxides (probable hematite) are common in sizes that vary from 15µm to 130µm. Pores and cavities are filled with spar calcite as patches, extending 50µm to 300µm long. Blastesis develops porphyroclasts inside the pores, including type δ and type θ[33]. Porphyroclasts present a sinistral relative shear direction or left-lateral motion. Fractures occur in three orientations filled with calcite and aragonite. The thicknesses of fractures vary from 15µm to 550µm, and extend from 20µm up to 3cm long. A group of fractures is deformed accordingly to the cinematic indicators from δ and θ[33] (Fig. S25e); the amplitudes in its fold crests measure 500µm to 3cm. Microcavities contain spar calcite, aragonite, and hematite, and vary in size from 200µm to 900µm (Fig. S25e). Shear planes are frequent (Fig. S25e). The edge of the sample is highly recrystallized and oxidize, possibly as a result of chemical weathering that allows hematite to precipitate.

The microfacies characteristics are similar to those in standard microfacies 2, 3, and 4, suggesting deep shelf-margin deposits within facies belt 3[32].

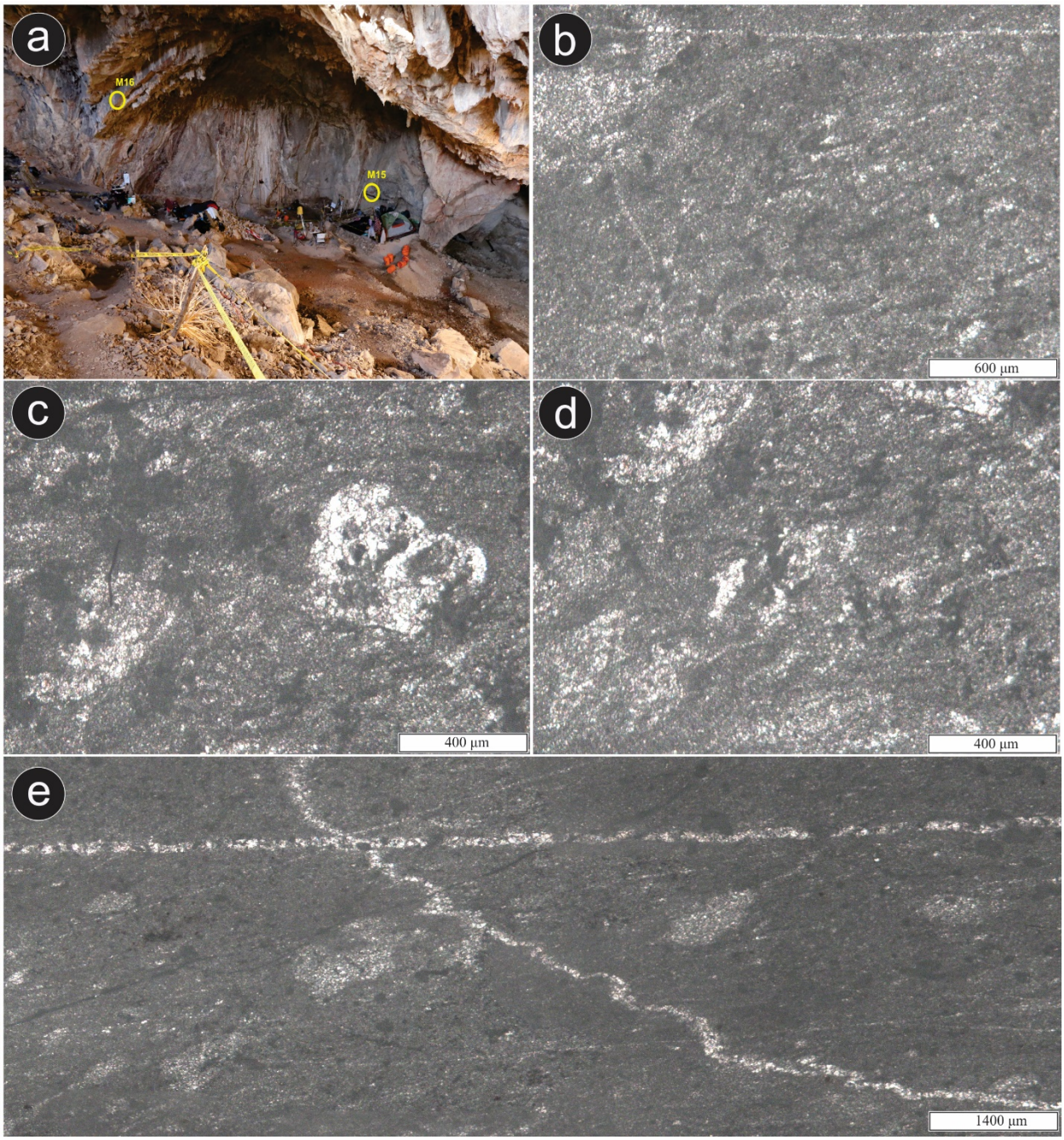


Fig. S25. **a**. Location (yellow circle) for the sample M-15 on the eastern wall of the main gallery at Chiquihuite Cave. **b**., **c**., **d**., and **e**. microphotographs in crossed polars of sample M-15; **b**. Wackstone-packstone with pellets and ooids highly recrystallized with microfractures filled by calcite; **c**. and **d**. Detail of a grade of crystallization in microcrystalline calcite, spar calcite, and other allochemical components such as in pellets; **e**. Panoramic view of a wackstone-packstone with pellets highly recrystallized, and calcite spar in fractures with three orientations.

7.16. Sample M-16

This sample has been extracted in 2019 from the roof of the main gallery, from an overhang above the western end of the dig X-12 (Fig. S26a).

Mudstone-Wackestone with pellets[29] or pelsparite-pelmicrite[30,31] highly recrystallized (Fig. S26b-e). The orthochemical content is microcrystalline calcite (50%) and spar calcite (50%). The

allochemical content includes pellets (<25%), foraminifera (globigerinids, Globotruncana sp., and Radiolaria), and some extrabasinal minerals like quartz and feldspar (Figs. S26c-d). Pellets are subrounded to rounded, equating in size and similar in form to a leaf, their mean size is 35µm, and occur in light and dark brown color. Cement is spar calcite, which is partly dolomitized and ankeritized (Fig. S26d). Some iron oxides like hematite are commonly filling pores or arranged in patches 10µm to 300µm in extent. Fractures array in four orientations filled by calcite and aragonite. The thickness of fractures varies from 20µm to 750µm, as long as 20µm to 3cm. One group of fractures is folded, with a spacing of 20µm to 3cm in between its fold crests; this cluster of fractures folds any other group of fractures (Fig. S26e). Iron oxides fill shear planes, which cut multiple fracture systems. Some microcavities hold iron and spar calcite, also deformed by the shear planes.

The microfacies characteristics correspond to microfacies 2, 3, and 4, indicating deep shelf-margin deposits as in facies belt 3[32].

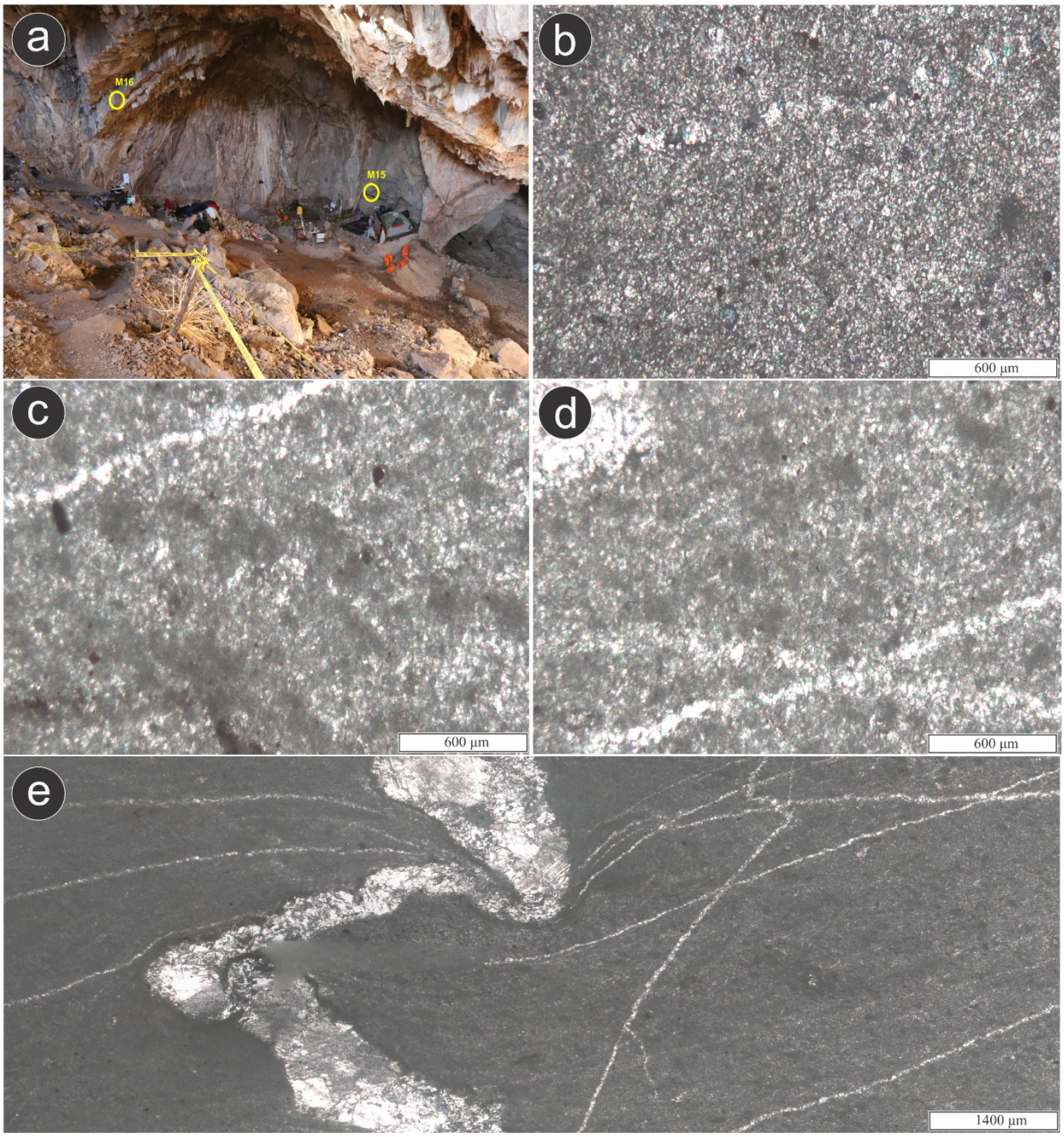


Fig. S26. **a**. Location (yellow circle) for sample M-16, inside the main chamber of Chiquihuite Cave. **b**., **c**., **d**., and **e**. microphrotographs in crossed polars for sample M-16; **b**. mudstone-wackstone with pellets highly recrystallized, showing fractures filled by calcite; **c**. and **d**. detail of the grade of crystallization in microcrystalline calcite and calcite spar, which is also present in pellets; **e**. photo mosaic showing several clusters of fractures filled by spar calcite, some fractures are folded and cut or deform other thinner microfractures.

7.17. Synthesis of appearance and similarities amongst samples.

The petrographical, microtextural, and qualitative analyses of fifteen samples (Table S8, Figs. S27, S28) contribute data for a cluster analysis using the methods of single linkage, by considering the Euclidian distances[34]. The cluster analysis differentiates among five groups related by physical features (such as texture or grade of recrystallization), textural characteristics (such as allochemical or orthochemical contents), and post-depositional structures (such as shear planes).

Group I –samples M-8, M-10, M-11, and M-12 show mild recrystallization, often the presence of shear planes and microfolds, their corresponding textures of Mudstone-Wackstone (Fig. S27).

Group II–samples M-3 and M-7 present low recrystallization, cavities, and pores are slightly developed; however, fractures are almost absent, restricting the formation of shear planes and microfolding (Fig. S27).

Group III–samples M-5, M-13, M-14, M15, and M-16 prove a high grade of recrystallization, well-developed shear planes, and microfolds with textures similar to Mudstone-Wackstone (Fig. S27).

Group IV–samples M-6 and M-9 exhibit high-grade of recrystallization, textures are Packstone-Grainstone, with scatter shear planes, and no cavities (Fig. S27).

Group V –samples M-2 and M-4 show mild to high recrystallization, pores, cavities remain sparse, and pellets are frequent (Fig. S27).

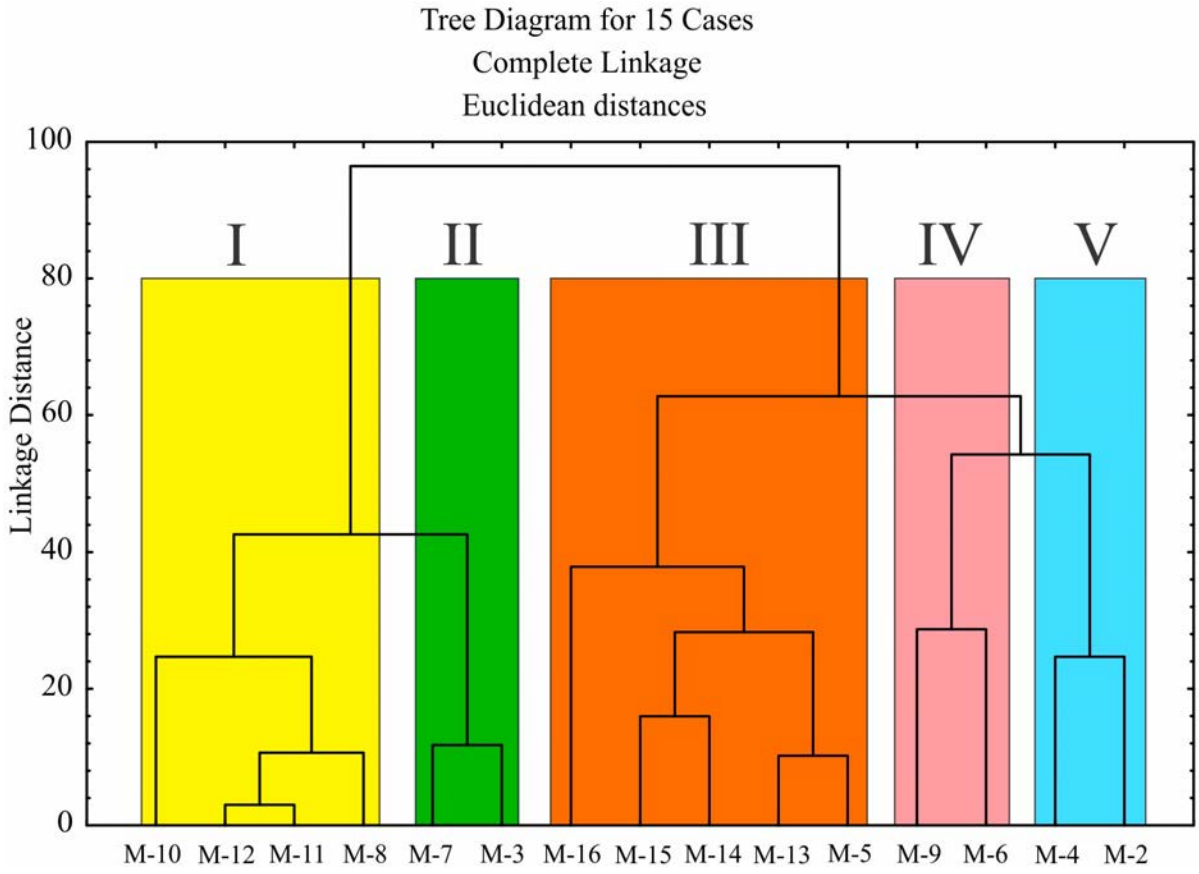


Fig. S27. Quantitative representation of assigning sampling units to groups. The petrographical descriptions, microtextural characterization, and quantitative analysis are the critical attributes for dissimilarities between groups.

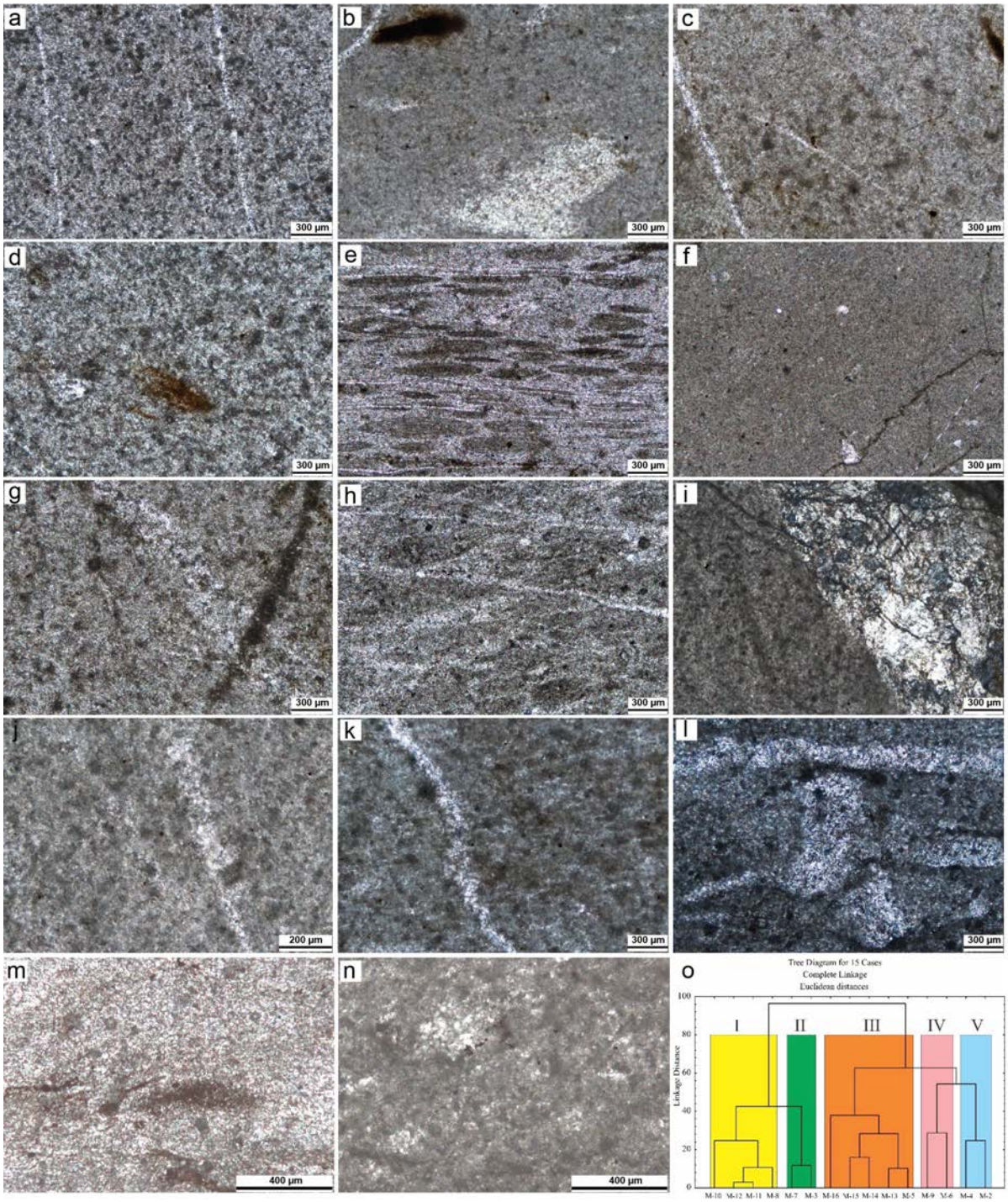
Table S8. List of similarities among samples collected in the Chiquihuite cave and nearby sites.

Folk (1959)	Dunhmam (1962)	Orthochemical content %	Allochemical content	Grade of recrystallization	Pores and cavities	Fractures	Deformational structures	ZF Wilson (1975)	Sample
Pelmicrite	Mudstone-Wackestone	Micrite (60%) Spatite (40%)	Pellets (60%)	90%	Absent (0)	Few (3)	Absent (0)	3	M-2

Pelmicrite-Pelspatite	Mudstone-Wackestone	Micrite (60%) Spatite (40%)	Pellets (15%)	20%	Absent (0)	Few (1)	Absent (0)	3	M-3
Pelmicrite	Mudstone-Wackestone	Micrite (50%) Spatite (50%)	Pellets (60%)	70%	Few (3)	Few (3)	Shear surfaces (1)	3	M-4
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (30%) Spatite (70%)	Pellets (35%)	90%	Absent (0)	Few (3)	Shear surfaces and micro folds (2)	3	M-5
Pelspatite-Pelmicrite	Packstone-Granistone	Micrite (30%) Spatite (70)	Pellets (45%) Ooids (30%)	95%	Absent (0)	Few (3)	Shear surfaces (1)	6	M-6
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (60%) Spatite (40%)	Pellets (15%)	30%	Few (3)	Moderately (6)	Shear surfaces and micro folds (2)	3	M-7
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (60%) Spatite (40%)	Pellets (35%)	50%	Absent (0)	Few (3)	Absent (0)	2	M-8
Pelspatite-Pelmicrite	Packstone-Grainstone	Micrite (30%) Spatite (70%)	Pellets (65%) Ooids (10%)	90%	Absent (0)	Few (3)	Shear surface (1)	6	M-9
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (50%) Spatite (50%)	Pellets (25%)	70%	Few (3)	Few (3)	Shear surface (1)	3	M-10
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (60%) Spatite (40%)	Pellets (45%)	50%	Absent (0)	Few (3)	Shear surfaces and micro folds (2)	3	M-11
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (60%) Spatite (40%)	Pellets (45%)	50%	Few (3)	Few (3)	Shear surfacees and micro folds (2)	3	M-12
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (30%) Spatite (70%)	Pellets (25%)	90%	Absent (0)	Few (3)	Shear surfaces, micro folds, and δ and θ cinematic indicators (4).	3	M-13
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (40%) Spatite (60%)	Pellets (15%) Ooids (10%)	80%	Absent (0)	Few (3)	Shear surfaces and micro folds (1)	3	M-14
Pelspatite-Pelmicrite	Wackestone-Packstone	Micrite (30%) Spatite (70%)	Pellets (20%) Ooids (15%)	80%	Few (1)	Few (3)	Shear surfaces and micro folds (3)	3	M-15
Pelspatite-Pelmicrite	Mudstone-Wackestone	Micrite (50%) Spatite (50%)	Pellets (10%)	90%	Few (2)	Few (4)	Shear surfaces and micro folds (4)	2	M-16

Fig. S28 (next page). Thin-section summary of results. **a**. M2 (gravel, stratum 1209). Mudstone-Wackestone with pellets (M-W-p), highly recrystallized, without deformation structures, fractures moderately developed, no cavities. **b**. M3 (gravel, 1212). M-W-p without deformation structures, low-grade of recrystallization, scarce fractures, cavities are absent. **c**. M4 (gravel, 1217). M- W-p highly recrystallized, accompanied by few deformation structures, including sparse fractures and cavities. **d**. M5 (gravel, 1219). M-W-p highly recrystallized, with moderate development of deformation structures, no cavities, scattered fractures. **e**. M6 (light grey pebble, outside slope). Packstone-Grainstone with pellets (P-G-p), lacking recrystallization, moderate development of deformation structures and fractures, without cavities or pores. **f**. M7 (dark grey pebble, outside slope). M-W-p, mild recrystallization, moderate development of deformation structures, fractures and, in lesser extent, cavities. **g**. M8 (greenish pebble, outside slope). M-W-p moderately recrystallized, without deformation structures or cavities, moderate development of fractures. **h**. M9 (artefact 1889-12698, dark limestone). P-G-p and ooids, highly recrystallized, deficient in deformation structures and fractures, in the absence of cavities. **i**. M10 (1866-12685, greenish limestone). M-W-p, mild-to-high recrystallization, scarce deformation structures, low in fractures, and sparse cavities. **j**. M11 (1899-12709, greenish limestone). M-W w/peloids, moderately recrystallized, moderate deformation structures and fractures; cavities are absent. **k**. M12 (1899-12710, pink-greenish limestone). M-W w/ peloids, moderately recrystallized, hardly any

deformations structures, moderate development of fractures and cavities. **l**. M13 (2110-12949, light grey limestone). M-W-p highly recrystallized, high presence of deformation structures, moderate development of fractures, free of cavities. **m**. M14 (limestone from the roof above the cave's entrance). M-W with pellets (Dunham, 1962), or pelsparite-pelmicrite (Folk, 1959; 1962), highly recrystallized. Cement is calcite spar, moderately dolomitized, and ankeritized. **n**. M15 (limestone sample from the eastern wall near the excavation). Wackstone-packstone with pellets and ooids (Dunham, 1962) or pelsparite-pelmicrite (Folk, 1959; 1962), highly recrystallized. Cement is spar calcite, partly dolomitized, and ankeritized. **o**. Euclidean linkage diagram showing the relationship between the 15 samples.



8. Commercial radiocarbon dating methods

8.1. PaleoResearch Institute

The following is an excerpt from a results report obtained for dated sample PRI-5414. It delineates the methods used by the PaleoResearch Institute.

A charcoal sample submitted for radiocarbon dating was identified and weighed prior to selecting a subsample for pre-treatment. Any remainder of the charred sample is curated permanently at PaleoResearch Institute. The subsample was vacuum freeze-dried, freezing out all moisture at -107 °C and < 10 millitorr. Then the sample was treated with cold pH 2 hydrochloric acid (HCl), followed by cold 6N HCl. The sample then was heated to approximately 110 °C while in 6N HCl. This step was repeated until the supernatant was clear. This step removes iron compounds and calcium carbonates that hamper humate compound removal. Next, the sample was subjected to 5% potassium hydroxide (KOH) to remove humates using both cold solutions and solutions that were heated. Once again, the sample was rinsed to neutral and re-acidified with pH 2 HCl between each KOH step. This step was repeated until the supernatant was clear, signaling removal of all humates, then was rinsed to neutral. After humate removal, the sample was made slightly acidic with pH2 HCl. Each sample was freeze-dried, then combined in a quartz tube with a specific ratio of cupric oxide (CuO) and elemental silver (Ag) in quantities based on the mass of carbon in the sample. The tubes were hydrogen flame-sealed under vacuum.

Standards and laboratory background wood samples were treated to the same acid and base processing as the charcoal sample of unknown age. A radiocarbon "dead" wood blank from the Grey Fossil site in Washington County, Tennessee, dated to the Hemphillian stage of the late Miocene, 4.5-7 MYA (currently beyond the detection capabilities of AMS) was used to calibrate the laboratory correction factor. In addition, standards of known age, such as the Third International Radiocarbon Inter-comparison (TIRI) Sample "B" (Belfast Pine) with a consensus age of 4503 ± 6, and TIRI Sample "J" (Bulston Crannog wood) with a consensus age of 1605 ± 8[35], are used to help establish the laboratory correction factor. After the requisite pre-treatment, a quantity similar to submitted samples of each wood standard was sealed in a quartz tube. Once all the wood standards, blanks, and submitted samples of unknown age were prepared and sealed in their individual quartz tubes, they were combusted at 820 °C, soaked for an extended period of time at that temperature, and allowed to cool slowly, enabling the chemical reaction that extracts carbon dioxide (CO_2) gas. Following this last step, the sample of unknown age, the wood standards, and the laboratory backgrounds were sent to The Center for Applied Isotope Studies in Athens, Georgia, where the CO_2 gas was processed into graphite. The graphitized samples

were placed in the target and run through the accelerator, generating numbers that are subsequently converted into radiocarbon dates.

8.2. Beta Analytic

The following is an excerpt from an email sent by R.E. Hatfield, President/Director of Beta Analytic, on June 14th, 2019, and received by C.F. Ardelean, delineating the pretreatment protocols used by the laboratory for samples Beta-345055, -436709, and -436710.

Bone Collagen (Beta-345055)

The pretreatment for the extraction of the bone collagen that was employed is composed of a proprietarily modified Longin Collagen Extraction Method (1971)[36], that we have developed in-house. The concentrations of the chemicals applied, duration and number of extractions are varied based on factors such as initial size, level of preservation, burial conditions (if known), and the observed level of reaction of the collagen extract to the pretreatment process as it is being performed. This is unique to each bone sample, so there is not a specific stepwise pretreatment regime, it must be modified for each sample based on our experience and observations.

In general, our process consists of an initial cleaning stage where the bone is washed and then physically cleaned by scraping as needed with a wire brush or abraded with a Dremel tool to remove any surface contamination (dirt, stains, surface debris, possible oils from prior handling, etc.). It is then placed on 0.2N Hydrochloric Acid (HCl) at ~ 21°C to dissolve the mineral fraction. After 12-24 hrs. in the initial HCl bath, the bones surface is again scraped to remove the outermost layers (size permitting) which may contain imbedded dirt or rootlet materials that may have penetrated below the bones surface during burial.

This material is generally discarded provided there is sufficient remaining bone for dating Over the course of several days, collagen was periodically scraped away as the surface mineral fraction dissolved. Once a sufficient amount of collagen was recovered, this step was terminated, and the collagen was rinsed to neutral. A solution of 1-2% alkali (50/50 wt/wt % NaOH) was carefully applied and reapplied under observation at room temperature until the solution remained clear (indicating effective removal of secondary organics such as humic acids). After rinsing to neutral, a final acid wash was applied to remove any adsorbed CO2. Throughout the process all roots, organic debris and minerals were eliminated. The purified collagen was then rinsed to neutral, dissected and microscopically examined for cleanliness and uniformity.

The clean gelatinous collagen extract is then dried by vacuum desiccation prior to combustion. The extracted collagen is then combusted to CO2 and the C13/12 ratio (and N15/14 if requested)

is checked to see if the value is consistent with the type of animal or human bone being analyzed. If the C13/12 ratio is consistent with the expected value (typically between -9 o/oo and -21 o/oo), and all other steps in the extraction process have proceeded normally, the CO2 from the combustion is then graphitized and AMS counted.

Organic Sediment (Beta-436709)

The sample was first visually inspected for size, homogeneity, debris, inclusions, clasts, grain size, organic constituents and potential contaminants. It was then dispersed in deionized H2O, homogenized through stirring and sonication and then sieved through a 180um sieve. The material passing through the sieve (< 180um) was used for the analysis. It was bathed in serial applications of 1.25 N HCl at 90C for a minimum of 2 hours at each application, to ensure the complete removal of any carbonates.

This was followed by serial deionized H2O rinses at 70C until neutrality was reached. Any debris or micro-rootlets smaller than < 180 um, were discarded during these rinses. After drying in an oven at 90C for 12- 24 hours, the dried sample was homogenized, and a representative sub-sample of the sediment was placed under a 45X microscope.

Concentrated HCl was applied to the representative sub-sample and the complete removal of any carbonate species was visually validated. Microscopic examination was performed on the remaining sample material to assess its characteristics and to determine the appropriate sub-sample size that would be suitable for d13C and AMS dating.

Charcoal (Beta-436710)

The sample was first visually inspected for size and durability. It was reduced to small particles (1-5 mm) through dissecting and crushing and saturated in de-ionized water at 70C. It was then soaked in 1N HCl for 1-2 hours and repeated, if needed, to eliminate any carbonates present. The sample was then rinsed to neutral with deionized H2O. A 1-2% alkali solution was then applied (50/50 wt/wt% NaOH) at 70C and the sample was allowed to soak for 2-4 hrs. The sample was rinsed multiple times with deionized H2O and the alkali applications and rinsing repeated until no color change was observed on the application of fresh NaOH. It was then rinsed to neutral with deionized H2O. A final hot acid wash (0.5-1.0 N HCL) was applied to ensure the alkali was neutralized and once again rinsed to neutral with deionized water. During this process all roots and organic debris were eliminated. The sample was dried at 100C or vacuum desiccated depending on its size and preservation level for 12-24 hours. It was then microscopically examined for cleanliness, uniformity and where applicable appropriately sub-sampled for the d13C and AMS measurements.

9. Optically stimulated luminescence (OSL) dating

The OSL tubes were opened under subdued amber laboratory lighting (low intensity LEDs with peak emission at 594 nm) and sample preparation involved removal of the light-exposed outer 2 cm from the ends and the extraction of sand-sized quartz mineral grains for OSL measurements from the inner, light-shielded core. Conventional preparation methods were used and all raw samples were then sieved in water to extract the 90-125, 125-180 and 180-250 µm grain size fractions. These were then treated with H2O2 (30%) to remove organics, HCl acid (10%) to dissolve carbonates and HF acid (45%) to dissolve feldspar minerals and to remove the outer (~10 µm) rind of quartz grains affected by alpha irradiation. A heavy liquid density separation using sodium polytungstate (2.62g/cm3) was then used to separate the quartz rich fraction from the heavy minerals. Finally, all samples were treated again with HCl (10%) to remove potential contaminant fluorides precipitated during the HF etching, followed by rinsing in demineralized water.

Due to the paucity of sand-sized quartz grains within the carbonate dominated cave sediments (see Supplementary Information SI1.1 and SI1.7.1.), OSL measurements were conducted on a wider grain size fraction (125-250µm) in order to increase the number of grains available for dating. Although, according to X-ray diffraction analyses (see Supplementary Information SI1.1) sand-sized quartz grains are the second most abundant mineral in the sediment (~6%), the vast majority of grains are dominated by calcite (~75-95%) with the common occurrence of illite/ muscovite (up to ~20%). Full elemental analysis by fusion ICP-MS also revealed very high concentrations of CaO3 (typically ~49%), high loss on ignition values (typically 39%), with low concentrations of SiO2 (typically ~6%) and consistent with a suite of mineral material derived primarily from the disintegration of limestone.

Grains were mounted either as circular multi-grain monolayers of 4-6mm diameter onto aluminium discs with a silicone oil adhesive (Viscasil 60,000) or as individual grains into gold-plated single-grain aluminium discs supplied by Risø National Laboratories (Denmark), capable of accommodating 100 individual grains inside circular cavities (ø = 300 µm).

Multi-grain OSL measurements were performed on a Lexsyg-Smart luminescence reader[35] manufactured by Freiberg Instruments (Germany). The instrument was fitted with a 90Sr/90Y ceramic disc β-source (~1.95GBq) allowing irradiations of the quartz grains with a dose rate of ~0.127Gy/sec and calibrated against a gamma-irradiated Risø National Laboratory standard[36]. For uniform optical excitation across the sample area, an OSL head unit fitted with 10 blue light emitting diodes (LEDs emitting at 458±10 nm; max. power 100mW/cm3) and ten infrared light emitting diodes (LEDs emitting at 850±10 nm; max. power 300mW/cm3) was used. The 370 nm quartz emission signal was detected using a combination of Hoya U340 and Delta BP 365/50EX optical filters mounted in front of a 25mm head-on Hamamatsu bi-alkaline cathode

photomultiplier tube (H7360-02 series; 280-650 nm with peak sensitivity at 420 nm and ~27% quantum efficiency). To detect the presence of feldspar contaminants, the 410 nm feldspar emission signal was detected using a filter combination comprising a Brightline HC414/46 and a Schott BG 39.

Single grain OSL measurements were made on an automated Risø TL/OSL DA-15 reader[37] fitted with a single-grain green laser attachment. Stimulation was provided by a 10mWNd:YVO, solid-state diode pumped laser emitting at 532nm (max. 103mW/cm3) and the emitted ultraviolet signal at ~370nm from individual quartz grains was detected through a 7.5mm Hoya U-340 filter mounted in front of an Electron Tubes Ltd 9235QA photomultiplier tube fitted with a blue-green sensitive bialkali photocathode. Laboratory doses used for constructing the dose response curves were provided by a 90Sr/90Y ceramic β-source house within the reader. This source was also calibrated against a gamma-irradiated Risø National Laboratory standard[36] and allowed irradiations of the quartz grains with a dose rate of ~0.043Gy/sec.

De estimates are based on a conventional single-aliquot regeneration (SAR) measurement protocol[38,39]. To detect the presence of infrared-sensitive minerals (e.g. feldspars) and to minimize potential contribution of residual feldspathic components to the quartz signal, each multigrain blue light stimulation measurement was also preceded by an infrared bleach at 50°C for 50 seconds[40,41]. No IRSL signal was detected in any of these measurements thereby confirming the absence of feldspar contaminants. For this reason, it was not considered necessary to adopt a similar double SAR measurement protocol for the subsequent single grain measurements.

The natural and regenerative doses were preheated to 230 °C for 10 seconds, and the fixed test doses used to correct for sensitivity changes were preheated to a reduced temperature of 200 °C, before optical stimulation. The choice of preheat combination followed a series of dose recovery experiments conducted at 180°C, 210°C, 220°C, 230°C, 240°C, 250°C and 260°C (Fig. S26). Due to the small quantities of material available for dating this test was done on a limited number of small sized multigrain aliquots (two aliquots per preheat step) obtained from sample X7232. The discs were bleached for 4 hours in direct sunlight and then given a laboratory beta dose equivalent to circa 18.9 Gy. OSL measurements were performed after a delay of several hours and the results (Fig. S29) suggest no noticeable dependency on temperature, returning values very close to the given laboratory dose. No results were obtained for the 200°C and 240°C temperature steps because these aliquots did not produce a measurable OSL signal. Indeed, besides the paucity of sand-sized quartz mineral grains within the cave sediment, the poor sensitivity of most grains presented a real challenge for the OSL dating. By far the most common reason for rejecting multi-grain or single-grain measurements (see Table S9) was caused by a lack of response to laboratory induced beta irradiation (low sensitivity), with initial Tn signals

found to be less than 3σ above background counts. A few single grain measurements were also affected by high recuperation (>15%) or a recycling ratio inconsistent with unity.

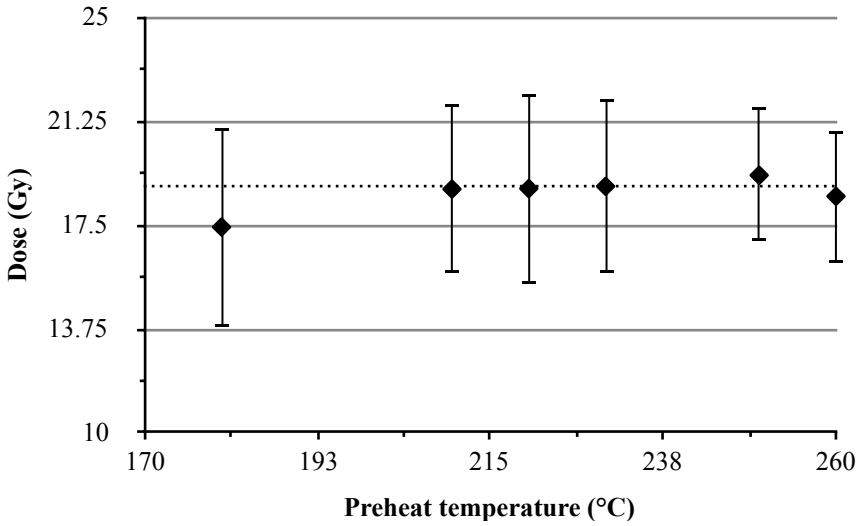


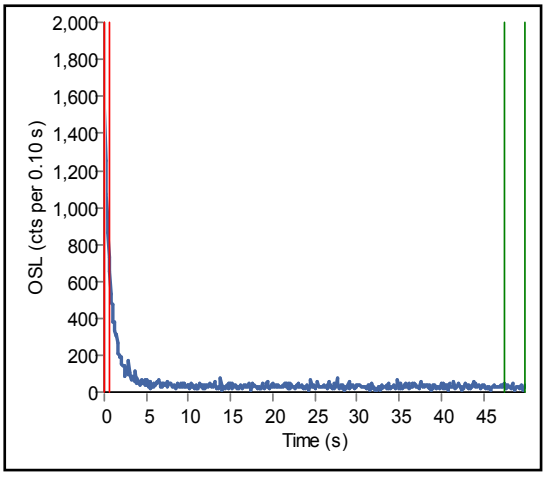
Fig. S29. Dose recovery test for increasing preheat temperatures (180-260°C) of multigrain aliquots prepared from sample X7232. The dotted line is at 18.9 Gy.

Table S9. Summary of the multi-grain and single-grain OSL measurements made, rejected and accepted for each sample. The main reason for rejecting individual measurements was almost invariably caused by a lack of response to laboratory induced beta irradiation (low sensitivity), with initial Tn signals found to be less than 3σ above background counts. A few single grain measurements were also affected by high recuperation (>15%) or a recycling ratio inconsistent with unity. In the case of sample X4135, the single-grain measurements also revealed the presence of 23 grains which featured a saturated OSL signal. No IRSL signal was detected in any of the measurements thereby confirming the absence of potential feldspar contaminants.

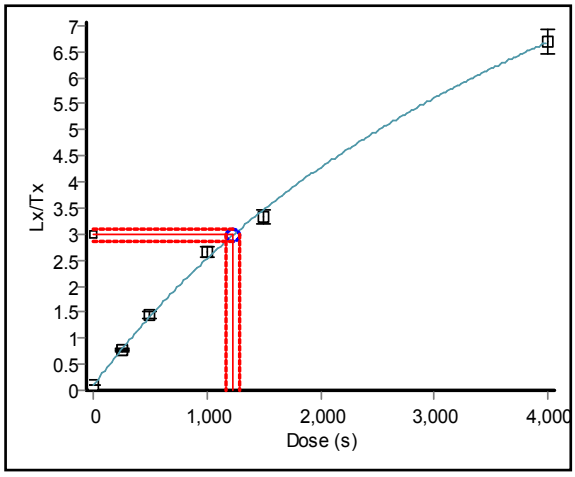
Sample	Type of measurement	Number of measurements	Sum of rejected aliquots/grains	Sum of accepted aliquots/grains	Total
X4135	multi-grain	11	1	10	76
	single-grain	3300	3234	66	
X7227	multi-grain	12	9	3	7
	single-grain	700	696	4	
X7229	multi-grain	12	10	2	11
	single-grain	500	491	9	
X7231	multi-grain	12	6	6	10
	single-grain	2300	2296	4	
X7232	multi-grain	4	2	2	12
	single-grain	2800	2790	10	
X7233	multi-grain	8	3	5	9
	single-grain	2900	2896	4	

The standard error on individual De measurements included an instrument reproducibility uncertainty of 1%, as well as a random 1% uncertainty arising from photon counting statistics. The total uncertainty on the final equivalent dose includes a further systematic component of 2% (added in quadrature) to account for uncertainties in the calibration of the laboratory beta sources. For multi-grain measurements, the equivalent dose was determined from the first second of the OSL decay curve using the final 5s as background noise (total stimulation time varied between 50-100s) and for single-grain measurements the De was determined from the first 0.1s of the decay curve using the last 0.2s as background noise (total stimulation time was 1s). Dose response curves (see inserts c & d in Figs. S30-S35) were fitted with the Analyst software package[42] using a double saturating exponential function. The distributions of single-grain and multi-grain quartz De measurements (see inserts e & f in Figs. S30-S35) and their associated data precision and error scatter are presented as abanico plots[43] which combine a radial plot with a histogram and kernel density estimate curve using the default function tool developed within the package 'Luminescence' for the statistical programming language 'R'[44,45]. Most samples with the exception of X4135, only contained minute quantities of quartz and unfortunately, the large majority of these grains were characterized by poor sensitivity. Due to the high number of rejected aliquots/grains and in order to retain sufficient measurements for De determination, the results from multi-grain and single-grain analyses were therefore combined to obtain a mean (unweighted) equivalent dose estimate used for age calculation. In the case of sample X4135, the OSL measurements revealed a much higher proportion of suitable grains. We hypothesize that the sampled sediment may have contained quartz grains of outside aeolian origin rather than being purely derived from the weathering of local bedrock or originating from vein quartz within the cave itself. The OSL measurements also demonstrated the presence of grains with a saturated OSL signal (n=23) and the single-grain De distribution of this sample is characterized by a higher degree of overdispersion (41%) which is also apparent in the kernel density estimate (KDE) plot (insert f in Fig. S30). When considered in combination, these traits may be indicative of issues pertaining either to partial bleaching, in-situ weathering, disturbance or mixing of sediment and perhaps even small scale microdosimetric effects within the coarse textured carbonate rich sedimentary matrix. The use of a minimum age model was considered to be more appropriate for this sample and indeed, the calculated minimum age estimate (27.79±4.34 ka) is in better agreement with the radiocarbon based chronology. However, in order to be consistent with the approach adopted for the dating of the other samples in this series, the mean De and the calculated OSL date are therefore also reported in Table S10.

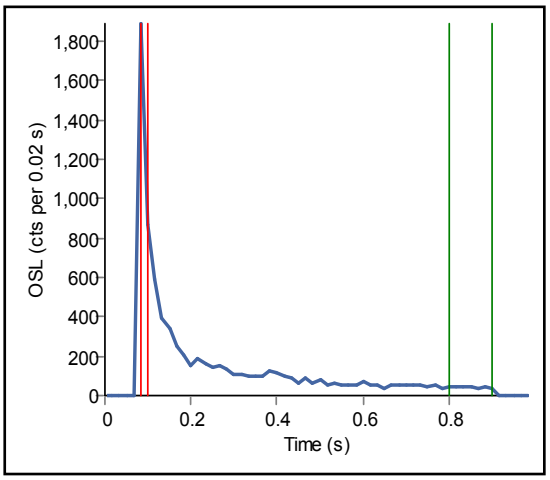
Sample X4135



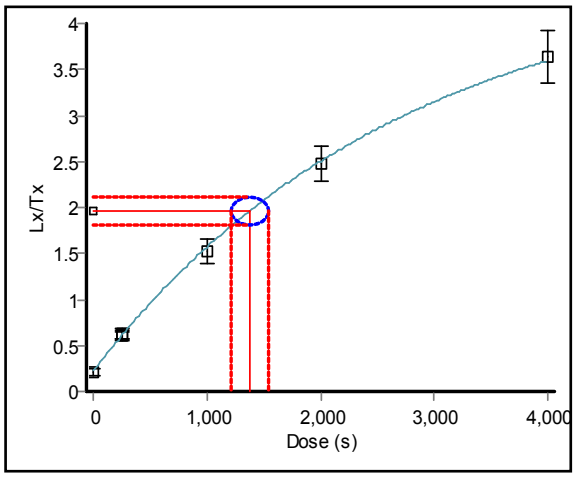
a) Multi-grain OSL decay curve.



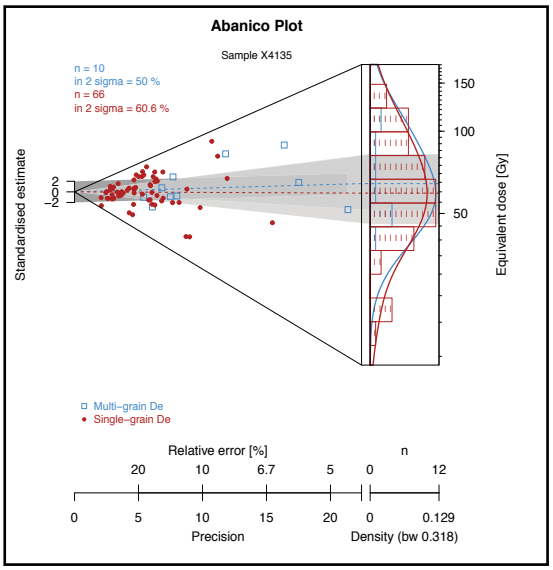
b) Multi-grain OSL growth curve (1sec=0.0425 Gy).



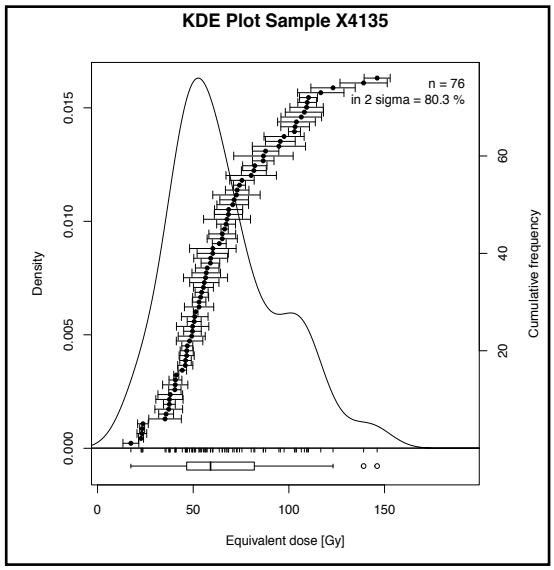
c) Single-grain OSL decay curve.



d) Single-grain OSL growth curve (1sec=0.0428 Gy).



e) Abanico plot of De distributions.

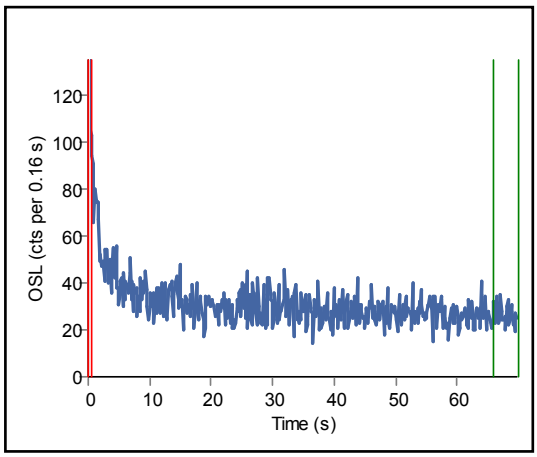


f) Kernel density plot of De distributions.

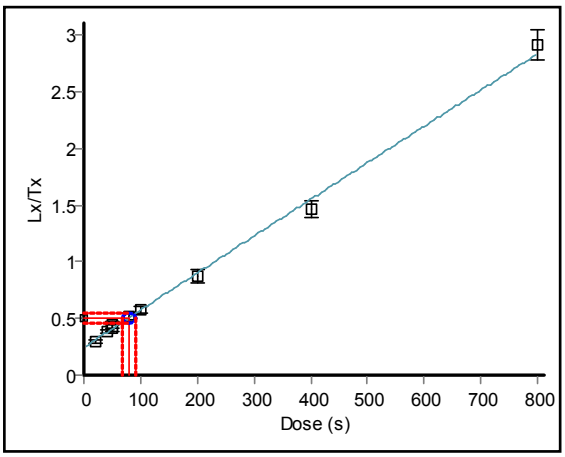
Fig. S30. **a**. and **c**. Example shine down curves of a multi-grain (n=1) and a single-grain (n=1) OSL measurement for sample X4135; **b**. and **d**. corresponding example growth curves (n=1) featuring the interpolated combined mean equivalent dose (De) as a central red line. The associated symmetric uncertainty (1σ) was obtained by using a Levenberg-Marquardt method to fit a linear plus exponential function within version 4.57 of the 'Analyst' software[42]; **e**. and **f**. plots of the mean De distributions of all

multi-grain (n=10) and all single-grain (n=66) quartz OSL measurements obtained for sample X4135. The De distribution is presented as an abanico plot (e) displaying the distribution of equivalent dose measurements (n=76) and their associated data precision and error scatter[45] and a kernel density plot (f). The former plot type combines a radial plot (bivariate plot on the left side) with a histogram and kernel density estimate curve (univariate plots on the right side) using the default function tool developed within version 0.8.6 of the package 'Luminescence'[44,45] for the statistical programming language 'R' (R Development Core Team). The 2σ dispersion range is shown in dark grey and the light grey polygon characterises the 1σ frequency distribution of the primary data (here the multi-grain De results).

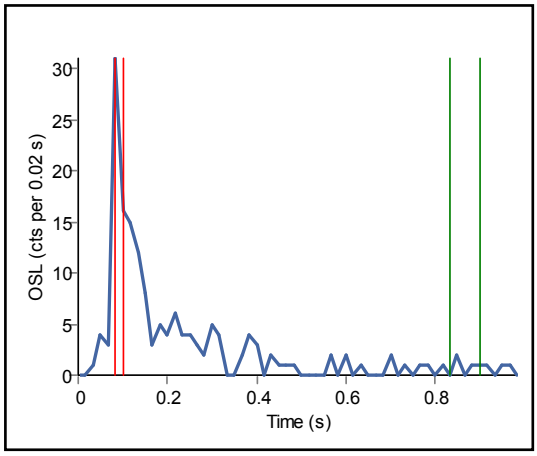
Sample X7227



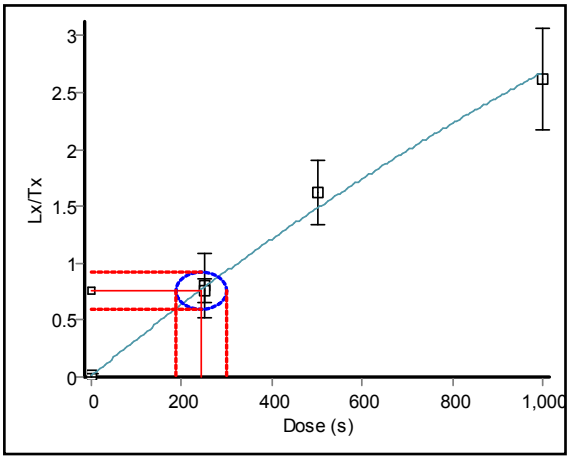
a) Multi-grain OSL decay curve.



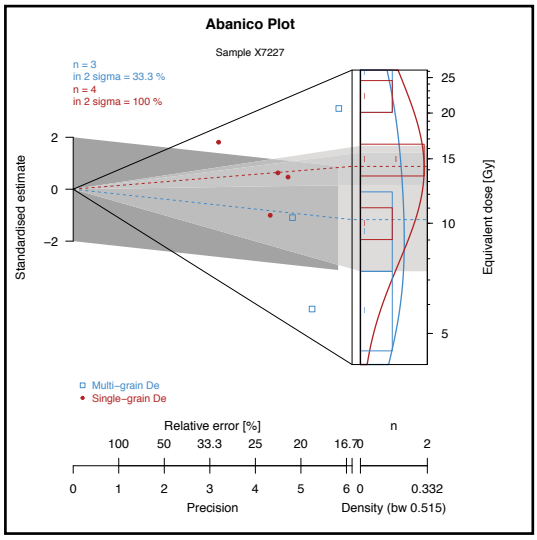
b) Multi-grain OSL growth curve (1sec=0.1268 Gy).



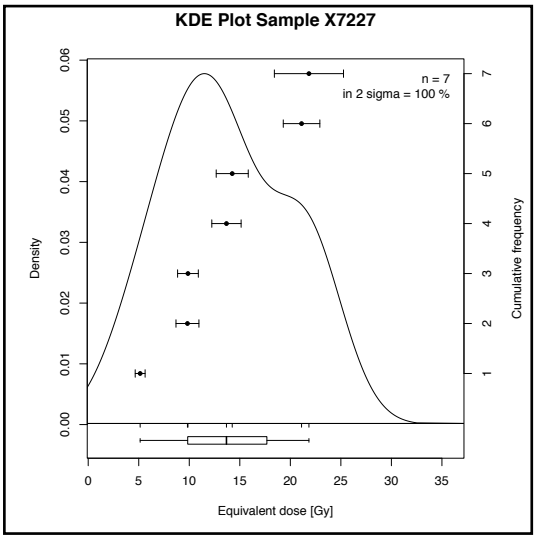
c) Single-grain OSL decay curve.



d) Single-grain OSL growth curve (1sec=0.0427 Gy).



e) Abanico plot of De distributions.

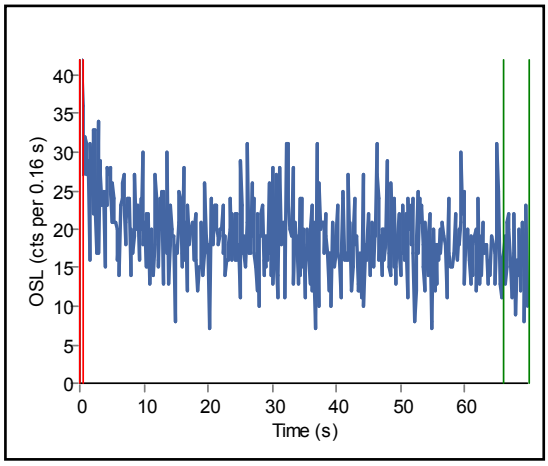


f) Kernel density plot of De distributions.

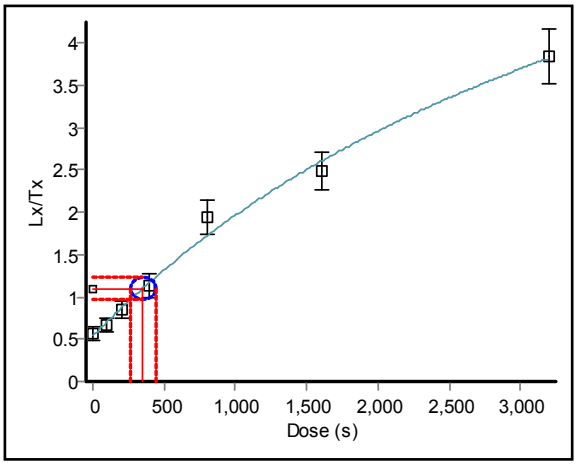
Fig. S31. **a**. and **c**. Example shine down curves of a multi-grain (n=1) and a single-grain (n=1) OSL measurement for sample X7227; **b**. and **d**. corresponding example growth curves (n=1) featuring the interpolated combined mean equivalent dose (De) as a central red line. The associated symmetric uncertainty (1σ) was obtained by using a Levenberg-Marquardt method to fit a linear plus exponential function within version 4.57 of the 'Analyst' software[42]; **e**. and **f**. plots of the mean De distributions of all multi-grain (n=3) and all single-grain (n=4) quartz OSL measurements obtained for sample X7227. The De distribution is presented as an abanico plot (**e**) displaying the distribution of equivalent dose

measurements (n=7) and their associated data precision and error scatter[43] and a kernel density plot (**f**). The former plot type combines a radial plot (bivariate plot on the left side) with a histogram and kernel density estimate curve (univariate plots on the right side) using the default function tool developed within version 0.8.6 of the package 'Luminescence'[44,45] for the statistical programming language 'R' (R Development Core Team). The 2σ dispersion range is shown in dark grey and the light grey polygon characterises the 1σ frequency distribution of the primary data (here the multi-grain De results).

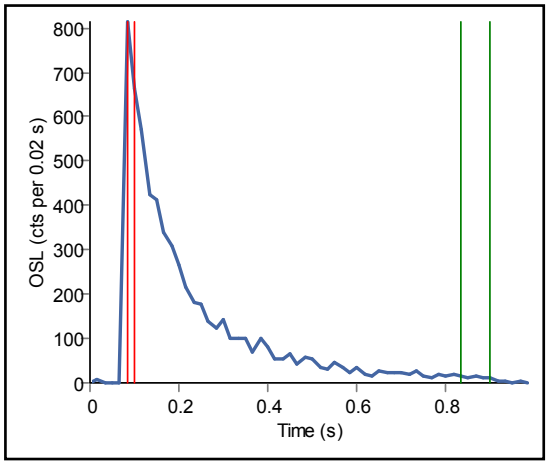
Sample X7229



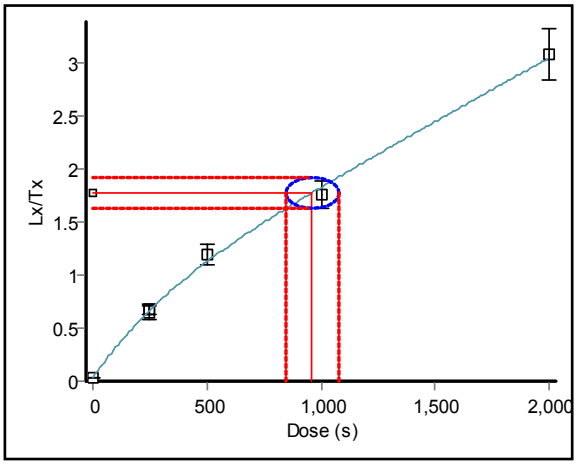
a) Multi-grain OSL decay curve.



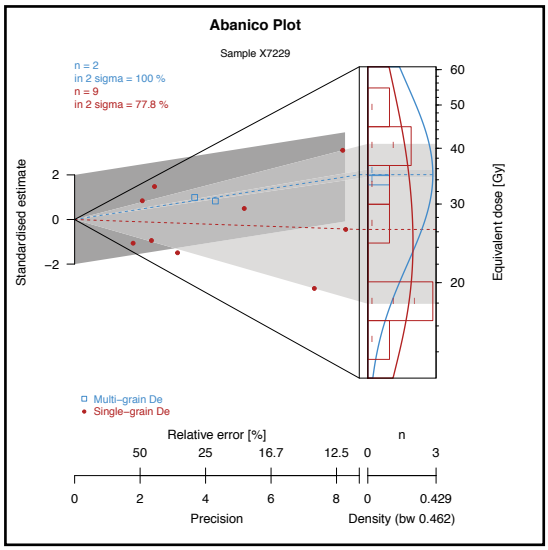
b) Multi-grain OSL growth curve (1sec=0.1192 Gy).



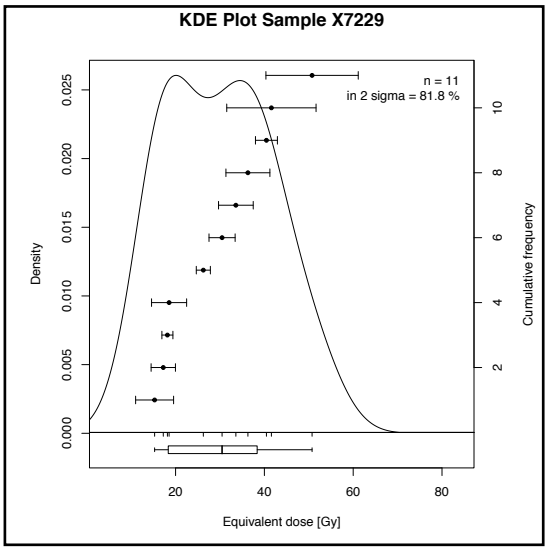
c) Single-grain OSL decay curve.



d) Single-grain OSL growth curve (1sec=0.0425 Gy).



e) Abanico plot of De distributions.

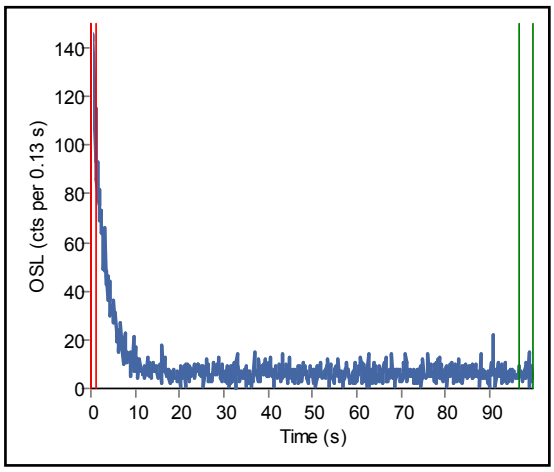


f) Kernel density plot of De distributions.

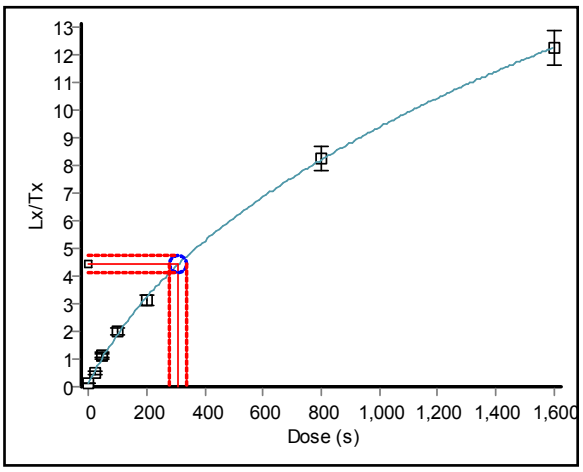
Fig. S32. **a**. and **c**. Example shine down curves of a multi-grain (n=1) and a single-grain (n=1) OSL measurement for sample X7229; **b**. and **d**. corresponding example growth curves (n=1) featuring the interpolated combined mean equivalent dose (De) as a central red line. The associated symmetric uncertainty (1σ) was obtained by using a Levenberg-Marquardt method to fit a linear plus exponential function within version 4.57 of the 'Analyst' software[42]; **e**. and **f**. plots of the mean De distributions of all multi-grain (n=2) and all single-grain (n=9) quartz OSL measurements obtained for sample X7229. The

De distribution is presented as an abanico plot (**e**) displaying the distribution of equivalent dose measurements (n=11) and their associated data precision and error scatter[43] and a kernel density plot (**f**). The former plot type combines a radial plot (bivariate plot on the left side) with a histogram and kernel density estimate curve (univariate plots on the right side) using the default function tool developed within version 0.8.6 of the package 'Luminescence'[44,45] for the statistical programming language 'R' (R Development Core Team). The 2σ dispersion range is shown in dark grey and the light grey polygon characterizes the 1σ frequency distribution of the primary data (here the multi-grain De results).

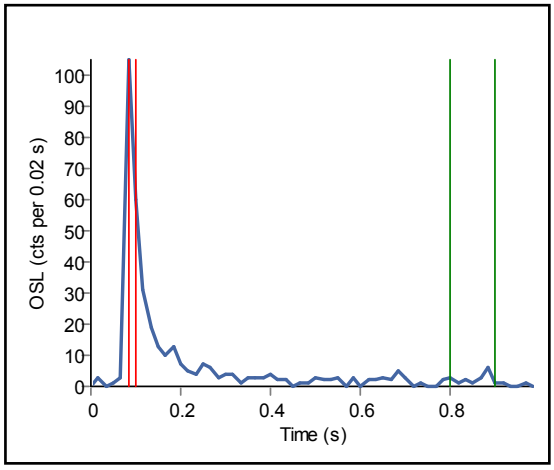
Sample X7231



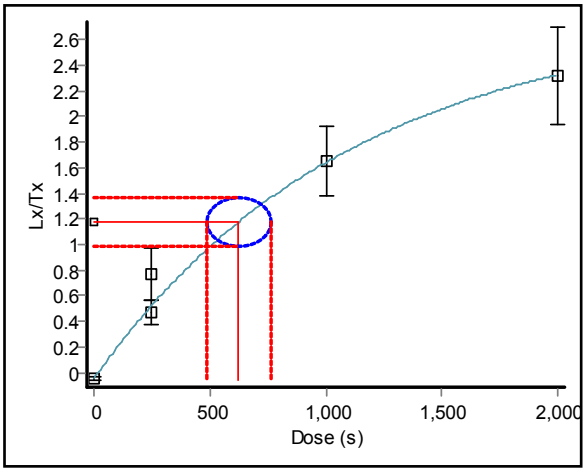
a) Multi-grain OSL decay curve.



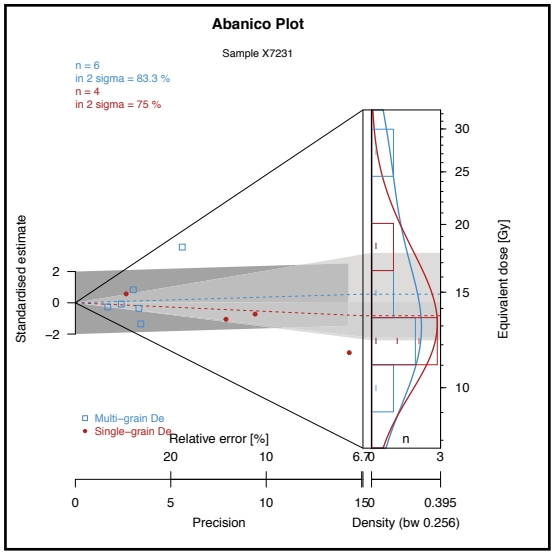
b) Multi-grain OSL growth curve (1sec=0.1183 Gy).



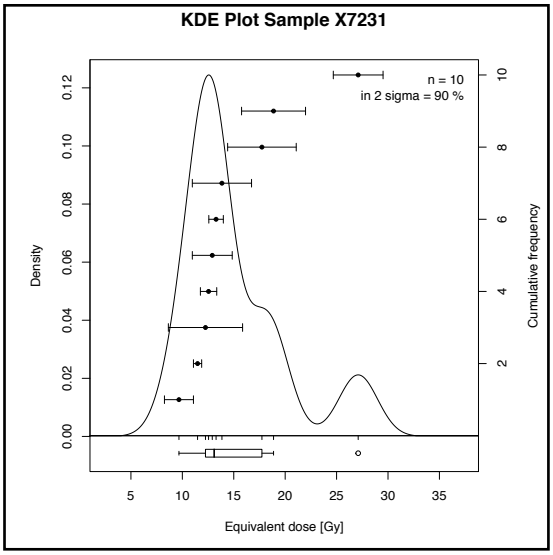
c) Single-grain OSL decay curve.



d) Single-grain OSL growth curve (1sec=0.0428 Gy).



e) Abanico plot of De distributions.

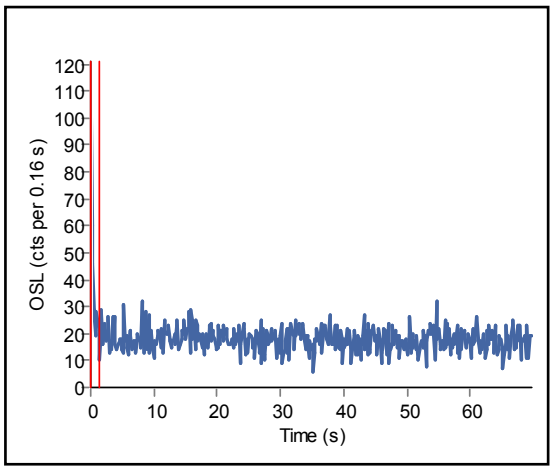


f) Kernel density plot of De distributions.

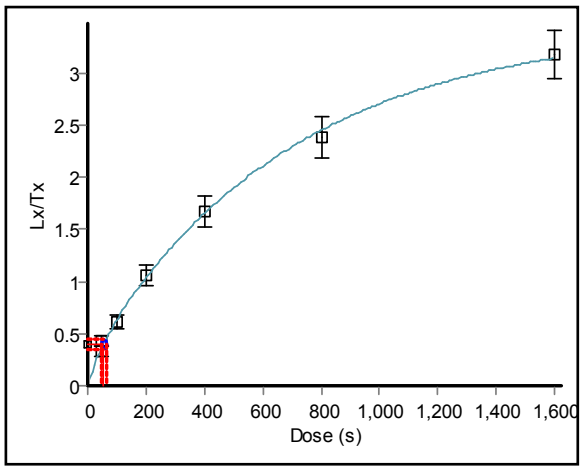
Fig. S33. **a**. and **c**. Example shine down curves of a multi-grain (n=1) and a single-grain (n=1) OSL measurement for sample X7231; **b**. and **d**. corresponding example growth curves (n=1) featuring the interpolated combined mean equivalent dose (De) as a central red line. The associated symmetric uncertainty (1σ) was obtained by using a Levenberg-Marquardt method to fit a linear plus exponential function within version 4.57 of the 'Analyst' software[42]; **e**. and **f**. plots of the mean De distributions of all multi-grain (n=6) and all single-grain (n=4) quartz OSL measurements obtained for sample X7231. The

De distribution is presented as an abanico plot (**e**) displaying the distribution of equivalent dose measurements (n=10) and their associated data precision and error scatter[43] and a kernel density plot (**f**). The former plot type combines a radial plot (bivariate plot on the left side) with a histogram and kernel density estimate curve (univariate plots on the right side) using the default function tool developed within version 0.8.6 of the package 'Luminescence'[44,45] for the statistical programming language 'R' (R Development Core Team). The 2σ dispersion range is shown in dark grey and the light grey polygon characterizes the 1σ frequency distribution of the primary data (here the multi-grain De results).

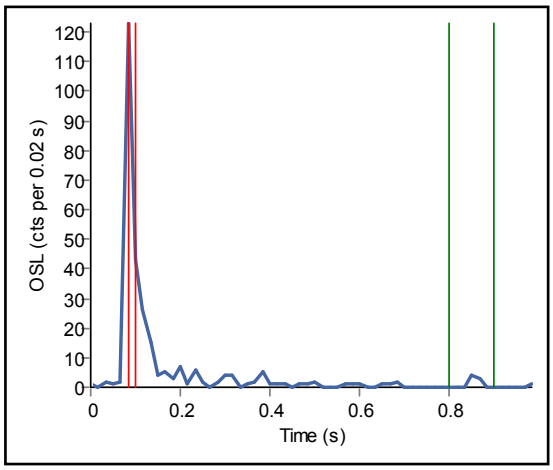
Sample X7232



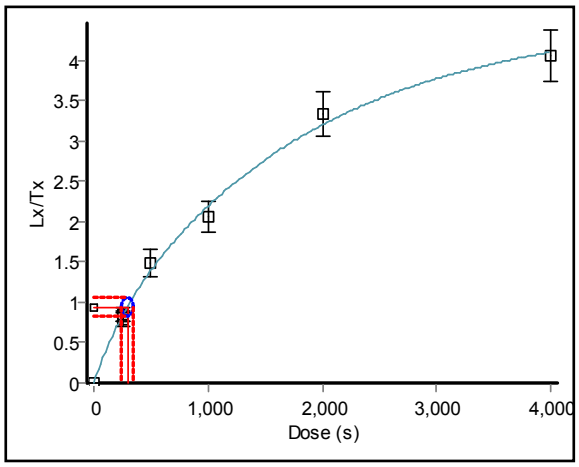
a) Multi-grain OSL decay curve.



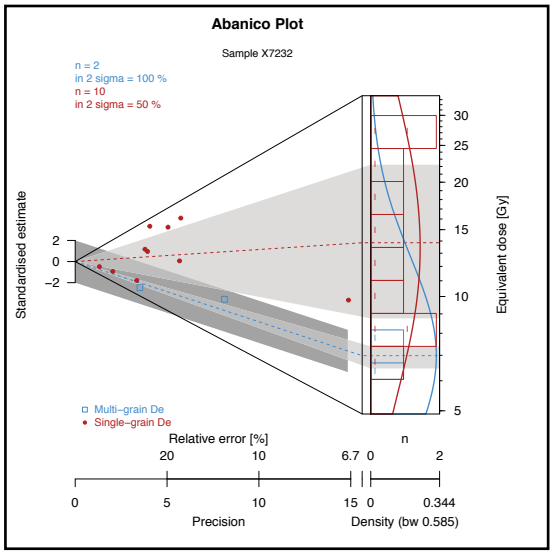
b) Multi-grain OSL growth curve (1sec=0.127 Gy).



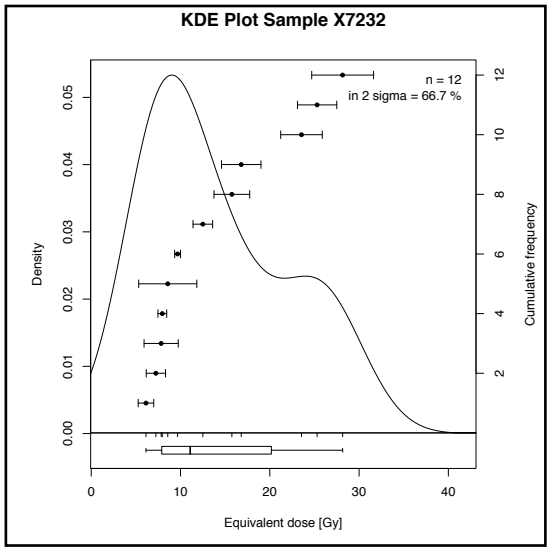
c) Single-grain OSL decay curve.



d) Single-grain OSL growth curve (1sec=0.0428 Gy).



e) Abanico plot of De distributions.

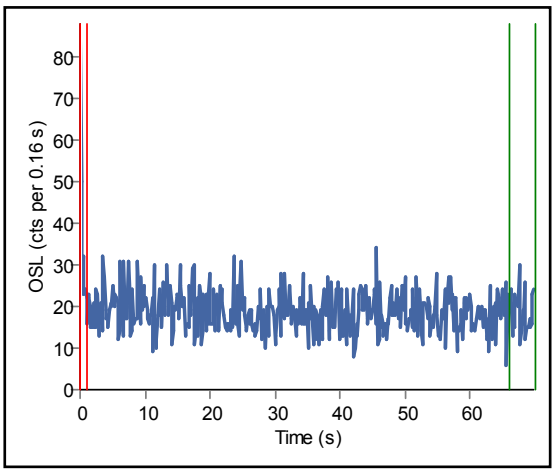


f) Kernel density plot of De distributions.

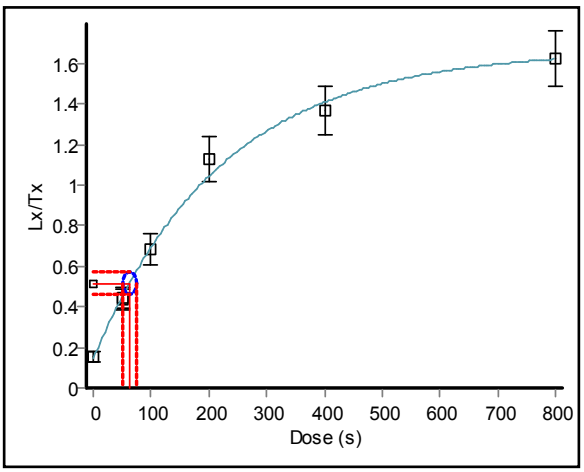
Fig. S34. **a**. and **c**. Example shine down curves of a multi-grain (n=1) and a single-grain (n=1) OSL measurement for sample X7232; **b**. and **d**. corresponding example growth curves (n=1) featuring the interpolated combined mean equivalent dose (De) as a central red line. The associated symmetric uncertainty (1σ) was obtained by using a Levenberg-Marquardt method to fit a linear plus exponential function within version 4.57 of the 'Analyst' software[42]; **e**. and **f**. plots of the mean De distributions of all multi-grain (n=2) and all single-grain (n=10) quartz OSL measurements obtained for sample X7232. The

De distribution is presented as an abanico plot (**e**) displaying the distribution of equivalent dose measurements (n=12) and their associated data precision and error scatter[43] and a kernel density plot (**f**). The former plot type combines a radial plot (bivariate plot on the left side) with a histogram and kernel density estimate curve (univariate plots on the right side) using the default function tool developed within version 0.8.6 of the package 'Luminescence'[44,45] for the statistical programming language 'R' (R Development Core Team). The 2σ dispersion range is shown in dark grey and the light grey polygon characterises the 1σ frequency distribution of the primary data (here the multi-grain De results).

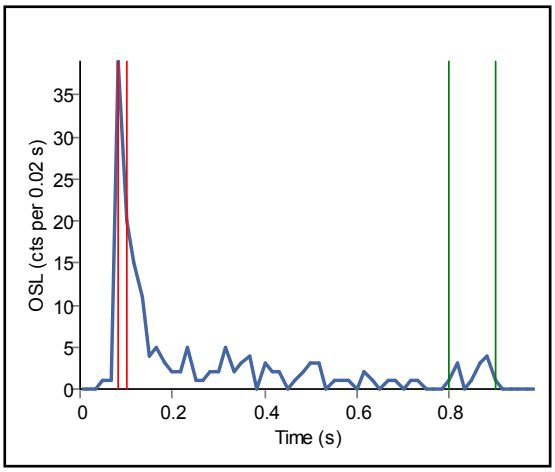
Sample X7233



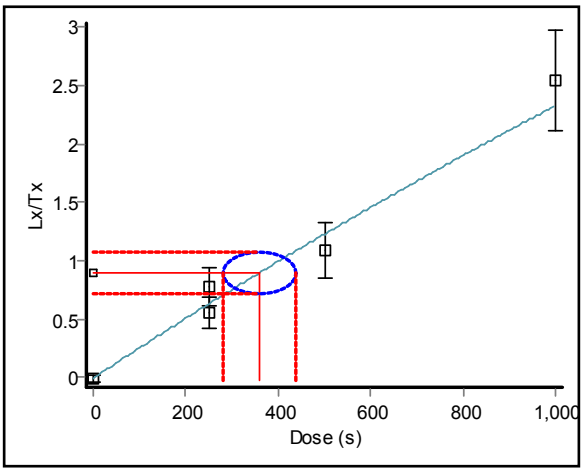
a) Multi-grain OSL decay curve.



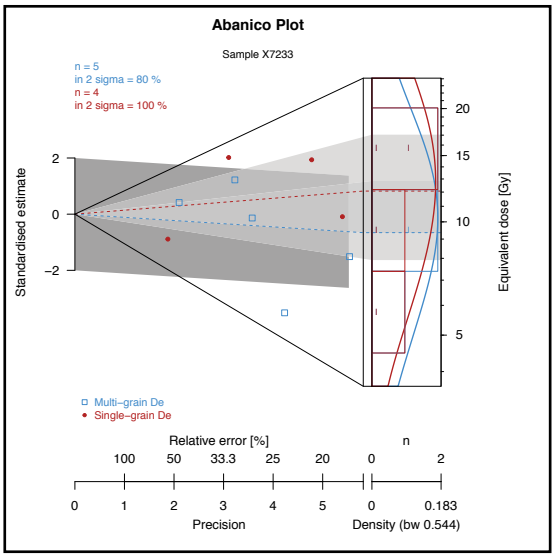
b) Multi-grain OSL growth curve (1sec=0.127 Gy).



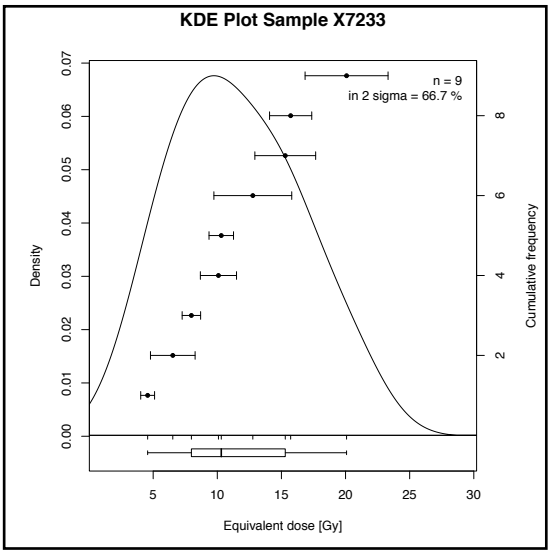
c) Single-grain OSL decay curve.



d) Single-grain OSL growth curve (1sec=0.0428 Gy).



e) Abanico plot of De distributions.



f) Kernel density plot of De distributions.

Fig. S35. **a**. and **c**. Example shine down curves of a multi-grain (n=1) and a single-grain (n=1) OSL measurement for sample X7233; **b**. and **d**. corresponding example growth curves (n=1) featuring the interpolated combined mean equivalent dose (De) as a central red line. The associated symmetric uncertainty (1σ) was obtained by using a Levenberg-Marquardt method to fit a linear plus exponential function within version 4.57 of the 'Analyst' software[42]; **e**. and **f**. plots of the mean De distributions of all multi-grain (n=5) and all single-grain (n=4) quartz OSL measurements obtained for sample X7233. The

De distribution is presented as an abanico plot (**e**) displaying the distribution of equivalent dose measurements (n=9) and their associated data precision and error scatter[43] and a kernel density plot (**f**). The former plot type combines a radial plot (bivariate plot on the left side) with a histogram and kernel density estimate curve (univariate plots on the right side) using the default function tool developed within version 0.8.6 of the package 'Luminescence'[44,45] for the statistical programming language 'R' (R Development Core Team). The 2σ dispersion range is shown in dark grey and the light grey polygon characterises the 1σ frequency distribution of the primary data (here the multi-grain De results).

Table S10. Radioactivity data, De determinations and calculated OSL age estimates for six sediment samples (n=6).

	Radioisotopes					Wet gamma	Wet beta	Cosmic	Total		
Sample	**K**	**Th**	**U**	**Rb**	**Water[a]**	**dose rate**	**dose rate**	**dose rate**	**dose rate[b]**	**De[c]**	**OSL age[d]**
X4135	0.71±0.04	4.5±0.2	3.3±0.2	41±2	3±2 *[3.3]*	0.65±0.14	0.98±0.05	0.011±0.002	1.64±0.15	*(65.74±3.21; n=76)*	*(40160±4160)*
										min. age model: 45.49±5.77 (*n=76*)	27790±4340
X7227	0.61±0.03	2.0±0.1	2.0±0.1	23±1	3±2	0.45±0.02	0.71±0.03	0.011±0.002	1.18±0.04	13.67±2.31 (*n=7*)	11620±2000
X7229	0.64±0.03	2.3±0.1	2.1±0.1	25±1	3±2 *[3.3]*	0.49±0.02	0.75±0.03	0.011±0.002	1.25±0.04	29.87±3.56 (*n=11*)	23940±2950
X7231	0.50±0.02	1.7±0.1	1.7±0.1	18±1	5±2	0.37±0.01	0.58±0.02	0.011±0.002	0.96±0.03	14.97±1.61 (*n=10*)	15560±1740
X7232	0.53±0.03	1.8±0.1	1.8±0.1	19±1	5±2 *[5.1]*	0.40±0.01	0.61±0.03	0.011±0.002	1.02±0.03	14.12±2.24 (*n=12*)	13870±2250
X7233	0.59±0.03	1.7±0.1	1.7±0.1	19±1	5±2	0.40±0.01	0.64±0.03	0.012±0.002	1.05±0.03	11.48±1.65 (*n=9*)	10960±1610

[a] Estimated mean water contents used for dose rate calculations. Values inserted in brackets were obtained from freshly collected sediment in December 2018.

[b] The total dose rate includes a small assumed internal dose rate of 0.03±0.02 Gy/ka to account for alpha and beta dose rates originating from small concentrations of ^{238}U and ^{232}Th within the quartz grains (de Corte et al. 2006; Vandenberghe et al 2008). The mean total uncertainty (at one sigma) was calculated as the quadratic sum of the random and systematic uncertainties. Dose rate calculations were obtained using DRAC (version 1.2) developed by Durcan et al. (2015)[51] with the conversion factors provided by Guérin et al. (2011)[50], the attenuation factors of Guérin et al. (2012)[52] and Brennan (1991)[53] and the beta-etch depth attenuation from Bell (1979)[54]. In the case of sample X4135, collected from the centre of a 30cm thick sedimentary unit, a scaled gamma dose rate was applied according to Aitken (1985)55 to allow for the contribution of 10.66% of the external gamma dose rate to originate from overlying and underlying sediments (based on mean concentrations of radioisotopes within samples X7227-X7233). Assuming an infinite matrix, the total environmental dose rate derived only from the concentrations of radionuclides within sample X4135 itself, would be slightly higher (ie 1.69±0.05Gy/ka) leading to a reduced age estimate of 26.92±3.41 ka.

[c] The equivalent dose was obtained using the mean (unweighted) De of the combined multigrain and single-grain measurements that passed the selection criteria (recuperation, recycling, absence of IR signal). In the case of sample X4135, a date based on a minimum age model was considered to be more appropriate (see KDE plot in Fig. S27, insert f) but the age estimate obtained using the mean De is also reported.

[d] The date is reported in years before 2017 and the uncertainty is expressed within one sigma (68% confidence interval).

No on site gamma-ray spectrometry measurements were made and the dose rate calculations presented in Table S10 are based on the concentrations of radioactive elements (K, Rb, Th and U) within the samples, as determined from elemental analysis performed on homogenized and pulverized subsamples (approximately 10g of sediment) by inductively coupled mass spectrometry (ICP- MS) and inductively coupled atomic emission spectroscopy (ICP-AES). An assumed internal (alpha and beta) dose rate of 0.03 ± 0.02 Gy/kyr based on published measurements for etched quartz[46,47] was included in the dose rate calculations. The concentrations of parent isotopes were converted to dose rates according to updated attenuation factors proposed by Guérin et al.[48] and the dose rate and OSL age estimates were obtained using the dose rate and age calculator (DRAC version 1.2) developed by Durcan et al.[49] with the attenuation factors of Guerin et al[50] and Brennan[51] as well as the beta-etch depth attenuation from Bell[52].

Concentrations of K, Th, U and Rb are relatively consistent between samples in this series (see Table S10). However, sample X4135 provided a much higher dose rate (1.64Gy/ka) due to

elevated concentrations of radionuclides. As mentioned above, this could be related to the occurrence of a higher proportion of exogenous mineral material blown into the cave from the outside environment. Whether or not this is the case will have to await confirmation from forthcoming studies. For this particular sample which was collected from the middle of a circa 30cm thick stratigraphic unit (stratum 1223), it may also be more appropriate to base the age calculation on a scaled gamma dose rate which takes into consideration contributions from neighboring sediments and based on averaged radioisotope concentrations derived from the other samples in the series. Using the conversion factors of Guerin et al.[48] and the fractional gamma dose table provided in Aitken[53], a scaled gamma dose rate of 0.664+/-0.146 Gy/ka can be calculated for this sample. This is considered to represent a best approximation of the external gamma dose rate affecting the quartz grains within sample X4135. It implies that only circa 89.34% of the gamma dose rate is derived from within the stratigraphic unit itself where the sample is located and which has an infinite matrix dose rate of ~0.72Gy/ka. Under this configuration, a further 10.66% would be contributed from the sedimentary units above and below. For the latter, we determined a lower mean infinite matrix dose rate of ~0.42Gy/ka.

The dose rate calculations are based on moisture contents which were determined from fresh sediment samples collected in 2019 and from stratigraphic units equivalent to those where the original OSL tubes were taken. The values ranged from 2 to 5% and were in overall good agreement with those initially recorded from the sediment contained in the OSL tubes (these varied from 1.5-3.7%) but which were suspected of having undergone a loss of pore water. Although the open sections may also have experienced some drying-up, the most recently recorded field values are considered to be the best approximation of the mean water content of the cave sediment throughout their burial period. In order to account for past and seasonal changes in the pore water content, an uncertainty of ± 2% was attached to the water content values.

The cosmic dose rate (0.011±0.001Gy/ka) was not directly measured on site but estimated using the RLumShiny function ‘Cosmic dose rate’ developed by Burow[54] which allows to calculate the cosmic dose rate taking into account multiple absorbers and corrections for the depths of the samples relative to the surface of the cave deposit[55], the thickness of the overlying limestone bedrock (~40m), the respective densities of the sediment (1.9±0.1g/cm3) and the limestone (2.4±0.2g/cm3), as well as the geomagnetic latitude (34.37deg) and the elevation of the site (2740m).

10. Environmental DNA

10.1. Sensitivity and specificity testing of the taxonomic assignments by Holi

We used in silico modelling of three genomes from three key organisms (*Homo sapiens*, *Ursus americanus* and *Juniper monosperma*) to estimate the specificity and sensitivity of the taxonomic assignment method used (Holi; see[56]). In order to keep the setup simple, we ran the modeling and taxonomic assignments of each organism separately. Importantly, the taxonomic identification (both the specificity and sensitivity) does not depend on the diversity of the pool from which the reads derive, but on the information contained by each read, i.e., length and position, as well as on the composition/representation of organisms in the database the reads are aligned to. To simulate the most common biases observed in ancient DNA datasets, such as post-mortem DNA fragmentation, miscoding lesions and sequencing errors[57], we used the program Gargammel[58]. To extract datasets similar to the observed in Gargammel, we first calculated the average read length across our data and converted these to relative proportions (Fig. S36). For each model, we randomly extracted fragmented reads corresponding to approximately 30x coverage for each genome and repeated 5 times for parallel simulations. We used the damage patterns observed in the mammal reads of UE1210 (UE1210_Mex_18_Lib4_seq2; Extended Data Fig. 2a) as template (the misincorporation.txt file generated by MapDamage 2.0[59]) for inserting deamination to the modeled reads (Fig. S38).

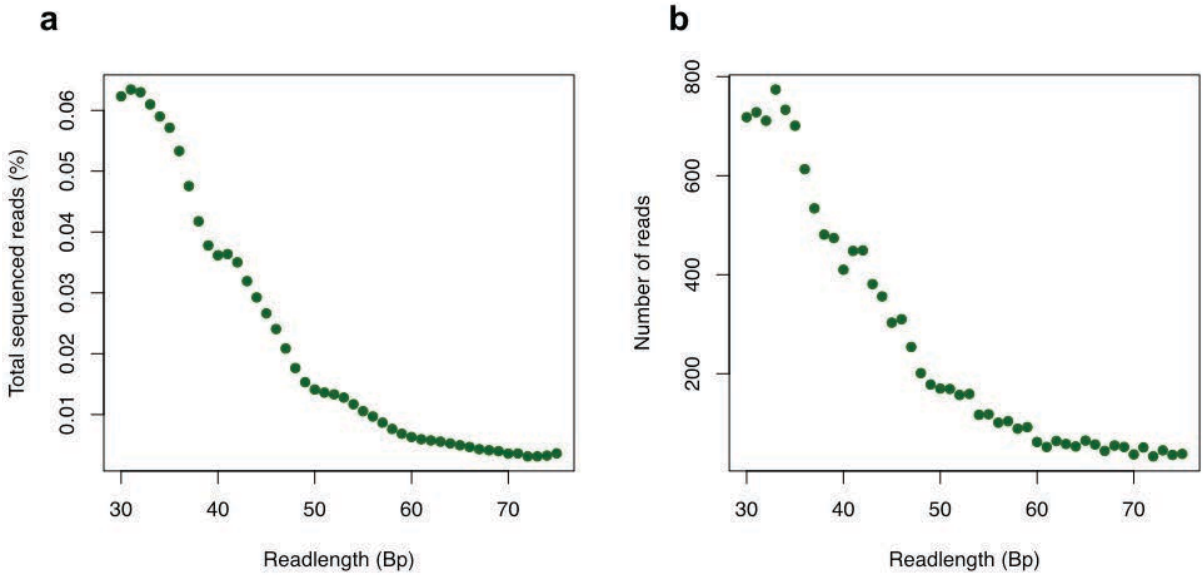


Fig. S36. Readlengths used in the simulations. **a.** average read lengths proportions extracted from all samples and used to generate fragmented DNA from the reference genomes. **b.** read length example of randomly extracted DNA from the human mitochondrial DNA (NC_012920.1, simulation human model 1) using the average library as input.

To test the sensitivity and the specificity of our taxonomic assignment method, we used 5 different customized databases (Table S11): DB1) the NCBI nt including the RefSeq (ver. 91),

DB2) the full mitochondrial genomes (mtDNA), or full plastid genomes (pDNA) from the RefSeq version 91. We next wanted to investigate the implications of removing the reference genome (DB3) and all the reference genomes from the given genus (DB4) from the full RefSeq mtDNA and pDNA databases. Lastly, we mapped all reads against the reference from which the reads were extracted from (DB5). Furthermore, to enable transparency of the consequences each addition of errors have on the taxonomic profiles, we split each model up into three scenarios using, first all extracted reads (scenario a), secondly extracted reads with deamination (scenario b) and thirdly extracted reads with deamination and sequencing errors (scenario c, see Fig. S37). All reads from each three scenarios (a,b,c) were hereafter mapped in parallel against the 5 databases. Each model was repeated 5 times (see Bash and R scripts below, in section 10.5).

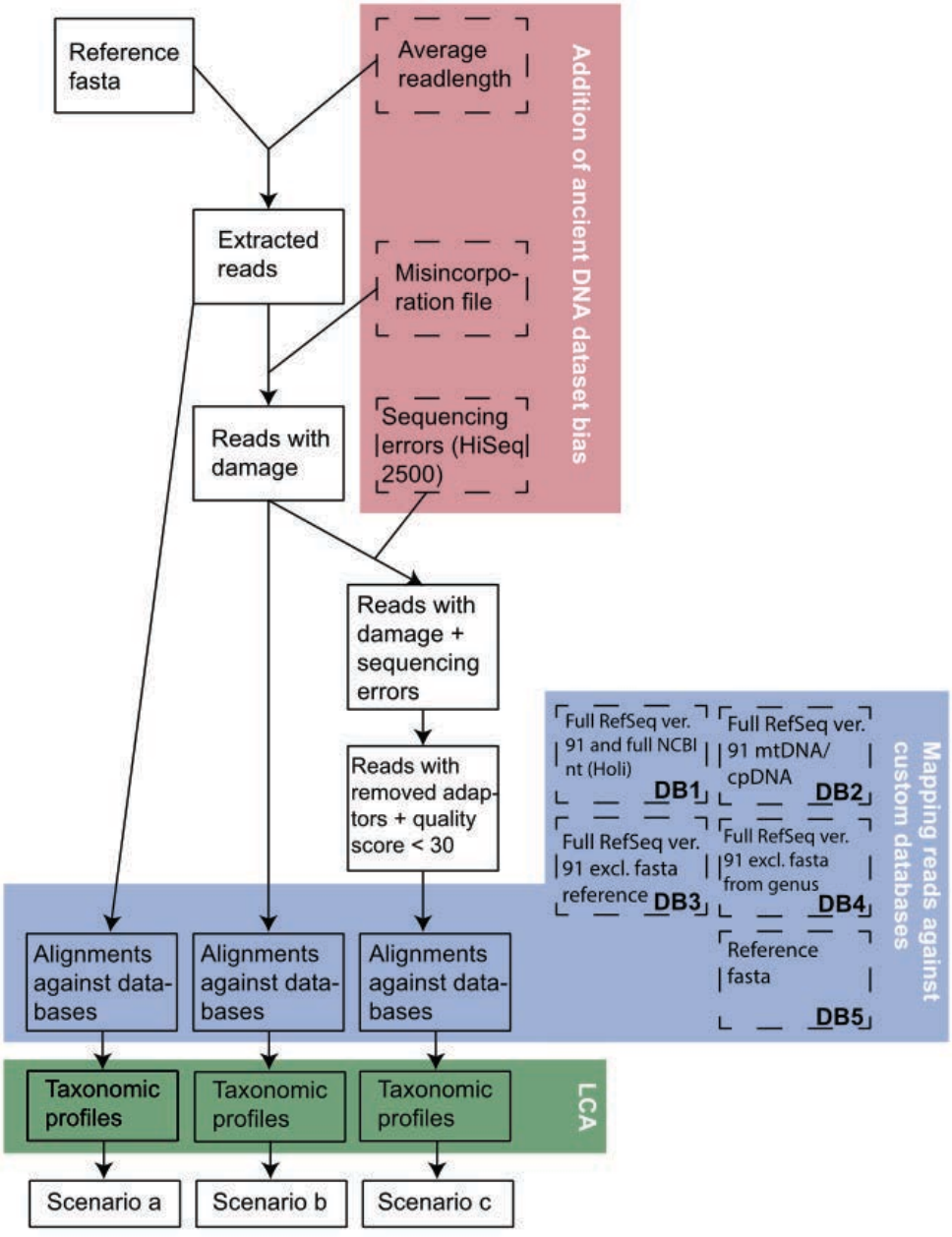


Fig. S37. Flow diagram for one model simulation, outlining the extraction and manipulation of the reads in gargammel and the eventual database alignments and taxonomic assignment for scenario a, b and c.

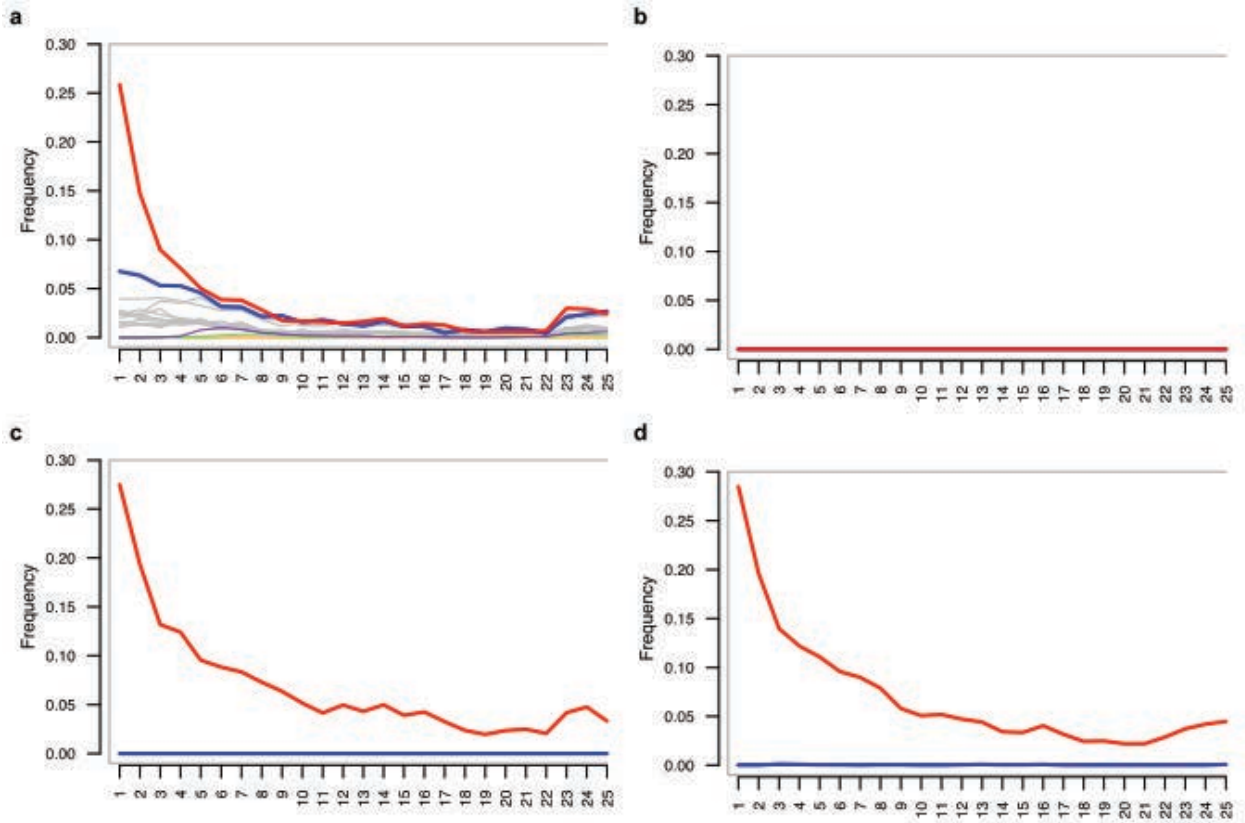


Fig. S38. DNA damage template and examples of the reads from the three steps of Human mtDNA model 5 **a.** Template DNA damage used as input to gargammel **b.** extracted reads from Homo sapiens model 1 (scenario a), **c.** reads with damage (scenario b), and **d.** reads with damage and sequencing errors (scenario c).

Table S11. NCBI accession numbers for the taxa used in the in-silico modeling and the excluded reference sequences.

Databases/Species	Holi (DB1)	Full mtDNA or cpDNA db (DB2)	mtDNA or cpDNA excluding reference fasta (DB3)	mtDNA or cpDNA excluding fasta from genus (DB4)	Fasta used as reference and DB5
Homo sapiens	RefSeq version 91, nt db (updated 21st Nov. 2018)	RefSeq version 91 mitochondrial genomes	RefSeq version 91 mitochondrial genomes excluding NC_012920.1	RefSeq version 91 mitochondrial genomes excluding NC_013993.1,NC_023100.1,NC_012920.1,NC_011137.1	NC_012920.1
Ursus americanus	RefSeq version 91, nt db (updated 21st Nov. 2018)	RefSeq version 91 mitochondrial genomes	RefSeq version 91 mitochondrial genomes excluding NC_003426.1	RefSeq version 91 mitochondrial genomes excluding NC_003426.1,NC_008753.1,NC_009331.1,NC_009971.1,NC_003427.1,NC_003428.1,NC_011112.1,NC_011117.1,NC_011118.1	NC_003426.1
Juniper monosperma	RefSeq version 91, nt db (updated 21st Nov. 2018)	RefSeq version 91 plastids genomes	RefSeq version 91 plastid genomes excluding NC_024022.1	RefSeq version 91 plastid genomes excluding NC_024022.1,NC_024023.1,NC_024024.1,NC_024021.1,NC_028190.1,NC_035068.1,NC_037430.1	NC_024022.1

We find the proportion of reads when mapping (scenario a) to the databases DB1-DB3 and DB5 similar to the number of extracted reads with only small differences (Table S12). While we find reads aligning to database DB4, in which all the references from the genus has been removed, to

decrease by ~35%. The small difference observed between the Holi database (DB1) and the full mtDNA database (DB2) are likely explained by the fact that the DB1 was divided into 35 smaller databases (~20Gb) due to its considerable size. The DB2, on the contrary, remained as one file. This, however, does not influence the taxa found (as shown below). The same observation is evident for scenario b. Here, we see a slight but overall reduction of the total reads mapping, which is due to the deamination inserted on the reads. These edits decrease the homologies between read and reference sequences. For scenario c, we see the same trends in the alignment rates although DB1 has a small increase in alignments equal to DB2. Identical patterns are found for the other modeled organisms of *Ursus* and *Juniper* (Tables S12 and S13).

All aligning reads (with 100% similarity between read and reference) in each model mapped against (DB1-DB5) were parsed, for each scenario, through an in-house naïve least common ancestor (LCA) algorithm and results were parsed and plotted in R (see R scripts below, in section 10.5). We find that a cut-off threshold must be applied (removing taxa with proportions <1%) across all databases for robust taxonomic assignments. Tables S14-16 show the resulting taxonomic profiles both before and after the cut-off threshold. The taxonomic profile for each model and scenario of *Homo sapiens* mtDNA are presented in Figs. S39-43 and Table S15. For the *Juniper* and the *Ursus* models, the resulting taxonomic profiles from scenario c ("the sequenced simulation") are presented in Figs. S45-48 and Tables S15-16, respectively.

Scenario a /Iteration	Total reads extracted	DB1	DB2	DB3	DB4	DB5
HomoSap.1	11835	11765	11835	11602	7711	11835
HomoSap.2	11835	11735	11835	11587	7729	11835
HomoSap.3	11835	11745	11835	11622	7787	11835
HomoSap.4	11835	11766	11835	11576	7738	11835
HomoSap.5	11835	11743	11835	11552	7716	11835

Scenario b /Iteration	Total reads extracted with deamination	DB1	DB2	DB3	DB4	DB5
HomoSap.1	11835	11618	11341	10787	6554	11219
HomoSap.2	11835	11575	11332	10754	6604	11233
HomoSap.3	11835	11565	11315	10789	6690	11218
HomoSap.4	11835	11638	11363	10777	6646	11250
HomoSap.5	11835	11587	11317	10725	6604	11224

Scenario c /Iteration	Total reads extracted with deamination and seq. errors (after removal of adaptors and > Q30)	DB1	DB2	DB3	DB4	DB5
HomoSap.1	11749	11173	11191	10648	6918	11074
HomoSap.2	11747	11132	11159	10604	6940	11018
HomoSap.3	11715	11141	11176	10643	6937	11051
HomoSap.4	11751	11169	11198	10647	6870	11070
HomoSap.5	11749	11189	11207	10658	6909	11083

Table S12. Extracted reads for the *Homo sapiens* modelling and reads aligned to DB1-5.

Scenario a /Iteration	Total reads extracted	DB1	DB2	DB3	DB4	DB5
Ursus_americanus.1	12155	12063	12155	9806	8576	12155
Ursus_americanus.2	12155	12063	12155	9738	8427	12155
Ursus_americanus.3	12155	12070	12155	9804	8521	12155
Ursus_americanus.4	12155	12062	12155	9872	8585	12155
Ursus_americanus.5	12155	12073	12155	9788	8523	12155

Scenario b /Iteration	Total reads extracted with deamination	DB1	DB2	DB3	DB4	DB5
Ursus_americanus.1	12155	11695	11724	8743	7457	11618
Ursus_americanus.2	12155	11702	11739	8607	7291	11638
Ursus_americanus.3	12155	11741	11764	8763	7454	11681
Ursus_americanus.4	12155	11738	11785	8719	7459	11692
Ursus_americanus.5	12155	11729	11742	8699	7380	11648

Scenario c /Iteration	Total reads extracted with deamination and seq. errors (after removal of adaptors and > Q30)	DB1	DB2	DB3	DB4	DB5
Ursus_americanus.1	8248	11599	11622	8834	7673	11500
Ursus_americanus.2	8346	11566	11587	8912	7765	11469
Ursus_americanus.3	8300	11600	11626	8919	7744	11509
Ursus_americanus.4	8362	11615	11632	8818	7626	11505
Ursus_americanus.5	8221	11571	11594	8889	7727	11480

Table S13. Extracted reads for the *Ursus americanus* modelling and reads aligned to DB1-5.

Scenario a /Iteration	Total reads extracted	DB1	DB2	DB3	DB4	DB5
Juniper.1	91903	91236	91903	88495	79425	91903
Juniper.2	91903	91229	91903	88387	79183	91903
Juniper.3	91903	91258	91903	88340	79061	91903
Juniper.4	91903	91279	91902	88507	79179	91903
Juniper.5	91903	91253	91903	88521	79335	91903

Scenario b /Iteration	Total reads extracted with deamination	DB1	DB2	DB3	DB4	DB5
Juniper.1	91903	89220	89680	84729	74255	89183
Juniper.2	91903	89214	89701	84678	73941	89190
Juniper.3	91903	89255	89694	84581	73834	89186
Juniper.4	91903	89225	89667	84731	73811	89154
Juniper.5	91903	89081	89525	84640	73967	89007

Scenario c /Iteration	Total reads extracted with deamination and seq. errors (after removal of adaptors and > Q30)	DB1	DB2	DB3	DB4	DB5
Juniper.1	91145	88097	88462	83568	73141	87844
Juniper.2	91125	87924	88356	83402	73080	87764
Juniper.3	91140	87970	88330	83446	73152	87714
Juniper.4	91142	87961	88360	83430	73017	87768
Juniper.5	91127	87947	88351	83481	73179	87728

Table S14. Extracted reads for the *Juniper monosperma* modelling and reads aligned to DB1-5.

10.1.1. Homo sapiens mtDNA scenario **a**

For the taxonomic profiles recovered by extracting reads from the human mtDNA (scenario a), without applying a minimum threshold (Fig. S39), we find that all reads aligning to DB1 and DB2 are assigned to *Homo sapiens* (or higher taxonomic levels to which *Homo sapiens* belongs). However, we find a difference in the alignment proportion between these two databases—namely that ~10% of the reads are assigned to species level using DB1 and ~50% of the reads are assigned to species level using DB2 (Fig. S39). At genus level, DB1 assigns ~6% of the reads while DB2 assigns ~33% of the reads. This discrepancy can be explained by the nature of the databases, where DB2 is limited to only contain mtDNA genomes and DB1 contains the full RefSeq and nt databases from NCBI. Reads aligned against DB1, therefore, have more references and larger reference genomes to align to, resulting in an increased possibility of matching to genomes that are non-human. This is reflected by the taxonomic profiles and the large proportion of reads from DB1 that are assigned to the root (~66%). While this can be interpreted as DB1 having lower sensitivity and specificity, we do not believe this to be the case as the source of these reads can derive from multiple organisms. Thus, for metagenomic data, Holi assigns to a higher but more precise taxonomic level. Furthermore, we observe that when the reference genome is absent in the database (DB3), we increase the alignments to *Homo* (~43%); ~25% reads are assigned to *Homo sapiens neanderthalensis*, ~3% to *Homo heidelbergensis,* and ~2% to *Homo sapiens ssp. Denisova*. We also see taxa distantly related to humans, however, these are taxa found in small proportions only (< 0.06%). In cases where all reference genomes from the *Homo* genus are absent, we find that the majority of the reads fall to *Pan* or *Gorilla* or higher taxonomical levels covering these genera (Hominidae). Again, taxa not related to Hominidae are found, but in low frequencies (< 0.3%). By applying a threshold ≥ 1%, dissimilarity between reads and database reference sequences, we assign all reads aligned to DB1-3 within the correct genus and within the correct family for DB4 (Fig. S40).

10.1.2. Homo sapiens mtDNA scenario **b**

By inserting deamination patterns on the extracted reads, we observe that the total amount of taxonomically assigned reads are decreased by 25-40%, depending on the database aligned to. This is a consequence of the decreased homology between reads and reference sequences as we are only considering similarities of 100% for taxonomic assignment in the LCA. However, despite this, we increase the number of unique taxonomic units, hence false-positives, by 60% (Fig. S41). In this scenario, we find ~27% of the reads assigned to species level when aligned against DB1 and ~48% for reads aligned against DB2. At genus level, ~5% and ~33% are assigned by DB1 and DB2, respectively. The proportion of reads assigned to root remains high (~50%) for DB1. For DB 1-3, we find that false-positives are emerging, although each with small proportions <0.4% (*Pan troglodytes* in DB3 is highest). This implies that if no cut-off threshold is applied, the taxonomic profiles will yield false positives even at family levels. We therefore

argue that it is important to apply a threshold to remove such false positives from the datasets. We applied a cut-off threshold of 1%, considering only taxa with ≥1% abundance, this results in that we assign all reads aligned to DB1-3 within the correct genus and within the correct family for DB4—similar to what we see in scenario a (Fig. S42).

10.1.3. Homo sapiens mtDNA scenario **c**

The previous two scenarios (a and b) show that the vast majority of the reads can be assigned correctly within *Homo sapiens* or higher taxonomic groups to which this species belongs, but also that false-positives emerge from DNA damage. In scenario c, we add an additional important source of errors, sequencing errors. Consequently, we observe a decrease in the total number of reads assigned (up to ~50% of the reads compared to scenario b). As for reads with deamination (scenario b), this is explained by an increase in the decreased homology (in this case) caused by the addition of sequencing errors. This resulted in a reduction in the number of unique taxonomic units by ~50%, which correlates with the number of totally assigned reads (Fig. S43). We find that taxonomic assignments without applying a cut-off threshold still remains robust at genus level for DB1-3, but not for DB4. The superorder of cingulate mammals, Cercopithecidae, are found to be the false-positive with highest abundance (0.19-1.33%; Table S15), and appear for the DB4 taxonomic profiles (database with no references from the genus *Homo*). We also find false-positive for DB1-3 with the highest abundance of the pale fork-marked lemur, *Phaner pallescens,* but this does not exceed 0.6% (which equals 3 reads of the total 11,749 reads aligned). It is noteworthy that only ~500 of the initial ~11,750 reads are assigned a taxonomic level for DB4 where the reference of the whole genus is absent. This implies that for genera in a true metagenomic sample, the proportion of false-positives generated will be proportionally small than for species with a reference sequence in the database of either species or genus levels. Considering also that other errors both prior and during laboratory processing can occur (such as oxidization and PCR errors), we argue that setting a cut-off threshold of ≥1% for metagenomic samples containing *Homo sapiens* mtDNA will produce robust taxonomic assignments to the genus *Homo* (Fig. S44 and Table S15). Importantly, in this case, the resulting taxonomic profiles of DB1 and DB2 become nearly identical—which is also observed for the models of *Ursus americanus* and *Juniper monosperma* (see below).

10.1.4. Ursus americanus mtDNA and *Juniper monosperma* scenario **c**

The factors driving the generation and proportional levels of false-positives identified in scenario a, b and c for *Homo sapiens* mtDNA, is identical for the two other modeled organisms, *Ursus americanus* and *Juniper monosperma*. Below, we present the resulting scenario c taxonomic profiles from both organisms (Figs. S45-48 and Tables S15-16). Significantly, the resulting taxonomic profiles of *Ursus americanus* yielded the seal, *Neomonachus schauinslandi,* as a false-positive in DB3 (1.09% and 1.06%, in model 1 and 5, respectively; see Table S16). DB1,

however, does not find seal as a false-positive at all, while DB2 finds it at, proportionally, very low levels (< 0.016%). Importantly, the proportions of false positives decrease when part of a metagenomic sample. For the modelled *Juniper monosperma,* we do not identify any false-positives above the cut-off threshold in any of the resulting taxonomic profiles (DB1-5; Fig. S48).

10.1.5. Conclusions

Considering the results above, we find that the taxonomic assignment of the ancient metagenomic sequence data can be robustly performed by using (i) the most comprehensive genetic reference databases available (without discriminating between organisms and/or environments), (ii) by preforming a naïve least-common-ancestor analysis considering only 100% similarities between read and reference, and (iii) by setting a minimum cut-off threshold (in this case ≥1%) to eliminate the false positives generated by the various sources of errors. To improve robustness for future studies, it will be important to perform similar modeling on the most abundant taxa found in order to increase understanding of how each genome behaves when subjected to fragmentation, miscoding lesions, sequencing errors, and mapping.

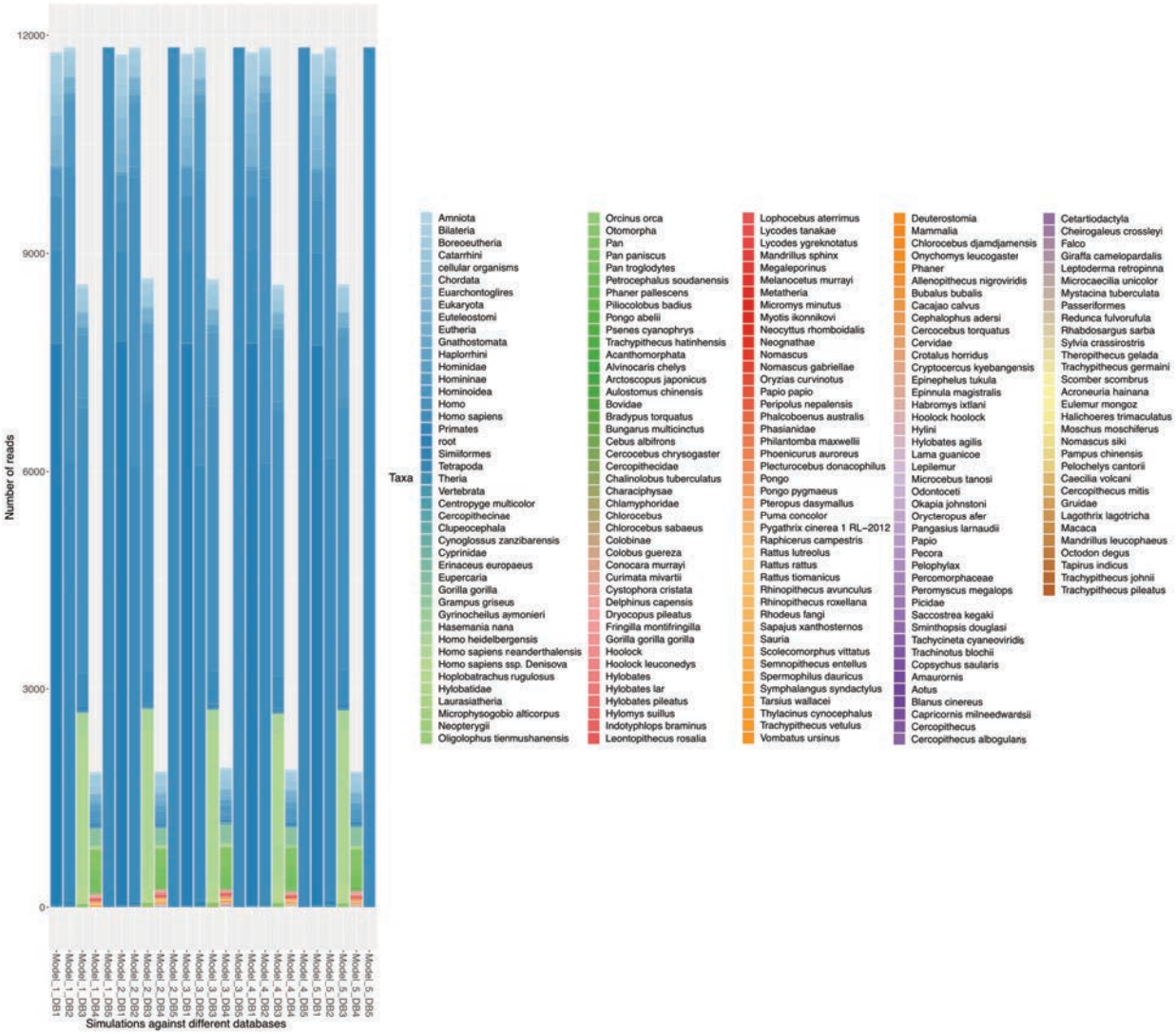


Fig. S39. Complete taxonomic profiles of extracted mtDNA reads from *Homo sapiens* with no cut-off threshold applied.

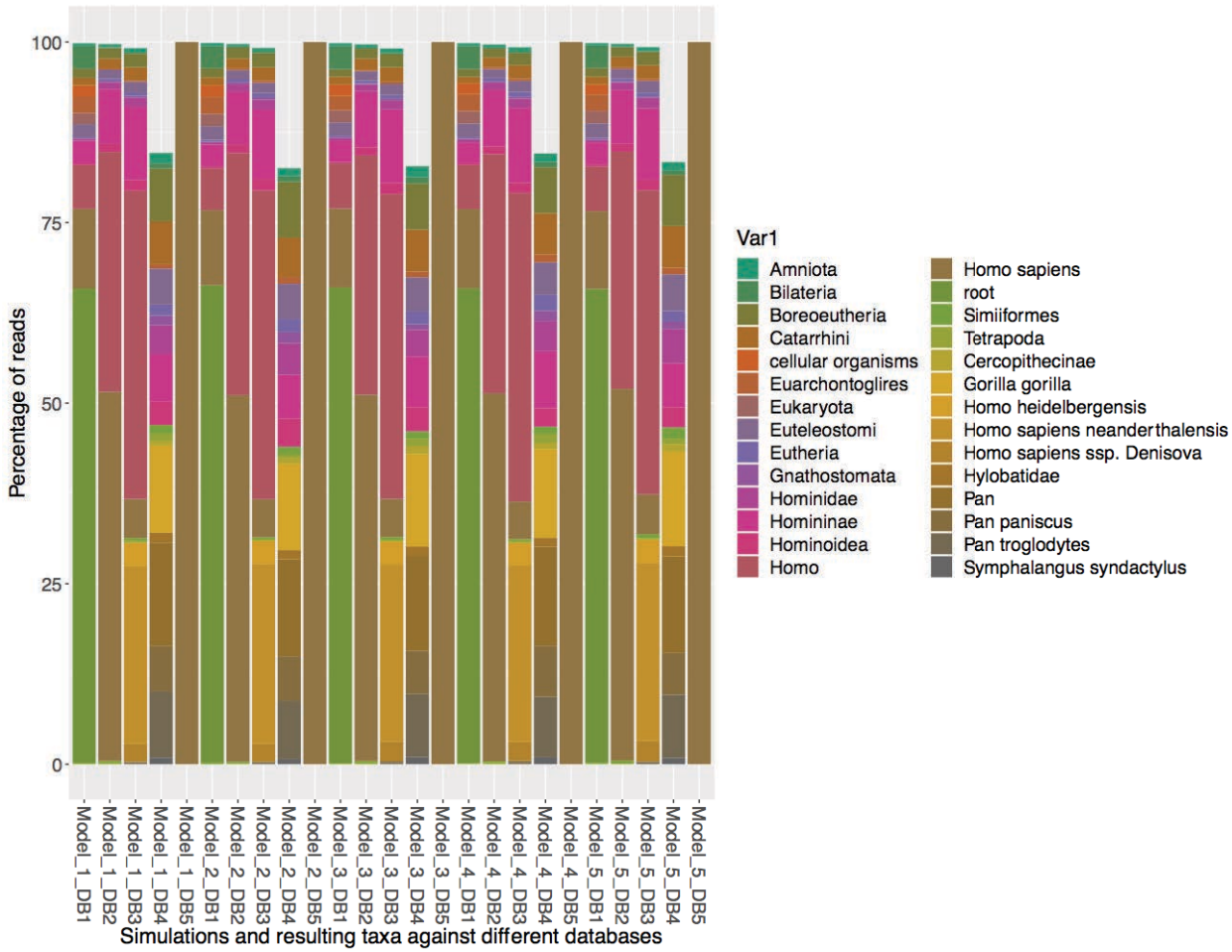


Fig S40. Taxonomic profiles from Homo sapiens models of extracted reads above the threshold 1%.

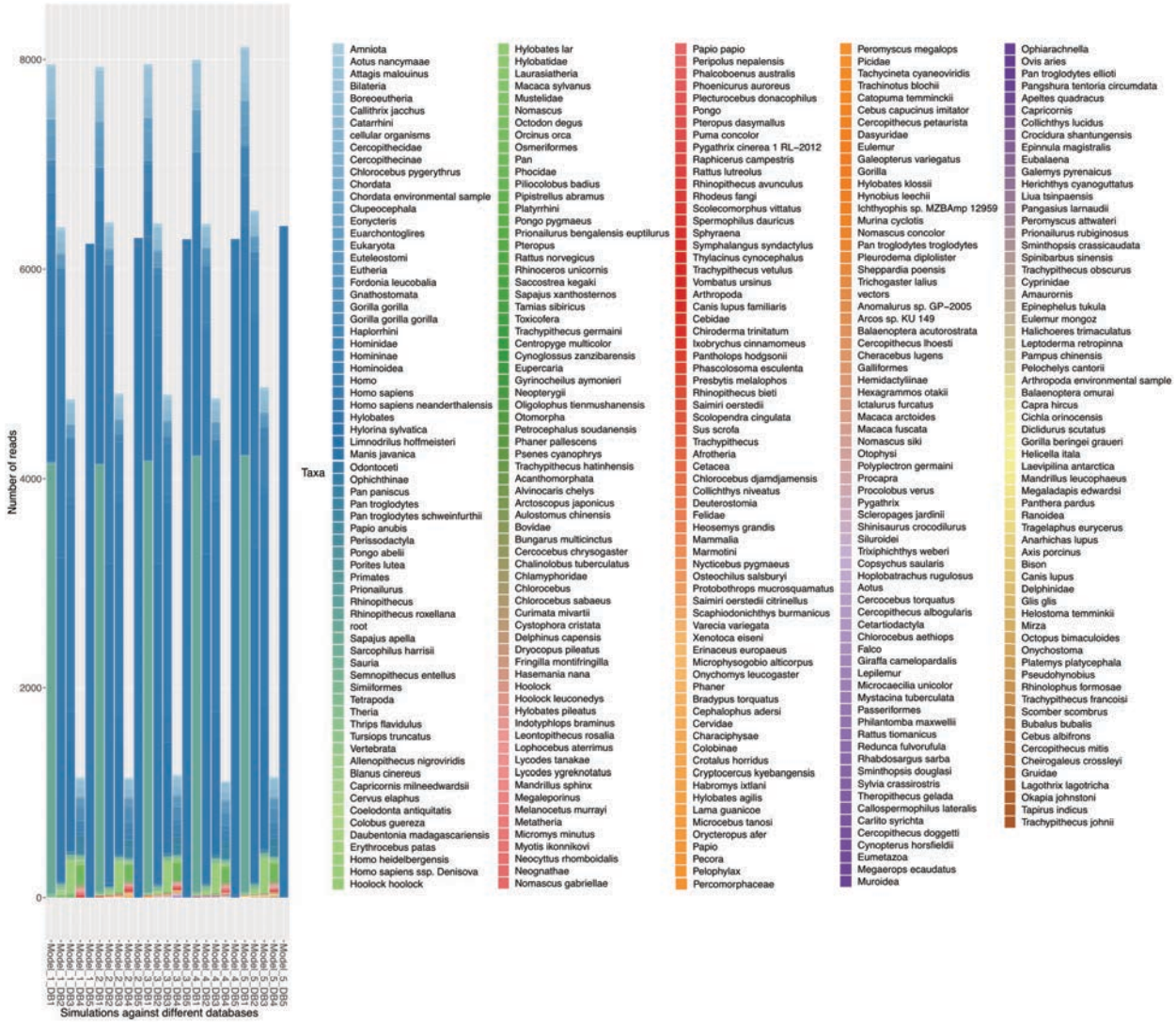


Fig. S41. Complete taxonomic profiles of damaged reads from *Homo sapiens* with no threshold applied.

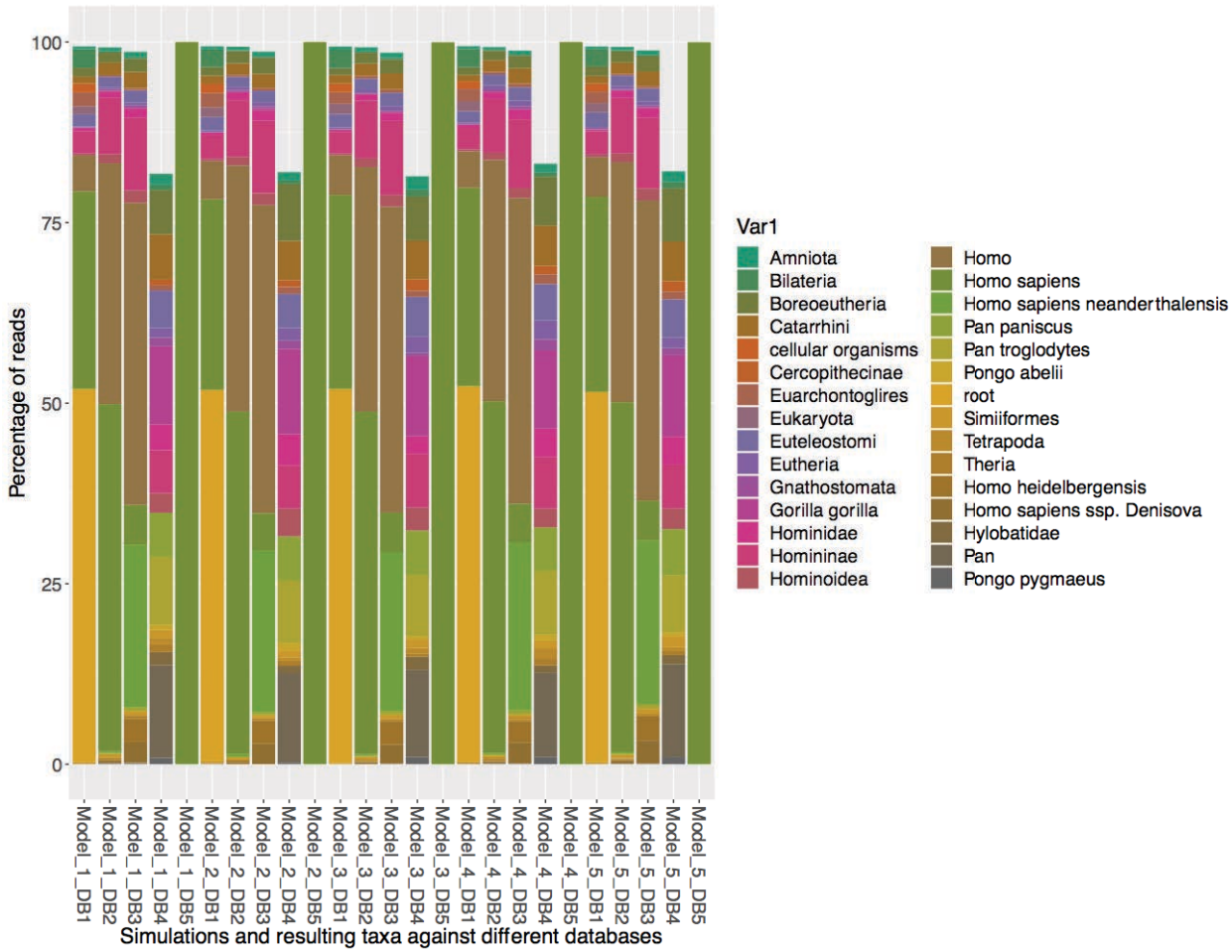


Fig. S42. Taxonomic profiles from *Homo sapiens* models of damaged reads above the threshold 1%.

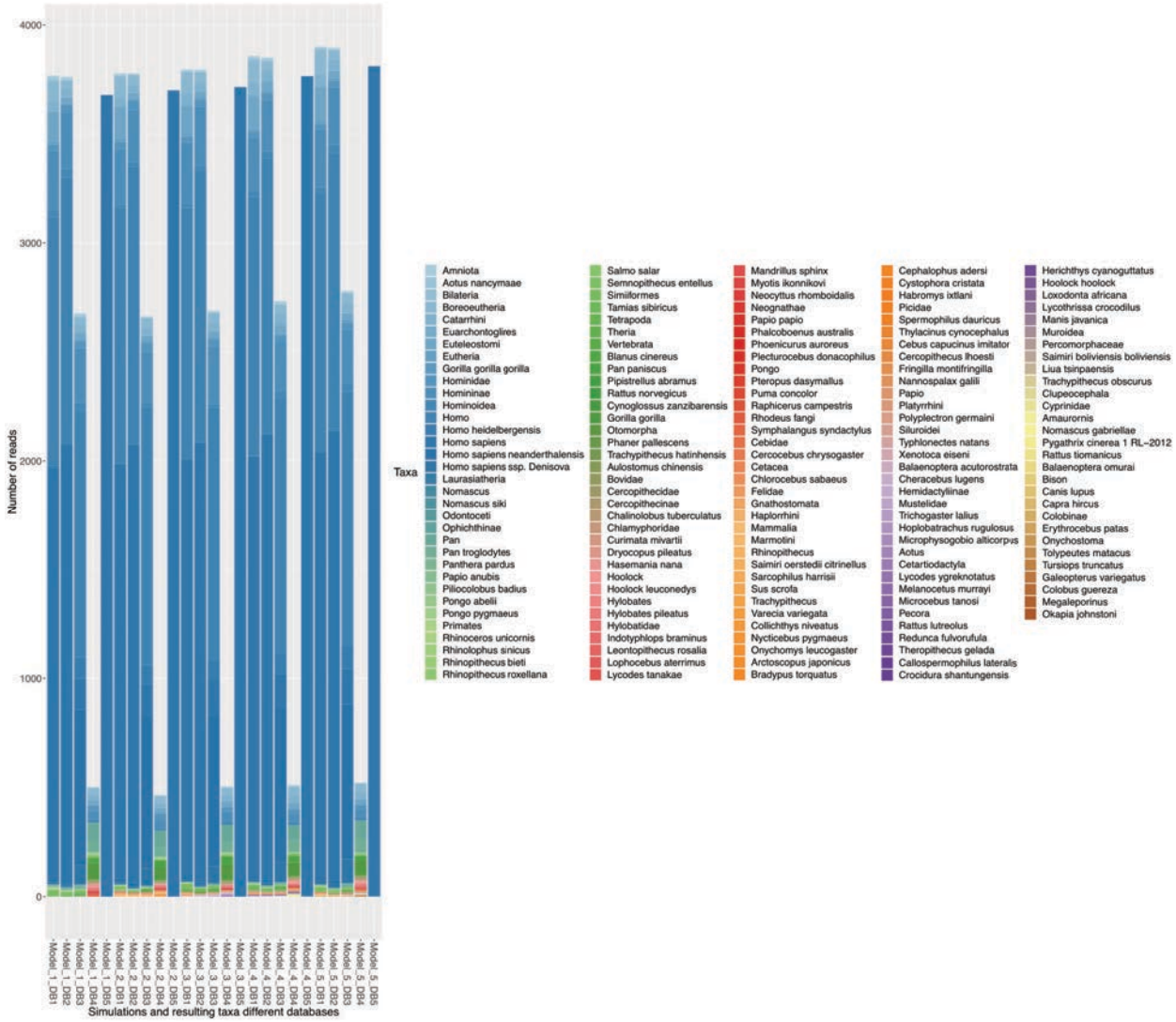


Fig. S43. Complete taxonomic profiles of damaged and 'sequenced' reads from *Homo sapiens* with no cut-off threshold applied.

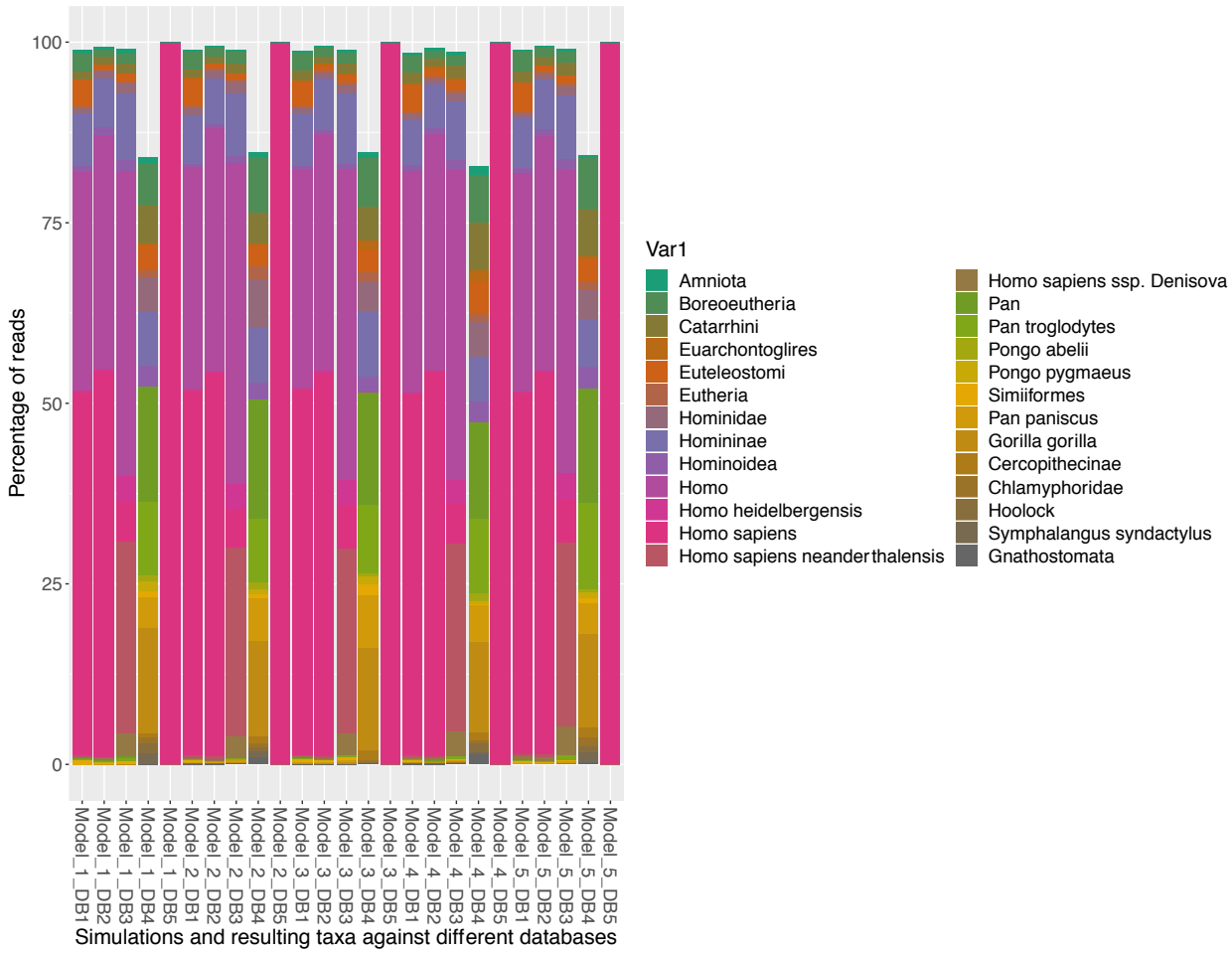


Fig. S44. Taxonomic profiles from *Homo sapiens* models of damaged and 'sequenced' reads above the threshold 1%.

	Model_1_DB1	Model_1_DB2	Model_1_DB3	Model_1_DB4	Model_1_DB5	Model_2_DB1	Model_2_DB2	Model_2_DB3	Model_2_DB4	Model_2_DB5
Amniota	0.424741173347491	0.398830098378091	0.560328726186029	0.793650793650794	0	0.185283218634198	0.185283218634198	0.2628614344724	0.642398286937901	0
Boreoeutheria	2.49535439341651	0.983780909332624	1.34478894284647	5.95238095238095	0	2.48808893594494	1.21757543673902	1.68982350732257	7.70877944325482	0
Catarrhini	1.16803822677056	1.01036958255783	1.38214419125887	5.15873015873016	0	1.13816834303864	0.92641609317099	1.23920390536968	4.28265524625268	0
Euarchontoglires	0.0530926466684364	0.0531773464504121	0.0747104968248039	0.396825396825397	0	0.0794070937003706	0.105876124933827	0.112654900488171	0.214132762312634	0
Euteleostomi	3.63684629678789	0.771071523530976	1.00859170713485	3.37301587301587	0	3.83800952685124	0.582319687136051	0.751032669921142	2.99785867237687	0
Eutheria	0.29200955666764	0.239298059026855	0.336197235711617	0.992063492063492	0	0.344097406034939	0.317628374801482	0.450619601952685	1.71306209850107	0
Hominidae	0.769843376692328	0.904014889657006	1.27007844602167	4.76190476190476	0	0.899947061937533	1.08523028057173	1.53961697333834	6.85224839400428	0
Hominoidea	7.167507300238892	6.833289018877960	9.301456854688080	7.539682539682540	0	6.749602964531	6.326098464796190	8.711978971085240	7.494646680942180	0
Homininae	0.876028670029201	1.036958255783040	1.456854688083680	2.777777777777780	0	0.423504499735310	0.608787718369508	0.863687570409313	2.355460385438970	0
Homo	30.289354924343	32.41159266152620	42.24878595442660	0	0	30.88935594494441	33.61566966649020	44.53623732632370	0	0
Homo heidelbergensis	0.238916910007964	0.265886732252061	3.660814344415390	0	0	0.132345156167284	0.132345156167284	3.342095381149080	0	0
Homo sapiens	50.25219007167510	53.41664509439	5.379155771385880	0	100	50.39703546850180	53.20275277924830	5.407435223432220	0	100
Homo sapiens neanderthalensis	0.238916910007964	0.239298059026855	26.63429211804260	0	0	0.449973530968767	0.476442562202223	26.17348854675180	0	0
Homo sapiens ssp. Denisova	0.238916910007964	0.239298059026855	3.212551363466570	0	0	0.211752249867655	0.211752249867655	3.004130679684570	0	0
Pan	0.0265463233342182	0.106354692900824	0.149420993649608	16.07142857142860	0	0	0	0	16.48822269807280	0
Pan troglodytes	0.132731616671091	0.132943366812603	0.410907732536421	10.11904761904760	0	0.026469031233456	0.026469031233456	0.187758167480285	8.779443254817990	0
Pongo abelii	0.0265463233342182	0.0265886732252061	0.037355248412401	0.793650793650794	0	0.026469031233456	0.026469031233456	0.037551633496057	1.070663811563170	0
Pongo pygmaeus	0.0796389700026546	0.0797660196756182	0.112065745237206	1.388888888888890	0	0	0	0	0.642398286937901	0
Similiformes	0.557472790018582	0.212709385801649	0.298841987299216	0.793650793650794	0	0.397035468501853	0.079407093700706	0.112654900488171	0.428265524625268	0
Pan paniscus	0	0.0265886732252061	0.074710496824803	4.365079365079360	0	0.052938062466913 7	0.079407093700706	0.225309800976342	5.995717344753750	0
Gorilla gorilla	0	0	0.037355248412401 9	14.48412698412700	0	0.026469031233456 9	0.079407093700706	0.112654900488171	13.27623126338330	0
Cercopithecinae	0	0	0	0.595238095238095	0	0.026469031233456 9	0	0	0.856531049250535	0
Chlamyphoridae	0	0	0	0.793650793650794	0	0	0	0	0.642398286937901	0
Hoolock	0	0	0	1.388888888888890	0	0	0	0	0.428265524625268	0
Symphalangus syndactylus	0	0	0	1.587301587301590	0	0.026469031233456 9	0.026469031233456 9	0.037551633496057 1	0.856531049250535	0
Gnathostomata	0	0	0	0	0	0.132345156167284	0.132345156167284	0.187758167480285	1.070663811563170	0

	Model_3_DB1	Model_3_DB2	Model_3_DB3	Model_3_DB4	Model_3_DB5	Model_4_DB1	Model_4_DB2	Model_4_DB3	Model_4_DB4	Model_4_DB5
Amniota	0.237154150197628	0.237216657880865	0.334696913350688	0.790513833992095	0	0.336962156557802	0.337750064951936	0.475667764361508	1.169590643274850	0
Boreoeutheria	2.476943346500856	1.159725882897312	1.524730383042020	6.719367588932810	0	2.436495593571 8	1.039230989082880	1.427003293084520	6.627680311890840	0
Catarrhini	1.343873517786560	1.054296257248290	1.450353291186310	4.743080309525700	0	1.425609123898390	1.195115614445310	1.719721917306990	6.627680311890840	0
Euarchontoglires	0.184453227931489	0.184501845018450	0.260319821494980	1.185770750988140	0	0.259201658890617	0.259807742270720	0.365898280278080	1.559454191033140	0
Euteleostomi	3.346508563899870	0.738007380073801	0.966902194124210	3.162055335968380	0	3.810264385692070	1.065211743309950	1.317233809001100	4.288499025341130	0
Eutheria	0.237154150197628	0.210859251449657	0.297508367422834	1.383399209486170	0	0.233281493001555	0.233826968043648	0.329308452250274	1.364522417154	0
Hominidae	0.711462450592885	0.843437005798629	1.190033496991340	3.952569169960470	0	0.751684810782789	0.831384775266303	1.170874496889860	4.678362573099410	0
Hominoidea	7.430830039525690	7.195571955719560	9.817776124953510	9.090909090909090	0	6.324520476931050	6.079501189134840	8.196121478229050	6.237816764132550	0
Homininae	0.474308300395257	0.474433315761729	0.632205280773522	2.173913043478260	0	0.725764644893727	0.883346323720447	1.244054152945480	2.923976608187130	0
Homo	30.30303030303030	32.84132841328410	43.02714763852730	0	0	30.71539657853810	32.70979475188360	42.99304793267470	0	0
Homo heidelbergensis	0.131752305665349	0.131787032156036	3.495723317218 3	0	0	0.155520995334370	0.155884645362432	3.293084522502740	0	0
Homo sapiens	50.56653491436 1	53.21560358460 73	5.987355894384530	0	100	49.97407983411090	53.15666406858920	5.561653860226860	0	99.97343957503320
Homo sapiens neanderthalensis	0.263504611330698	0.263574064312072	25.73447378207510	0	0	0.285121824779679	0.311769290724864	25.97877789974390	0	0
Homo sapiens ssp. Denisova	0.210803689064559	0.210859251449657	2.937895128300480	0	0	0.336962156557802	0.337750064951936	3.402854006586170	0	0
Pan	0	0	0	15.61264822313390	0	0.025920165889061 7	0.051961548454143 9	0.073179656055616 5	13.25536062378170	0
Pan troglodytes	0.052700922266139 7	0.079072219293621 5	0.260319821494980	9.486166007905140	0	0.129060829445308	0.129903871135360	0.329308452250274	10.33138401559450	0
Pongo abelii	0	0	0	0.395256916996047	0	0	0	0	1.169590643274850	0
Pongo pygmaeus	0	0	0	1.185770750988140	0	0.025920165889061 7	0.025980774227072 0	0.036589828027808 3	0.389863547758285	0
Similiformes	0.527009222661397	0.289931470743279	0.409074005206396	1.383399209486170	0	0.259201658890617	0.051961548454143 9	0.073179656055616 5	0.194931773879142	0
Pan paniscus	0.105401844532279	0.105429625724829	0.297508367422834	7.312252964426880	0	0.025920165889061 7	0.025980774227072 0	0.219538966166850	5.068226120857 7	0
Gorilla gorilla	0.131752305665349	0.131787032156036	0.223131275567125	14.22924901185770	0	0.025920165889061 7	0.025980774227072 0	0.073179656055616 5	12.47563352826510	0
Cercopithecinae	0.026350461133069 8	0.052714812862414 3	0.074377091855708 4	1.383399209486170	0	0.051840331781234	0.051961548454143 9	0.073179656055616 5	1.169590643274850	0
Chlamyphoridae	0	0	0	0.197628458498024	0	0	0	0	0.389863547758285	0
Hoolock	0	0	0	0.197628458498024	0	0.025920165889061 7	0.025980774227072 0	0.036589828027808 3	1.169590643274850	0
Symphalangus syndactylus	0	0	0	0	0	0	0	0	0.389863547758285	0
Gnathostomata	0.026350461133069 8	0.026357406431207 2	0.037188545927854 2	0.197628458498024	0	0.181441161223432	0.181865419589504	0.256128796194658	1.364522417154	0

	Model_5_DB1	Model_5_DB2	Model_5_DB3	Model_5_DB4	Model_5_DB5
Amniota	0.205128205128205	0.205391527599487	0.287769784172662	0.381679389312977	0
Boreoeutheria	2.76923076923077	1.10397946084724	1.54676258992806	7.06106870229008	0
Catarrhini	1.461538461538460	1.258023106546860	1.690647482014390	6.488549618320610	0
Euarchontoglires	0.076923076923076 9	0.077021822849807 5	0.107913669064748	0.381679389312977	0
Euteleostomi	3.974358974358970	0.821566110397946	1.043165467625 9	3.053435114503820	0
Eutheria	0.333333333333333	0.333761232349166	0.467625899280576	1.145038167938930	0
Hominidae	0.641025641025641	0.795892169448010	1.115107913669060	4.198473282442750	0
Homininae	6.794871794871790	6.777920410783050	8.920863309352520	6.488549618320610	0
Hominoidea	0.769230769230769	0.975609756097561	1.366906474820140	3.053435114503820	0
Homo	30.20512820512820	32.50320924261870	42.05035971223020	0	0
Homo heidelbergensis	0.179487179487179	0.179717586649551	3.741007194244 8	0	0
Homo sapiens	49.97435897435900	52.96534017971760	5.863309352517990	0	99.97376016793490
Homo sapiens neanderthalensis	0.333333333333333	0.333761232349166	25.53956834533237	0	0
Homo sapiens ssp. Denisova	0.564102564102564	0.564826700898588	3.956834532374 1	0	0
Pan	0	0.025673940949935 8	0.035971223021582 7	15.83969465648850	0
Pan troglodytes	0.102564102564103	0.128369704749679	0.611510791366906	12.02290076335880	0
Pongo abelii	0	0	0	0.381679389312977	0
Pongo pygmaeus	0	0	0	0.763358778625954	0
Similiformes	0.358974358974359	0.179717586649551	0.251798561151079	0.763358778625954	0
Pan paniscus	0.025641025641025 6	0.051347881899871 6	0.179856115107914	4.198473282442750	0
Gorilla gorilla	0.025641025641025 6	0.025673940949935 8	0.071942446043165 5	12.97709923664120	0
Cercopithecinae	0	0	0	1.335877862595420	0
Chlamyphoridae	0	0	0	1.335877862595420	0
Hoolock	0.025641025641025 6	0.025673940949935 8	0.035971223021582 7	0.763358778625954	0
Symphalangus syndactylus	0.025641025641025 6	0.025673940949935 8	0.035971223021582 7	1.335877862595420	0
Gnathostomata	0.051282051282051 3	0.051347881899871 6	0.071942446043165 5	0.381679389312977	0

Table S15. Resulting taxa for scenario c for *Homo sapiens* model plotted in Fig. S11.

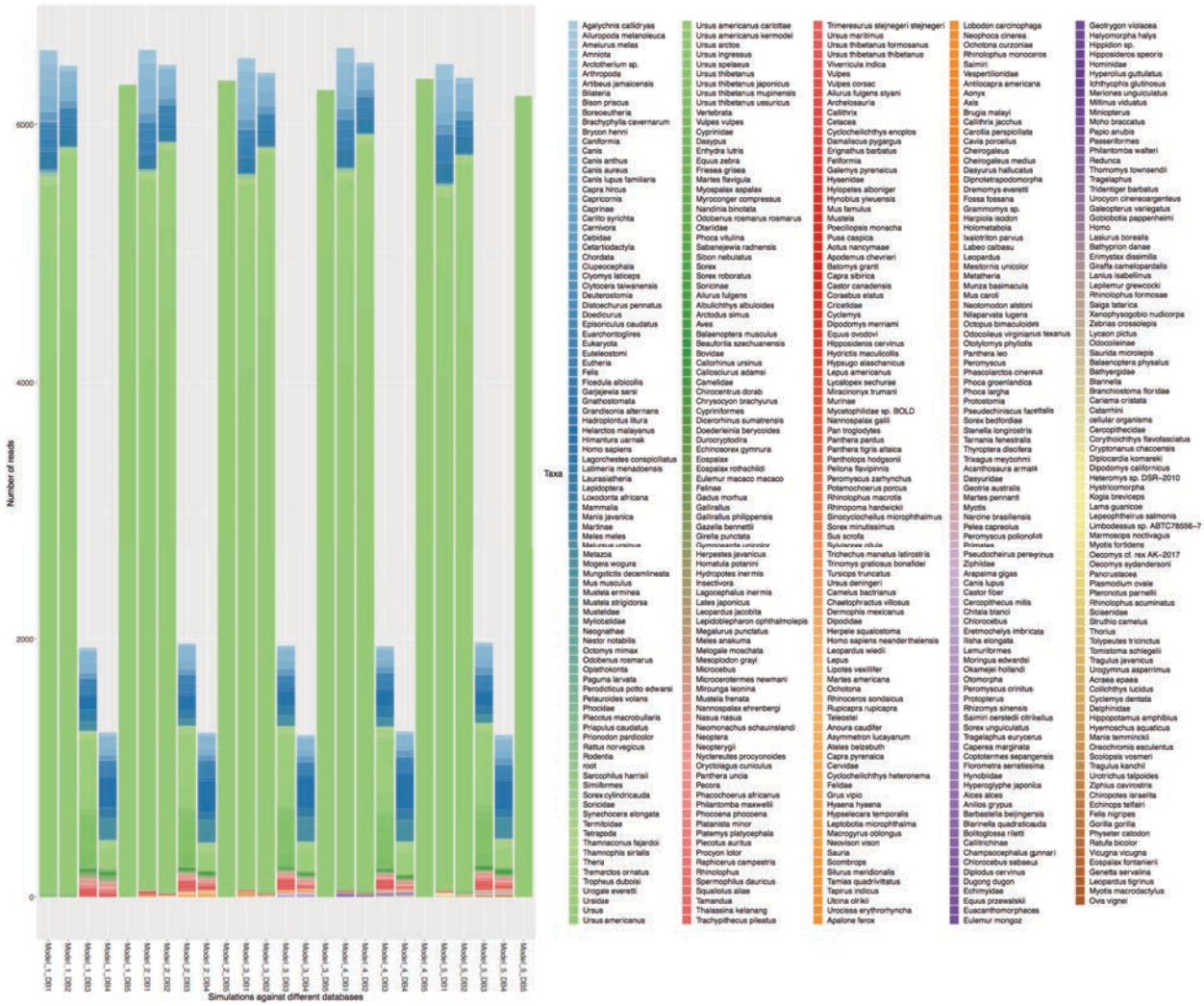


Fig. S45. Complete taxonomic profiles of damaged and 'sequenced' reads from *Ursus americanus* with no cut-off threshold applied.

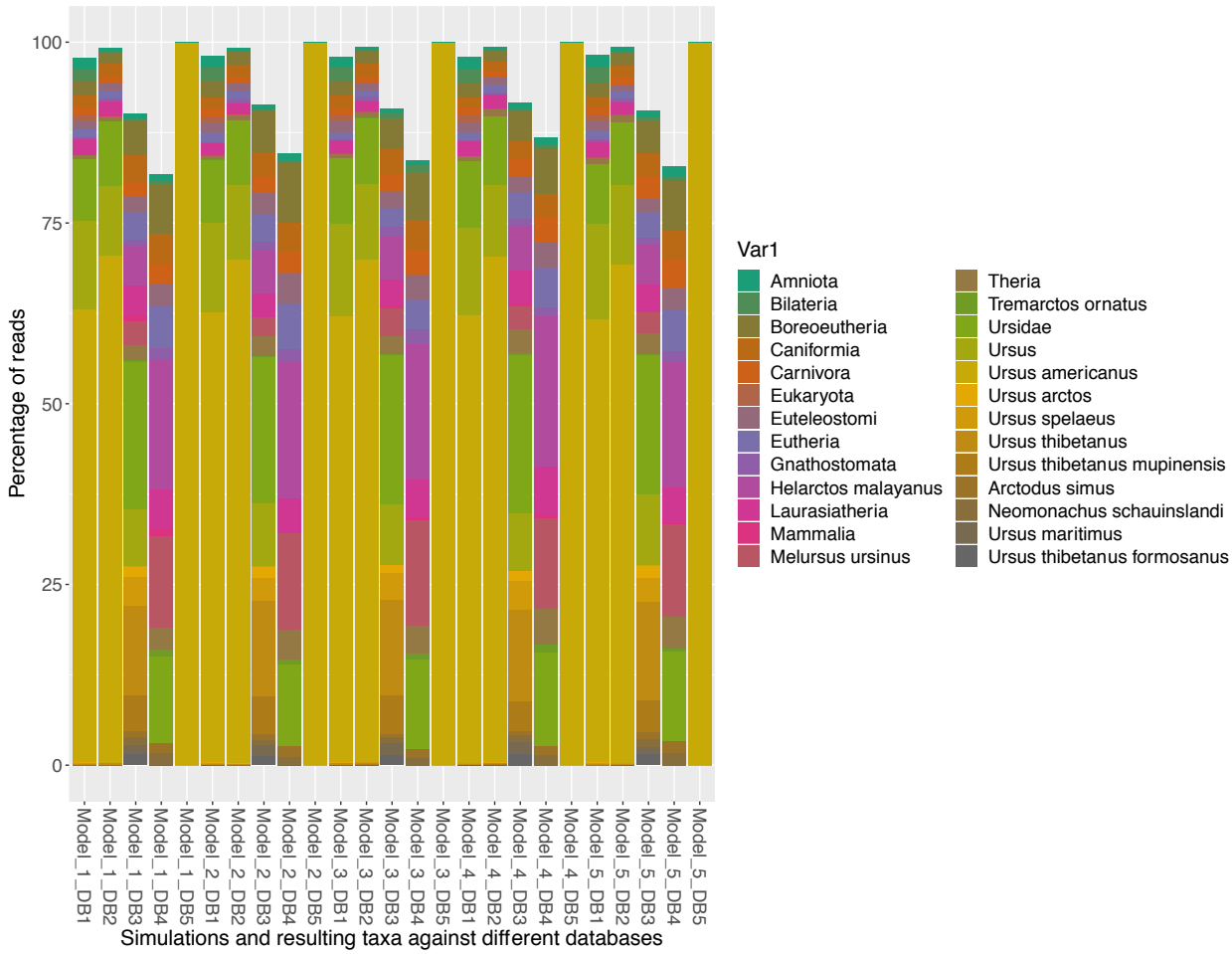


Fig. S46. Taxonomic profiles from *Ursus americanus* models of damaged and 'sequenced' reads (scenario c) above the threshold 1%.

	Model_1_DB1	Model_1_DB2	Model_1_DB3	Model_1_DB4	Model_1_DB5	Model_2_DB1	Model_2_DB2	Model_2_DB3	Model_2_DB4	Model_2_DB5
Amniota	1.50524555268359	0.387356678029129	0.569063631660631	0.782472613458529	0	1.50501672240803	0.371459526389104	0.610687022900763	0.86342229199372	0
Bilateria	1.88535806598753	0.201425472575147	0.362131401965856	0.54773082942097	0	1.85466707205838	0.123819842129701	0.254452926208651	0.392464678178964	0
Boreoeutheria	1.8245400638589	1.61140378060118	4.75944128297982	6.96400625978091	0	2.29553055640012	1.84182015167931	5.59796437659033	8.16326530612245	0
Caniformia	1.74851756119811	1.85931205453982	3.72478013450595	4.1471048513302	0	1.65703861356035	1.71800030954961	3.51145038167939	4.00313971742543	0
Carnivora	1.03390603618671	0.836690424542919	2.06932229694775	2.8169014084507	0	1.03374885983582	0.727441572511995	2.08651399491094	2.98273155416013	0
Eukaryota	0.714611525011403	0	0	0	0	0.805716023107327	0	0	0	0
Euteleostomi	1.17074654097613	0.945150294391075	2.12105535437144	3.05164319248826	0	1.45941015506233	1.17628850023216	2.90076335877863	4.3956043956044	0
Eutheria	1.14033753991181	1.14657576696622	3.77651319192964	5.71205007824726	0	1.23137731833384	1.2227209410308	3.96946564885496	6.12244897959184	0
Gnathostomata	0.212863007450205	0.356368143786799	1.03466114847387	1.56494522691706	0	0.152021891152326	0.371459526389104	1.17048346055598	1.80533751962323	0
Helarctos malayanus	0.0456135015964726	0.0619770684846607	5.27677185721676	17.9968701095462	0	0.0456065673456978	0.092864881597276	5.85241730279898	18.9167974882261	0
Laurasiatheria	2.12863007450205	1.82832352029749	4.29384376616658	5.39906103286385	0	1.67224080267559	1.4703606252902	3.3587786259542	4.63108320251177	0
Mammalia	0.0304090010643315	0.216919739696312	0.724262803931712	1.09546165884194	0	0	0.0154774802662127	0.0508905852417303	0.0784929356357928	0
Melursus ursinus	0.0760225026607876	0.0619770684846607	3.31091567511164	12.5978090766823	0	0.0456065673456978	0.0619099210648506	2.59541984732824	13.4222919937206	0
Theria	0.471339516496883	0.619770684846607	2.06932229694775	3.12989045383412	0	0.532076619033141	0.820306454109271	2.69720101781171	4.16012558869702	0
Tremarctos ornatus	0.0304090010643315	0.0309885342423303	0.258665287118469	0.860719874804382	0	0	0	0.254452926208651	0.549450549450549	0
Ursidae	8.57533830013684	8.87821506042764	20.3310915675116	11.9718309859155	0	8.69565217391304	8.93050611360471	20.1526717557252	11.38147566719	0
Ursus	12.3004409305154	9.73039975209173	7.91515778582514	0	0	12.4353906962603	10.4008667388949	8.70229007633588	0	0
Ursus americanus	62.551315189296	70.0495816547877	0	0	100	62.2529644268775	69.5867512768921	0	0	100
Ursus arctos	0.136840504789418	0.0464828013634955	1.50025866528712	0	0	0.076019045576163	0.046432440798638	1.52671755725191	0	0
Ursus spelaeus	0.106431503725103	0.108459869848156	3.87997930677703	0	0	0.076010945576163	0.0619099210648506	3.25699745547074	0	0
Ursus thibetanus	0.167249505853733	0.170436938332817	12.4676668391102	0	0	0.212830647613256	0.154774802662127	13.2315521628499	0	0
Ursus thibetanus mupinensis	0.0608180021286301	0.0774713356058258	4.9663735126746	0	0	0.0152021891152326	0.0154774802662127	5.13994910941476	0	0
Arctodus simus	0	0	0.775995861355406	1.4866979655712	0	0	0.0154774802662127	0.865139949109415	1.49136577708006	0
Neomonachus schauinslandi	0	0	1.08639420589757	1.84319248826291	0	0	0	0.763358778625954	1.17739403453689	0
Ursus maritimus	0	0	1.34505949301604	0	0	0.0152021891152326	0	1.42493638676845	0	0
Ursus thibetanus formosanus	0	0	1.50025866528712	0	0	0.0152021891152326	0.0154774802662127	1.32315521628499	0	0

	Model_3_DB1	Model_3_DB2	Model_3_DB3	Model_3_DB4	Model_3_DB5	Model_4_DB1	Model_4_DB2	Model_4_DB3	Model_4_DB4	Model_4_DB5
Amniota	1.45839729812711	0.359318856428683	0.512820512820513	0.636435958631663	0	1.65326861823146	0.370484717505403	0.720164609053498	0.933852140077821	0
Bilateria	1.98035001535155	0.281206061552882	0.82051282051282	1.11376292760541	0	2.01729106628242	0.123494905835134	0.360082304526749	0.544747081712062	0
Boreoeutheria	1.82683451028554	1.57787845649117	4.25641025641026	6.52346857597454	0	1.9414530562718	1.45106514356283	4.16666666666667	6.30350194552529	0
Caniformia	1.84218606079214	1.81221684111858	3.43589743589744	4.05727923627685	0	1.42575458819961	1.29669651126891	2.46913580246914	3.11284046692607	0
Carnivora	1.04390543444888	0.906108420559288	2.41025641025641	3.50039777247414	0	1.07689974215077	0.802716887928373	2.31481481481481	3.50194552529183	0
Eukaryota	0.752225974823457	0	0	0	0	1.01622933414227	0	0	0	0
Euteleostomi	1.51980350015352	1.07795656926605	2.35897435897436	3.50039777247414	0	1.31958137418474	1.09601728926682	2.41769547325103	3.57976653696498	0
Eutheria	0.936444580902671	0.827995625683487	2.61538461538462	4.05727923627685	0	1.09206734415289	1.06514356282803	3.54938271604938	5.36964980544747	0
Gnathostomata	0.13816395455941	0.453054210279644	1.33333333333333	2.0884168655529	0	0.106173214014864	0.26242667489966	0.874485596707819	1.32295719844358	0
Helarctos malayanus	0.107460853546208	0.124980471801281	5.84615384615385	18.6953062848051	0	0.0606704080084939	0.0771843161469589	6.22427983539095	20.7782101167315	0
Laurasiatheria	1.73472520724593	1.42165286673957	3.74358974358974	5.25059665871122	0	1.95662065827393	1.82154986106823	4.57818930041152	6.61478599922179	0
Mammalia	0	0.0624902359006405	0.205128205128205	0.318217979315831	0	0	0.0926211793763507	0.308641975308642	0.466926070038911	0
Melursus ursinus	0.0614062020264047	0.0937359538509608	3.8974358974359	14.6380270485282	0	0.0606704080084939	0.0463105896881754	3.29218106995885	12.6070038910506	0
Theria	0.583358919250844	0.749882830807686	2.46153846153846	3.81861575178998	0	0.621871682087062	0.972522383451683	3.24074074074074	4.90272373540856	0
Tremarctos ornatus	0	0	0.205128205128205	0.875099443118536	0	0.0151676020021235	0.0154368632293918	0.308641975308642	1.16731517509728	0
Ursidae	9.1802272029475	9.20168723636932	20.5128205128205	12.4105011933174	0	9.23706961929319	9.61716579191108	21.8106995884774	12.9182879377432	0
Ursus	12.6957322689592	10.3890017184815	8.51282051282051	0	0	12.1340816016988	9.78697128743439	8.02469135802469	0	0
Ursus americanus	61.7285845870433	69.5360099984377	0	0	100	61.8686485666616	69.9907378820624	0	0	100
Ursus arctos	0.0767577525330058	0.0468676769254804	1.07692307692308	0	0	0.0910056120127408	0.0308737264587836	1.33744855967078	0	0
Ursus spelaeus	0.0460546515198035	0.109357912826121	3.69230769230769	0	0	0.0151676020021235	0.0926211793763507	4.11522633744856	0	0
Ursus thibetanus	0.230273257599017	0.156225589751601	13.2307692307692	0	0	0.182011224025482	0.123494905835134	12.6028806584362	0	0
Ursus thibetanus mupinensis	0.0153515505066012	0.0624902359006405	5.33333333333333	0	0	0.0151676020021235	0.0308737264587836	4.06378600823045	0	0
Arctodus simus	0.0153515505066012	0.0468676769254804	0.512820512820513	1.11376292760541	0	0.0455028060063704	0.0463105896881754	0.668724279835391	1.24513618677043	0
Neomonachus schauinslandi	0	0.0156225589751601	0.717948717948718	1.11376292760541	0	0	0	0.925925925925926	1.40077821011673	0
Ursus maritimus	0.0307031010132023	0.0312451179503203	1.69230769230769	0	0	0.0151676020021235	0.0617474529175671	1.64609053497942	0	0
Ursus thibetanus formosanus	0	0	1.43589743589744	0	0	0.0303352040042469	0.0308737264587836	1.54320987654321	0	0

	Model_5_DB1	Model_5_DB2	Model_5_DB3	Model_5_DB4	Model_5_DB5
Amniota	1.73186949126334	0.55022795157994	0.861632032437912	1.35242641209228	0
Bilateria	1.94835317767125	0.172928784782267	0.456158134820071	0.636435958631663	0
Boreoeutheria	2.0411319004175	1.69784625058953	4.56158134820071	6.92124105011933	0
Caniformia	1.40714396165146	1.57207986165697	3.34515965534719	4.05727923627685	0
Carnivora	1.14427091387042	1.10045590315988	2.83831728332489	3.89817024661893	0
Eukaryota	0.881397866089377	0	0	0	0
Euteleostomi	1.19066027524354	0.738877534978777	1.92600101368474	2.94351630867144	0
Eutheria	1.2989021184475	1.21050149347587	3.59858084135834	5.64836913285601	0
Gnathostomata	0.123704963661667	0.314415972331394	0.912316269640142	1.43198090692124	0
Helarctos malayanus	0.154631204577084	0.141487187549128	5.57526609224531	17.4224343675418	0
Laurasiatheria	2.14937374362146	1.68212545197296	3.64926507856057	4.93237867939539	0
Mammalia	0	0.0471623958497092	0.15205271160669	0.238663484486874	0
Melursus ursinus	0.0773156022885418	0.110045590315988	2.93968575772935	12.7287191726333	0
Theria	0.819545384258543	0.880364722527905	2.83831728332489	4.45505171042164	0
Tremarctos ornatus	0	0	0.15205271160669	0.397772474144789	0
Ursidae	8.33462192670481	8.78792642666247	19.2600101368474	12.4105011933174	0
Ursus	13.1591155095098	10.8630718440497	9.88342625443487	0	0
Ursus americanus	61.3112726148137	68.9828643295079	0	0	100
Ursus arctos	0.123704963661667	0.0471623958497092	1.67257982767359	0	0
Ursus spelaeus	0.0463893613731251	0.110045590315988	3.29447541814496	0	0
Ursus thibetanus	0.231946806865625	0.188649583398837	13.6340598073999	0	0
Ursus thibetanus mupinensis	0.0154631204577084	0.0157207986165697	4.46021287379625	0	0
Arctodus simus	0	0	0.912316269640142	1.67064439140811	0
Neomonachus schauinslandi	0	0.0157207986165697	1.06436898124683	1.67064439140811	0
Ursus maritimus	0.0154631204577084	0.0157207986165697	0.963000506842372	0	0
Ursus thibetanus formosanus	0	0	1.62189559047136	0	0

Table S16. Resulting taxa for scenario c for *Ursus americanus* model plotted in Fig. S13.

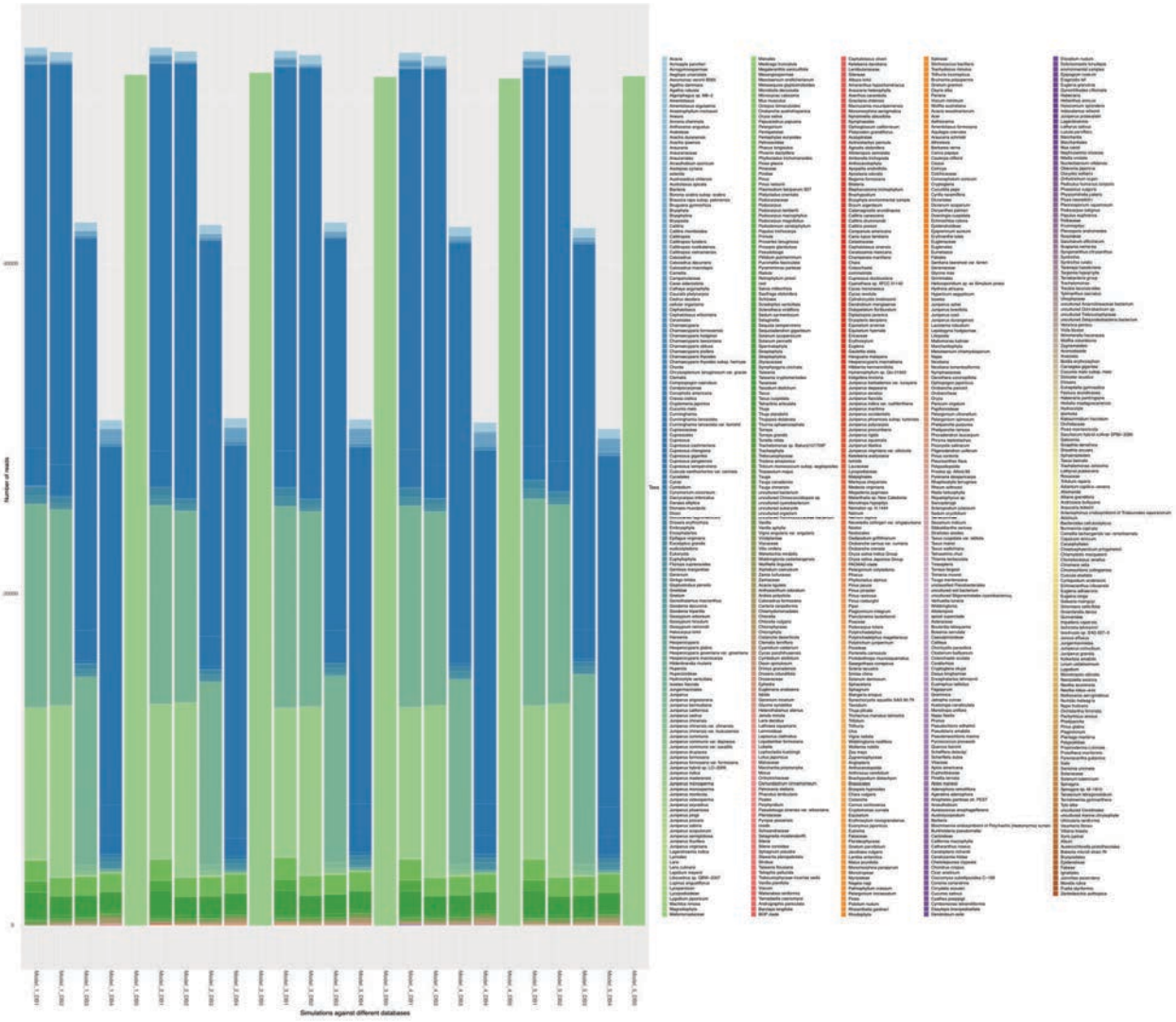


Fig. S47. Complete taxonomic profiles of damaged and ‘sequenced’ reads from *Juniper monosperma* with no cut-off threshold applied.

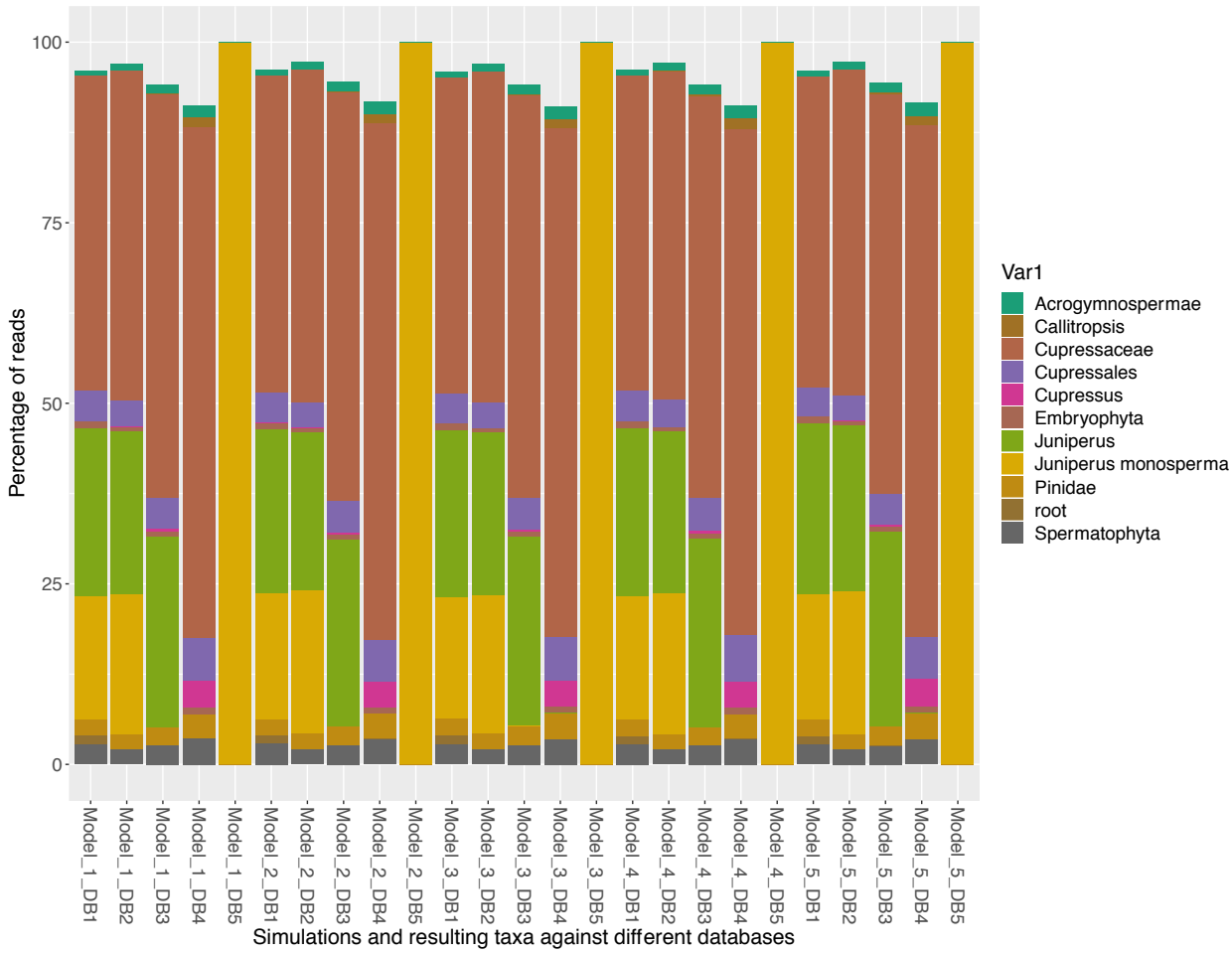


Fig. S48. Taxonomic profiles from *Juniper monosperma* models of damaged and 'sequenced' reads above the threshold 1%.

	Model_1_DB1	Model_1_DB2	Model_1_DB3	Model_1_DB4	Model_1_DB5	Model_2_DB1	Model_2_DB2	Model_2_DB3	Model_2_DB4	Model_2_DB5
Acrogymnospermae	0.680490103675778	0.945451789537505	1.16456887425009	1.59638849815172	0	0.735224809124328	1.01861143192533	1.26543428476993	1.72717200026071	0
Callitropsis	0.01131100848256362	0.01136815779003 01	0.167039171862134	1.24963198011057	0	0.011311150996051	0.01135997879470 63	0.188870788771632	1.24486736622564	0
Cupressaceae	43.6531573986805	45.7246253244662	55.8169627102694	70.7710425594557	0	44.009802997455	46.0552473635383	56.6022145099983	71.5766147428795	0
Cupressales	4.17907634307257	3.57339093199947	4.3641924479473	5.97010042853872	0	4.05127721745688	3.506446787966	4.29681044455462	5.83327901974842	0
Cupressus	0.06786050895381 72	0.08526118342522 59	0.338783672509117	3.64421472733815	0	0.081063248185 5029	0.10223980915 2356	0.351771844087164	3.46086163071107	0
Embryophyta	0.982092365692743	0.587354819151557	0.722267968474297	0.974843796002486	0	0.918088415496277	0.533919003351194	0.663408645560356	0.889656520889005	0
Juniperus	23.2026390197926	22.4843214157146	26.3757205034702	0	0	22.6863983410312	21.9758789783592	25.8634936374153	0	0
Juniperus monosperma	17.1008482563619	19.4660755224616	0	0	100	17.4342539353379	19.8042296987712	0	0	100
Pinidae	2.19604147031103	2.04816309517043	2.54087754381837	3.35961267951192	0	2.27542652464888	2.12620936440919	2.64419104280284	3.54233200808186	0
root	1.17436380772856	0	0	0	0	1.08964087095862	0	0	0	0
Spermatophyta	2.85956644674835	2.14858182231569	2.64674744147747	3.60168798455952	0	2.90885097558677	2.13188935380654	2.64891281252213	3.55210845336636	0

	Model_3_DB1	Model_3_DB2	Model_3_DB3	Model_3_DB4	Model_3_DB5	Model_4_DB1	Model_4_DB2	Model_4_DB3	Model_4_DB4	Model_4_DB5
Acrogymnospermae	0.711272534665078	1.01697492729105	1.25408813910261	1.7083129185 04	0	0.78871530411042	1.05267165400796	1.30448164019035	1.78113703581991	0
Callitropsis	0.00378336454609084	0.00570266314368 81	0.145878920495988	1.32941368610158	0	0.00379190050053087	0.00380713075590581	0.17282606122304	1.40322050607953	0
Cupressaceae	43.7962279855475	45.8304027981067	55.7986870897155	70.4458598726115	0	43.5348096465949	45.4647554870272	55.5955396671323	70.1314492277358	0
Cupressales	4.15980931842688	3.59647955595263	4.38813204395191	5.99706026457619	0	4.28105566509935	3.79570936363809	4.649731290987	6.34571146894512	0
Cupressus	0.04918373909918 09	0.05512574372231 83	0.256464553775205	3.58321084435734	0	0.05119065675716 67	0.05901052671654 01	0.30066999692227 6	3.61485376273414	0
Embryophyta	0.942057771976619	0.534149447792119	0.658808028046399	0.898252490609179	0	1.0257090853936	0.567262482629966	0.696039205473614	0.946434439697667	0
Juniperus	23.0501484970584	22.4494839089855	26.2370297169478	0	0	23.1362809039891	22.3440504064112	26.1606572125287	0	0
Juniperus monosperma	16.8378639123773	19.2084703556561	0	0	99.9980501121185	17.135598361899	19.5953020006472	0	0	100
Pinidae	2.37973629949114	2.18411998403254	2.69405425754688	3.59300996243671	0	2.27703625056879	2.05013991205528	2.53557139136817	3.38153138350312	0
root	1.18419310292643	0	0	0	0	1.21151220991961	0	0	0	0
Spermatophyta	2.82617331592986	2.14800311745585	2.65640808451566	3.57994447166422	0	2.74912786288488	2.14912531170883	2.66104784677668	3.59513637857378	0

	Model_5_DB1	Model_5_DB2	Model_5_DB3	Model_5_DB4	Model_5_DB5
Acrogymnospermae	0.770687369816323	1.044202678028	1.29737678478251	1.79760319573901	0
Callitropsis	0.00757432304487786	0.00760803408399 27	0.18262890754708	1.35818908122503	0
Cupressaceae	43.2077258095058	45.1175441265977	55.3602770267065	70.7556591211718	0
Cupressales	3.85154326832039	3.45214546561169	4.24078554148285	5.85885486018642	0
Cupressus	0.07763681120999 81	0.09510042604990 87	0.310706323229448	3.74500665778961	0
Embryophyta	0.922173830713 88	0.56679853925745 6	0.697310374270 67	0.955392809587217	0
Juniperus	23.7170990342738	22.9153986600 86	26.9626678051326	0	0
Juniperus monosperma	17.3111153190684	19.8075167376 75	0	0	100
Pinidae	2.29312630183677	2.13595556908095	2.65879227740619	3.59520639147803	0
root	1.1380420374929	0	0	0	0
Spermatophyta	2.81575459193335	2.11122945830797	2.62084341350031	3.56524633821571	0

Table S17. Resulting taxa for scenario c for *Juniper monosperma* model plotted in Fig. S15.

10.2. Testing the sensitivity and specificity for assigning sequenced DNA from the whole genome of *Homo sapiens*

Although the mitochondrial DNA and the chloroplast DNA (from animals and plants, respectively) represents different parts of the genomic pool present in an ancient environmental sample, it remains important to test entire genomes. This can be especially important for determining the robustness of identifying ancient human DNA in any sample. We therefore extracted 100.000 reads from the human genome (NCBI RefSeq assembly: GCF_000001405.38) using an identical approach to the above mtDNA and cpDNA modelling (outlines in Fig. S37). Using the read distributions shown in Fig. S35 as template, all reads were added deamination and sequencing errors (see Fig. S38). The model was repeated 5 times. The resulting taxonomic profiles for scenario c (the sequenced and damaged reads) were aligned to the Holi database (DB1) and against the reference genome (DB5) (see Fig. S49 and S50).

10.2.1. Discussion and conclusion

We find that inserting damage and sequencing errors on the reads generates a tail of false-positives (see Fig. S49). However, the proportion of each false-positives does not exceed the cut-off threshold of ≥ 1% (back-ground noise) and therefore not present in the final taxonomic profiles presented in Fig. S50. We find that ~21% of all reads can be assigned to species level and ~26% to family level, while no reads were assigned to genus level. The proportions are highly consistent between all 5 model iterations (see Table S18).

The taxonomic assignments are therefore showing the same robustness as found with the mtDNA and cpDNA above.

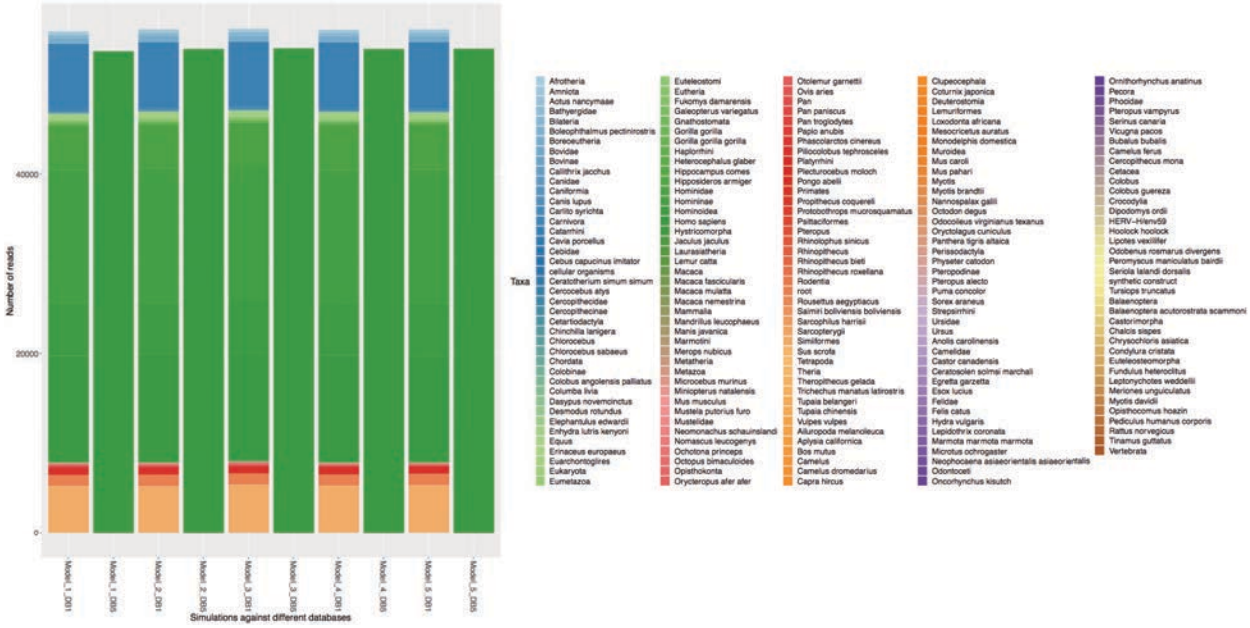


Fig. S49. Complete taxonomic profiles of damaged and 'sequenced' reads from *Homo sapiens* with no cut-off threshold applied.

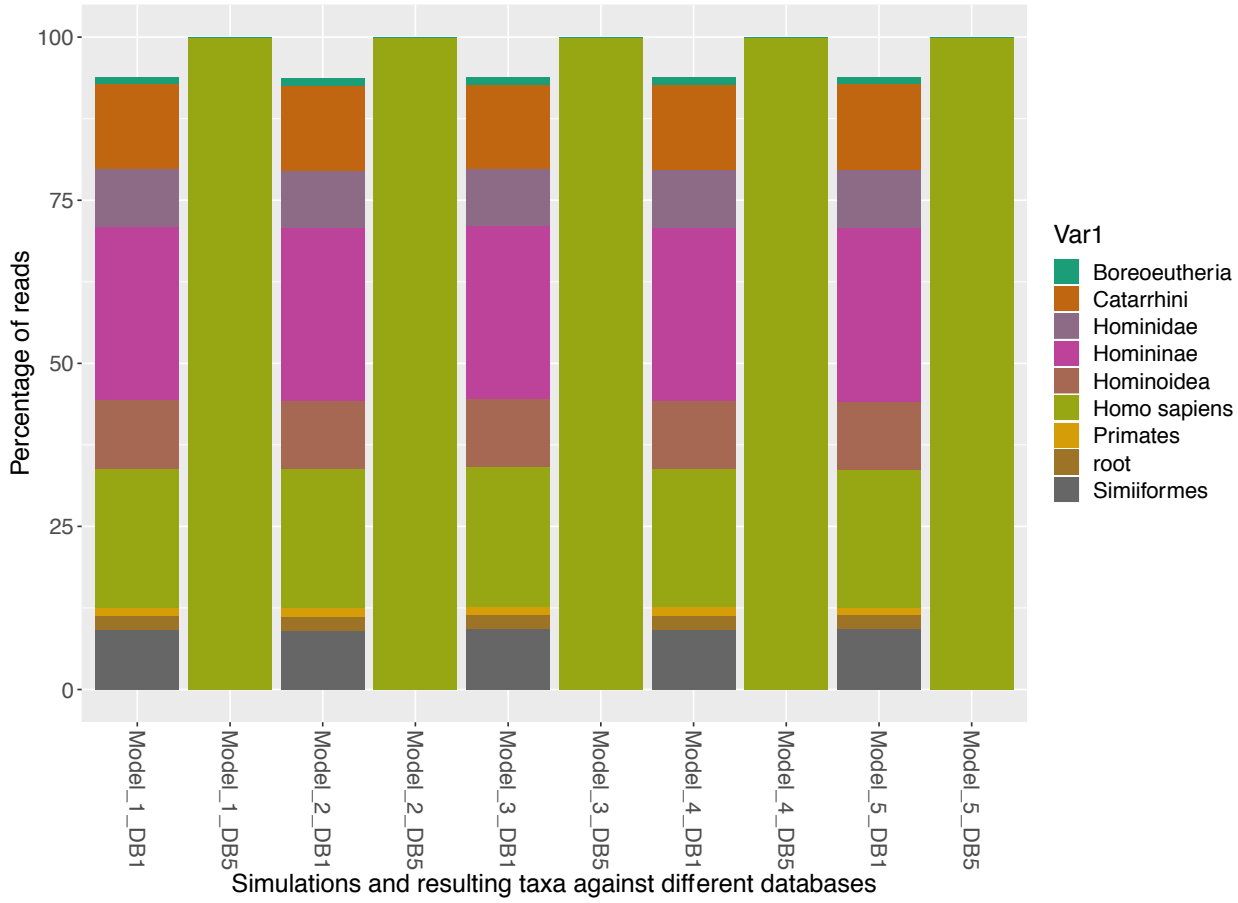


Fig. S50. Taxonomic profiles from *Homo sapiens* models of damaged and 'sequenced' reads above the threshold >1%.

Table S18. Resulting taxa for scenario c for *Homo sapiens* model plotted in Fig. S17.

	Model_1_DB1	Model_1_DB5	Model_2_DB1	Model_2_DB5	Model_3_DB1	Model_3_DB5	Model_4_DB1	Model_4_DB5	Model_5_DB1	Model_5_DB5
Boreoeutheria	1.02924855905202	0	1.15686274509804	0	1.11467236467236	0	1.1294293972808	0	1.03225116328823	0
Catarrhini	13.1493215909498	0	13.0641711229947	0	12.8828347578348	0	13.0428576526425	0	13.1589738104152	0
Hominidae	8.82289764794329	0	8.76114081996435	0	8.7695868945869	0	8.9533597402134	0	8.89090941505767	0
Homininae	26.4454229764078	0	26.5846702317291	0	26.3710826210826	0	26.5175034792849	0	26.5960671052397	0
Hominoidea	10.6200551319228	0	10.3618538324421	0	10.5448717948718	0	10.3432894408165	0	10.4205665793086	0
Homo sapiens	21.2562202412917	100	21.3618538324421	100	21.4084757834758	100	21.3128501587981	100	21.1798684280901	100
Primates	1.35681810045466	0	1.31016042780749	0	1.28383190883191	0	1.36316597080969	0	1.18913907757038	0
root	2.03343715318799	0	2.13547237076649	0	2.13319088319088	0	2.10184491310709	0	2.11976966001676	0
Simiiformes	9.19163713170802	0	9.03386809269162	0	9.29665242165242	0	9.14249009741998	0	9.25282130823127	0

10.3. Extraction buffer test and taxonomic profile comparison

Given that previous work suggested that a phosphate-based buffer performs better in the extraction of DNA from cave sediments than the Bulat buffer[60], we tested the efficiency of the two in the extraction batch 1# of the samples (see details in method section) (see Table S19). Post extraction using a Qubit 3.0 ® to estimate the DNA concentration in the extracts, we found the two buffers to yield significantly different quantities of DNA, by a factor ~10 (Fig. S50).

Table S19. Metadata for all samples including controls.

Layer_extract_library ID	Extraction batch	Sequencing run	Raw reads after trimming, Q30 and lenght >29 Bp	Number of reads after low complexity removal and duplicate removal
all_merged_blanks	1,2,3	1,2,3	19688458	778132
UE1201_Mex_70_Lib_18	3	3	279911398	7333921
UE1204_Mex_11_Lib4_seq2	1	2	36167727	25617270
UE1204_Mex_12_Lib4_seq2	1	2	57862405	36279265
UE1204_Mex_13_Lib1_seq1	1	1	34855986	18051278
UE1204_Mex_13_Lib1_seq2	1	2	17035590	13341076
UE1204_Mex_14_Lib1_seq1	1	1	51643646	37750321
UE1204_Mex_29_Lib4_seq2	2	2	66019892	45639421
UE1204A_Mex_67_Lib_15	3	3	251238816	4085097
UE1204C_Mex_66_Lib_14	3	3	243140108	10366352
UE1206_Mex_65_Lib_13	3	3	123154614	7829486
UE1207A_Mex_62_Lib_10	3	3	234661165	70596213
UE1207C_Mex_61_Lib_9	3	3	166962110	78821562
UE1208_Mex_60_Lib_8	3	3	257580994	84812252
UE1210_Mex_1_Lib4_seq2	1	2	34766119	21618940
UE1210_Mex_18_Lib4_seq2	2	2	114865759	66482464
UE1210_Mex_2_Lib4_seq2	1	2	33848329	23928848
UE1210_Mex_3_Lib1_seq1	1	1	35394821	20647086
UE1210_Mex_3_Lib1_seq2	1	2	25147994	16925928
UE1210_Mex_4_Lib_45	3	3	155806416	74601786
UE1210_Mex_4_Lib1_seq1	1	1	48287199	32816580
UE1210_Mex_59_Lib_7	3	3	56688240	36650737
UE1212_Mex_22_Lib4_seq2	2	2	44194035	32899980
UE1212_Mex_24_Lib4_seq2	2	2	134803562	85036702
UE1212_Mex_5_Lib4_seq2	1	2	69154901	46341233
UE1212_Mex_58_Lib_6	3	3	76604702	50145443
UE1212_Mex_6_Lib1_seq1	1	1	43693190	25464374
UE1212_Mex_6_Lib1_seq2	1	2	30770066	22091288
UE1212_Mex_7_Lib1_seq1	1	1	91491535	66906070
UE1215_Mex_57_Lib_5	3	3	229937042	108160803
UE1217_Mex_56_Lib_4	3	3	145708355	72629586
UE1218_Mex_10_2_Lib4_seq2	1	2	26526636	21625731
UE1218_Mex_10_Lib4_seq2	1	2	122284460	27916781
UE1218_Mex_55_Lib_3	3	3	75755024	40784446
UE1218_Mex_9_Lib1_seq1	1	1	96884231	59710215
UE1218_Mex_9_Lib1_seq2	1	2	122284460	85124874
UE1222_Mex_54_Lib_2	3	3	38	31
UE1223_Mex_53_Lib_1	3	3	81656058	20834332

The samples were then multiplexed and sequenced in parallel to examine the origin of this difference. At this point, we found that the initial difference in concentration was also reflected by the lengths of the sequenced reads between samples (Fig. S53). The Bulat buffer, yielding the lowest quantity of DNA, was found to yield shorter sequence reads than the sodium-phosphate based buffer. Therefore, we next compared the taxonomic profiles of each of the samples using principal component analysis (PCA). Within amniota, we find that the first principal component (PC1) was associated with differences between stratum 1204 explaining 19.9% of the variation. While PC2 is characterized by the differences between stratum 1201 and the rest of the layers, explaining 18% of this variation (Fig. S53a). A similar trend is observed for the plants (Fig. S53b) in which PC1 is associated with the differences between stratum 1204 and the remaining layers explaining 20.6% of this variation. On the PC2, stratum 1204C is falling distant to the other samples with a slight lower explanation value of 13.4%. It is important to note that stratum 1201 is absent in the plant PCA model due to too few reads classified as plants.

We next plotted the raw taxonomic profiles prior to removal of taxa without DNA damage (see Fig. S54 and S55) to investigate the complete diversity of all samples, but excluding taxa found in the controls. Firstly, we observe that variation occur in the proportion of the identified taxa between each library from the same sample and between samples from the same strata. However, the taxa identified are highly similar. This variation is likely explained by the nature of the subsampling and the depositional environment from which the samples come from, as DNA and tissue have not been completely homogenized during these processes. It can also be connected to the fact that DNA from one organism potentially aggregates together around the same sediment particles, resulting in an uneven distribution within a larger sample. The animal taxonomic profiles (Fig. S54 in stratum 1201 show distinctively difference from the rest of the samples, and are characterized by 50% horse (*Equus sp.)* DNA while two of five samples from stratum 1204 also have a high abundance (>50%) of vole (*Microtus sp.).* Deer mice (*Peromyscus sp.*) are also found with high abundance in two samples from strata 1204 and 1210, respectively, while bear (*Ursus sp.*) is found to have higher abundance in five different samples in which two were extracted with the sodium-phosphate buffer. Whether this is due to intra sample variation or is the result of the buffer type remains inconclusive, and further replication and comparison is needed in order to establish this. In the plant taxonomic profiles we find a similar however slightly different pattern. Stratum 1204 is clearly different from the other layers below in the stratigraphy, containing high proportions of the grass genus *Zea*. Interestingly the Zea reads did not show ancient DNA characteristics and are contained only in strata 1204 and not present in the layers immediately below (e.g. UE1204C and UE1207) which is another indication that leaching is not occurring within the strata.

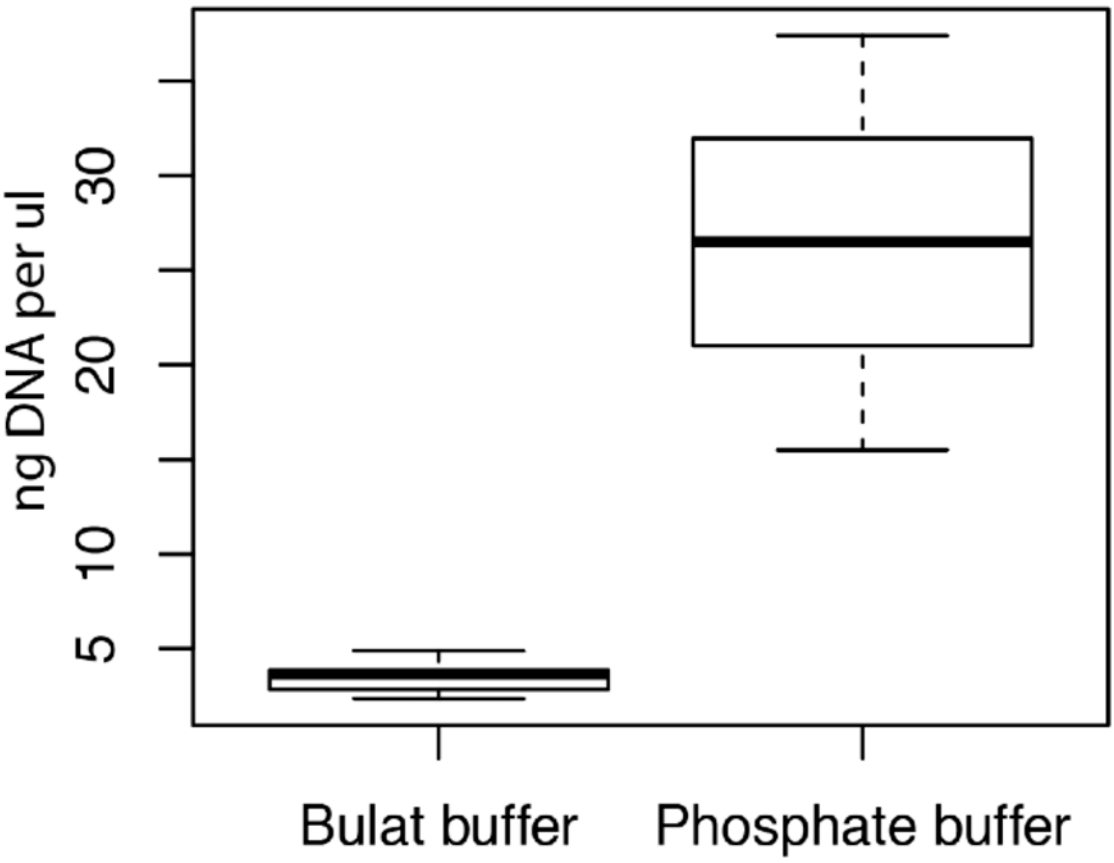


Fig. S51. DNA concentration, measured with an Qubit 2.0 (ng/ul), of the Bulat buffer (n = 11) and the Sodium phosphate buffer (n = 3) from extraction batch 1#. Whiskers represent largest and smallest observation less than or equal to upper and lower hinge + 1.5 * IQR. (see Table S19).

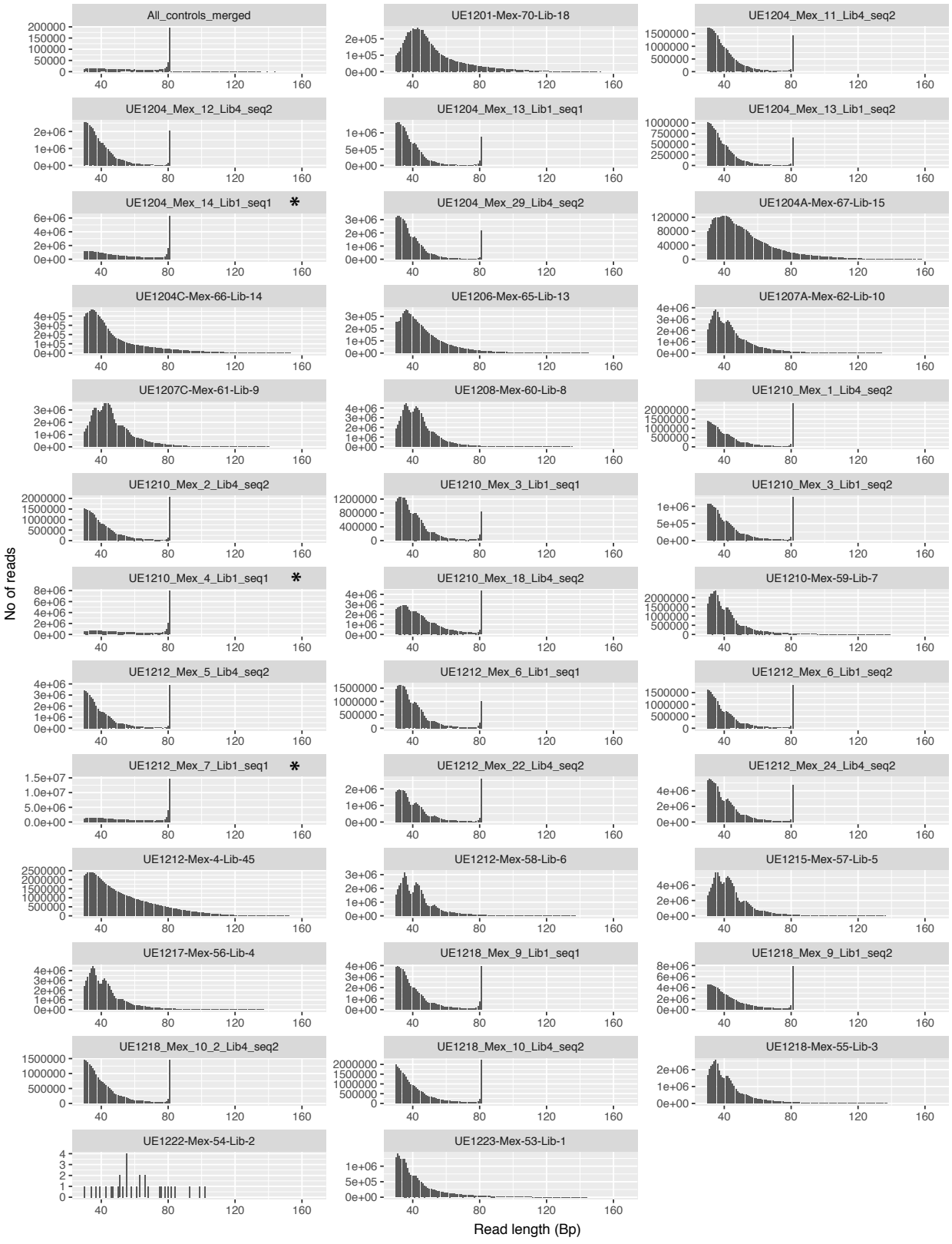


Fig. S52. Read length distributions of the final quality controlled and duplicate removed fastq files for all the samples and all merged controls. Samples extracted using the sodium-phosphate buffer are marked *. The remaining have been extracted using the Bulat buffer. Each header contain layer name, extraction ID and library ID, the corresponding metadata can be found in Table S15, and Supplementary Metadata file.

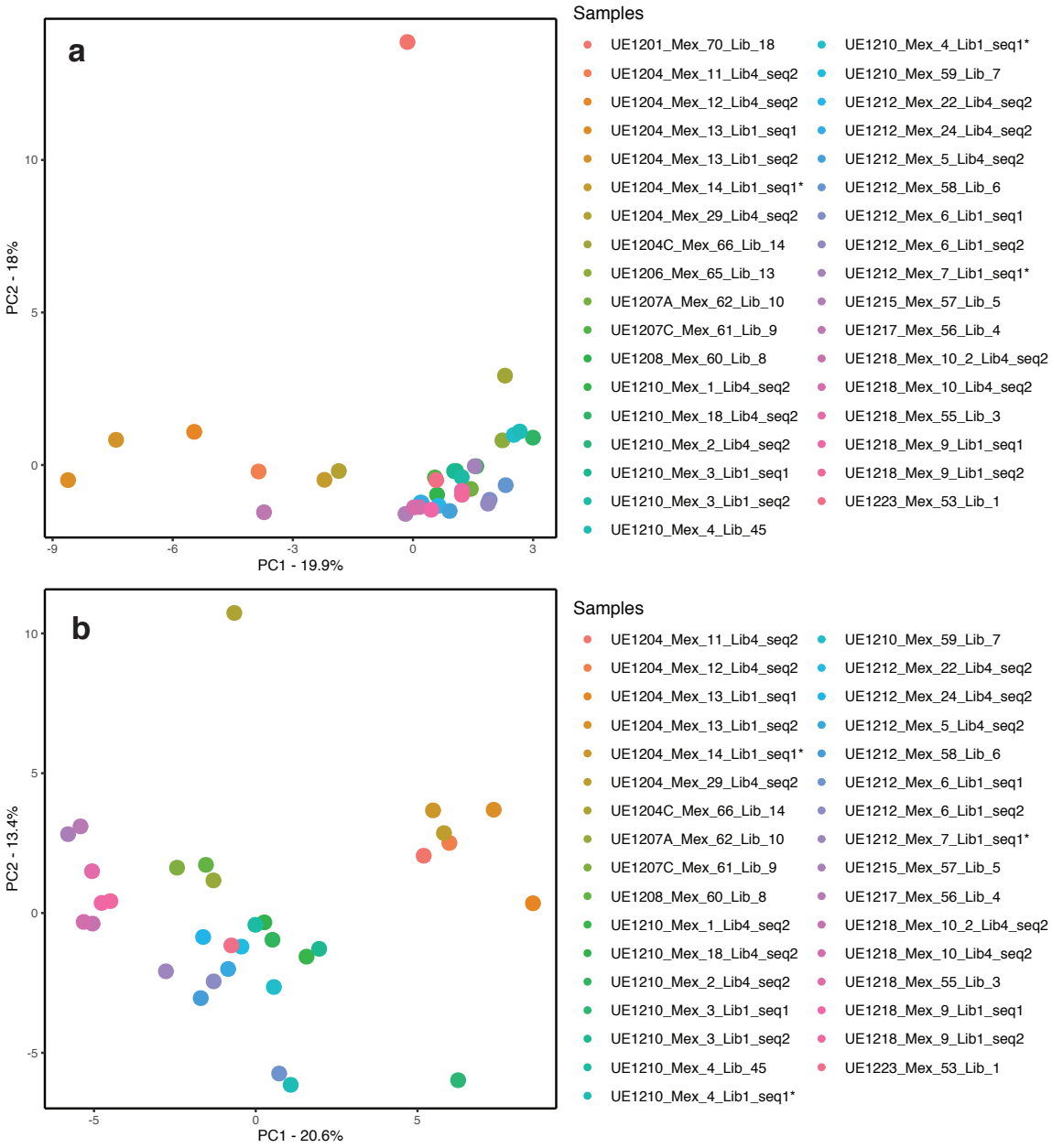


Fig. S53. Ordination analysis of the taxonomic composition in each individual layer, extract and library divided into kingdoms (see also Fig. S50). **a** Amniota (animals) and **b** Viridiplantae (plants).

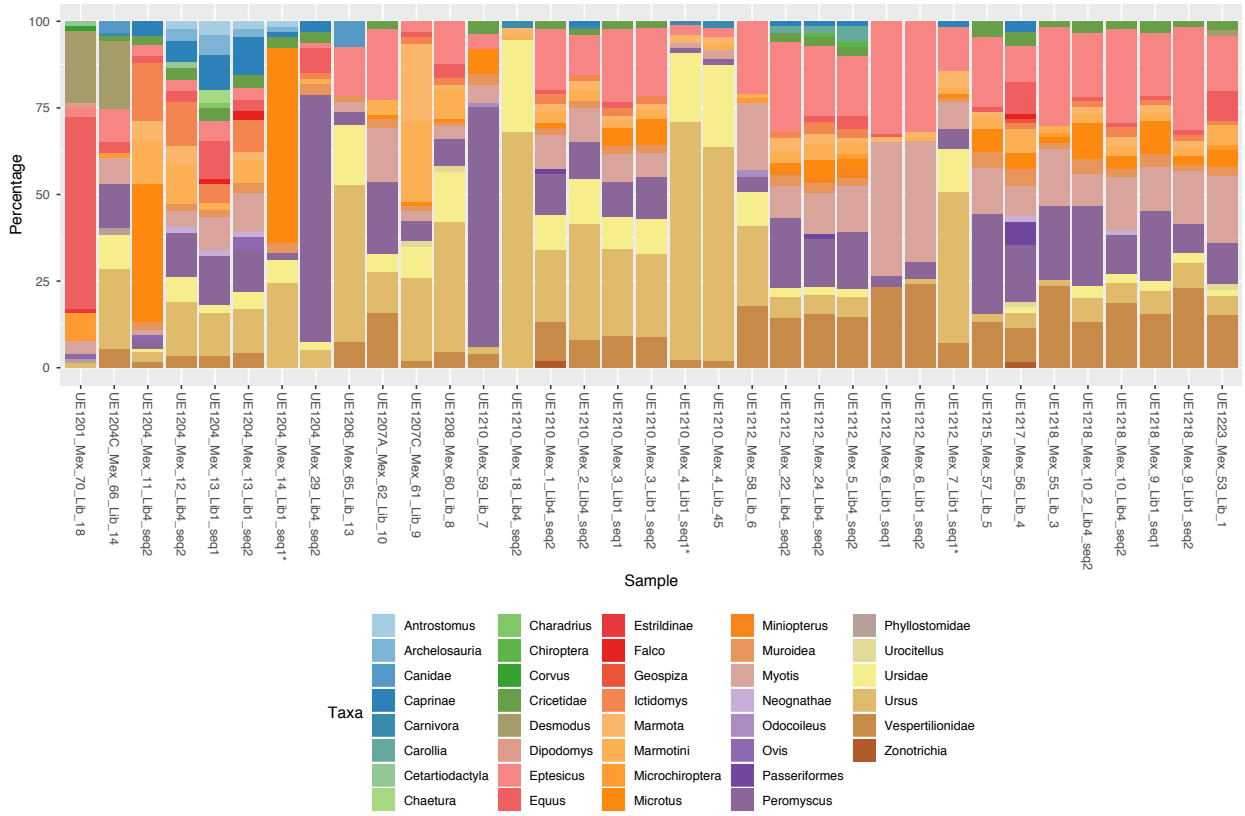


Fig. S54. Animal (Amniota) taxonomic profiles of each sub sample from the different eDNA samples. Sample names marked * are the sodium-phosphate based buffers.

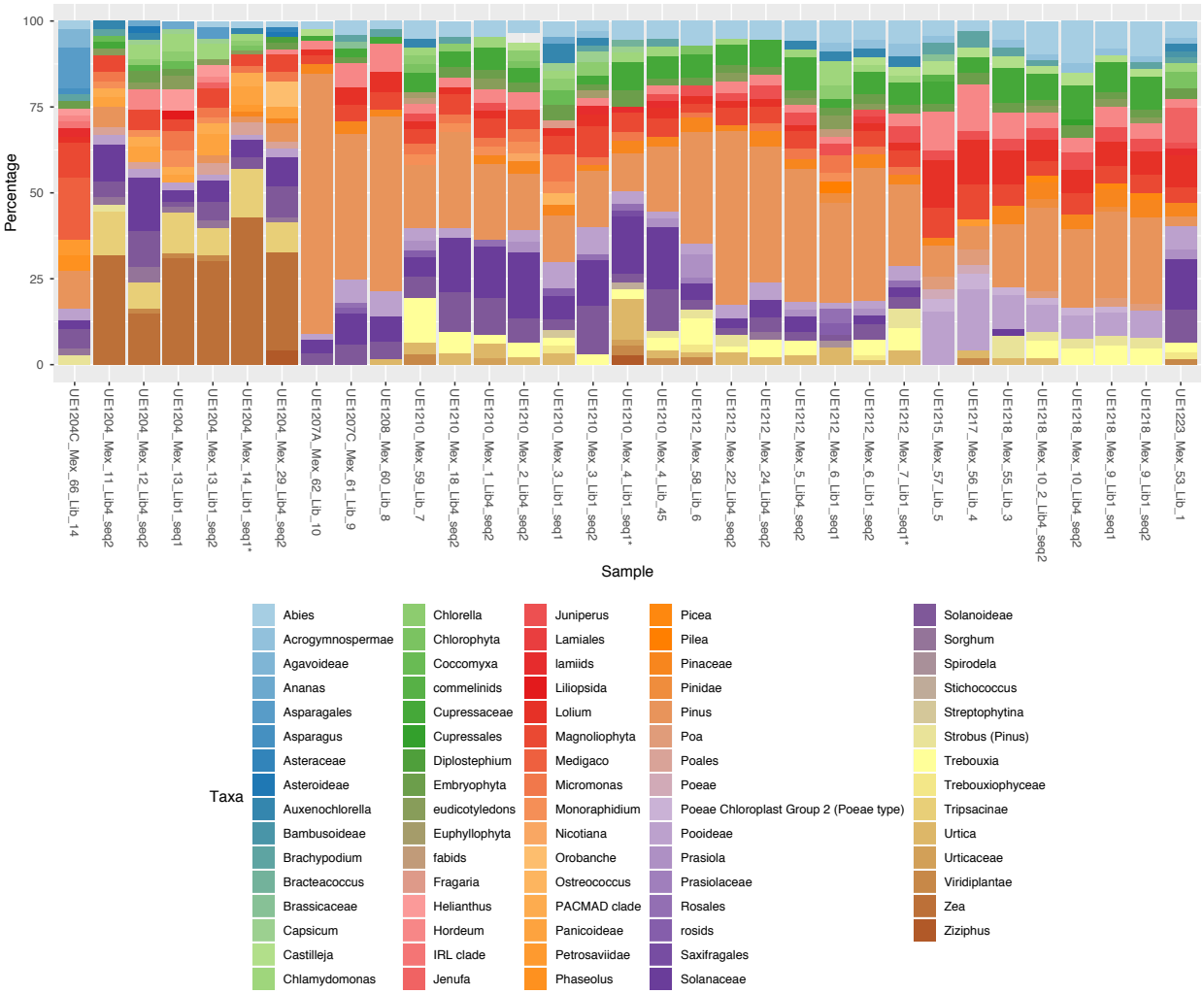


Fig. S55. Plant (Viridiplantae) taxonomic profiles of each sub sample from the different eDNA samples. Sample names marked * are the sodium-phosphate based buffers.

To further investigate the observed differences in DNA yield and read lengths between the two buffers, we counted the total number of reads assigned and reads assigned to the different kingdoms, and plotted these as a percentage of total reads sequenced and total reads assigned compared on the basis of the buffer used (Fig. S56 and S57).

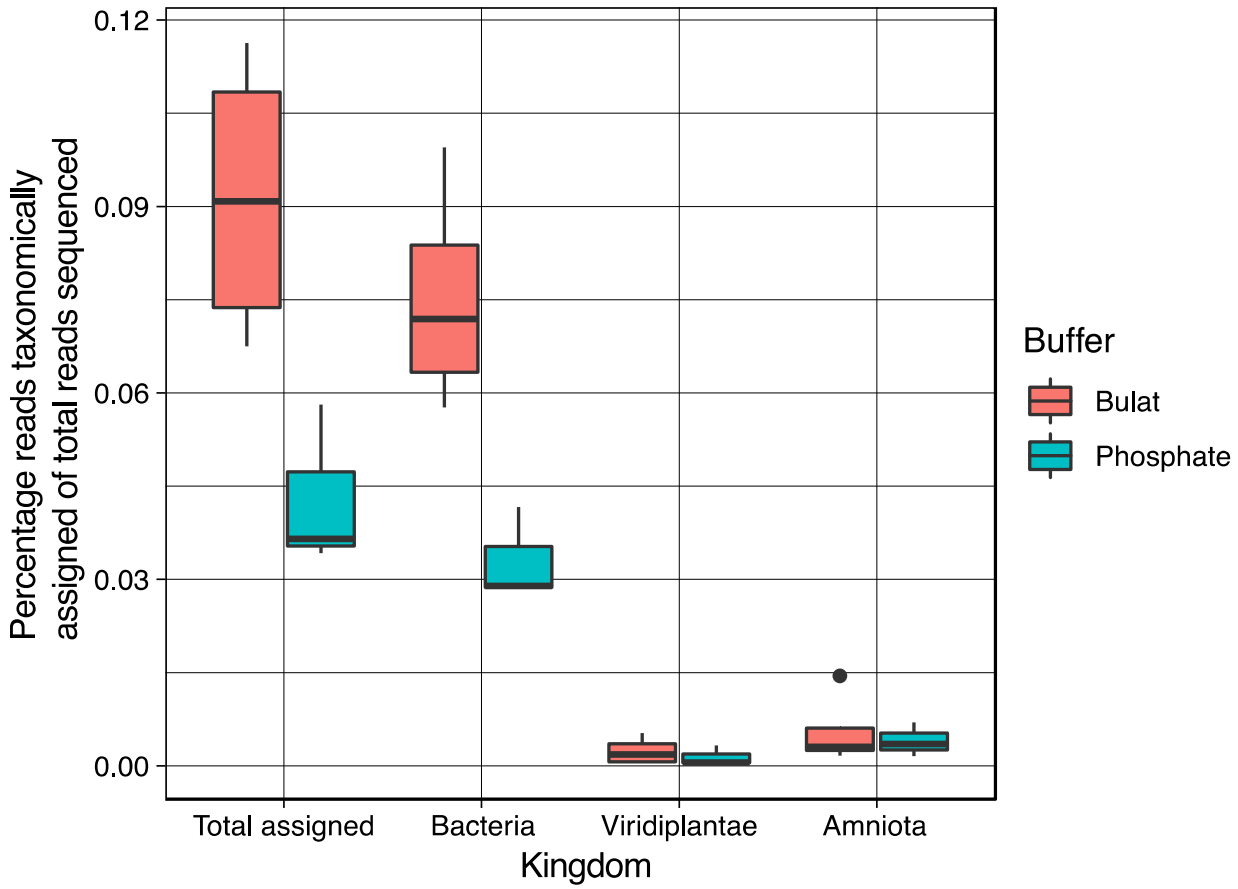


Fig. S56. Boxplot of the total number of assigned reads for the Bulat buffer (n = 11) and the Sodium phosphate buffer (n = 3), split into kingdoms as a proportion of the total reads sequenced for each sample from the extraction buffer comparison (sequencing batch #1). Whiskers represent largest and smallest observation less than or equal to upper and lower hinge + 1.5 * IQR.

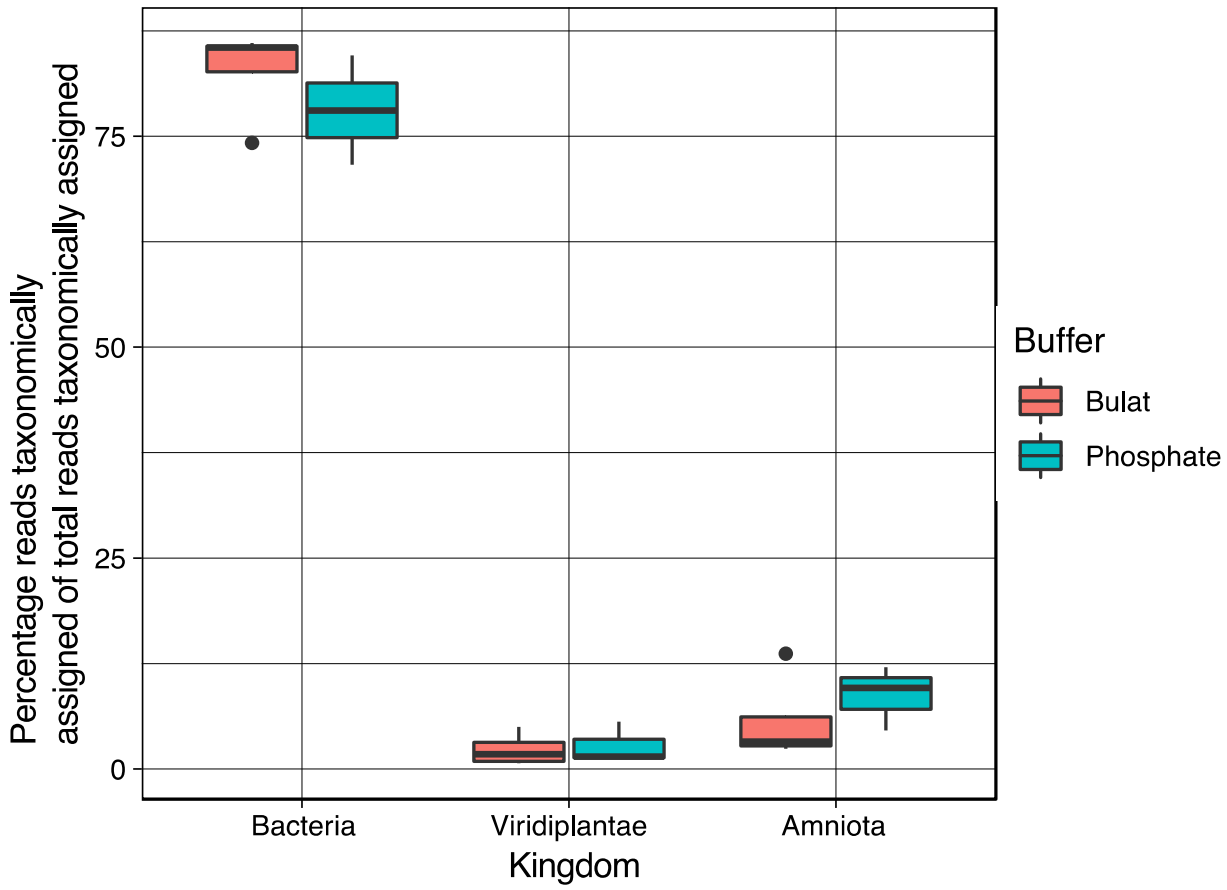


Fig. S57. Boxplot of the assigned reads totally and split to kingdoms as a proportion of the total reads taxonomically assigned for each sample from the Bulat buffer (n = 11) and the Sodium phosphate buffer (n = 3), (sequencing batch #1). Whiskers represent largest and smallest observation less than or equal to upper and lower hinge + 1.5 * IQR.

We find that the proportion of the reads assigned to each three kingdoms as a percentage to the total number of assigned reads, is highly similar between the buffers (Fig. S56). However, we find the proportion of reads, of the totally sequenced, to show differences in the proportion of reads assigned to bacteria (Fig. S57). There is a significant lower proportion of reads assigned to bacteria and the totally assigned reads for the sodium-phosphate buffer (Fig. S56 and S57). Given that the sodium-phosphate buffer also yields longer reads, a plausible explanation could be that this are unknown bacterial sequences presumably of modern origin. Importantly, the sodium-phosphate extraction control yielded no detectable DNA concentration and it is therefore likely that genetically unknown bacterial colonies dormant/living in the sediments grow during the sodium-phosphate incubation and are hereafter co-extracted. An alternative explanation could be that the sodium-phosphate buffer is more prone to release longer DNA fragments from the sediment substrate and that these derive from genetically unknown bacterial species. While this is of high importance, the true source of the long fragments is outside the scope of this study and further investigation is needed to determine this. As a last attempt to understand the observed differences between the two buffers, we split the barplots (Fig. S56 and S57) by samples (see Fig. S58 and S59). We find a clear correlation between total assigned and the proportion assigned to bacteria (Fig. S57), while we find no direct pattern for Viridiplantae nor metazoans.

In conclusion, due to the fact that the sodium-phosphate buffer yields longer reads (potentially not of ancient origin) and that the proportion of total assigned reads are less than the Bulat buffer, we decided to process and sequence the remaining samples (sequencing batch #2 and #3) using the Bulat buffer (see Table S17).

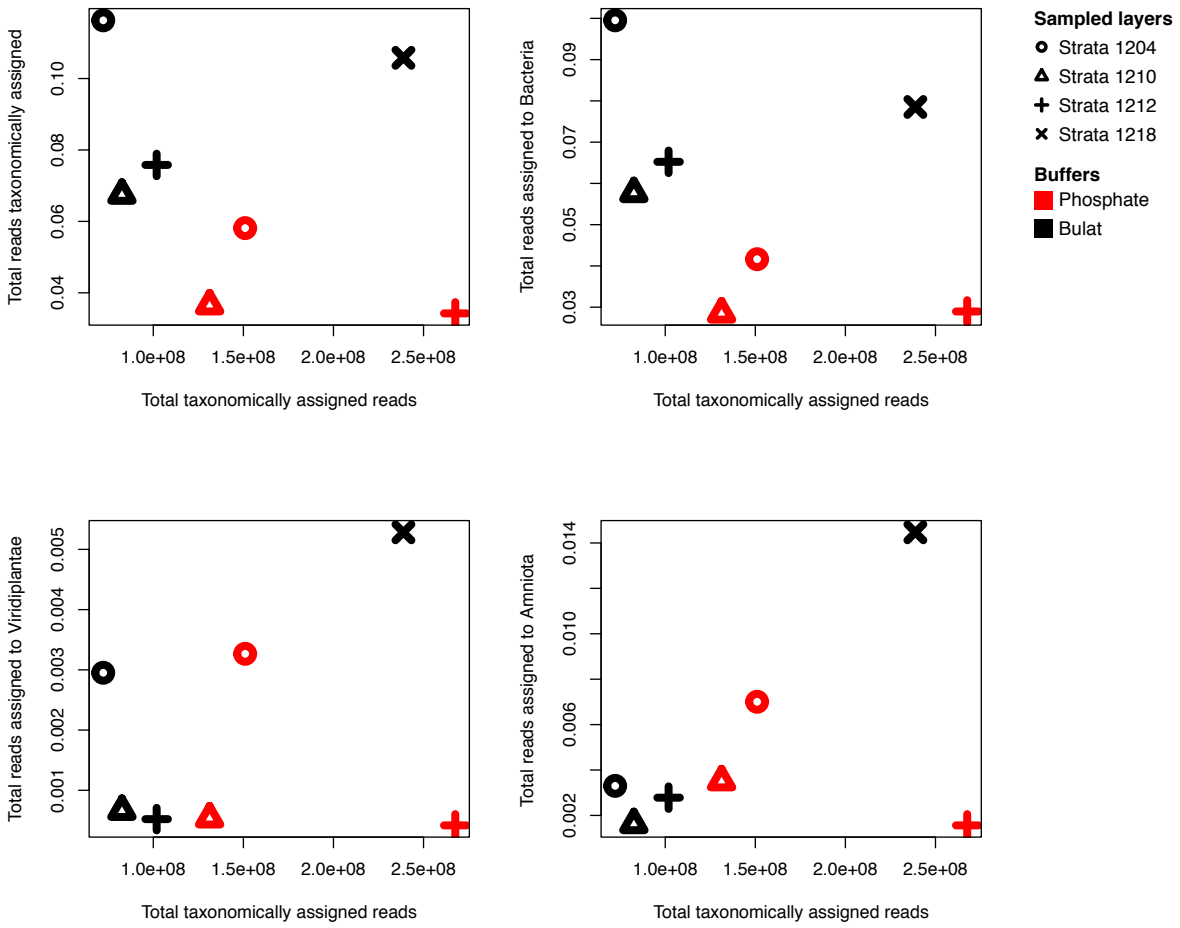


Fig. S58. Assigned reads relative to total sequenced reads on the three kingdoms split by buffer and sample from the extraction buffer comparison (sequencing batch #1).

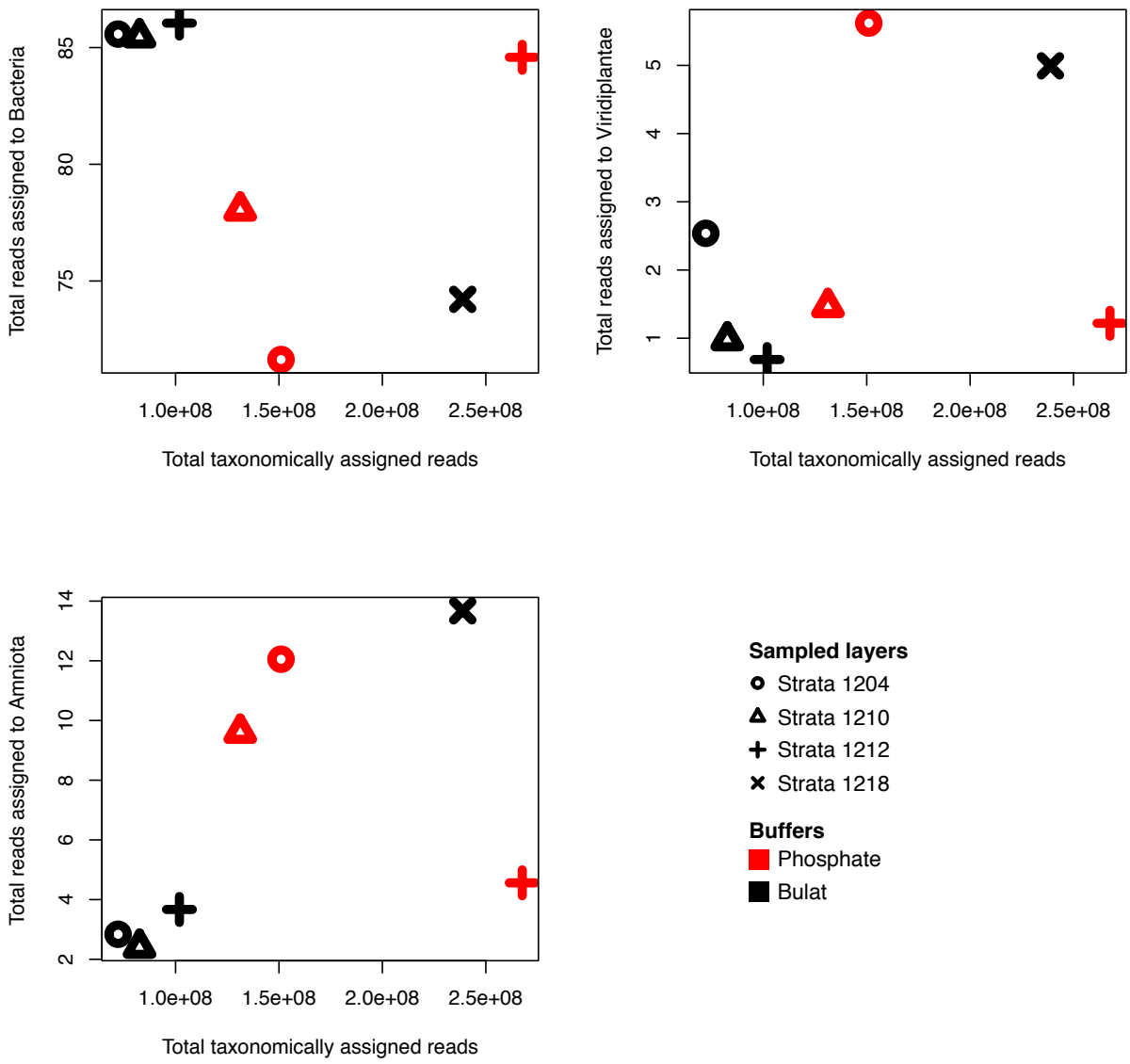


Fig. S59. Assigned reads relative to total assigned reads on the three kingdoms split by buffer and sample.

10.4. DNA damage

10.4.1. Ancient DNA authenticity

DNA deamination levels of each taxa was calculated and a cut-off threshold of ≥0.10 (10%) was used as confirmation of ancient authenticity. We employed a conservative approach to calculate DNA damage for each genus identified through the 'Holi' pipeline by parsing reads exclusively assigned within each genus and here after used the species level identifications found by 'Holi' to identify reference genomes to align reads within each genus (Supplementary Metadata file). When possible, we used whole genomes. Otherwise, full chloroplast or mitochondrial genomes were employed. Reads from the quality controlled fastq files were realigned against the species reference genomes and MapDamage2.0[59] was used to calculate the deamination patterns. In addition, we calculated the number of reads and their edit distances together with the length distribution for each taxa found (Supplementary Information files SI2 and SI3). We find that the

organisms in all layers exhibit a high degree of DNA damage and fragmentation, but that the DNA damage varies within each layer as well as variation between all layers (Fig. S60 and Fig S61) as found in previous studies[56]. We also find that DNA damage is not directly correlating with age, which implies that other factors are influencing the degree of deamination occurring on the deposited DNA. The deamination is a hydrolysation process and is therefore dependent on the water (or hydrogen) availability, absence of this could reduce the amount of DNA damage. Other factors such as type of tissue, depositional micro-environment and other taphonomic dynamics are also key to this process. It is therefore expected to have variation within ancient metagenomic samples. While the read length distribution for each of the taxa assessed (Supplementary Information files SI2 and SI3) show signs of high fragmentation which is another characteristic of ancient DNA. The edit distances to each reference genome showed more variation. Especially within the algae the DNA found was less similar to the references than for the other plants. Within the amniota, we observe similar patterns, where *Desmodus* in strata UE1204C and UE1223 is less similar to the reference genome compared to the upper stratum UE1201. Several factors may influence the observed results, firstly the reference genome available rarely holds the genetic variation within population, secondly current populations which is often used to generate references might have accumulated mutations over the course of time. Thirdly, too few reads does not seem to represent the whole variation. Lastly, it is likely that the DNA derives from an extinct to extant closely related species that have not been sequenced, and which contains conserved genes shared with other closely related species. A combination of read length distribution, edit distances and deamination patterns are therefore key to understand the authenticity of the taxa identified and if it is ancient of origin. We evaluated all taxa based on these parameters (Supplementary Information files SI2 and SI3 & Metadata file). One plant taxon, *Urtica*, did not have a full chloroplast genome and only short genes have been published and were therefore removed from the dataset.

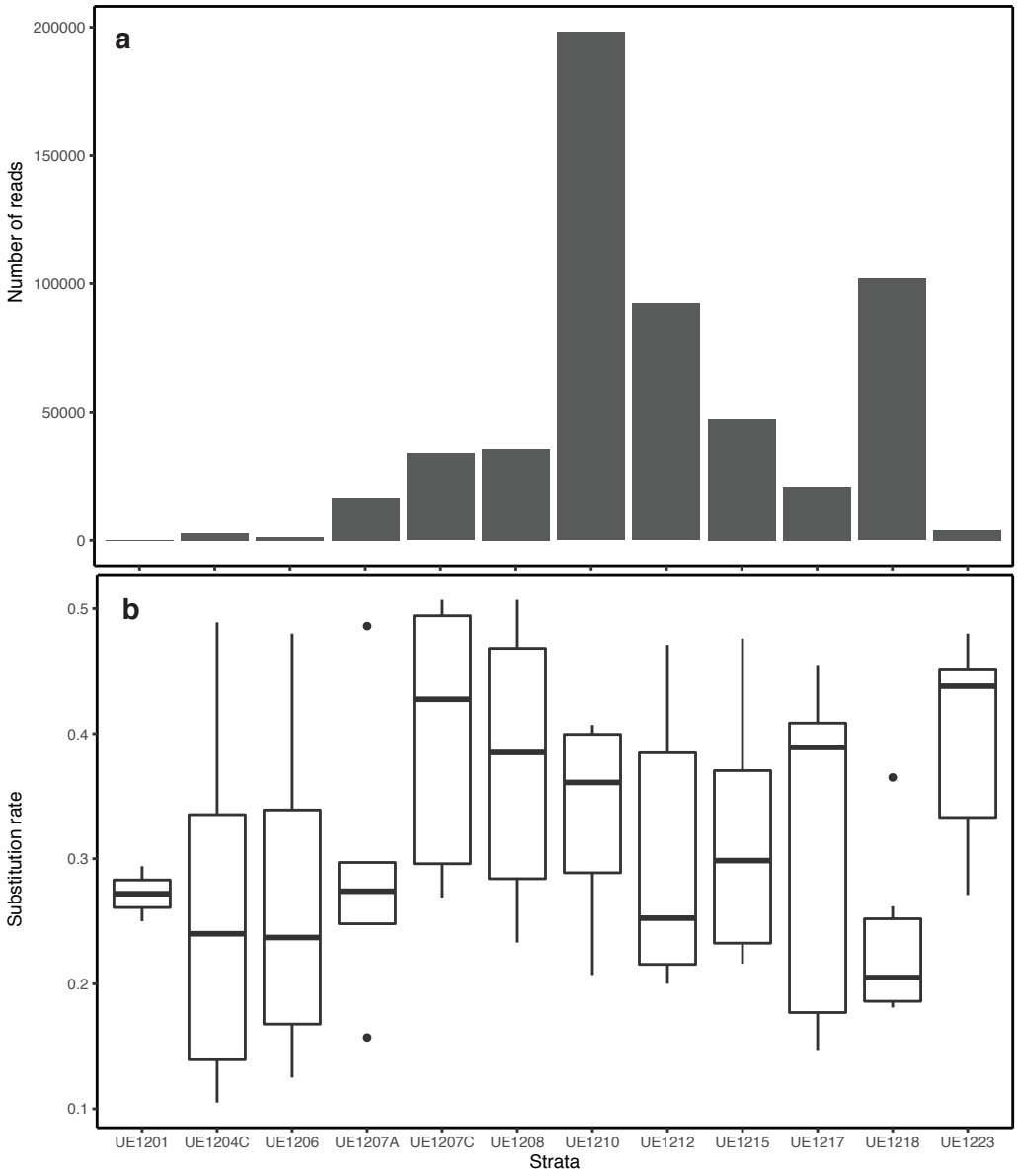


Fig S60. Positions' specific substitutions due to DNA damage from the 3' end, e.g., deamination (C-T) due to damage levels for Amniota (animals) see detailed table and individual plots in Supplementary Metadata file, Supplementary file SI3. The bar plot (**a**) shows the cumulative number of reads across each taxa that have been used to calculate the DNA damage for the different strata UE1201 (n = 2), UE1204C (n = 8), UE1206 (n = 4), UE1207A (n = 5), UE1207C (n = 10), UE1208 (n = 10), UE1210 (n = 6), UE1212 (n = 6), UE1215 (n = 8), UE1217 (n = 11), UE1218 (n = 9), UE1223 (n = 9).The box plot (**b**) displays the corresponding C-T transitions (substitution rate) at the 3' excluding taxa below the threshold criteria of ≥ 10% substitution rate. Whiskers represent largest and smallest observation less than or equal to upper and lower hinge + 1.5 * IQR.

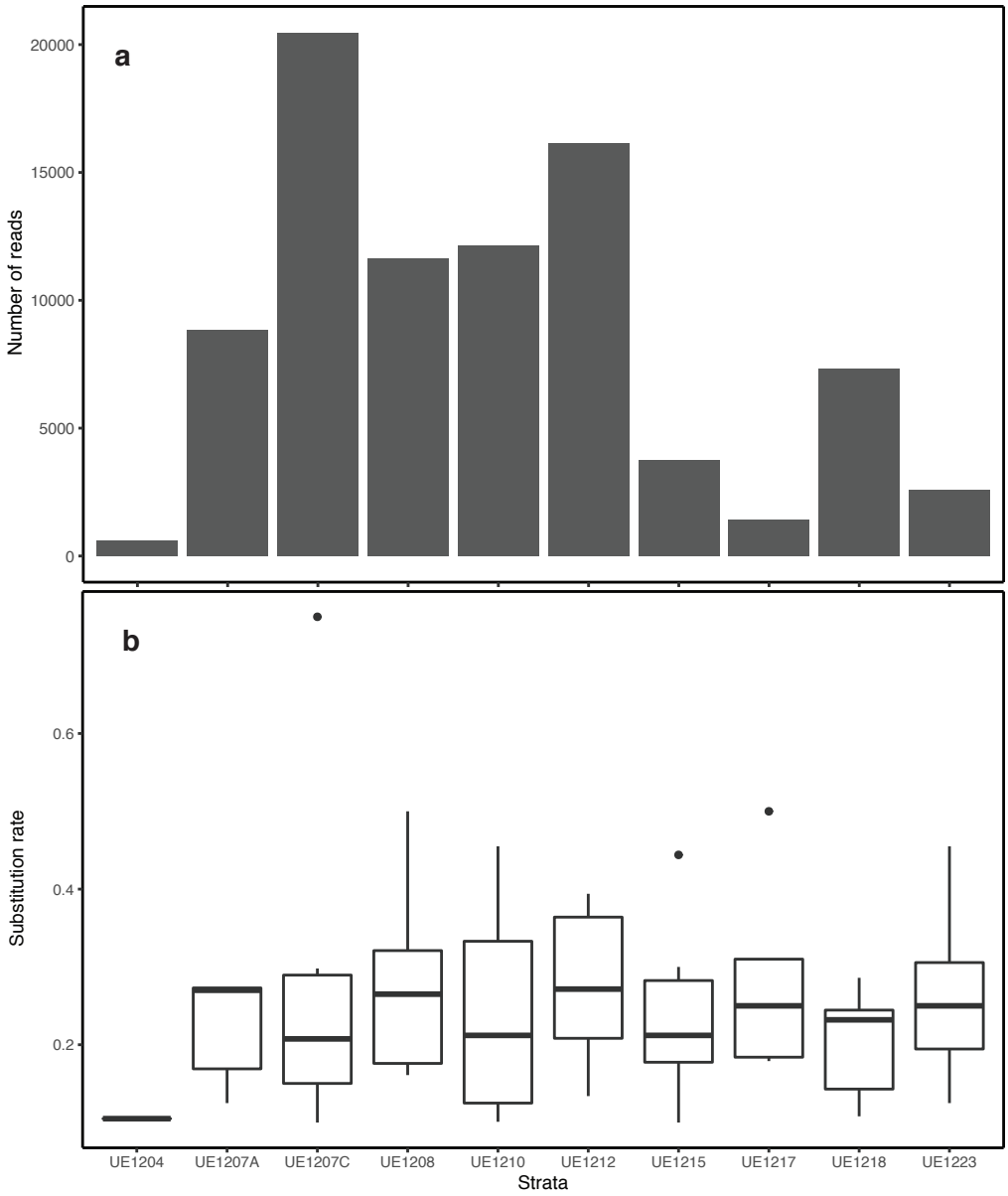


Fig S61. Positions' specific substitutions due to DNA damage from the 3' end, e.g., deamination (C-T) due to damage levels for Viridiplantae (plants); see detailed table and individual plots in Supplementary Metadata file, Supplementary file SI2. The bar plot (**a**) shows the cumulative number of reads across each taxa that have been used to calculate the DNA damage for the different strata UE1204C (n = 1), UE1207A (n = 5), UE1207C (n = 6), UE1208 (n = 5), UE1210 (n = 5), UE1212 (n = 8), UE1215 (n = 6), UE1217 (n = 5), UE1218 (n = 7), UE1223 (n = 7). The box plot (**b**) displays the C-T transitions (substitution rate) at the 3' parsing excluding taxa below the threshold criteria of ≥ 10% substitution rate. Whiskers represent largest and smallest observation less than or equal to upper and lower hinge + 1.5 * IQR.

We investigated the presence of ancient human DNA out by mapping sequencing reads of each sample against two different reference indices (see Methods). We first determined the presence of mitochondrial (MT) sequences by mapping against a reference index containing all mitochondrial genomes contained in the RefSeq database (release 92). Reads mapping uniquely and with high quality (MQ25) to a single MT reference contig were extracted and assessed for genomic coverage and ancient DNA damage. We find that only the sample from UE1210 contains sufficient human MT reads for analysis, with a total of 189 reads mapping at MQ25 and covering ~58% of the MT genome (contig NC_012920.1, Supplementary Metadata file). However, rates of characteristic ancient DNA damage substitutions (5' C>T or 3' G>A) were

indistinguishable from other substitution types, indicating that the reads originated likely from contaminating modern human DNA. This contrasts with reads mapped to the American black bear MT (contig NC_003426.1) from the same sample, which showed similar genomic coverage but elevated rates of 5' C>T and 3' G>A substitutions, consistent with authentic ancient DNA (Supplementary Metadata file, Fig. S62).

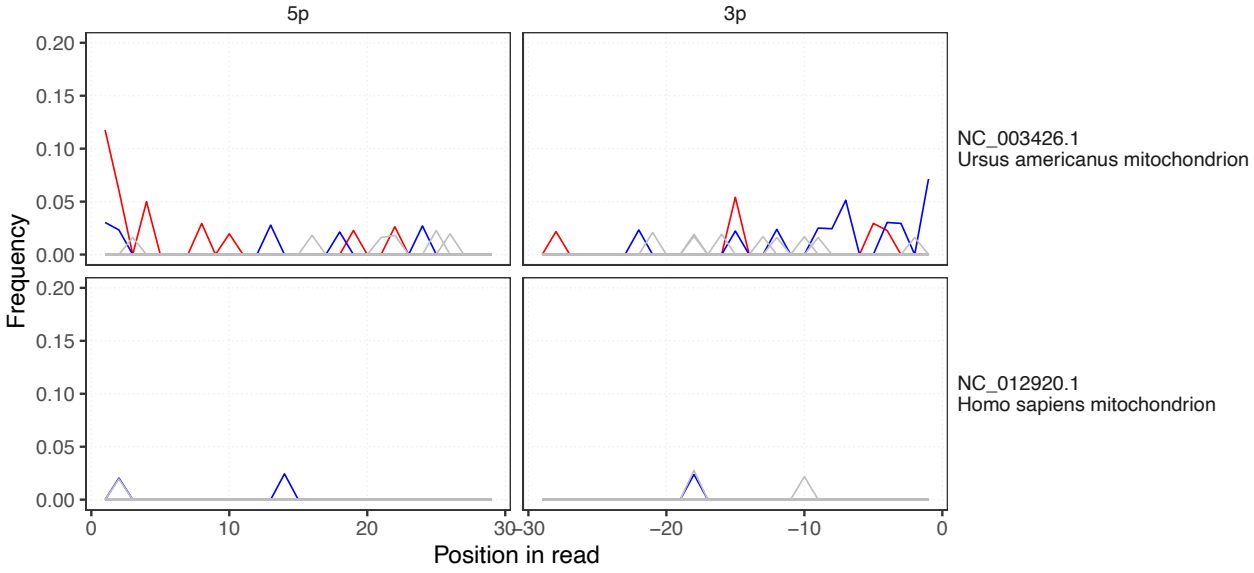


Fig. S62. Human DNA damage plots for mitochondrial reads from sample UE1210, mapping uniquely to either American black bear (top) or human (bottom). Panels show position specific substitutions from the 5' (left) and the 3' (right) end of the reads. Colors represent different substation types: C >T (red), G > A (blue), all other substitutions (grey).

When using the full human genome as a reference index, we find reads mapping from all samples, with coverage ranging from 595 up to 32,727 reads at MQ25 (Supplementary Metadata file). For the majority of samples rates of ancient DNA damage substitutions were ≤ 0.01, again suggesting their modern origin. However, three of the strata (UE1210, UE1212, UE1215) exhibit elevated rates, ranging from 0.03 up to 0.07 (Supplementary Metadata file). As remnant human background contamination present in the reagents used in the laboratory preparations will dilute or decrease the DNA damage signal of ancient human reads if present in only low quantities, we considered those samples as putative candidates for follow up.

A further complication for the ancient DNA authentication stems from the possibility of spurious mapping of reads that originate from DNA sequences conserved between humans and closely related species. If ancient DNA sequences of a closely related species are present in sufficient numbers in the sample, their spurious alignment to the human genome can create a false-positive signal of ancient DNA damage. To investigate whether this was the case for our samples, we re-calculated the substitution rates restricting to reads assigned to Old world monkeys (Hominidae) using the 'Holi' pipeline, parsing reads with mismatches ≤ 5 for all samples and controls. Substitution rates of 5' C>T and 3' G>A changes were indistinguishable from the other types after this filtering step (Fig S63-S65), suggesting spurious mapping as the main culprit for the elevated rates observed before.

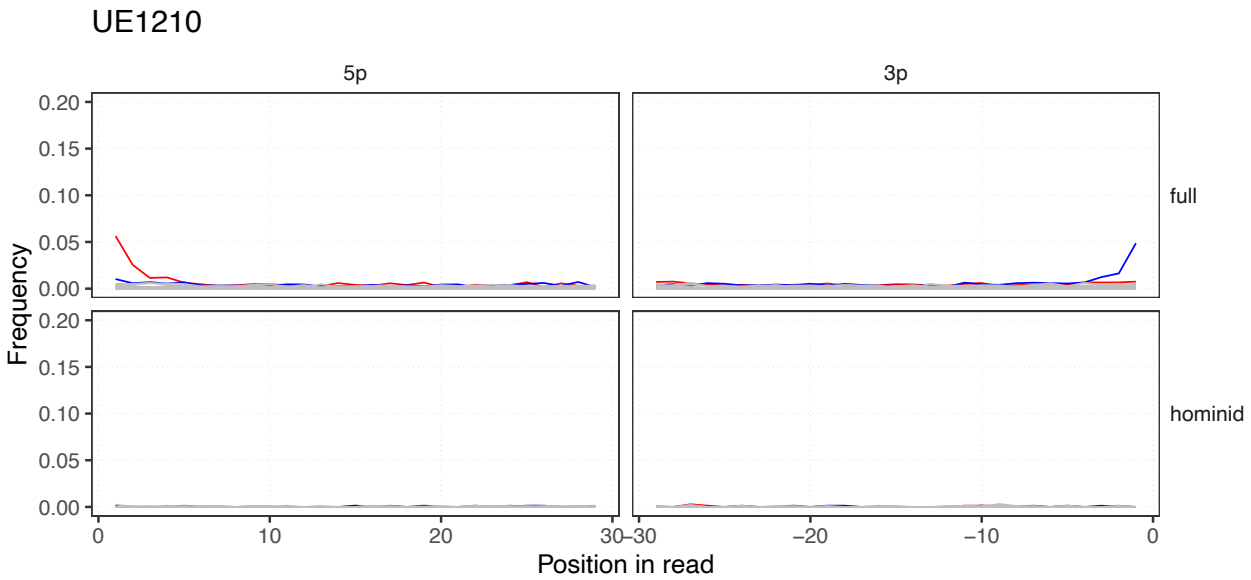


Fig. S63. DNA damage rates for sample UE1210, inferred using either all MQ25 reads (top) or MQ25 reads assigned to Old World monkeys using Holi (bottom). Panels show position specific substitutions from the 5' (left) and the 3' (right) end of the reads. Colors represent different substation types: C >T (red), G > A (blue), all other substitutions (grey).

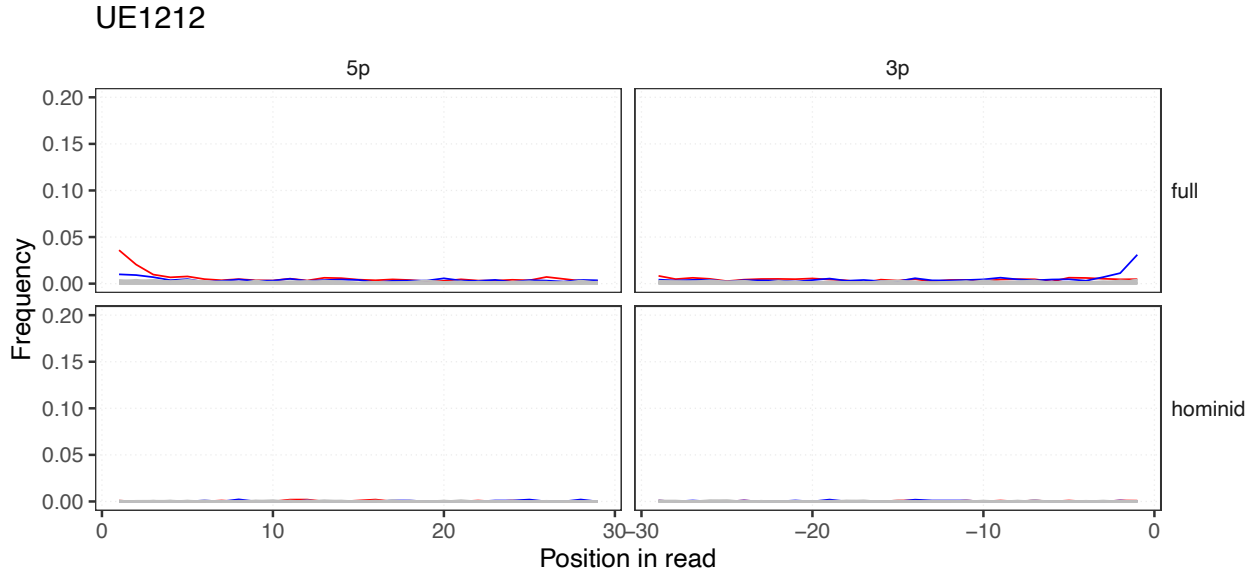


Fig. S64. DNA damage rates for sample UE1212, inferred using either all MQ25 reads (top) or MQ25 reads assigned to Old World monkeys using Holi (bottom). Panels show position specific substitutions from the 5' (left) and the 3' (right) end of the reads. Colors represent different substation types: C >T (red), G > A (blue), all other substitutions (grey).

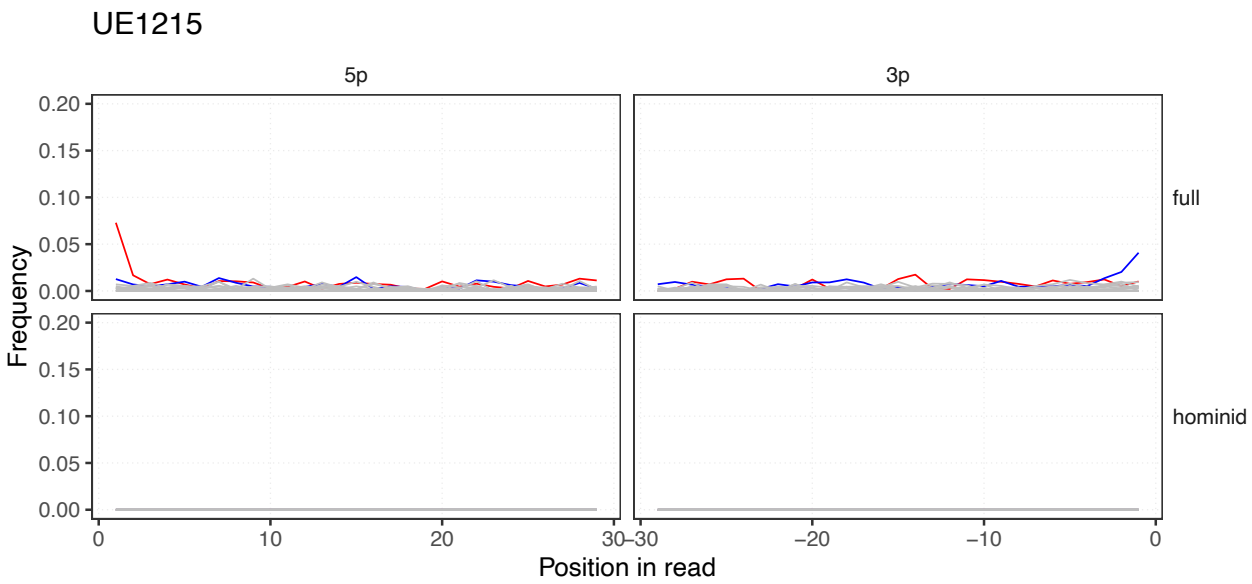


Fig. S65. DNA damage rates for sample UE1215, inferred using either all MQ25 reads (top) or MQ25 reads assigned to Old World monkeys using Holi (bottom). Panels show position specific substitutions from the 5' (left) and the 3' (right) end of the reads. Colors represent different substation types: C >T (red), G > A (blue), all other substitutions (grey).

10.5. Modelling scripts

10.5.1. Bash scripts

10.5.1.1. cpDNA and mtDNA models

```
# list of dependencies
Bowtie2
Samtools
Gargammel
MapDamage2.0
ngsLCA

########## extracting fasta and generating db's

###Homo S.
# Extracts one reference fasta (NC_012920.1 Homo sapiens mitochondrion, complete genome)
PATH/software/bbmap/filterbyname.sh in=refseq91_mtDNA.fa out=Homo_sapiens_mtDNA.fa names=NC_012920.1 include=t ow=t

# removes Homo S. (NC_012920.1 Homo sapiens mitochondrion, complete genome)
PATH/software/bbmap/filterbyname.sh in=refseq91_mtDNA.fa out=refseq91_missin_1_Homo_S_mtDNA.fa names=NC_012920.1 include=f ow=t

# removes all Homo species
PATH/software/bbmap/filterbyname.sh in=refseq91_mtDNA.fa out=refseq91_no_Homo_Sp_mtDNA.fa names=NC_013993.1,NC_023100.1,NC_012920.1,NC_011137.1 include=f ow=t

#setting folder structure
mkdir hum_mtDNA_model
mkdir hum_mtDNA_model/endo
mkdir hum_mtDNA_model/cont
mkdir hum_mtDNA_model/bact
cp PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA.fa hum_mtDNA_model/endo
samtools faidx hum_mtDNA_model/endo/Homo_sapiens_mtDNA.fa

# simulating and extacting read libraries and adding damage and seq errors
counter=1
while [ $counter -le 5 ]
do
echo $counter
bname='HomoSap'
basefolder='hum_mtDNA_model'
fasta='Homo_sapiens_mtDNA.fa'
```

```
ReadNo='11835'
echo $bname
### 30X fragment retrieval   # number of reads needed to generate depends on the length of the genome and the average read length extracted if 30X covereage is needed divide the total bases in the genome with averageread length and times 30. In this case ReadNo = 16569Bp/42*30 = 11835
PATH/software/gargammel/src/fragSim -tag e -n $ReadNo -m 0 -M 1000 -f PATH/Mex_Cave/model_simulations/gargammel/average_readDist.txt $basefolder/endo/$fasta > $basefolder/$bname.e.$counter.fa
#### adding DNA deamination
PATH/software/gargammel/src/deamSim -mapdamage PATH/Mex_Cave/model_simulations/gargammel/misincorporation.txt single $basefolder/$bname.e.$counter.fa > $basefolder/$bname.d.$counter.fa
#adding adaptors
PATH/software/gargammel/src/adptSim -arts $basefolder/$bname.f.$counter.fasta $basefolder/$bname.d.$counter.fa
#adding sequencing errors
PATH/software/gargammel/art_src_MountRainier/art_illumina -ss HS25 -amp -na -qs -qs2 -i $basefolder/$bname.f.$counter.fasta -l 80 -c 1 -qs 0 -qs2 0 -o $basefolder/$bname.g.$counter
((counter++))
done
for file in $basefolder/$bname.g.?.fq
do
# removing adaptors
AdapterRemoval --file1 $file --mm 3 --minlength 30 --basename $file --trimns --trimqualities --minquality 30
for infile in $basefolder/*.truncated; do bname=$(basename $infile); echo $bname; bname2=$(echo $bname | sed 's/.fq.truncated*/.trunc.fq/'); echo $bname2 ; mv $basefolder/$bname $basefolder/$bname2 ; done
done

######################## fq mapping

for infile in $(pwd)/*trunc.fq
do
bname=$(basename $infile)
echo $bname
bname2=$(echo $bname | sed 's/.fq*/_holi/')
basepath=$(pwd)/
basefolder=$basepath
echo $basepath
echo $bname2
mkdir $basepath$bname2
cd $bname2


# mapping against fasta reference
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA
do
echo Mapping $bname against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against cp reference db
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/refseq91_mtDNA
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against no Homo Sapiens db
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/refseq91_missin_1_Homo_S_mtDNA
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against no Homo at all db
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/refseq91_no_Homo_Sp_mtDNA
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against nt_db
for DB in PATH/database/ncbi_nt/nt.?
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done
```

```
# mapping against RefSeq

for DB in PATH/database/refseq/vert_other/vert_other.?
do
echo Mapping $bname.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/vert_mam/vert_mam.?
do
echo Mapping $bname.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/invert/invert.?
do
echo Mapping $bname.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

cd $basepath
done

######################################### fasta mapping
for infile in $(pwd)/*fa
do
bname=$(basename $infile)
echo $bname
bname2=$(echo $bname | sed 's/.fa*/_holi/')
basepath=$(pwd)/
basefolder=$basepath
echo $basepath
echo $bname2
mkdir $basepath$bname2
cd $bname2

# mapping against fasta reference
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA
do
echo Mapping $bname against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against cp reference db
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/refseq91_mtDNA
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against no Homo Sapiens db
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/refseq91_missin_1_Homo_S_mtDNA
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against no Homo at all db
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/refseq91_no_Homo_Sp_mtDNA
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done


# mapping against nt_db
for DB in PATH/database/ncbi_nt/nt.?
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/vert_other/vert_other.?
do
echo Mapping $bname.fq against $DB
```

```
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/vert_mam/vert_mam.?
do
echo Mapping $bname.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/invert/invert.?
do
echo Mapping $bname.fq against $DB
nice -n 5 bowtie2 --threads 40 -k 5000 -x $DB -f ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

cd $basepath
done


######################################## ngs Fastq

for infile in *trunc.fq
do
bname=$(basename $infile)
echo $bname
bname2=$(echo $bname | sed 's/.fq*/_holi/')
bname3=$(echo $bname | sed 's/.fq*/.holi.metagenome.txt/')
bname4=$(echo $bname | sed 's/.fq*/.mtDB.metagenome.txt/')
bname5=$(echo $bname | sed 's/.fq*/.RefFA.metagenome.txt/')
bname6=$(echo $bname | sed 's/.fq*/.noGenus.metagenome.txt/')
bname7=$(echo $bname | sed 's/.fq*/.noFA.metagenome.txt/')
basepath='PATH/Mex_Cave/model_simulations/gargammel/data/hum_mtDNA_model'
basefolder=$basepath
echo $basepath
echo $bname2
echo $bname3
#mkdir $basepath$bname2
cd $bname2

samtools merge -@ 40 -f -n tmp.bam.merged *.bam
samtools sort -n -T PATH/TMP -O bam -o file.sort.bam -@ 40 tmp.bam.merged
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -
acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam file.sort.bam -outnames $bname.holi
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -
acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *refseq91_mtDNA.bam -outnames $bname.mtDB
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -
acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *Homo_sapiens_mtDNA.bam -outnames $bname.RefFA
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -
acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *refseq91_missin_1_Homo_S_mtDNA.bam -outnames $bname.noFA
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -
acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *refseq91_no_Homo_Sp_mtDNA.bam -outnames $bname.noGenus
cut -f9 -d":" $bname.holi.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.1.tax_counts
cut -f9 -d":" $bname.holi.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.1.taxNam
cut -f9 -d":" $bname.mtDB.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.2.tax_counts
cut -f9 -d":" $bname.mtDB.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.2.taxNam
cut -f9 -d":" $bname.RefFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.3.tax_counts
cut -f9 -d":" $bname.RefFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.3.taxNam
cut -f9 -d":" $bname.noGenus.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.4.tax_counts
cut -f9 -d":" $bname.noGenus.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.4.taxNam
cut -f9 -d":" $bname.noFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.5.tax_counts
cut -f9 -d":" $bname.noFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.5.taxNam
paste -d"," $bname.1.taxNam $bname.1.tax_counts > $bname3
paste -d"," $bname.2.taxNam $bname.2.tax_counts > $bname4
paste -d"," $bname.3.taxNam $bname.3.tax_counts > $bname5
paste -d"," $bname.4.taxNam $bname.4.tax_counts > $bname6
paste -d"," $bname.5.taxNam $bname.5.tax_counts > $bname7
cd $basepath
done


######################################## ngs Fasta

for infile in *.fa
do
bname=$(basename $infile)
echo $bname
bname2=$(echo $bname | sed 's/.fa*/_holi/')
bname3=$(echo $bname | sed 's/.fa*/.holi.metagenome.txt/')
bname4=$(echo $bname | sed 's/.fa*/.mtDB.metagenome.txt/')
```

```
bname5=$(echo $bname | sed 's/.fa*/.RefFA.metagenome.txt/')
bname6=$(echo $bname | sed 's/.fa*/.noGenus.metagenome.txt/')
bname7=$(echo $bname | sed 's/.fa*/.noFA.metagenome.txt/')
basepath='PATH/Mex_Cave/model_simulations/gargammel/data/hum_mtDNA_model'
basefolder=$basepath
echo $basepath
echo $bname2
echo $bname3
#mkdir $basepath$bname2
cd $bname2

samtools merge -@ 40 -f -n tmp.bam.merged *.bam
samtools sort -n -T PATH/TMP -O bam -o file.sort.bam -@ 40 tmp.bam.merged
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam file.sort.bam -outnames $bname.holi
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *refseq91_mtDNA.bam -outnames $bname.mtDB
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *Homo_sapiens_mtDNA.bam -outnames $bname.RefFA
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *refseq91_missin_1_Homo_S_mtDNA.bam -outnames $bname.noFA
PATH/software/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/software/names.dmp.gz -nodes PATH/software/nodes.dmp.gz -acc2tax PATH/software/nucl_gb.accession2taxid.gz -bam *refseq91_no_Homo_Sp_mtDNA.bam -outnames $bname.noGenus
cut -f9 -d":" $bname.holi.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.1.tax_counts
cut -f9 -d":" $bname.holi.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.1.taxNam
cut -f9 -d":" $bname.mtDB.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.2.tax_counts
cut -f9 -d":" $bname.mtDB.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.2.taxNam
cut -f9 -d":" $bname.RefFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.3.tax_counts
cut -f9 -d":" $bname.RefFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.3.taxNam
cut -f9 -d":" $bname.noGenus.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.4.tax_counts
cut -f9 -d":" $bname.noGenus.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.4.taxNam
cut -f9 -d":" $bname.noFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.5.tax_counts
cut -f9 -d":" $bname.noFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.5.taxNam
paste -d"," $bname.1.taxNam $bname.1.tax_counts > $bname3
paste -d"," $bname.2.taxNam $bname.2.tax_counts > $bname4
paste -d"," $bname.3.taxNam $bname.3.tax_counts > $bname5
paste -d"," $bname.4.taxNam $bname.4.tax_counts > $bname6
paste -d"," $bname.5.taxNam $bname.5.tax_counts > $bname7
cd $basepath
done


############ metadata summary stats
ll *trunc.fq | cut -f9 -d" " | cut -f1,3 -d "." > readsA_total_modelID.txt
wc -l *e.?.fa | grep '.f' | cut -f3 -d" " | awk '{print $1/2}' > readsB_extracted.txt
wc -l *d.?.fa | grep '.f' | cut -f3 -d" " | awk '{print $1/2}' > readsC_deaminated.txt
wc -l *trunc.fq | grep '.f' | cut -f3 -d" "| awk '{print $1/4}' > readsD_post_adaptorRem.txt


for file in HomoSap.e.*_holi/*Homo_sapiens_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsF_extracted_mapping2mtDNA.txt; done &
for file in HomoSap.e.*_holi/*missin_1_Homo_S_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsG_extracted_mapping_no_mtDNA.txt; done &
for file in HomoSap.e.*_holi/*_no_Homo_Sp_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsH_extracted_mapping_no_Genus.txt; done &
for file in HomoSap.e.*_holi/*refseq91_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsI_extracted_mapping_refseq_mtDNA.txt; done &
for file in HomoSap.e.*_holi/file.sort.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsJ_extracted_mapping2holi_db.txt; done &

for file in HomoSap.d.*_holi/*Homo_sapiens_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsK_deaminated_mapping2cpDNA.txt; done &
for file in HomoSap.d.*_holi/*missin_1_Homo_S_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsL_deaminated_mapping_no_mtDNA.txt; done &
for file in HomoSap.d.*_holi/*_no_Homo_Sp_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsM_deaminated_mapping_no_Genus.txt; done &
for file in HomoSap.d.*_holi/*refseq91_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsN_deaminated_mapping_refseq_mtDNA.txt; done &
for file in HomoSap.d.*_holi/file.sort.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsO_deaminated_mapping2holi_db.txt; done &

for file in HomoSap.g.*_holi/*Homo_sapiens_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsP_deamSeqErr_mapping2cpDNA.txt; done &
for file in HomoSap.g.*_holi/*missin_1_Homo_S_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsQ_deamSeqErr_mapping_no_mtDNA.txt; done &
for file in HomoSap.g.*_holi/*_no_Homo_Sp_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsR_deamSeqErr_mapping_no_Genus.txt; done &
for file in HomoSap.g.*_holi/*refseq91_mtDNA.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsS_deamSeqErr_mapping_refseq_mtDNA.txt; done &
for file in HomoSap.g.*_holi/file.sort.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsT_deamSeqErr_mapping2holi_db.txt; done &
```

```
paste reads* > read_metadata_merged.txt

############ checking damage profiles for fastas and fastqs

for file in *.5.trunc.fq
do
db=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA
fasta=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA.fa
echo $(basename $file)
echo $(basename $db)
echo $(basename $fasta)
bowtie2 -x $db -U $file --threads 30 --no-unal | samtools view -bS -> $(basename $file).$(basename $db).bam

nice -n 5 samtools sort $(basename $file).$(basename $db).bam -o $(basename $file).$(basename $db).sorted.bam
nice -n 5 mapDamage -i $(basename $file).$(basename $db).sorted.bam -r $fasta --merge-reference-sequences --forward
done

for file in *5.fa
do
db=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA
fasta=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA.fa
echo $(basename $file)
echo $(basename $db)
echo $(basename $fasta)
bowtie2 -x $db -f $file --threads 30 --no-unal | samtools view -bS -> $(basename $file).$(basename $db).bam

,
nice -n 5 samtools sort $(basename $file).$(basename $db).bam -o $(basename $file).$(basename $db).sorted.bam
nice -n 5 mapDamage -i $(basename $file).$(basename $db).sorted.bam -r $fasta --merge-reference-sequences --forward
done


###### metaG files and read_metadata_merged.txt downloaded locally for processing in R
```

10.5.1.2. Whole genome human modelling

```
########## extracting fasta and generating db's


#setting folder structure
mkdir Homo_sapiens_GRCh38_p12
mkdir Homo_sapiens_GRCh38_p12/endo
mkdir Homo_sapiens_GRCh38_p12/cont
mkdir Homo_sapiens_GRCh38_p12/bact
cp PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_GRCh38_p12.fa Homo_sapiens_GRCh38_p12/endo
samtools faidx Homo_sapiens_GRCh38_p12/endo/Homo_sapiens_GRCh38_p12.fa

# simulating and extacting read libraries and adding damage and seq errors
counter=1
while [ $counter -le 5 ]
do
echo $counter
bname='Homo_sapiens_GRCh38_p12'
basefolder='Homo_sapiens_GRCh38_p12'
fasta='Homo_sapiens_GRCh38_p12.fa'
ReadNo='100000'
echo $bname
### 30X fragment retrieval  # number of reads needed to generate depends on the length of the genome and the average read length extracted if 30X covereage is needed divide the total bases in the genome with averageread length and times 30. In this case ReadNo = 16569Bp/42*30 = 11835
PATH/software/gargammel/src/fragSim -tag e -n $ReadNo -m 0 -M 1000 -f PATH/Mex_Cave/model_simulations/gargammel/average_readDist.txt $basefolder/endo/$fasta > $basefolder/$bname.e.$counter.fa
#### adding DNA deamination
PATH/software/gargammel/src/deamSim -mapdamage PATH/Mex_Cave/model_simulations/gargammel/misincorporation.txt single $basefolder/$bname.e.$counter.fa > $basefolder/$bname.d.$counter.fa
#adding adaptors
PATH/software/gargammel/src/adptSim -arts $basefolder/$bname.f.$counter.fasta $basefolder/$bname.d.$counter.fa
#adding sequencing errors
PATH/software/gargammel/art_src_MountRainier/art_illumina -ss HS25 -amp -na -qs -qs2 -i $basefolder/$bname.f.$counter.fasta -l 80 -c 1 -qs 0 -qs2 0 -o $basefolder/$bname.g.$counter
((counter++))
done
for file in $basefolder/$bname.g.?.fq
do
# removing adaptors
AdapterRemoval --file1 $file --mm 3 --minlength 30 --basename $file --trimns --trimqualities --minquality 30
```

```
for infile in $basefolder/*.truncated; do bname=$(basename $infile); echo $bname; bname2=$(echo $bname | sed 's/.fq.truncated*/.trunc.fq/');
echo $bname2 ; mv $basefolder/$bname $basefolder/$bname2 ; done
done


###################### fq mapping
for infile in $(pwd)/*trunc.fq
do
bname=$(basename $infile)
echo $bname
bname2=$(echo $bname | sed 's/.fq*/_holi/')
basepath=$(pwd)/
basefolder=$basepath
echo $basepath
echo $bname2
mkdir $basepath$bname2
cd $bname2


# mapping against fasta reference
pwd
for DB in PATH/Mex_Cave/model_simulations/reference_dbs/homo_sapiens_GRCh38_p12
do
echo Mapping $bname against $DB
bowtie2 --threads 60 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done


# mapping against nt_db
for DB in PATH/database/ncbi_nt/nt.?
do
echo Mapping adap2_kmer2_$bname.pp.rmdup.fq against $DB
bowtie2 --threads 60 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

# mapping against RefSeq

for DB in PATH/database/refseq/vert_other/vert_other.?
do
echo Mapping $bname.fq against $DB
bowtie2 --threads 60 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/vert_mam/vert_mam.?
do
echo Mapping $bname.fq against $DB
bowtie2 --threads 60 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/vert_mam/vert_mam.??
do
echo Mapping $bname.fq against $DB
bowtie2 --threads 60 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

for DB in PATH/database/refseq/invert/invert.?
do
echo Mapping $bname.fq against $DB
bowtie2 --threads 60 -k 5000 -x $DB -U ../${bname} --no-unal | samtools view -bS - > $bname.$(basename $DB).bam
done

cd $basepath
done


######################################### ngs Fastq

for infile in *trunc.fq
do
bname=$(basename $infile)
echo $bname
bname2=$(echo $bname | sed 's/.fq*/_holi/')
bname3=$(echo $bname | sed 's/.fq*/.holi.metagenome.txt/')
bname5=$(echo $bname | sed 's/.fq*/.RefFA.metagenome.txt/')

basepath='PATH/Mex_Cave/model_simulations/gargammel/data/Homo_sapiens_GRCh38_p12'
basefolder=$basepath
echo $basepath
echo $bname2
```

```
echo $bname3
#mkdir $basepath$bname2
cd $bname2

samtools merge -@ 40 -f -n tmp.bam.merged *.bam
samtools sort -n -T PATH/TMP -O bam -o file.sort.bam -@ 40 tmp.bam.merged
PATH/ngsLCA/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/ngsLCA/ngsLCA/ncbi_tax_dump_files/names.dmp.gz -nodes PATH/ngsLCA/ngsLCA/ncbi_tax_dump_files/nodes.dmp.gz -acc2tax PATH/ngsLCA/ngsLCA/ncbi_tax_dump_files/nucl_gb.accession2taxid.gz -bam file.sort.bam -outnames $bname.holi
PATH/ngsLCA/ngsLCA/ngsLCA -editdistmin 0 -editdistmax 0 -names PATH/ngsLCA/ngsLCA/ncbi_tax_dump_files/names.dmp.gz -nodes PATH/ngsLCA/ngsLCA/ncbi_tax_dump_files/nodes.dmp.gz -acc2tax PATH/ngsLCA/ngsLCA/ncbi_tax_dump_files/nucl_gb.accession2taxid.gz -bam *GRCh38_p12.bam -outnames $bname.RefFA
cut -f9 -d":" $bname.holi.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.1.tax_counts
cut -f9 -d":" $bname.holi.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.1.taxNam

cut -f9 -d":" $bname.RefFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f1 -d" "  > $bname.3.tax_counts
cut -f9 -d":" $bname.RefFA.lca | sort | uniq -c | sort -k1 | perl -pe 's/^\s+//' | cut -f2,3,4,5,6,7,8,9,10,11,12,13 -d" " > $bname.3.taxNam

paste -d"," $bname.1.taxNam $bname.1.tax_counts > $bname3

paste -d"," $bname.3.taxNam $bname.3.tax_counts > $bname5

cd $basepath
done


############ metadata summary stats
ll *trunc.fq | cut -f9 -d" " | cut -f1,3 -d "." > readsA_total_modelID.txt

wc -l *trunc.fq | grep '.f' | cut -f3 -d" "| awk '{print $1/4}' > readsD_post_adaptorRem.txt

for file in Homo_sapiens_GRCh38_p12.g*_holi/*homo_sapiens_GRCh38_p12.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsP_deamSeqErr_mapping2cpDNA.txt; done &
for file in Homo_sapiens_GRCh38_p12.g*_holi/file.sort.bam; do samtools view $file | cut -f1 | uniq | wc -l >> readsT_deamSeqErr_mapping2holi_db.txt; done &

paste reads* > read_metadata_merged.txt

############# checking damage profiles for fastas and fastqs

for file in *.5.trunc.fq
do
db=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA
fasta=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA.fa
echo $(basename $file)
echo $(basename $db)
echo $(basename $fasta)
bowtie2 -x $db -U $file --threads 30 --no-unal | samtools view -bS -> $(basename $file).$(basename $db).bam

nice -n 5 samtools sort $(basename $file).$(basename $db).bam -o $(basename $file).$(basename $db).sorted.bam
nice -n 5 mapDamage -i $(basename $file).$(basename $db).sorted.bam -r $fasta --merge-reference-sequences --forward
done


for file in *5.fa
do
db=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA
fasta=PATH/Mex_Cave/model_simulations/reference_dbs/Homo_sapiens_mtDNA.fa
echo $(basename $file)
echo $(basename $db)
echo $(basename $fasta)
bowtie2 -x $db -f $file --threads 30 --no-unal | samtools view -bS -> $(basename $file).$(basename $db).bam

'

nice -n 5 samtools sort $(basename $file).$(basename $db).bam -o $(basename $file).$(basename $db).sorted.bam
nice -n 5 mapDamage -i $(basename $file).$(basename $db).sorted.bam -r $fasta --merge-reference-sequences --forward
done


###### metaG files and read_metadata_merged.txt downloaded locally for processing in R
```

10.5.2. R scripts

10.5.2.1. Plotting the mt and cpDNA models

```
################ Barplot for the mt and cpDNA models ###############
multmerge = function(mypath){
  filenames=list.files(path=mypath, full.names=TRUE)
  datalist = lapply(filenames, function(x){read.csv(file=x,header=F)})
  Reduce(function(x,y) {merge(x,y, by=c("V1"),all=TRUE, fill=0, split=FALSE, verbose=TRUE)}, datalist)
}

mydata=multmerge("/PATH/model_simulations/hum_mtDNA_model_metagenome/sequenced/")

filenames=list.files(path="/PATH/model_simulations/hum_mtDNA_model_metagenome/sequenced/", full.names=F)
colnames(mydata) <- c("Taxa", "Model_1_DB1", "Model_1_DB2","Model_1_DB3","Model_1_DB4","Model_1_DB5", "Model_2_DB1",
"Model_2_DB2","Model_2_DB3","Model_2_DB4","Model_2_DB5","Model_3_DB1",
"Model_3_DB2","Model_3_DB3","Model_3_DB4","Model_3_DB5","Model_4_DB1",
"Model_4_DB2","Model_4_DB3","Model_4_DB4","Model_4_DB5","Model_5_DB1",
"Model_5_DB2","Model_5_DB3","Model_5_DB4","Model_5_DB5")

mydata[is.na(mydata)]=0 #if this one fails it might have text in the number of reads coloumn

require(reshape2)
test_long2 <- melt(mydata, id.vars=c("Taxa"))

require(ggplot2)
require(RColorBrewer)
colourCount =length(unique(test_long2$Taxa))
getPalette = colorRampPalette(brewer.pal(12, "Paired"))
p <- ggplot(test_long2, aes(fill=Taxa, y=value, x=variable)) +
      geom_bar(stat="identity") + theme(legend.position="none", axis.text.x = element_text(angle = -90, hjust = 0, vjust = 0.5)) +
scale_fill_manual(values = getPalette(colourCount))
p
p <- ggplot(test_long2, aes(fill=Taxa, y=value, x=variable)) +
  geom_bar(stat="identity") + theme( axis.text.x = element_text(angle = -90, hjust = -0.5)) + scale_fill_manual(values = getPalette(colourCount))

#saved as 20x23
p + theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust = 0.5, size=14), axis.text.y = element_text(size=14), axis.title.x =
element_text(size=16), axis.title.y = element_text(size=16), legend.position = "right", legend.direction = "horizontal",
legend.text=element_text(size=14), legend.title=element_text(size=16)) + xlab("Simulations and resulting taxa different databases") +
ylab("Number of reads")

######### stacked percentage barplot
#ggplot(test_long2, aes(fill=Taxa, y=value, x=variable)) +
  #geom_bar(stat="identity", position="fill") + theme(legend.position="bottom", axis.text.x = element_text(angle = -90, hjust = 0, vjust = 0.5)) +
scale_fill_manual(values = getPalette(colourCount))


############# calculating percentage and removing threshold
b1=as.matrix(mydata[,seq(2,26)])
rownames(b1)<-mydata$Taxa

b2 <- prop.table(b1, margin=2)*100 # makes proportion table, needs 2 margins e.g. header and 1st row names
colSums(prop.table(b1, margin=2)*100) # should give 100 for each coloumn

tmp <- b2[apply(b2[,1:25], MARGIN = 1, function(x) any(x > 1.0)), ]
filenames=list.files(path="/PATH/model_simulations/hum_mtDNA_model_metagenome/sequenced/", full.names=F)
require(reshape2)
tmp2 <- melt(tmp, keep.rownames = TRUE)

###### plotting remaining taxa
colourCount =length(unique(c(rownames(tmp))))
getPalette = colorRampPalette(brewer.pal(8, "Dark2"))
p <- ggplot(tmp2, aes(fill=Var1, y=value, x=Var2)) +
  geom_bar(stat="identity") +  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust = 0.5, size=14), axis.text.y = element_text(size=14),
axis.title.x = element_text(size=16), axis.title.y = element_text(size=16), legend.position = "right", legend.direction = "vertical",
legend.text=element_text(size=14), legend.title=element_text(size=16)) + xlab("Simulations and resulting taxa against different databases") +
ylab("Percentage of reads")
p + scale_fill_manual(values = getPalette(colourCount))

install.packages("gridExtra")
library(gridExtra)
grid.table(tmp)
```

10.5.2.2. Plotting WGS human model

```
################ Barplot for the models

multmerge = function(mypath){
  filenames=list.files(path=mypath, full.names=TRUE)
```

```
  datalist = lapply(filenames, function(x){read.csv(file=x,header=F)})
  Reduce(function(x,y) {merge(x,y, by=c("V1"),all=TRUE, fill=0, split=FALSE, verbose=TRUE)}, datalist)
}

mydata=multmerge("/PATH/model_simulations/wgs_human/")

filenames=list.files(path="/PATH/model_simulations/wgs_human/", full.names=F)
colnames(mydata) <- c("Taxa", "Model_1_DB1", "Model_1_DB5", "Model_2_DB1","Model_2_DB5","Model_3_DB1","Model_3_DB5","Model_4_DB1", "Model_4_DB5","Model_5_DB1","Model_5_DB5")

mydata[is.na(mydata)]=0 #if this one fails it might have text in the number of reads coloumn

require(reshape2)
test_long2 <- melt(mydata, id.vars=c("Taxa"))

require(ggplot2)
require(RColorBrewer)
colourCount =length(unique(test_long2$Taxa))
getPalette = colorRampPalette(brewer.pal(12, "Paired"))
p <- ggplot(test_long2, aes(fill=Taxa, y=value, x=variable)) +
   geom_bar(stat="identity") + theme(legend.position="none", axis.text.x = element_text(angle = -90, hjust = 0)) + scale_fill_manual(values = getPalette(colourCount))

#saved as 15x30
p + theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust = 0.5, size=14), axis.text.y = element_text(size=14), axis.title.x = element_text(size=16), axis.title.y = element_text(size=16), legend.position = "right", legend.direction = "horizontal", legend.text=element_text(size=14), legend.title=element_text(size=16)) + xlab("Simulations against different databases") + ylab("Number of reads")


############# calculating percentage and removing threshold
b1=as.matrix(mydata[,seq(2,11)])
rownames(b1)<-mydata$Taxa

b2 <- prop.table(b1, margin=2)*100 # makes proportion table, needs 2 margins e.g. header and 1st row names
colSums(prop.table(b1, margin=2)*100) # should give 100 for each coloumn

tmp <- b2[apply(b2[,1:10], MARGIN = 1, function(x) any(x > 1.0)), ]
filenames=list.files(path="/PATH/model_simulations/wgs_human/sequenced/", full.names=F)
require(reshape2)
tmp2 <- melt(tmp, keep.rownames = TRUE)

###### plotting remaining taxa saved as 10x13 pdf
colourCount =length(unique(c(rownames(tmp))))
getPalette = colorRampPalette(brewer.pal(8, "Dark2"))
p <- ggplot(tmp2, aes(fill=Var1, y=value, x=Var2)) +
  geom_bar(stat="identity") +  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust = 0.5, size=14), axis.text.y = element_text(size=14), axis.title.x = element_text(size=16), axis.title.y = element_text(size=16), legend.position = "right", legend.direction = "vertical", legend.text=element_text(size=14), legend.title=element_text(size=16)) + xlab("Simulations and resulting taxa against different databases") + ylab("Percentage of reads")
p + scale_fill_manual(values = getPalette(colourCount))


## making table
install.packages("gridExtra")
library(gridExtra)
grid.table(tmp)
```

References

1. Brock, F., Higham, T., Ditchfield, P. & Ramsey, C. B. Current Pretreatment Methods for AMS Radiocarbon Dating at the Oxford Radiocarbon Accelerator Unit (Orau). *Radiocarbon* 52, 103–112 (2010).

2. Bronk Ramsey, C. Dealing with Outliers and Offsets in Radiocarbon Dating. Radiocarbon 51, 1023–1045 (2009).

3. Bronk Ramsey, C. Development of the radiocarbon program OxCal. Radiocarbon 43, 355–363 (2001).

4. Bronk Ramsey, C. OxCal 4.3 Manual. OxCal 4.3 Manual https://c14.arch.ox.ac.uk/oxcalhelp/hlp_contents.html.

5. Barba, L. A., Ortiz, A., Link, K. F., López Lujan, L. & Lazos, L. Chemical analysis of residues in floors and the reconstruction of ritual activities at the Templo Mayor, Mexico. in *Archaeological Chemistry: Organic, Inorganic, and Biochemical Analysis* (ed. Orna, M. V.) 139–156 (American Chemical Society, 1996).

6. Middleton, W. D. et al. The Study of Archaeological Floors: Methodological Proposal for the Analysis of Anthropogenic Residues by Spot Tests, ICP-OES, and GC-MS. *Journal of Archaeological Method and Theory* 17, 183–208 (2010).

7. Barba, L. Chemical residues in lime-plastered archaeological floors. *Geoarchaeology: An International Journal* 22, 439–452 (2007).

8. Barba, L., Ortiz, A., Blancas, J., Hernández, M. & Obregón, M. Residuos químicos y el contexto de enterramiento. *Estudios Interdisciplinarios sobre un mamut y su contexto* 79–96 (2015).

9. Piperno, D. R. Phytoliths: *A Comprehensive Guide for Archaeologists and Paleoecologists.* (AltaMira Press, 2006).

10. Juggins, S. C2 Software. (2010).

11. Piperno, D. R. & Pearsall, D. M. The silica bodies of tropical American grasses: morphology, taxonomy, and implications for grass systematics and fossil phytolith identification. *Smithsonian Contributions to Botany* 85, 1–40 (1998).

12. Fredlund, G. G. & Tieszen, L. T. Modern Phytolith Assemblages from the North American Great Plains. *J. Biogeogr.* 21, 321–335 (1994).

13. Gallego, L. & Distel, R. A. Phytolith assemblages in grasses native to central Argentina. *Ann. Bot.* 94, 865–874 (2004).

14. Lu, H. & Liu, K.-B. Phytoliths of common grasses in the coastal environments of southeastern USA. *Estuar. Coast. Shelf Sci.* 58, 587–600 (2003).

15. Barboni, D. & Bremond, L. Phytoliths of East African grasses: An assessment of their environmental and taxonomic significance based on floristic data. *Rev. Palaeobot. Palynol.* 158, 29–41 (2009).

16. Mercader, J. et al. Poaceae phytoliths from the Niassa Rift, Mozambique. *J. Archaeol. Sci.* 37, 1953–1967 (2010).

17. Kondo, R., Childs, C. W. & Atkinson, I. A. E. *Opal Phytoliths of New Zealand*. (Manaaki Whenua Press, 1994).

18. Jattisha, P. I. & Sabu, M. Phytoliths as a tool for the identification of some chloridoideae grasses in Kerala. *ISRN Botany* 2012 (2012).

19. Collinvaux, P. A., De Oliveira, P. E. & Moreno, E. Amazon: Pollen Manual and Atlas: Pollen Manual and Atlas. (CRC Press, 2014).

20. Grimm, E. C. & Troostheide, C. D. Tilia 2.00, program for plotting palynological diagrams. Springfield: Illinois State Museum (1994).

21. Hooghiemstra, H. *Vegetational and climatic history of the high plain of Bogotá, Colombia: A Continuous Record of 3.5 Million Years (Dissertationes Botanicae)*. (Lubrecht & Cramer Ltd, 1984).

22. Roubik, D. W. & Moreno, J. E. *Pollen and spores of Barro Colorado Island*. Monographs in Systematic Botany 36 (Missouri Botanical Garden, 1991).

23. Markgraf, V. & d'Antoni, H. L. *Pollen flora of Argentina.* (University of Arizona Press, 1978).

24. Villaseñor, J. L. Checklist of the native vascular plants of Mexico. *Rev. Mex. Biodivers.* 87, 559–902 (2016).

25. Johnston, I. M. Plants of Coahuila, eastern Chihuahua, and adjoining Zacatecas and Durango, Volumes I-IV. J*ournal of the Arnold Arboretum* 24; 25, (1944).

26. González-Tagle, M. A., Schwendenmann, L., Pérez, J. J. & Schulz, R. Forest structure and woody plant species composition along a fire chronosequence in mixed pine–oak forest in the Sierra Madre Oriental, Northeast Mexico. *For. Ecol. Manage.* 256, 161–167 (2008).

27. Miall, A. Alluvial deposits. in *Facies Models 4* (eds. James, N. P. & Dalrymple, R. W.) 105–137 (Geological Association of Canada, 2010).

28. Garzanti, E. & Vezzoli, G. A Classification of Metamorphic Grains in Sands Based on their Composition and Grade. *J. Sediment. Res.* 73, 830–837 (2003).

29. Dunham, R. J. Classification of Carbonate Rocks According to Depositional Texture. in *Classification of Carbonate Rocks—A Symposium* (ed. Ham, W. E.) 108–121 (The American Association of Petroleum Geologists, 1962).

30. Folk, R. L. Spectral subdivision of limestone types: classification of carbonate rocks. *American Association of Petroleum Geologists Memoir* 1, 62–82 (1962).

31. Folk, R. L. Practical petrographic classification of limestones. *American Association of Petroleum Geologist Bulletin* 43, 1–38 (1959).

32. Wilson, J. L. *Carbonate Facies in Geologic History*. (Springer-Verlag, 1975).

33. Passchier, C. W. & Trouw, R. A. J. *Microtectonics*. (Springer-Verlag, 1998).

34. Gulliksen, S. & Scott, M. Report of the TIRI Workshop, Saturday 13 August 1994. *Radiocarbon* 37, 820–821 (1995).

35. Richter, D., Richter, A. & Dornich, K. Lexsyg smart—a luminescence detection system for dosimetry, material research and dating application. *Geochronometria* 42, (2015).

36. Hansen, V., Murray, A., Buylaert, J.-P., Yeo, E.-Y. & Thomsen, K. A new irradiated quartz for beta source calibration. Radiat. Meas. 81, 123–127 (2015).

37. Bøtter-Jensen, L., Bulur, E., Duller, G. A. T. & Murray, A. S. Advances in luminescence instrument systems. *Radiat. Meas.* 32, 523–528 (2000).

38. Murray, A. S. & Wintle, A. G. Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiat. Meas.* 32, 57–73 (2000).

39. Wintle, A. G. & Murray, A. S. A review of quartz optically stimulated luminescence characteristics and their relevance in single-aliquot regeneration dating protocols. *Radiat. Meas.* 41, 369–391 (2006).

40. Banerjee, D., Murray, A. S., Bøtter-Jensen, L. & Lang, A. Equivalent dose estimation using a single aliquot of polymineral fine grains. *Radiat. Meas.* 33, 73–94 (2001).

41. Wallinga, J., Murray, A. S. & Bøtter-Jensen, L. Measurement of the dose in quartz in the presence of feldspar contamination. *Radiat. Prot. Dosimetry* 101, 367–370 (2002).

42. Duller, G. A. T. The Analyst software package for luminescence data: overview and recent improvements. *Ancient TL* 33, 35–42 (2015).

43. Dietze, M. et al. The abanico plot: Visualising chronometric data with individual standard errors. *Quat. Geochronol.* 31, 12–18 (2016).

44. Kreutzer, S. et al. Introducing an R package for luminescence dating analysis. *Ancient TL* 30, 1–8 (2012).

45. Kreutzer, S. et al. Luminescence: Comprehensive Luminescence Dating Data Analysis. R package version 0.8.6. Luminescence: Comprehensive Luminescence Dating Data Analysis. (2018). https://CRAN.R-project.org/package=Luminescence.

46. Corte, F. D. et al. Relative and k0-standardized INAA to assess the internal (Th, U) radiation dose rate in the 'quartz coarse-grain protocol' for OSL dating of sediments: Unexpected observations. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 564, 743–751 (2006).

47. Vandenberghe, D., De Corte, F., Buylaert, J.-P., Kučera, J. & Van den haute, P. On the internal radioactivity in quartz. *Radiat. Meas.* 43, 771–775 (2008).

48. Guérin, G., Mercier, N. & Adamiec, G. Dose-rate conversion factors: update. *Ancient TL* 29, 5–8 (2011).

49. Durcan, J. A., King, G. E. & Duller, G. A. T. DRAC: Dose Rate and Age Calculator for trapped charge dating. *Quat. Geochronol.* 28, 54–61 (2015).

50. Guérin, G., Mercier, N., Nathan, R., Adamiec, G. & Lefrais, Y. On the use of the infinite matrix assumption and associated concepts: A critical review. *Radiat. Meas.* 47, 778–785 (2012).

51. Brennan, B. J., Lyons, R. G. & Phillips, S. W. Attenuation of alpha particle track dose for spherical grains. *Int. J. Rad. Appl. Instrum.* D 18, 249–253 (1991).

52. Bell, W. T. Attenuation factors for the absorbed radiation dose in quartz inclusions for thermoluminescence dating. *Ancient TL* 8, 12 (1979).

53. Aitken, M. J. *Thermoluminescence dating.* (1985). (Academic Press, 1985).

54. Burow, C. calc_CosmicDoseRate(): Calculate the cosmic dose rate. Function version 0.5.2. in Luminescence: Comprehensive Luminescence Dating Data Analysis. R package version 0.8.6. (Kreutzer, S. et al.) (2018). https://CRAN.R-project.org/package=Luminescence.

55. Prescott, J. R. & Hutton, J. T. Cosmic ray contributions to dose rates for luminescence and ESR dating: Large depths and long-term time variations. *Radiat. Meas*. 23, 497–500 (1994).

56. Pedersen, M. W. et al. Postglacial viability and colonization in North America's ice-free corridor. *Nature* 537, 45–49 (2016).

57. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U. S. A.* 104, 14616–14621 (2007).

58. Renaud, G., Hanghøj, K., Willerslev, E. & Orlando, L. gargammel: a sequence simulator for ancient DNA. *Bioinformatics* 33, 577–579 (2017).

59. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 29, 1682–1684 (2013).

60. Slon, V. et al. Neandertal and Denisovan DNA from Pleistocene sediments. *Science* 356, 605–608 (2017).