10-30-2020

Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau

Dongju Zhang

Huan Xia

Fahu Chen

Bo Li

Viviane Slon

See next page for additional authors

This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Zhang, Dongju, Huan Xia, Fahu Chen, Bo Li, Viviane Slon, Ting Cheng, Melinda A. Yang, et al. “Denisovan DNA in Late Pleistocene Sediments from Baishiya Karst Cave on the Tibetan Plateau.” Science 370, no. 6516 (October 30, 2020): 584–87. https://doi.org/10.1126/science.abb6320.

Authors
Dongju Zhang, Huan Xia, Fahu Chen, Bo Li, Viviane Slon, Ting Cheng, Ruowei Yang, and Melinda A. Yang

Title: Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau

Authors: Dongju Zhang[1,2,3]*, Huan Xia[1], Fahu Chen[2,1], Bo Li[4,5]*, Viviane Slon[6], Ting Cheng[1], Ruowei Yang[7,8], Zenobia Jacobs[4,5], Qingyan Dai[7], Diyendo Massilani[6], Xuke Shen[1], Jian Wang[1,9], Xiaotian Feng[7], Peng Cao[7], Melinda A. Yang[10], Juanting Yao[1], Jishuai Yang[1], David B Madsen[1,11], Yuanyuan Han[1], Wanjing Ping[7], Feng Liu[7], Charles Perreault[12,13], Xiaoshan Chen[1], Matthias Meyer[6], Janet Kelso[6], Svante Pääbo[6]* and Qiaomei Fu[7,8,14]*

Affiliations:

[1]Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, 730000, China.

[2]Key Laboratory of Alpine Ecology (LAE), CAS Center for Excellence in Tibetan Plateau Earth Sciences and Institute of Tibetan Plateau Research, Chinese Academy of Sciences (CAS), Beijing, 100101, China.

[3]Frontier Center for Eco-environment and Climate Change in Pan-third Pole Regions, Lanzhou University, Lanzhou, 730000, China.

[4]Centre for Archaeological Science, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, New South Wales, 2522, Australia.

[5]Australian Research Council (ARC) Centre of Excellence for Australian Biodiversity and Heritage, University of Wollongong, Wollongong, New South Wales, 2522, Australia.

[6]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, 04103, Germany.

[7]Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, CAS, Beijing, 100044, China.

[8]Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing, 100044, China.

[9]School of Earth Sciences, Lanzhou University, Lanzhou,730000, China.

[10]Department of Biology, University of Richmond, Richmond, VA 23173, United States.

[11]Department of Anthropology, University of Nevada-Reno, Reno, NV 89557, United States.

[12]School of Human Evolution and Social Change, Arizona State University, Tempe AZ 85281, United States.

[13]Institute of Human Origins, Arizona State University, Tempe AZ 85281, United States.

[14]University of Chinese Academy of Sciences, 100049, China.

*Correspondence to: djzhang@lzu.edu.cn; bli@uow.edu.au; paabo@eva.mpg.de; fuqiaomei@ivpp.ac.cn.

Abstract: A late Middle Pleistocene mandible from Baishiya Karst Cave (BKC) on the Tibetan Plateau has been inferred to be a Denisovan, an Asian hominin related to Neandertals, based on an amino acid substitution in its collagen. Here we describe the stratigraphy, chronology and mitochondrial DNA extracted from the sediments in BKC. We recover Denisovan mitochondrial DNA from sediments deposited ~100 and ~60 thousand years ago (ka), and possibly as recently as ~45 ka. The long-term occupation of BKC by Denisovans suggests that they may have adapted to life at high altitudes and may have contributed such adaptations to modern humans on the Tibetan Plateau.

One Sentence Summary: Sediment DNA reveals long-term presence of Late Pleistocene Denisovans in Baishiya Karst Cave on the Tibetan Plateau

Main Text:

Denisovans are an extinct hominin group initially identified from a genome sequence determined from a fragment of a phalanx found at Denisova Cave in the Altai Mountains in southern Siberia (*1-3*). Subsequent analyses of the genome have showed that Denisovans diverged from Neandertals ~400 thousand years ago (ka) (*4*) and that at least two distinct Denisovan populations mixed with ancestors of present-day Asians (*2-9*). Thus, they are assumed to have been widely dispersed across Asia. However, physical remains of Denisovans in Siberia have been restricted to a fragmentary phalanx (*1*), three teeth (*2, 10, 11*), and a cranial fragment (*12*), all of which were found at Denisova Cave.

Recently, a half mandible from the Baishiya Karst Cave (BKC), Xiahe County, Gansu, China, dated to at least 160 ka, was identified as a Denisovan (*13*). However, this identification of the Xiahe mandible as a Denisovan is based on a single amino acid position and is therefore tenuous. Here we report the results of ongoing archaeological and chronological investigations and sedimentary DNA analyses from BKC. We find evidence for the long-term presence of Denisovans in BKC and provide stratigraphic and chronological context for their occupation in the cave.

BKC (35.45 N, 102.57 E, 3,280 masl) is a limestone cave located in the northeastern margin of the Tibetan Plateau (Fig. 1A and Fig. S1A). In 2018 and 2019, three 1 x 2 m^2 units (T1, T2 and T3) were plotted for excavation in the entrance chamber, which is about 60 m long, 8 m wide and 5 m high (Fig. 1B, and Fig. S1, B and C) (*14*). The second unit (T2) exposed intact cultural strata that are truncated in the southeastern part of the trench by a large pit (H1) dug during the historical period (*14*) , 780–700 cal yr BP (calibrated years before present) (Figs. S2 and S3, Table S3). Ten stratigraphic layers were identified mainly on the basis of sedimentary characteristics (Figs. S2B and S3) (*14*). Most layers are poorly sorted, composed of a silt matrix with abundant angular clasts of autogenic limestone gravels. The latter originates from the reworking of eroded parent bedrock, sediments by colluviation or spalling of material from the cave walls and roof (see details in Supplementary Materials). Stone artifacts and animal fossils were recovered from all layers (Figs. S11 and S12)(*14*). A total of 1,310 stone artifacts and 579 animal bone fragments were recorded and collected. Preliminary analysis of the stone artifact assemblage suggests they were made mostly from local metamorphic quartz sandstone and hornstone stream cobbles using a simple core and flake technology (Fig. S11). Remains of small and medium-size animals dominate the fossil assemblage in Layers 6–1, including gazelles, marmots, and foxes, whereas large animals, such as rhinoceros, large bovids and hyenas, dominate Layers 10–7 (Fig. S12).

We constructed a numerical chronology for the T2 sequence from optical dating of 12 sediment samples and radiocarbon dating of 14 bone fragments (Fig. 2 and Fig. S3, tables S3, S9 and S10). The age estimates were used to develop a Bayesian model for the depositional chronology of the site and to provide an age framework for hominin occupation (Fig. 2 and Table S12). Details of sample locations and collection, preparation, measurement, and data-analysis procedures are provided, together with the measured and modelled ages and related data (*14*). The deposits in Layers 10–4 have a stratigraphically coherent chronology, limited age variation within layers and equivalent dose (D_e) distributions that show minimal evidence for mixing. Layer 10 accumulated between 190 ± 34 ka and 129 ± 20 ka (here and below, we give modelled age estimates and total uncertainties at 95.4% probability), followed by relatively fast accumulation of Layers 9–6 until 96 ± 5 ka. No age was obtained for Layer 5. We modelled a time interval with a duration of 24–39 ka for Layer 5. The sedimentary features of Layer 5 (Table S1) are indicative of a fluvial environment within BKC and may represent an erosional event that removed deposits dated to between ~100 and ~60 ka. Layer 4 was deposited from 66 ± 6 ka to 47 ± 2 ka. A depositional hiatus with a duration of ~7–18 ka was identified in the middle of Layer 4 between ~60–50 ka, suggesting that sediments in Layer 4 may have been deposited in two broad pulses. Layer 3 accumulated from 46 ± 2 ka to 33 ± 1 ka, followed by Layer 2 until 17 ± 12 ka. Layers 2 and 3 are more complex; radiocarbon ages vary significantly within a layer (Table S3), and single-grain D_e values are broadly distributed (Figs. S17, S18, S19 and S26).

To test whether ancient DNA was preserved in the cave, we extracted DNA (*15*) from eight sediment samples (100–250 mg each) collected from the middle of each layer (except Layers 1 and 5) (Fig. 2A and Fig. S3, Table S18). Aliquots of each extract were converted to DNA libraries and enriched for mammalian and human mitochondrial DNA (mtDNA) using probes for 242 mammalian mtDNAs (*15*) and for human mtDNA (*16*). For each library, the number of DNA fragments sequenced ranged from 0.07 to 1.7 million. From these, we obtained between 10 and 27,150 unique fragments mapping to mammalian mitochondrial genomes. All sampled layers, except Layers 8 and 9, contained mammalian mtDNA. In Layers 4, 6, 7, and 10, we detected DNA from animal species that have not been present in the area since about 10,000 years ago, including extinct hyenas and rhinoceros (*17*), species for which bones were also identified in Layer 10 (Figs. S12 and S27). This confirms that ancient DNA is preserved in the cave.

We then assessed whether ancient hominin DNA was present in each library by determining the frequency of apparent cytosine (C) to thymine (T) substitutions at the ends of DNA sequences; characteristic of ancient DNA where terminal C residues tends to become deaminated. The libraries from Layers 2, 3, 4, and 7 have between 15.6% and 50% C->T terminal substitutions indicating the presence of ancient hominin DNA (Table S19). We thus prepared additional DNA extracts from Layers 2, 3, 4, and 7 (Table S18). To determine which hominin groups may have contributed mtDNA to these samples, we examined sequences for substitutions found to be unique to modern humans, Neandertals, Denisovans, and a ~430,000-

year-old hominin individual from Sima de los Huesos (*Sima*) (*18*) in phylogenetic analyses of mtDNAs as described (*14, 19*).

From the 24 libraries, between 31% (368/1186) and 95% (601/635) of the mtDNA fragments that covered informative positions matched the Denisovan state (Fig. 3 and Fig. S28), while 0–14% (1/7) matched the Neandertal state (Fig. 3 and Fig. S28), 0–3.7% (5/135) the *Sima* state (Table S19), and 0–67% (338/508) the modern human state (Fig. 3 and Fig. S28). Restricting the analysis to DNA fragments with first and last three C->T substitutions indicating cytosine deamination (Fig. 3 and Fig. S28) decreased the proportion of fragments matching the modern human state to 0–43% (3/7), and increased the proportion matching the Denisovan state to 71–100%. To reduce the influence of modern human contamination, we restricted subsequent analyses to deaminated mtDNA fragments and excluded two DNA libraries where modern human mtDNA fragments were slightly deaminated albeit much less so than Denisovan mtDNA fragments (*14*).

By merging deaminated hominin mtDNA fragments from libraries for each layer, we arrive at an average mtDNA coverage for Layers 2, 3, 4, and 7 of 0.37-fold, 1.5-fold, 40-fold, and 1.3-fold, respectively. DNA recovered from sediments may be derived from multiple different individuals, and this is the case at least in Layer 4 where we have sufficient information to estimate the number of mtDNA fragments present (*14*). However, to gauge the average relationships of mtDNA in each layer we called a consensus mtDNA sequence for each layer using positions covered by at least two different DNA fragments, excluding positions covered by only two fragments and where they differ. We also required that more than two-thirds of the fragments covering each position must carry an identical base at positions covered by more than two fragments (*18*). These sequences covered 7%, 36%, 99% and 26% of the mtDNA, respectively (Table S15).

We then estimated phylogenetic trees using previously published mtDNA sequences from four Denisovans from Denisova Cave (*Denisova 2*, *Denisova 3*, *Denisova 4* and *Denisova 8*) and the individual from Sima de los Huesos. The composite consensus mtDNA from Layer 4 that is of comparatively high quality falls within the mtDNA variation of Denisovans, forming a clade with *Denisova 3* and *4* to the exclusion of *Denisova 2* and *8* (Fig. 4). When the consensus mtDNA sequences that are of lower quality are analyzed separately (Fig. 4), the mtDNA sequences from Layers 2 and 3 form a clade with the Layer 4 mtDNA, whereas the consensus Layer 7 mtDNA diverges earlier from the lineage leading to *Denisova 3* and *4*. Thus, the mtDNA sequences from BKC form a clade (100% posterior support) with the mtDNA sequences for *Denisova 3* and *4* (*20, 21*). The depositional age for the lower part of Layer 4 (~60 ka) (Fig. 2) is comparable to the date of *Denisova 3* (76–52 ka) and *Denisova 4* (84–55 ka) (*20, 21*). Besides, the depositional age for Layer 7 (108–97 ka) (Fig. 2) is older than those for *Denisova 3* and *4*, but younger than the ages for *Denisova 2* (194–122 ka) and *Denisova 8* (136–105 ka) (*20, 21*). Although Denisovan mtDNA is present in Layers 3 and 2, it is tenuous to associate them to their corresponding depositional ages (~30–50 ka), given the reworked nature of the layers. Therefore, whether the BKC Denisovans had survived until the arrival of modern humans on the Tibetan Plateau by 30–40 ka (*22*) remains an open question.

In conclusion, the stratigraphic, chronological and sedimentary DNA results presented show that Denisovans occupied BKC at ~100 and ~60 ka ago. This confirms that Denisovans were widely distributed in Asia during the late Pleistocene. Together with the older Xiahe mandible, it suggests that they had presumably adapted to the high altitude environments on the Tibetan Plateau over a long period of time. It is tempting to speculate that the genetic adaptations to high altitudes seen in modern Tibetans and perhaps associated with a haplotype introgressed from Denisovans (*23*) may have evolved during the extended occupation of this high altitude environment by Denisovans. Deeper investigations at BKC and other Paleolithic sites in a broader region surrounding the Tibetan Plateau may help to understand the relationship and evolution of Denisovans, modern humans and possible other archaic humans in East Asia.

References and Notes:

1. J. Krause *et al.*, The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature* **464**, 894-897 (2010).
2. D. Reich *et al.*, Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053-1060 (2010).
3. M. Meyer *et al.*, A high-coverage genome sequence from an archaic Denisovan individual. *Science* **338**, 222-226 (2012).
4. K. Prufer *et al.*, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43-49 (2014).
5. D. Reich *et al.*, Denisova Admixture and the First Modern Human Dispersals into Southeast Asia and Oceania. *The American Journal of Human Genetics* **89**, 516-528 (2011).
6. P. Qin, M. Stoneking, Denisovan Ancestry in East Eurasian and Native American Populations. *Molecular Biology and Evolution* **32**, 2665-2674 (2015).
7. S. Sankararaman, S. Mallick, N. Patterson, D. Reich, The Combined Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans. *Curr Biol* **26**, 1241-1247 (2016).

8. B. Vernot *et al.*, Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235-239 (2016).
9. S. R. Browning, B. L. Browning, Y. Zhou, S. Tucci, J. M. Akey, Analysis of Human Sequence Data Reveals Two Pulses of Archaic Denisovan Admixture. *Cell* **173**, 53-+ (2018).
10. S. Sawyer *et al.*, Nuclear and mitochondrial DNA sequences from two Denisovan individuals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15696-15700 (2015).
11. V. Slon *et al.*, A fourth Denisovan individual. *Sci. Adv.* **3**, e1700186 (2017).
12. T. B. Viola *et al.*, in *American Association of Physical Anthropologists annual meeting*. (Cleveland, 2019).
13. F. Chen *et al.*, A late Middle Pleistocene Denisovan mandible from the Tibetan Plateau. *Nature* **569**, 409-412 (2019).
14. Materials and methods are available as supplementary materials at the Science website.
15. V. Slon *et al.*, Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605-608 (2017).
16. Q. Fu *et al.*, DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2223-2227 (2013).
17. G. Zong, W. Chen, X. Huang, Q. Xu, *Cenozoic Mammals and Environment of Hengduan Mountains Region*. (Beijing: China Ocean Press, 1996).
18. M. Meyer *et al.*, A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* **505**, 403-406 (2014).
19. M. Meyer *et al.*, Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature* **531**, 504-507 (2016).
20. K. Douka *et al.*, Age estimates for hominin fossils and the onset of the Upper Palaeolithic at Denisova Cave. *Nature* **565**, 640-644 (2019).
21. Z. Jacobs *et al.*, Timing of archaic hominin occupation of Denisova Cave in southern Siberia. *Nature* **565**, 594-599 (2019).
22. X. Zhang *et al.*, The earliest human occupation of the high-altitude Tibetan Plateau 40 thousand to 30 thousand years ago. *Science* **362**, 1049-1051 (2018).
23. E. Huerta-Sanchez *et al.*, Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194-197 (2014).
24. B. I. G. D. C. Members, Database Resources of the BIG Data Center in 2018. *Nucleic. Acids. Res.* **46**, D14-D20 (2018).
25. G. Miehe *et al.*, How old is the human footprint in the world's largest alpine ecosystem? A review of multiproxy records from the Tibetan Plateau from the ecologists' viewpoint. *Quaternary Science Reviews* **86**, 190-209 (2014).
26. Y. Xie, *Atlas of quartz sand surface textural features of China micrographs*. (China Ocean Press, Beijing, 1984).
27. C. B. Ramsey, T. Higham, A. Bowles, R. Hedges, Improvements to the pretreatment of bone at Oxford. *Radiocarbon* **46**, 155-163 (2004).
28. F. Brock, T. Higham, P. Ditchfield, C. B. Ramsey, Current pretreatment methods for AMS radiocarbon dating at the Oxford Radiocarbon Accelerator Unit. *Radiocarbon* **52**, 103-112 (2010).
29. P. J. Reimer *et al.*, IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years cal BP. *Radiocarbon* **55**, 1869-1887 (2013).
30. M. J. Aitken, *Thermoluminescence Dating*. (Academic press, London, 1985).
31. M. J. Aitken, *An Introduction to Optical Dating*. (Oxford University Press, Oxford, 1998).

32. D. J. Huntley, D. I. Godfreysmith, M. L. W. Thewalt, Optical dating of sediments. *Nature* **313**, 105-107 (1985).
33. R. G. Roberts *et al.*, Optical dating in archaeology: thirty years in retrospect and grand challenges for the future. *Journal of Archaeological Science* **56**, 41-60 (2015).
34. G. A. T. Duller, Luminescence dating of Quaternary sediments: recent advances. *J Quaternary Sci* **19**, 183-192 (2004).
35. F. Preusser *et al.*, Luminescence dating: basics, methods and applications. *Quaternary Science Journal* **57**, 95-149 (2008).
36. A. G. Wintle, Luminescence dating: where it has been and where it is going. *Boreas* **37**, 471-482 (2008).
37. N. Porat, G. Faerstein, A. Medialdea, A. Murray, Re-examination of common extraction and purification methods of quartz and feldspar for luminescence dating. *Ancient TL* **33**, 22-30 (2015).
38. L. Bøtter-Jensen, C. E. Andersen, G. A. T. Duller, A. S. Murray, Developments in radiation, stimulation and observation facilities in luminescence measurements. *Radiation Measurements* **37**, 535-541 (2003).
39. M. Ballarini, A. G. Wintle, J. Wallinga, Spatial variation of dose rate from beta sources as measured using single grains. *Ancient TL* **24**, 1-8 (2006).
40. L. Bøtter-Jensen, V. Mejdahl, Assessment of beta-dose-rate using a GM multicounter system. *Nuclear Tracks and Radiation Measurements* **14**, 187-191 (1988).
41. Z. Jacobs, R. G. Roberts, An improved single grain OSL chronology for the sedimentary deposits from Diepkloof Rockshelter, Western Cape, South Africa. *Journal of Archaeological Science* **63**, 175-192 (2015).
42. B. J. Brennan, Beta doses to spherical grains. *Radiation Measurements* **37**, 299-303 (2003).
43. R. K. Smedley, G. A. T. Duller, N. J. G. Pearce, H. M. Roberts, Determining the K-content of single-grains of feldspar for luminescence dating. *Radiation Measurements* **47**, 790-796 (2012).
44. X. Rui *et al.*, Luminescence chronology of the Palaeolithic–Neolithic transition in the Yujiagou site at the Nihewan Basin, northern China. *J Quaternary Sci* **34**, 125-137 (2019).
45. D. J. Huntley, M. R. Baril, The K content of the K-feldspars being measured in optical dating or in thermoluminescence dating. *Ancient TL* **15**, 11-13 (1997).
46. H. Zhao, S. Li, Internal dose rate to K-feldspar grains from radioactive elements other than potassium. *Radiation Measurements* **40**, 84-93 (2005).
47. B. Li, S. Li, A. G. Wintle, H. Zhao, Isochron measurements of naturally irradiated K-feldspar grains. *Radiation Measurements* **42**, 1315-1327 (2007).
48. D. J. Huntley, R. G. V. Hancock, The Rb contents of the K-feldspars being measured in optical dating. *Ancient TL* **19**, 43-46 (2001).
49. E. Rhodes, J.-L. Schwenninger, Dose rates and radioisotope concentrations in the concrete calibration blocks at Oxford. *Ancient TL* **25**, 5-8 (2007).
50. N. Mercier, C. Falguères, Field gamma dose-rate measurement with a NaI(Tl) detector: re-evaluation of the "threshold" technique. *Ancient TL* **1**, (2007).
51. R. Roberts *et al.*, Optical and radiocarbon dating at Jinmium rock shelter in northern Australia. *Nature* **393**, 358-362 (1998).
52. G. A. T. Duller, L. Bøtter-Jensen, A. S. Murray, A. J. Truscott, Single grain laser luminescence (SGLL) measurements using a novel automated reader. *Nucl Instrum Meth B* **155**, 506-514 (1999).
53. Z. Jacobs, G. A. T. Duller, A. G. Wintle, Interpretation of single grain De distributions and calculation of De. *Radiation Measurements* **41**, 264-277 (2006).

54. Z. Jacobs, R. G. Roberts, Advances in optically stimulated luminescence dating of individual grains of quartz from archeological deposits. *Evol. Anthropol.* **16**, 210-223 (2007).
55. G. A. T. Duller, Single-grain optical dating of Quaternary sediments: why aliquot size matters in luminescence dating. *Boreas* **37**, 589-612 (2008).
56. J. M. Olley, G. G. Caitcheon, R. G. Roberts, The origin of dose distributions in fluvial sediments, and the prospect of dating single grains from fluvial deposits using optically stimulated luminescence. *Radiation Measurements* **30**, 207-217 (1999).
57. M. D. Bateman *et al.*, Detecting post-depositional sediment disturbance in sandy deposits using optical luminescence. *Quat. Geochronol.* **2**, 57-64 (2007).
58. C. Tribolo, N. Mercier, M. Rasse, S. Soriano, E. Huysecom, Kobo 1 and L'Abri aux Vaches (Mali, West Africa): Two case studies for the optical dating of bioturbated sediments. *Quat. Geochronol.* **5**, 317-323 (2010).
59. Z. Jacobs, G. A. T. Duller, A. G. Wintle, Optical dating of dune sand from Blombos Cave, South Africa: II - single grain data. *J. Hum. Evol.* **44**, 613-625 (2003).
60. Z. Jacobs, E. H. Hayes, R. G. Roberts, R. F. Galbraith, C. S. Henshilwood, An improved OSL chronology for the Still Bay layers at Blombos Cave, South Africa: further tests of single-grain dating procedures and a re-evaluation of the timing of the Still Bay industry across southern Africa. *Journal of Archaeological Science* **40**, 579-594 (2013).
61. J. K. Feathers, V. T. Holliday, D. J. Meltzer, Optically stimulated luminescence dating of Southern High Plains archaeological sites. *Journal of Archaeological Science* **33**, 1651-1665 (2006).
62. K. J. Thomsen, A. S. Murray, M. Jain, L. Botter-Jensen, Laboratory fading rates of various luminescence signals from feldspar-rich sediment extracts. *Radiation Measurements* **43**, 1474-1486 (2008).
63. N. Blegen *et al.*, Distal tephras of the eastern Lake Victoria basin, equatorial East Africa: correlations, chronology and a context for early modern humans. *Quaternary Science Reviews* **122**, 89-111 (2015).
64. B. Li, S. Li, A reply to the comments by Thomsen et al. on "Luminescence dating of K-feldspar from sediments: A protocol without anomalous fading correction". *Quat. Geochronol.* **8**, 49-51 (2012).
65. R. Sohbati *et al.*, Luminescence dating of Pleistocene alluvial sediments affected by the Alhama de Murcia fault (eastern Betics, Spain) - a comparison between OSL, IRSL and post-IR IRSL ages. *Boreas* **41**, 250-262 (2012).
66. J. P. Buylaert *et al.*, A robust feldspar luminescence dating method for Middle and Late P leistocene sediments. *Boreas* **41**, 435-451 (2012).
67. X. Rui, B. Li, Y. Guo, The effect of residual signal on dose measurements using MET-pIRIR signals from K-feldspar. *Quat. Geochronol.* **58**, 101065 (2020).
68. R. F. Galbraith, R. G. Roberts, G. M. Laslett, H. Yoshida, J. M. Olley, Optical dating of single and multiple grains of quartz from Jinmium rock shelter, northern Australia, part 1, Experimental design and statistical models. *Archaeometry* **41**, 339-364 (1999).
69. B. Li, R. G. Roberts, Z. Jacobs, On the dose dependency of the bleachable and non-bleachable components of IRSL from K-feldspar: Improved procedures for luminescence dating of Quaternary sediments. *Quat. Geochronol.* **17**, 1-13 (2013).
70. J. P. Buylaert *et al.*, IRSL and post-IR IRSL residual doses recorded in modern dust samples from the Chinese Loess Plateau. *Geochronometria* **38**, 432-440 (2011).
71. B. Li, Z. Jacobs, R. G. Roberts, R. Galbraith, J. Peng, Variability in quartz OSL signals caused by measurement uncertainties: Problems and solutions. *Quat. Geochronol.* **41**, 11-25 (2017).

72. B. Li, Z. Jacobs, R. G. Roberts, S. Li, Single-grain dating of potassium-rich feldspar grains: Towards a global standardised growth curve for the post-IR IRSL signal. *Quat. Geochronol.* **45**, 23-36 (2018).
73. B. Li, R. G. Roberts, Z. Jacobs, S. Li, Y. Guo, Construction of a 'global standardised growth curve' (gSGC) for infrared stimulated luminescence dating of K-feldspar. *Quat. Geochronol.* **27**, 119-130 (2015).
74. H. G. Balian, N. W. Eddy, Figure-of-merit (FOM), an improved criterion over the normalized chi-squared test for assessing goodness-of-fit of gammaray spectral peaks. *Nucl. Instrum. Methods* **145**, 389–395 (1977).
75. J. Peng, V. Pagonis, B. Li, On the intrinsic accuracy and precision of the standardised growth curve (SGC) and global-SGC (gSGC) methods for equivalent dose determination: A simulation study. *Radiation Measurements* **94**, 53-64 (2016).
76. J. Peng, B. Li, Single-aliquot regenerative-dose (SAR) and standardised growth curve (SGC) equivalent dose determination in a batch model using the R Package 'numOSL'. *Ancient TL* **35**, 32-53 (2017).
77. R Core Team, R: A language and environment for statistical computing. Vienna, Austria. (2016).
78. J. Peng, Z. Dong, F. Han, H. Long, X. Liu, R package numOSL: numeric routines for optically stimulated luminescence dating. *Ancient TL* **31**, 41-48 (2013).
79. B. Guralnik *et al.*, Radiation-induced growth and isothermal decay of infrared-stimulated luminescence from feldspar. *Radiation Measurements* **81**, 224-231 (2015).
80. B. Li, R. G. Roberts, Z. Jacobs, S. Li, Potential of establishing a 'global standardised growth curve' (gSGC) for optical dating of quartz from sediments. *Quat. Geochronol.* **27**, 94-104 (2015).
81. A. S. Murray, A. G. Wintle, Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiation Measurements* **32**, 57-73 (2000).
82. R. G. Roberts, R. F. Galbraith, J. M. Olley, H. Yoshida, G. M. Laslett, Optical dating of single and multiple grains of quartz from jinmium rock shelter, northern Australia, part 2, Results and implications. *Archaeometry* **41**, 365-395 (1999).
83. G. A. T. J. R. M. Duller, Distinguishing quartz and feldspar in single grain luminescence measurements. *Radiation Measurements* **37**, 161-165 (2003).
84. B. Li, Z. Jacobs, R. G. Roberts, Investigation of the applicability of standardised growth curves for OSL dating of quartz from Haua Fteah cave, Libya. *Quat. Geochronol.* **35**, 1-15 (2016).
85. L. A. Gliganic, T. J. Cohen, M. Slack, J. K. Feathers, Sediment mixing in aeolian sandsheets identified and quantified using single-grain optically stimulated luminescence. *Quat. Geochronol.* **32**, 53-66 (2016).
86. G. A. T. Duller, Improving the accuracy and precision of equivalent doses determined using the optically stimulated luminescence signal from single grains of quartz. *Radiation Measurements* **47**, 770-777 (2012).
87. L. A. Gliganic, Z. Jacobs, R. G. Roberts, Luminescence characteristics and dose distributions for quartz and feldspar grains from Mumba rockshelter, Tanzania. *Archaeol Anthropol Sci* **4**, 115-135 (2012).
88. K. J. Thomsen *et al.*, Testing single-grain quartz OSL methods using sediment samples with independent age control from the Bordes-Fitte rockshelter (Roches d'Abilly site, Central France). *Quat. Geochronol.* **31**, 77-96 (2016).
89. Y. Guo *et al.*, New ages for the Upper Palaeolithic site of Xibaimaying in the Nihewan Basin, northern China: implications for small-tool and microblade industries in north-east Asia during Marine Isotope Stages 2 and 3. *J Quaternary Sci* **32**, 540-552 (2017).

90. B. Li, Z. Jacobs, R. G. Roberts, Validation of the LnTn method for De determination in optical dating of K-feldspar and quartz. *Quat. Geochronol.* **58**, 101066 (2020).
91. R. F. Galbraith, R. G. Roberts, Statistical aspects of equivalent dose and error calculation and display in OSL dating: An overview and some recommendations. *Quat. Geochronol.* **11**, 1-27 (2012).
92. Y. Hu, Q. Ruan, J. Liu, B. Marwick, B. Li, Luminescence chronology and lithic technology of Tianhuadong Cave, an early Upper Pleistocene Paleolithic site in southwest China. *Quat. Res.* **94**, 121-136 (2020).
93. Y. Hu *et al.*, Late Middle Pleistocene Levallois stone-tool technology in southwest China. *Nature* **565**, 82-85 (2019).
94. Z. Jacobs, B. Li, N. Jankowski, M. Soressi, Testing of a single grain OSL chronology across the Middle to Upper Palaeolithic transition at Les Cates (France). *Journal of Archaeological Science* **54**, 110-122 (2015).
95. J. M. Olley, R. G. Roberts, H. Yoshida, J. M. Bowler, Single-grain optical dating of grave-infill associated with human burials at Lake Mungo, Australia. *Quaternary Science Reviews* **25**, 2469-2474 (2006).
96. S. J. Fleming, The pre-dose technique: a new thermoluminescence dating method. *Archaeometry* **15**, 13-30 (1973).
97. B. Li, Z. Jacobs, R. G. Roberts, S. Li, Extending the age limit of luminescence dating using the dose-dependent sensitivity of MET-pIRIR signals from K-feldspar. *Quat. Geochronol.* **17**, 55-67 (2013).
98. A. P. Derevianko *et al.*, Eds., *Interdisciplinary Studies of Chagyrskaya Cave – Middle Paleolithic site of Altai*, (Institute of Archaeology and Ethnography, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 2018).
99. Y. Guo, B. Li, H. Zhao, Comparison of single-aliquot and single-grain MET-pIRIR De results for potassium feldspar samples from the Nihewan Basin, northen China. *Quat. Geochronol.* **56**, 101040 (2020).
100. P. J. Rousseeuw, C. Croux, Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association* **88**, 1273-1283 (1993).
101. P. J. Rousseeuw, M. Debruyne, S. Engelen, M. Hubert, Robustness and Outlier Detection in Chemometrics. *Critical Reviews in Analytical Chemistry* **36**, 221-242 (2006).
102. R. Powell, J. Hergt, J. Woodhead, Improving isochron calculations with robust statistics and the bootstrap. *Chem Geol* **185**, 191-204 (2002).
103. B. Li, Z. Jacobs, R. G. Roberts, S. Li, Review and assessment of the potential of post-IR IRSL dating methods to circumvent the problem of anomalous fading in feldspar luminescence. *Geochronometria* **41**, 178-201 (2014).
104. A. S. Murray, K. J. Thomsen, N. Masuda, J. P. Buylaert, M. Jain, Identifying well-bleached quartz using the different bleaching rates of quartz and feldspar luminescence signals. *Radiation Measurements* **47**, 688-695 (2012).
105. C. B. Ramsey, Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337-360 (2009).
106. C. B. Ramsey, Dealing with outliers and offsets in radiocarbon datin. *Radiocarbon* **51**, 1023-1045 (2009).
107. J. Dabney *et al.*, Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl. Acad. Sci. U.S.A.*, (2013).
108. M. T. Gansauge *et al.*, Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res* **45**, e79 (2017).

109. I. Glocke, M. Meyer, Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. *Genome research* **27**, 1230-1237 (2017).
110. M. T. Gansauge, M. Meyer, Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nature protocols* **8**, 737-748 (2013).
111. M. Kircher, S. Sawyer, M. Meyer, Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012).
112. T. Maricic, M. Whitten, S. Paabo, Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* **5**, e14004 (2010).
113. J. Dabney, M. Meyer, Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques* **52**, 87-94 (2012).
114. M. Margulies *et al.*, Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).
115. G. Renaud, U. Stenzel, J. Kelso, leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res* **42**, e141 (2014).
116. R. M. Andrews *et al.*, Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**, 147 (1999).
117. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
118. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
119. D. H. Huson, A. F. Auch, J. Qi, S. C. Schuster, MEGAN analysis of metagenomic data. *Genome Res* **17**, 377-386 (2007).
120. A. W. Briggs *et al.*, Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14616-14621 (2007).
121. S. Sawyer, J. Krause, K. Guschanski, V. Savolainen, S. Paabo, Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131 (2012).
122. M. Hajdinjak *et al.*, Reconstructing the genetic history of late Neanderthals. *Nature* **555**, 652-656 (2018).
123. Q. Fu *et al.*, A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* **23**, 553-559 (2013).
124. F. Ronquist, J. P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574 (2003).
125. D. Darriba, G. L. Taboada, R. Doallo, D. Posada, jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* **9**, 772 (2012).
126. S. Kreutzer *et al.*, Introducing an R package for luminescence dating analysis. *Ancient TL* **30**, 1-8 (2012).
127. E. Garrison, *Techniques in Archaeological Geology*. Natural Science in Archaeology (Springer, ed. second edition, 2016), pp. 345.

Acknowledgments: We thank G. Dong, Z. Wang, J. Brantingham, D. Rhode for participating in early investigation of BKC, Z. Li and D. Lin for help in the interpretation of sedimentary data, M. Qiu for taking photos of the stone artifacts and bones, S. Pang, Z. Jiang and Z. Jia for measuring the cave. We are grateful to the support of National Cultural Heritage Administration of China, Provincial Cultural Heritage Administration of Gansu, Gansu Provincial Archaeological Institute and local government for our archaeological excavation in BKC. We are also grateful to the support from the Baishiya Temple and the local residents in Ganjia town； **Funding:** This study was funded by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (2019QZKK0601) and the Strategic Priority Research Program (XDB26000000, XDA20040000) of CAS; NSFC (41771225) to D.Z.; NSFC (91731303, 41925009, 41672021, 41630102), Tencent Foundation through the EXPLORER PRIZE and "Research on the roots of Chinese civilization"of Zhengzhou University (XKZDJC202006) to Q.F.; The Strategic Innovation Fund of the Max Planck Society to S.P.; Australian Research Council Future Fellowships to B.L. (FT140100384) and Z.J. (FT150100138); **Author contributions:** D.Z. and F.C. designed the study. D.Z., F.C., H.X., T.C., X.S., J.W., J.Y., J.Y., D.B.M., C.P., Y.H., and X.C. carried out field investigation, excavated the site, and carried out sampling processes. D.Z., F.C., H.X., B.L., D.B.M. and C.P. conducted stratigraphic and taphonomy analysis. B.L., Z.J., H.X. and T.C. performed the OSL and radiocarbon dating. Q.F., R.Y., Q.D., X.F., P.C., W.P., F.L. performed the ancient DNA experiments and analysis. S.P., M.M, J.K., V.S., D.M. helped for discussing and interpreting the genetic data. D.Z., Q.F., S.P., B.L., M.Y., Z.J. and F.C. wrote the paper with contributions of all authors; **Competing interests:** The authors declare no competing interests; **Data and materials availability:** All relevant data are available in the main text or the accompanying Supplementary Materials. The new mitochondrial consensus files reported in this paper have been deposited in the Genome Warehouse in National Genomics Data Center (*24*), Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences, under accession number PRJCA002765 that is publicly accessible at https://bigd.big.ac.cn/gwh. Artifacts and animal fossils referred to in this study are curated in Lanzhou University.

List of Supplementary Materials:

Materials and Methods

Figs. S1 to S29

Tables S1 to S16

Online excel files Tables S17 to S19

References 25-127

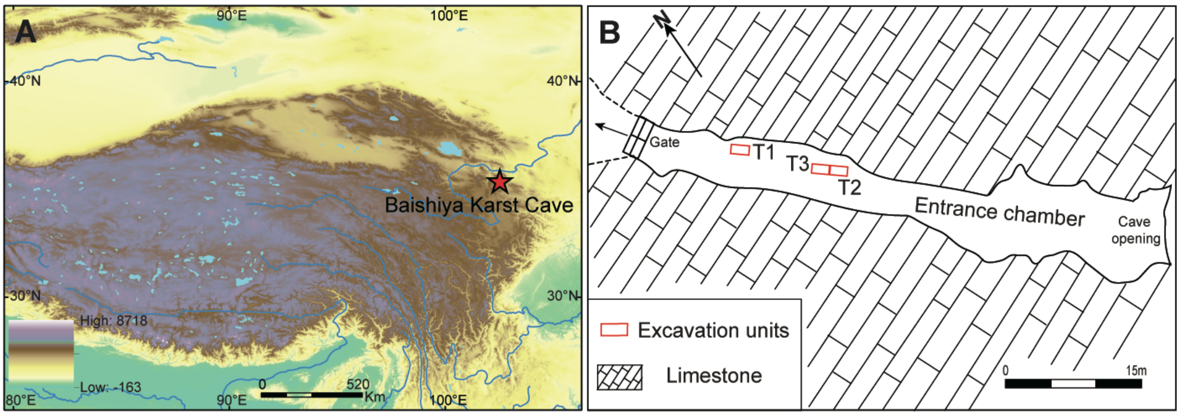

Fig. 1. Location of Baishiya Karst Cave on the Tibetan Plateau. (A) Regional maps showing the location of the site. (B) Plan view of the entrance chamber and locations of excavation units (T1, T2 and T3). The gate separates the entrance chamber from other chambers further inside the cave**.**

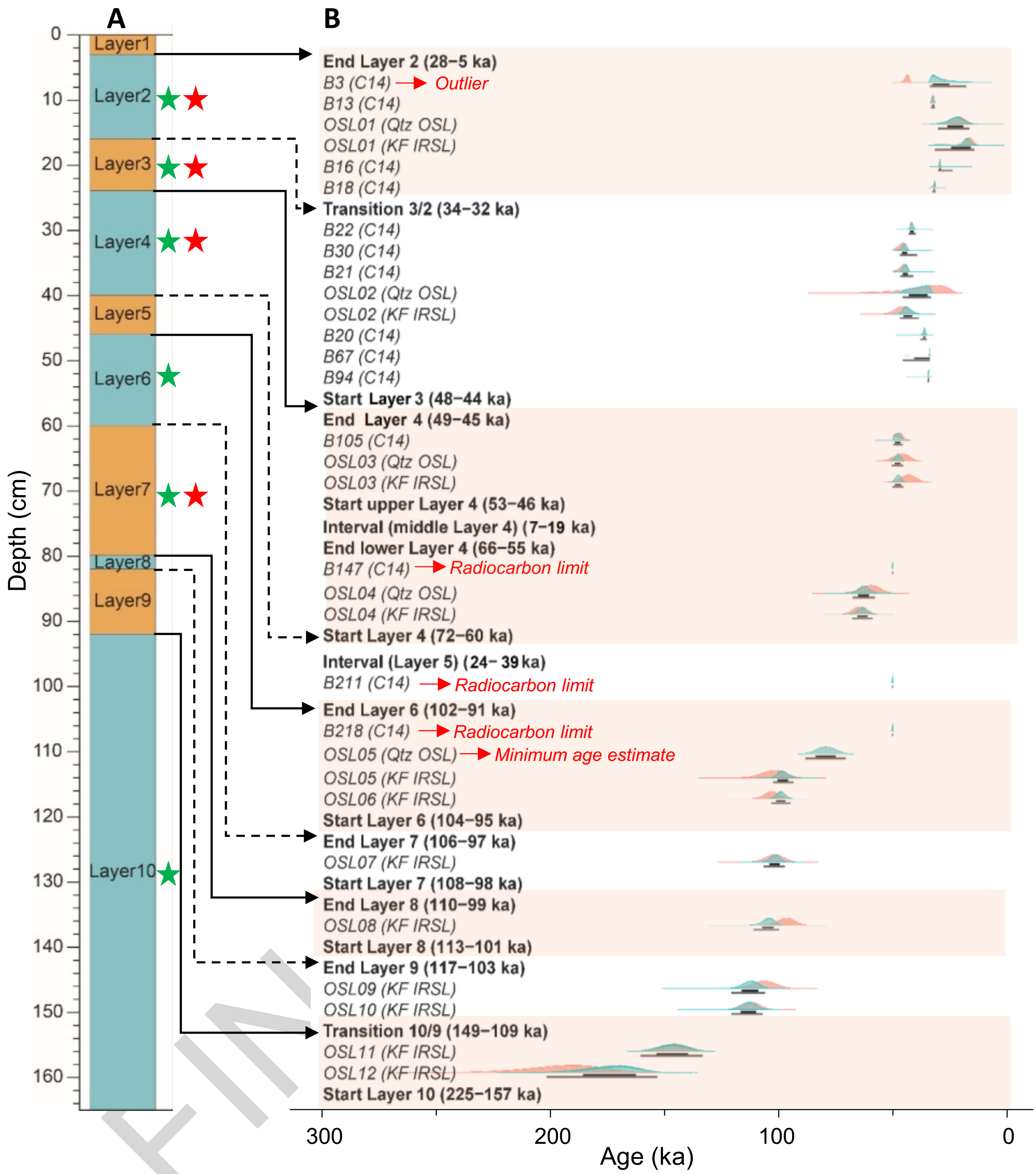

Fig. 2. Stratigraphy and dating results of T2. (A) Composite schematic stratigraphy of excavation area T2. The alternating colors are for illustration purposes only. The positions of sedimentary DNA samples from which Denisovan and animal DNA were found are shown as red stars and green stars, respectively. (B) Bayesian modelling results for all radiocarbon and optical ages. Red probability distributions represent the unmodelled ages (likelihoods) and green distributions represent the modelled ages (posterior probabilities). The narrow and wide bars beneath each distribution represent the 68.2% and 95.4% probability ranges of the

modelled ages. Start and end ages have been modelled for each layer and phase, with age ranges (95.4% confidence interval, random-only errors) given in parentheses.

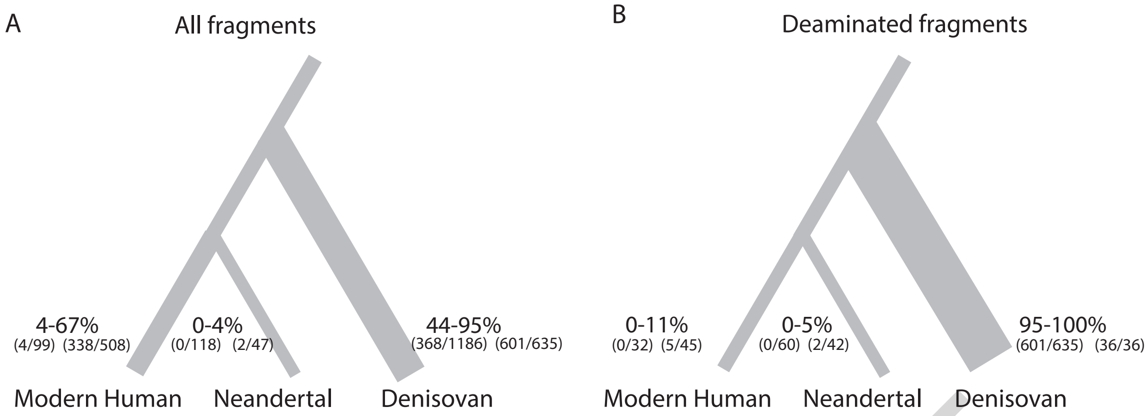

Fig. 3. Lineage inferences using modern human, Neandertal and Denisovan branch-specific substitutions for all fragments (A) and deaminated fragments (B) from Layer 4. Ranges for the percentage of lineage-matching sites for all libraries from Layer 4 are given. The fractions give the absolute number of sequenced fragments carrying derived lineage-specific alleles over the total number of fragments covering positions where such alleles occur. The lineage inferences for Layers 2, 3 and 7 are in Fig. S28.

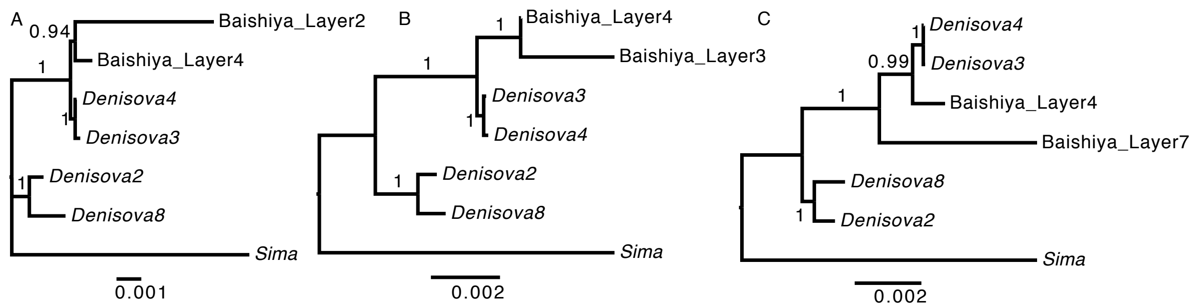

Fig. 4. MtDNA phylogenetic trees for sediment samples from Layers 4 and 2 (A), Layers 4 and 3 (B) or Layers 4 and 7 (C) of the Baishiya Karst Cave as well as mtDNA from four Denisovans from Denisova Cave and a ~430,000- year-old hominin from Sima de los Huesos in Spain. Consensus sequences with deaminated fragments were used for BKC samples, except for Layer 2 mtDNA, which is from all fragments from low contamination libraries and deaminated fragments from potentially contaminated libraries (see "decision" column in Table S11). The phylogeny was estimated with a Bayesian approach under a GTR+Gamma model of sequence evolution.

Science
AAAS

Supplementary Materials for

Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau

Dongju Zhang*, Huan Xia, Fahu Chen, Bo Li*, Viviane Slon, Ting Cheng, Ruowei Yang, Zenobia Jacobs, Qingyan Dai, Diyendo Massilani, Xuke Shen, Jian Wang, Xiaotian Feng, Peng Cao, Melinda A. Yang, Juanting Yao, Jishuai Yang, David B. Madsen, Yuanyuan Han, Wanjing Ping, Feng Liu, Charles Perreault, Xiaoshan Chen, Matthias Meyer, Janet Kelso, Svante Pääbo* and Qiaomei Fu*

Correspondence to: djzhang@lzu.edu.cn; bli@uow.edu.au; paabo@eva.mpg.de; fuqiaomei@ivpp.ac.cn

This PDF file includes:

Materials and Methods

Location and setting of BKC

Baishiya Karst Cave (BKC) (35.45 N, 102.57 E, 3,280 meters above sea level [masl]) is located in the northern Ganjia Basin on the northeastern margin of the Tibetan Plateau (Fig. 1A). The Yangqu River flows through the basin from west to east, with elevations ranging from as high as 4,636 masl at the top of Dalijia Mountain on the northern margin of the basin to a low of 2,900 masl at the Yangqu River's confluence with the Daxia River, the first order tributary of the Yellow River. The cave lies ~30 m above a small tributary (Jiangla Gully) of the Yangqu River, which cuts through a vertically tilted limestone massif locally termed Dalijia Mountain (Fig. S1A). Loess is widely distributed throughout this region on the mountains and in the basin. The cave is named after the white limestone cliff in which it is formed; "Baishiya" means "white cliff" in Chinese. BKC is as long as 300 m based on our explorations, but perhaps much longer according to local informants. We explored six chambers, including the entrance chamber where the archaeological excavations, presented in this study, were conducted. The entrance chamber is ~60 m long, 8 m wide, 5 m high (Fig. 1B, Figs. S1B and S1C) and relatively dry, but dripping water and small puddles occur in some of the other chambers.

Average annual precipitation in this region is ~350 mm and the average annual temperature is ~4.7°C. In January, the coldest month, the regional temperatures average -8.3°C, and during July, the warmest month, ~13.7°C. The temperature inside the cave remains relatively constant, e.g., it remains at 8–10°C during December and January when the outside can be as low as -18°C. The back of the entrance chamber is substantially warmer than the front part; the other chambers may be even warmer.

At present, BKC and the foothills of Dalijia Mountain are dominated by alpine meadow (Fig. S1A), composed of a variety of grasses, sedges and herbs. These vegetation communities have been greatly disturbed by grazing and may have been substantially different in the past, particularly during shifting climatic regimes (*25*). Riparian brush communities, composed primarily of *Prunus sp.*, *Salix sp.*, and *Berberidaceae sp.*, are present along the small stream in front of BKC. Isolated refugia of conifers, such as pine and cypress, are present on the steeper slopes of Dalijia Mountain.

BKC is a holy cave to Tibetan Buddhists and is managed by the local Baishiya Temple. Monks from the temple and local Tibetans use the cave for praying, meditation and other ritual purposes (Fig. S1B). The cave is also a pilgrimage and tourist destination, heavily visited during the summer months.

BKC has been known to archaeologists from Lanzhou University since 2005, when the Xiahe mandible arrived in Lanzhou University and where it has since been curated. A local monk reportedly found the mandible inside BKC several decades earlier. Since 2011 an archaeology team led by Lanzhou University has explored BKC and the Ganjia Basin where many other karstic caves are also found. During these field investigations, most of the caves were surveyed, but very few stone artifacts were found. During a survey in 2016, the team found indisputable stone artifacts in BKC, providing further evidence for ancient human occupation of the cave, in addition to the Xiahe mandible. To better understand the archaeological context of prehistoric human occupation in BKC, Lanzhou University applied for permission and permits to conduct the first ever excavations inside the cave in 2018 and 2019.

<u>Excavation and sampling procedures</u>

Three 1 x 2 m^2 excavation units (T1, T2 and T3) were established in the back half of the entrance chamber (Fig. 1B and Fig. S1C) in 2018 (T1 and T2) and 2019 (T3). The upper 60 cm of deposit in T1, located ~48 m from the cave entrance, is mixed, containing stone artifacts, bone fragments and modern artifacts. No cultural remains were found below a depth of 60 cm. Excavations stopped when a thick flowstone and rock were encountered at a depth of ~1 m. T2, located ~37 m from the cave entrance, preserves secure, intact and thicker cultural deposits. During the 2018 excavations, 10 cultural layers were identified down to a depth of 165 cm; bedrock was not reached. T3 and T2 are connected to each other, sharing the same set of stratigraphic layers (Fig. S2). Since detailed work on T3 is still ongoing, this study focusses on the data obtained from the 2018 excavation (mainly from T2).

Sediments in T2 were excavated in 5 cm-thick standard levels, within natural stratigraphic layers where possible. All identifiable artifacts, bone fragments and charcoal pieces were measured and collected separately, and their proveniences were recorded using a total station. After excavation, a column (~20 cm wide) from the southwest wall of T2 (Fig. S2A) was cleaned to collect sediment samples for multiple analyses, including all the sediment DNA samples, most of the OSL samples (OSL1–OSL8 and OSL11) and serial samples (in 5 cm intervals within each layer) for grain size, carbonate content, X-ray diffraction (XRD), clast morphology and scanning electron microscopy (SEM) analyses. The preliminary test excavations in 2018 was conducted over a short time period. Further detailed excavations and analyses, including sediment micromorphology still needs to be done. Eight sedimentary DNA samples were collected from the middle of each layer, except Layers 1 and 5 (Figs. S2 and S3). Sterile masks and gloves were worn during sample collection to avoid contamination by modern DNA. Twelve samples were collected for optical dating, including ten from Layers 2–4, 6–8 and 10 in the southwest wall and one each from Layer 9 in the northwest and northeast walls (Fig. S3).

<u>Stratigraphic and sedimentary features of T2</u>

Stratigraphy

The 165 cm deep exposed deposits in T2 were divided into Layers 1–10 (top to bottom), based on sedimentary texture, composition, color and inclusions (Figs. S2 and S3). Most of the layers are slightly inclined towards the cave entrance. Cultural materials, including stone artifacts and bone fragments were found in each of the layers. A bowl-shaped pit (H1) of historical origin was found in the eastern part of T2 (Figs. S2 and S3). This pit reached as deep as Layer 10 and was overlain by Layer 1. It contained a few historical period pottery sherds, and a large number of out-of-context stone artifacts and animal bones, probably derived from digging and refilling of the pit. Radiocarbon dating of a bone fragment from H1 yielded an age of ~700 years (Fig. 2 and Table S3), confirming that this pit was created very recently by human activities. Another bowl-shaped pit (H3) was found in the western part of T3 (Fig. S2) that also contain a few historical period pottery sherds and out-of-context stone artifacts and animal bones. No radiocarbon date has been obtained from this pit, but the presence of pottery sherds indicate that it was likely also dug during the historic period.

Detailed descriptions of each of Layers 1–10 are provided in Table S1. As shown in Fig. S2B, each strata can be readily distinguished. Most of the layer boundaries are visually abrupt and clear, except those between Layers 2 and 3 and Layers 9 and 10, which appear to be gradual. The clear and sharp boundaries between most of the layers below Layer 3 indicate minimal post-depositional disturbances between layers. The single-grain optical dating results (see optical dating section below) suggest some movement of younger fine-grained sediments into lower, older sediments in the upper stratigraphic layers, but this downward movement of fine-grained sediments is not noticeable below Layer 3. Layer 5 contains small gravels, rounded clay aggregates, slightly rounded calcite crystal fragments and limestone fragments (Fig. S4A and B, S7E), which may be indicative of a fluvial environment within the karstic cave system.

Except Layer 1, most layers are generally compact, especially Layer 4 and those below. Most of the layers are laterally continuous, except Layer 3 which appears to be discontinuous in the northeastern part of T2 (Fig. S3). Unconformities between some layers (due to erosional events or depositional gaps) can be observed, i.e., Layer 3 appears to truncate Layer 4 in T3 (Fig. S2B). A discontinuous flowstone layer (1–2 cm thick) was found at the base of Layer 3 in the western part of T2 and T3 (Fig. S4C). This flowstone layer is weakly crystalized but highly cemented with clay and silts (Fig. S4D). A lamella structure could be seen on the cross-section of the flowstone (Fig. S4E). Microscopic observation (Fig. S4F) and chemical analysis indicate that it is mainly composed of carbonate particles (Measured carbonate content is ~62%).

Sedimentology

During the 2018 excavation season, bulk sediment samples (n = 29) were collected at 5 cm intervals in a column from T2 (Fig. S2A). A sub-sample (~0.35 g) of the fine sediments, defined here as <1 mm in diameter, were measured with a Malvern 2000 laser particle size analyzer after organic matter and carbonates were removed with 30% H_2O_2 and 10% HCl, respectively. A different sub-sample (0.2 g) of the fine sediment from each sample were ground to a fine powder and their carbonate contents were measured using a GMY-3A Carbonate Content Automatic Tester (China, Hainan Oil Scientific Research Apparatus Company). Twelve samples, one each from Layers 1–8 and two each from Layers 9 and 10, were selected for sieving, and XRD analysis. The bulk samples (varies from ~8g to ~300 g) were first dry-sieved using a 10-mm diameter mesh to separate larger clasts (>10 mm) from smaller clasts (1–10 mm) and fine sediments (<1 mm). About 10 g of the <10 mm diameter size fraction was then wet-sieved using a series of sieves with mesh sizes of 1 mm, 2 mm, 4 mm, 6 mm, 8 mm and 10 mm diameter. The fine sediment fractions (<1 mm) of these 12 samples were then sub-sampled (~2 g), ground and measured for XRD analysis using a X' Pert Pro MPD (Holland, PANalytical). Sub-samples for 6 of these samples (one each from Layers 3–8) were also submerged in 30% H_2O_2 and 10% HCl to remove organic matter and carbonates, respectively, for collection of SEM images using an Apreo S (America, Fei).

Sieving data shows that clasts >10 mm make up more than 70% of the total weight of each sample, except for the sample from Layer 8 (Fig. S5G). The larger clasts (>10 mm) are made up predominantly of limestone fragments, lithic debris, calcite crystals, calcareous nodules, bone fragments, pebbles and cobbles. Calcite crystals, including rhombohedral, scalenohedral and prismatic forms, dominate this clast fraction (Fig. S4G). Most large cobbles (~7 cm) and few pebbles (~4 cm) are well-rounded metamorphic quartz sandstones and hornstones; these are lithologically consistent with the stone artifacts found in the cave and the gravels found in the riverbed at the front of the cave. This may suggest that these cobbles and pebbles were brought into the cave by hominins. The rest of the >10 mm fraction is poorly sorted and angular (Fig. S7). The <10 mm fraction is mainly composed of fine sediments (<1 mm) and clasts (6–10 mm), with a small component of clasts in the size range of 1–6 mm (Fig. S5A-F). The coarser particles (1–10 mm) mainly consist of calcite crystals with a small amount of lithic debris, bone fragments, calcareous nodules and limestone fragments, all of which are angular and poorly sorted (Fig. S6).

Grain size frequency distributions of the fine sediment fraction (<1 mm) for all samples from the sequence show a similar pattern (Fig. S9), indicating that the deposits are poorly sorted (σ_{Φ} = 1.7–2.1) (Fig. S8B). The sediments are composed of coarse silts (~70%; 35–45 µm modal size), fine silts (~20%; ~10 µm modal size) and clays (< 10%; 1 µm modal size) (Fig. S8A and S9). This is consistent with the characteristics of typical loess in China. SEM images of randomly selected sediment grains from Layers 2–8 show that they are predominantly composed of grains with a variety of shapes, ranging from irregular shapes and dish-shaped depressions (*26*) to very angular grains with sharp edges (Fig. S10). This suggests that the fine sediment (<1 mm) in the cave might derive from a variety of provenances. The quartz grains with dish-shaped depressions are probably from loess that was washed or blown into the cave. Angular grains imply a limited transportation distance. XRD analysis indicates that the fine components of these samples are dominated by quartz, calcite and feldspar (Table S2). The carbonate content of each sample is within the range ~20–40% and show no systemic variation among different layers (Fig. S8D); these are presumably dominated by the calcite and limestone fragments in the sediments.

The stratigraphic and sedimentological evidence indicates that most of the BKC sediments are composed of a poorly sorted silt matrix that includes various amounts of angular clasts of autogenic limestone gravel that could originate from either the reworking of eroded parent bedrock by colluviation or spalling of material from the cave walls and roof. The exception is Layer 5, which appears to be a fluvial deposit. The larger sub-rounded clasts in many layers appear to be present as the result of human transport. Overall, the stratigraphic integrity of the deposits appears to be intact, except where disturbed by large and well-defined pit structures (i.e., H1 and H3).

Lithic and fauna assemblages

A total of 1,310 stone artifacts were recorded and collected from T2. Of these, 188 were from the historic H1 pit. Preliminary analyses suggest the stone artifact assemblage is dominated by simple flakes, cores and shatter (Fig. S11). Delicately retouched tools are relatively rare but many flakes are used without retouch. Cortex is common on flakes and cores, indicating a local raw material source, most likely from the riverbed in front of the cave. Simple core and flake technologies were used for flake production. Stone raw materials are dominated by metamorphic quartz sandstones and hornstones which are very common in the nearby riverbed.

A total of 579 animal bone fragments were recorded and collected from T2 (Fig. S12). Of these, 48 were from the historic H1 pit. Preliminary identifications show that the bone assemblage, except those from H1, is dominated by highly fragmented limb and axial bones, and isolated teeth. Layers 6 and above are dominated by small to middle-bodied animals, including gazelles, marmots, foxes and birds. Layer 7 and below are dominated by megafauna, such as rhinoceros, hyenas and large bovids. Cut marks and percussion marks are common, even on hyena bones. Bones of domesticated cattle and sheep/goat were found in H1, further supporting the historical age of H1.

Radiocarbon dating

Fifteen bone samples from T2 were selected for radiocarbon dating and measured with an accelerator mass spectrometer (AMS). One sample from Layer 3 and two from Layer 4 were pre-processed and the targets prepared at Lanzhou University and measured on the AMS at Peking University. Other samples, including one from H1, four from Layer 2, five from Layer 3, two from Layer 4 and one each from Layers 5 and 6, were processed and measured at the University of Oxford. Chemical pretreatment, target preparation and AMS measurements of all bone samples were based on the methods described in Ramsey *et al. (27)* and Brock *et al.* (*28*). Before target preparation, the percent carbon content and the C:N atomic weight ratio of the bone collagen were measured. Values were all within the normal range. The $\delta^{13}C$ values obtained during AMS measurements were used to correct for isotopic fractionation. The calibrated ages were generated in OxCal v.4.3 using the IntCal13 calibration curve (*29*). The results of radiocarbon dating are summarized in Table S3.

Optical dating

Optical dating determines the time elapsed since minerals (such as quartz and feldspar) were last exposed to sunlight or heat (*30-32*). This technique is based on the time-dependent accumulation of trapped charges in the crystal lattice of mineral grains as a result of ionising radiation from the decay of naturally occurring radioactive elements (mainly uranium, thorium and potassium) and cosmic rays. The radiation dose received by minerals after burial, the so-called equivalent dose (D_e), can be estimated by measuring the optically stimulated luminescence (OSL) and infrared (IR) stimulated luminescence (IRSL) signals. The burial age is estimated by dividing the D_e by the environmental dose rate; the latter can be determined from measurements of the radioactivity of the sample and the material surrounding it. Given its capability of dating beyond the upper limit of radiocarbon dating (~50 ka), it has been widely applied to estimate the burial ages for sediments from archaeological sites dating back to the last ~0.5 million years (*33-36*).

Sample collection, preparation and measurement facilities

All optical dating samples were collected by hammering stainless steel tubes (5 cm in diameter and 25 cm long) into a cleaned profile wall. Both tube-ends were sealed to avoid light exposure during transportation. Additional sediment samples (~100 g) were also collected from each sample position for radioactivity measurements and estimation of the environmental dose rate. Sample bags were tightly sealed to retain its current sediment moisture content.

All optical dating samples were prepared under subdued red light following routine laboratory procedures (*31*). Samples were first treated with HCl acid and H_2O_2 solution to remove carbonates and organic matter, respectively. They were then wet sieved to obtain a range of different sand-sized grain fractions. Grains 90–212 μm in diameter were retained, and K-feldspar and quartz extracted using sodium polytungstate solutions at three different densities (2.70, 2.62 and 2.58 g/cm^3). Quartz grains were purified using 40% HF acid for ~40 min to dissolve any remaining feldspar grains and to remove the outer layer of the grains that were irradiated by alpha particles. K-feldspar grains were etched in 10% HF for ~10 min to clean the surfaces of the grains and remove the outer alpha-irradiated layers (*37*). After HF-etching, both the quartz and feldspar grains were rinsed with 10% HCl acid to remove any precipitated fluorides. The etched grains were then dried and sieved again to remove grains <90 μm in diameter.

Both quartz and K-feldspar grains were measured on automated Risø TL-DA-20 luminescence readers equipped with focused green (532 nm) and infrared (830 nm) lasers for single-grain stimulation (*38*). Luminescences were detected using Electron Tubes Ltd 9235QA or 9107Q-AP-TTL-03 photomultiplier tubes. The UV emissions of quartz OSL signals were detected through Hoya U-340 filters. For feldspar IRSL, Schott BG-39 and Corning 7-59 filters were used to detect the violet/blue emissions. Individual grains were placed in aluminium discs, each of which were drilled with 100 holes 300 μm in diameter and 300 μm deep (*38*). Irradiations were carried out within each luminescence reader using calibrated $^{90}Sr/^{90}Y$ beta sources. The beta dose rates delivered to individual grain-hole positions were calibrated to take account of any spatial variations (*39*), using a range of known gamma-irradiated quartz standards. Solar bleaching tests were conducted using a Dr Hönle solar simulator (model: UVACUBE 400).

Environmental dose rate determination

The environmental dose rate of etched mineral grains consists of several components: alpha (internal to a grain), beta (internal and external to a grain), gamma and cosmic-ray dose rates. Since the excavation is located inside the cave chamber, ~30 m from the cave entrance (Fig. S1C), the contribution of cosmic rays are considered negligible. The internal alpha dose received by quartz and K-feldspar from U and Th inside the mineral grains are also assumed to be negligible.

For estimation of the external beta dose rate, a low-level beta counting technique was used. This involves measurement of dried, homogenised and powdered sediment samples with a Risø GM-25-5 multi-counter system (*40*); measurement and analytical details are provided in Jacobs and Roberts (*41*). The effect of grain size on attenuation of the beta dose was taken into account (*42*).

K-feldspar could have a significant internal beta dose rate contribution due to the radioactive decay of ^{40}K and ^{87}Rb inside a grain. Previous studies suggested that values of between ~10–13% can be assumed for the internal K concentration of K-rich feldspar grains (*21, 43-47*). We measured K-rich feldspar fractions from OSL3 and OSL10 with X-ray powder diffraction (XRD) (X' Pert-Pro MPD, Holland Panalytical) to check the mineralogical composition of our samples. The K-feldspar fractions were ground into powder and then measured using anode material Cu, step size 0.0167, operating voltage 40 kV and current 40 mA. The XRD results (Fig. S13) show that the K-feldspar fractions are dominated by K-rich feldspar (orthoclase, ~70%), with a smaller contribution from Na-rich feldspar (~10–15%) and quartz (~10–15%). So, we used values of 12 ± 1% and 400 ± 100 ppm for internal K and Rb concentrations (*48*), because our samples are dominated by K-rich feldspar, and we determined K-feldspar D_e values for only the brightest grains (see the next section). The pIRIR signals from Na-rich feldspar grains are usually dimmer compared to K-rich feldspar grains (*43, 45*).

Gamma rays can penetrate ~30 cm through most sediment and rock, and should ideally be measured directly for each sample to take into account any spatial heterogeneity in the gamma radiation field. This is especially important for samples collected from inhomogeneous stratigraphic units. We were only able to measure in-situ gamma dose rates for four samples (OSL3, OSL6, OSL7 and OSL8). Measurements were for 30 min with a 2 inch detector that contain a 1-inch NaI (Tl) crystal placed in the sample hole left after removal of the OSL sample tube. The detector was calibrated using the Oxford University doped concrete blocks, each with a known radioactivity (*49*). Gamma dose rates were determined using the 'threshold' technique (*50*). For the remaining samples, we estimated their gamma dose rates by constructing a gamma dose rate gradient for the sediment profile. We first obtained estimates of U, Th and K concentrations for each sample using a combination of ICP-MS (for U and Th) and ICP-OES (for K). We then assumed that samples from each layer have the same U, Th and K concentrations, and calculated the gamma dose rate contribution of individual layers to other layers using Appendix H in ref (*30*). The modelled and in-situ gamma dose rates are compared for the top 1.3 m of deposit (Fig. S14); gamma dose rates are consistent, confirming the reliability of the model and our use of the modelled gamma dose rates for the rest of the samples.

External beta and gamma dose rates were corrected for long-term water content. Water content for each sample was estimated from sediment collected in two consecutive years, but different seasons (winter and summer); identical results were obtained, suggesting that water contents are relatively stable in the cave. We used the current measured field water contents of all samples, which ranged from 5–21%, and assigned a ± 25% relative error at 1σ to capture any likely fluctuations over the burial period. The environmental dose rate data for all samples are provided in Table S4.

D_e measurements

Both quartz and K-feldspar grains were measured for the uppermost five samples (OSL1–OSL5) collected from Layers 2–4 and 6. For the deeper layers (Layers 7–10), only K-feldspar grains were measured; all quartz samples were in dose saturation. For D_e determination, we measured individual K-feldspar grains and small aliquots of quartz. Using a microsope, K-feldspar grains were individually loaded into each 300 μm diameter grain hole on a single grain disc. This was achieved using a needle tip and static electricity to pick the grain up and drop it into each hole individually. This was a painstaking approach, but essential to ensure that our data sets were based on true single-grain measurements. For quartz, we filled each grain hole on a single grain disc with several (4–8) grains, assuming a probability that only one grain will emit an OSL signal (*51, 52*). This approach allows us to optimize: 1) identification and elimination of grains that exhibit aberrant luminescence characteristics (*53-55*), 2) detection of partial bleaching of samples (*56*) and, 3) investigation of depositional and post-depositional disturbance (*57-61*). All three factors can contribute towards calculation of erroneous age estimates.

D_e determination – single-grain K-feldspar

We used a two-step post-IR IRSL (pIRIR) procedure (*62, 63*) to determine the D_e for individual K-feldspar grains (Table S5). In this procedure, all grains on a single-grain disc were first stimulated simultaneously with infrared diodes at 200°C for 200 s to minimize or remove the unstable signal affected by anomalous fading (*64*). The grains were then individually stimulated with an infrared laser at 275°C for 1.5 s to acquire their pIRIR signals used for D_e determinations. The same preheat of 320°C for 60 s was applied prior to measurement of the natural (L_n), regenerative (L_x) and test dose (T_x) signals. A fixed test dose of ~50 Gy was used for all samples. At the end of each measurement cycle, all grains on a disc were exposed to a 'hot bleach' stimulating grains with infrared diodes at 325°C for 200 s to minimise any optical transfer and carry-over of residual signals to the next measurement cycle.

One of four grain size fractions were measured for each sample: 90–125, 125–250, 90–180 or 180–250 μm. Care was taken to ensure that each hole contained only one grain. Representative single-grain pIRIR decay curves of the test dose signal (T_n) are shown for two samples (OSL3 and OSL10) in Fig. S15A and B. The pIRIR signals appear to decay to a stable and negligible level after ~1 s of optical stimulation. The net pIRIR intensity was calculated by subtracting the signal integrated over the last 0.2 s of optical stimulation from the signal integrated over the first 0.1 s. We used the net T_n intensities as indicators of brightness or sensitivity of the grains. The pIRIR T_n intensity distribution of individual grains for all samples are shown in Fig. S15C. It shows that the pIRIR T_n intensities range from several tens to thousands of counts per 0.1 s of optical stimulation time; most grains have a brightness of <10,000 cts per 0.1 s. The cumulative light sum plots for these samples are shown in Fig. S15D, demonstrating that <10% of the grains produced >80% of the total luminescence.

pIRIR performance tests

The pIRIR procedure was tested using three performance tests: residual dose, dose recovery and anomalous fading. We first conducted a residual dose test to determine whether the pIRIR signals of our samples can be reset or ‘bleached’ to a negligible level. The natural signal of 100 grains from two samples (OSL2 and OSL6) was first bleached using a solar simulator for ~8 hrs. These grains were then measured using the pIRIR procedure (Table S5) to estimate their residual doses. Both samples yielded similar results, and mean residual doses of ~8 Gy were obtained from grains that gave detectable pIRIR signals. The residual doses are also highly variable from grain to grain, ranging from zero to up to ~30 Gy (Fig. S15J). However, we decided not to subtract residual doses from the K-feldspar D_e values for age estimations, for two reasons: 1) the mean measured residual doses are relatively small compared to the D_e values for most of our samples (<5%), expect for the uppermost sample, which is highly dispersed. However, we used a minimum age model for this sample (see next section), that would bias the results to those grains with negligible residual doses; 2) previous studies suggest that residual doses are age and dose dependent, and that a modern sample usually has a negligible residual (*65-67*), indicating that either (a) the resetting of the non-bleachable signal under natural processes are more efficient than under laboratory conditions, or (b) that the residual signal is a ‘laboratory artifact’ resulting from thermal transfer or recuperation due to optical bleaching and subsequent preheating of the sample. Unfortunately, we are not able to test and verify this, because we currently have no ‘modern K-feldspar grains’ from the site, and the upper-most samples are either mixed or poorly bleached, so it is impractical to study the dose dependency of the residual signal.

For the dose recovery test, the same sample (OSL6) was used. The natural signals of 900 grains were first bleached using a solar simulator for ~8 hrs. A 200 Gy surrogate ‘natural’ dose was given to 300 grains and a 500 Gy surrogate ‘natural’ dose to 600 grains. The grains were then measured using a simplified SAR procedure, which involved only two SAR measurement cycles. The first cycle measured the surrogate ‘natural’ dose (200 or 500 Gy) and the second a regenerative-dose of the same size (200 or 500 Gy). We then calculated the ratio between the ‘natural’ (L_n/T_n) and subsequent regenerative-dose (L_1/T_1) signals. A successful dose recovery is indicated by ratios consistent with unity. Totals of 186 (200 Gy) and 314 (500 Gy) grains were accepted. The respective individual $(L_n/T_n)/(L_1/T_1)$ ratios are shown in Fig. S15K for the 500-Gy test. Weighted mean ratios of 1.06 ± 0.01 and 1.08 ± 0.01 were obtained using the central age model (CAM) (*68*), corresponding to 1.01 ± 0.01 and 1.03 ± 0.01, respectively, after correcting for the residual dose signal (0.053 ± 003) obtained from the residual test. We used the intensity-subtraction method (*69*) for residual dose correction, rather than the conventionally-adopted dose-subtraction method (*70*). The results suggest that the pIRIR procedure can recover the given dose successfully.

We also conducted an anomalous fading test to confirm that the 200°C infrared stimulation can effectively reduce the fading component to a negligible level (*64*). After D_e measurements of OSL6, 200 grains were given an additional regenerative dose of the same size as the dose used for calculation of the recycling ratio. The grains were then preheated, stored in the dark at room temperature for ~3 days prior to measurement of their pIRIR signals. The delay period corresponds to ~2.1 decades compared to the prompt measurements of the pIRIR signals during the SAR sequence. We then calculated the ratios of the pIRIR signals measured after the delay to that measured promptly during the SAR sequence. It is expected that this delayed-to-prompt signal ratio should be equal to unity if there is negligible anomalous fading of the pIRIR signal. The ratios for individual grains are displayed as a radial plot in Fig. S15L. It shows that the majority of ratios are consistent with unity at 2σ, with some of them randomly distributed around unity. The CAM ratio is 1.01 ± 0.01, suggesting that the pIRIR signal fades negligibly in our samples. For this reason, we did not make any fading correction to the calculated ages.

SGC and L_nT_n method for K-feldspar

We used the L_nT_n method (*71*) to determine D_e values for K-feldspar from our samples. This involves the establishment of a standardised growth curve (SGC) and analysis of re-normalized L_n/T_n values for individual grains. In this method, all 'saturated' grains are included, so a full and untruncated dose distribution is obtained. This allows reliable estimation of D_e values beyond the conventional limit of $2D_0$ when using standard procedures. This method is suitable for K-feldspar, because grains share a similar dose response curve (DRC) shape from which a single SGC can be established (*72, 73*). It has previously been used to date sediments from Denisova Cave (*21*).

To determine the suitability of this method for our samples, we first tested whether a SGC exists. A total of 2,700 grains from 10 samples (OSL1–OSL4 and OSL6–OSL8, OSL10–OSL12) were measured using a full SAR procedure that involved a series of regenerative doses up to 1900 Gy. A SGC must be established using only those grains that have reliable DRCs (*72*). So, we used the criteria of Li et al. (*72*) to reject grains if: (1) their initial T_n signal is <3σ above the corresponding background, or the relative standard error of the net T_n intensity is >25%; (2) the ratio of the L_x/T_x value of the zero regenerative dose to that obtained from the maximum regenerative-dose is >5%; and (3) the L_x/T_x data points are too scattered, e.g., they were associated with poor recycling ratios (>10% from unity), or have a figure-of-merit (FOM) value (*74, 75*) of >10% or a reduced chi squared (RCS) value of >5 (*76*). The selection of grains based on the above criteria was achieved using the function calSARED() provided in the R-package 'numOSL' (*77, 78*).

The numbers of grains rejected for each criteria are summarized in Table S6. A total of 1,420 grains passed all three broad criteria and were used to establish a SGC using the least-square normalization (LS-normalization) method and a general order kinetics (GOK) function (*79*). The L_x/T_x values for all accepted grains, before and after LS-normalization, are shown in Fig. S15E and F. The LS-normalization procedure significantly reduced the between-grain variation in L_x/T_x values. The ratios between the LS-normalized L_x/T_x values and expected values on the SGC are indicated in Fig. S15G. About 90% of the ratios are statistically consistent with unity at 2σ, which implies that most of the accepted grains share the same DRC. The reliability of the established SGC was further tested by comparing the D_e values for individual grains based on their full SAR DRCs with those obtained using the SGC; the latter was obtained by projecting the LS-normalized L_n/T_n ratios onto the SGC (*80*). Fig. S15H and I shows that the SAR and SGC D_e values for the same grains are statistically indistinguishable; ~98% are consistent at 2σ. This further validate the SGC curve established for our samples. We, therefore, measured the rest of the samples using a simplified SAR procedure that include only two measurement cycles, one for measuring L_n/T_n and the other for measuring an additional L_x/T_x used for normalization.

***D_e* determination – small single aliquot quartz**

All small quartz aliquots were measured using a single aliquot regenerative-dose (SAR) procedure (*81, 82*) (see Table S4). The procedure involves measuring the natural (L_n) and a series of regenerative-dose (L_x) OSL signals. A fixed test dose (~10–20 Gy) was given after the measurements of each L_n and L_x, and the induced OSL signals (T_n and T_x) were used to monitor any sensitivity change that may have occurred during the SAR cycles. A preheat of 260°C for 10 s was applied prior to measurement of the natural and regenerative dose signals, and a cutheat of 180°C was used prior to measurement of all test dose signals. Each green laser stimulation was conducted at 125°C for 1.5 s. Net OSL intensity was calculated using the integral over the first 0.1 s of optical stimulation minus a background estimated from the last 0.2 s of the signal. A duplicate regenerative dose and a zero dose were included in the procedure to check on the efficacy of the sensitivity correction and the extent of any thermal transfer of charge induced by the preheats. The OSL IR-depletion ratio test (*83*) was also applied to check for any feldspar contamination, using an infrared exposure of 40 s at 50°C.

Quartz OSL characteristics

In contrast to K-feldspar, >99% of individually-measured quartz grains emit insufficient OSL signals; individual grains either have net T_n signals that are <3σ above its corresponding background signal, or relative standard errors of the net T_n signal that are >25%. To increase the likelihood of finding grains that luminesce and increase measurement efficiency, we included up to eight 90–125 µm diameter quartz grains in each grain-hole on a standard single grain disc. This is equivalent to small single-aliquot measurements, but we measured the small aliquots in the same way we would measure single grains with a green laser beam. We measured a total of 18,800 small aliquots; 4000, 5500, 4700, 3000 and 1600 for OSL1–OSL5, respectively. This resulted in ~5–9% of aliquots from OSL1–OSL4 and ~40% from OSL5 with sufficient sensitivities from which D_e values could be estimated (Table S7). The probability of finding two bright grains in the same hole is small (<2%), so we assume that most of the measured OSL signals from these small aliquots arise from only one grain, essentially making the small-aliquot results equivalent to single-grain results (*84*).

Fig. S16A and B shows typical OSL decay curves for selected aliquots from OSL1 and OSL3. A range of OSL decayed rates was observed, but most decayed to a stable and negligible level after 0.5 s of optical stimulation time. If we rank the aliquots from brightest to dimmest according to their net T_n signals and calculate their cumulative light sum, we observed that <5% of aliquots contribute to >90% of the total signal for OSL1–OSL4 (Fig. S16C), and that OSL5 has ~20% of aliquots contributing to >80% of the total signal (Fig. S16C). The net T_n signal varies significantly between aliquots, ranging from a few tens to several thousands of counts per 0.1 s of optical stimulation time, but the majority of aliquots have T_n signals that are

<1000 cts per 0.1 s (Fig. S16C). DRC shapes also vary significantly between aliquots. Representative DRCs for 6 aliquots each from OSL1 and OSL3 are shown in Fig. S16D and E. Some signals saturate as early as ~50 Gy whilst others show no sign of saturation up to 300 Gy.

Dose recovery test for quartz

The suitability of the SAR procedure for quartz was tested for OSL3 using a simplified dose recovery test (*68*). Natural grains were first bleached with blue LEDs at room temperature for 100 s, then stored for 1 day before another room-temperature blue LED bleach for 100 s. The bleached grains were then given a laboratory dose of 100 Gy. The simplified SAR procedure involves measurement of only two SAR cycles of the same size dose (100 Gy) to calculate a sensitivity-corrected signal-recovery ratio. The signal-recovery ratios are shown as a radial plot in Fig. S16F. Most ratios are statistically consistent with unity at 2σ. A weighted mean ratio of 1.01 ± 0.02 and overdispersion (OD) value of 12 ± 1% was obtained using the CAM, suggesting that the SAR procedure is suitable for our samples.

Rejection criteria for quartz

We applied a series of criteria to reject aliquots that may yield unreliable results. The first three criteria are identical to those for K-feldspar and include rejection of grains if: (1) their initial T_n signal is <3σ above the corresponding background, or the relative standard error on the net T_n intensity is >25%; (2) the ratio of the L_x/T_x value of the zero regenerative dose to that obtained from the maximum regenerative dose is >5%; or (3) the L_x/T_x data points are too scattered, e.g., they were associated with poor recycling ratios or OSL IR-depletion ratios (>10% from unity), or they had a FOM value (*74, 75*) >10% or a RCS >5 (*76*). Two additional criteria were also applied to quartz. Aliquots were rejected if: (4) D_e was obtained by extrapolation of the DRC fitted with a GOK function, or L_n/T_n was statistically consistent with, or above, the saturation level of the corresponding DRC, so that a finite D_e or D_e error could not be obtained; or (5) both negative and positive D_e values that were statistically consistent with zero at 2σ, representing contamination by modern grains. Table S7 provides details for all samples and reasons for rejecting aliquots. Only very few aliquots had recuperation values >5% (criterion 2). In contrast, a significant proportion (~50% or more) had poor DRCs (criterion 3). Amongst grains with satisfactory DRCs, 9–20% are 'saturated' (criterion 4). All 5 samples contain 'modern' aliquots, representing a large proportion of D_e values for OSL1 (25 of 100) and OSL2 (40 of 64), but only small proportions for OSL3 (3 of 56), OSL4 (10 of 113) and OSL5 (5 of 343). This is consistent with previous observations (*85*) showing that intrusion of younger (or modern) grains into older sediments decreases exponentially with depth. In the end, only 0.4–3.4% of measured aliquots for OSL1–OSL4 and 21% for OSL5 yielded acceptable D_e values.

Quartz SAR D_e distributions

The distribution of accepted D_e values for all samples are shown as radial plots in Fig. S17. All samples show significant scatter, with OD values ranging between 56 ± 3% (OSL5) and 179 ± 26% (OSL2). The D_e distributions show two key features: (1) OSL1 and OSL2 both have a substantial number of aliquots with low D_e values. We used the minimum age model (MAM) to calculate their minimum D_e values, keeping in mind that the lower end of their D_e distributions are truncated because all negative and positive D_e values consistent with zero (e.g., 'modern' contamination) were rejected (Table S7). It can also be seen that the low D_e values form a continuum of values from zero. These low D_e values can, therefore be interpreted as the result of post-depositional mixing due to recent and historical activities that cannot easily be disentangled. (2) All samples also have a high D_e component at, or close to, the saturation limit of quartz OSL. Each sample have between 9% and 20% of aliquots with natural signals that are 'saturated' and for which D_e values cannot be calculated and presented in the radial plots (Table S7). Previous studies showed that rejection of 'saturated' grains may result in truncation of the upper end of the D_e distributions and underestimation of D_e (*86-89*). Values that fall between these two extremes may just represent grains with early dose saturation characteristics (i.e., low D_0 values) (*90*), or may be further evidence of post-depositional mixing. Reliable results for the depositional age of these samples can, therefore, not be obtained from the present datasets with any certainty. The stratigraphic integrity of OSL1 and OSL2, in particular, is compromised. We consider the maximum D_e value obtained using the maximum age model (MAX) to represent the best estimate for sediment deposition of these samples, keeping in mind that this may still be an underestimate. An additional uncertainty (i.e., σ_b) of 25% was added in quadrature to the relative standard error of each individual D_e value prior to running the MIN and MAX models (*91*). The OD and D_e values and age estimates are summarized in Table S8.

L_nT_n distribution of quartz

To reveal the full extent of the quartz D_e distributions, we applied the L_nT_n method described and tested in (*21, 90, 92, 93*). This method requires construction of SGCs, but quartz grains or aliquots are usually characterized by their significant between-grain variation in DRC shapes and do not share a common SGC (Fig. S16D and E) (*84, 94*). We applied the method of Li et al. (*84*) who showed that grains can be grouped according to their DRC shapes. This involves calculating the ratios of L_x/T_x values from a relatively large regenerative dose (~300 Gy) and a smaller regenerative dose (~80 Gy) for all aliquots that passed the first 3 rejection criteria described above. This provides a quantitative indicator of the saturation characteristics of the corresponding DRC for each aliquot; a higher ratio means later saturation and a smaller ratio means earlier saturation. We then applied the FMM to the distribution of L_xT_x ratios to divide the ratios into a number (k) of groups, assuming that each group is statistically consistent with each other (i.e., over-dispersion $\sigma_b = 0$). A *k* value giving rise to the smallest Bayes Information Criterion (BIC) was chosen as the optimal estimate for the number of groups (*91*). Between 4 and 6 groups provided an optimum fit to the data for each sample (Fig. S18). Most ratios fall between ~1 to ~2.5, indicating a wide range of saturation characteristics.

We then applied the LS-normalization method (*84*) to establish a SGC for each group using a GOK function (*79*). The DRCs before and after LS-normalization, together with the SGCs established for each group, are shown in Figs. S19–S23. To test the reliability of the groups, we compared the ratios of the measured L_x/T_x ratios obtained from full SAR measurements and expected L_x/T_x values obtained from the SGCs. Most ratios (>90%) are consistent with unity at 2σ (e.g., Fig. S19C–H).

To compare and analyze the results from the grains sharing the same SGCs, L_n/T_n values of all accepted grains, including those identified as 'saturated' and grains with zero doses (Table S7), were re-normalized using the scaling factors obtained from the LS-normalization process. The LS-normalized L_n/T_n values for individual DRC groups are shown as radial plots in Figs. S19–S23. It should be noted that any LS-normalized natural signals (L_n/T_n) with negative values are not counted or shown in the radial plots. Also, the number of LS-normalized L_n/T_n values shown in Fig. S19–S23 are larger than the number of SAR D_e values shown in Fig. S17 for the same samples. This is because no D_e value can be estimated for 'saturated' grains using the SAR procedure and zero-dose grains were rejected, so neither are shown in the SAR D_e radial plots.

As shown in Figs. S19–S23, except for OSL1 and OSL2 whose LS-normalized L_n/T_n values have broad distributions, all other samples appear to have a dominant component of high L_n/T_n values, with some much smaller values also present in the distributions of different groups and samples. We interpret the components with the highest L_n/T_n ratios to be representative of the depositional age of the sediments and the lower D_e values representing post-depositional contamination by younger grains.

We used three different statistical models to estimate the weighted mean LS-normalized L_n/T_n values for each group: the maximum age model (MAX), central age model (CAM) and finite mixture model (FMM). CAM was used for distributions consistent with no mixing (e.g., Group 2 of OSL4, Fig. S22). For distributions that display discrete components (e.g., Groups 1 and 2 for OSL3, Fig. S21), we used the FMM to identify the number of discrete components and their corresponding weighted mean values. For distributions that show continuous mixtures (e.g., Group 3 of sample OSL1, Fig. S19), we used the MAX (*95*). An additional uncertainty (i.e., σ_b) of 15% was added in quadrature to the relative standard error of each individual re-normalized L_nT_n value prior to running the MAX model (*90, 91*). For some groups (e.g., Groups 4 and 6 of OSL1, Fig. S19), there is an insufficient number of aliquots (<10) for reliable statistical analysis.

To obtain final D_e values, the best estimate re-normalized L_n/T_n values for each group was projected onto its corresponding SGC. The results are summarized in Table S9. The D_e values obtained for each group from the same sample are statistically consistent at 2σ, so were combined. The final D_e estimates and their corresponding 95% confidence intervals are also provided in Table S9.

L_nT_n distributions of K-feldspar

The re-normalized L_n/T_n values for individual K-feldspar grains are shown as radial plots in Fig. S24. The three uppermost samples (OSL1–OSL3) is associated with the largest OD values (54%, 34% and 35%, respectively), consistent with the large scatter also observed for the equivalent quartz data sets. In contrast, no ‘modern’ or historical contamination was detected. These three samples also contain a proportion of grains with very high L_n/T_n values not observed in the L_n/T_n distributions of the quartz aliquots. Since it is unlikely that modern quartz grains (from the surface) and old K-feldspar grains (from older layers) would preferentially intrude into the upper layers, we provide three possible explanations below; these are not necessarily mutually exclusive.

1) The first may explain the presence of low-D_e grains in the quartz D_e distributions and a lack thereof in the K-feldspar D_e distributions in Layers 2 and 3 as a result of downward movement of grains from the surface/Layer 1. Here we assume that the young intrusive grains derive from the surface or Layer 1 that have been heated by human activities. This cave is known to have been used as a place of religion over the last few hundred years. There is evidence of extensive burning in the cave that may or may not be associated with it being a Buddhist temple and used frequently. It has previously been shown that heating may result in the desensitization of the optical signal of K-feldspar grains and an increase in sensitivity of the quartz OSL signal. So, heating at high temperatures has an opposite effect on the luminescence signals emitted by the two mineral types (*96, 97*). We speculate that sensitized ('bright') modern quartz grains will show up clearly within a sample of non-sensitized quartz grains from the deeper and unburnt layers that have generally low sensitivities. In contrast, the desensitized ('dim') K-feldspar grains will have signals that are not detectable above instrumental background and will be absent within a sample of otherwise bright feldspar grains. We observed a similar pattern at Denisova cave for samples collected from burnt layers in the East and Main chambers (*21*).

2) The second may explain the presence of younger quartz grains and older feldspar grains in Layers 2 and 3 with or without burning as a result of downward movement of grains from the surface/Layer 1. There may still be sufficient quartz single grains with bright enough signals in a multi-grain aliquot, without being burnt, to produce the same pattern. This assumes, that the quartz grains on the surface or Layer 1 has been exposed to sufficient light to have their signals reset. In contrast, K-feldspar on the surface were not exposed to adequate amounts of light, resulting in high residual doses. This would be especially true for grains liberated from the cave walls or roof or limestone clasts. It is well-known that the quartz OSL signal is much easier to reset than the K-feldspar pIRIR signal.

3) The third may explain the presence of high-D_e values in the K-feldspar D_e distributions and a lack thereof in the quartz D_e distributions, but does not address the young quartz D_e values. The stratigraphic and sedimentological features of Layer 2 and 3 show that both layers are compact and probably intact (Table S1). So, it argues against large-scale post-depositional mixing, and the inclusion of large particles, such as bones. There is a large range of radiocarbon ages on bones from these layers (especially Layer 2) that can be best explained as a result of deposition of reworked materials from the inner parts of the cave. The relatively steep topography of the entrance chamber (Fig. S1C) could have facilitated the reworking and

transportation of the materials down slope and subsequent re-deposition onto the relatively flat area where T2 is located. So, we speculate that this re-deposition of sediment down slope may include grains with a range of ages and that for the feldspar grains the transportation time and exposure to sufficient light to reset the signal was not sufficient, resulting in partially bleached grains and a broad D_e distribution. It seems to have been sufficient for the quartz grains, which is more impacted by the post-depositional intrusion of young grains from the surface or Layer 1.

At this stage we cannot distinguish which, if any, is the true reason responsible for the observed D_e distribution patterns for samples OSL1 and OSL2. Further studies on modern samples in the cave and perhaps micromorphological investigations of sediments should be able to provide further evidence to address these hypotheses in the future.

The lower samples (OSL4–OSL12) have most of their L_n/T_n values randomly distributed around a central value with OD values that range from 18–38%. Samples with OD values in the upper part of the range (27–38%) contain a few very old or young outliers (e.g., OSL6, OSL7, OSL9, OSL10 and OSL12).

Previous studies reported a relationship between D_e or L_n/T_n values and sensitivity (T_n intensity), where weighted mean D_e or L_n/T_n values increase steadily as a function of T_n intensity until a 'plateau' is reached (*21, 98*). The weighted mean re-normalized L_n/T_n ratios as a function of T_n intensity are shown in Fig. S25 for those samples with OD values of $<20\%$ (OSL4, OSL5, OSL7, OSL9 and OSL11). All samples show a similar pattern; weighted mean L_n/T_n ratios increase with T_n and reach a plateau at ~200–500 cts/0.1 s of optical stimulation time (Fig. S25A–E), suggesting that dimmer grains have smaller D_e values compared to brighter grains. This pattern was not observed for the dose recovery test (Fig. S25F). There are two possible explanations: 1) dimmer grains have lower K concentrations than brighter grains (*43, 45*), so receive less internal dose and yield smaller D_e values, and/or 2) dimmer grains are more affected by a small amount of anomalous fading (*99*) that is not detectable using a laboratory fading test. Regardless of the reason, the brighter grains should be preferred for D_e estimation. We, therefore, applied a common T_n threshold of 500 cts/0.1 s to all L_n/T_n distributions.

The L_n/T_n distributions for all samples, after application of the T_n threshold, are shown as radial plots in Fig. S26. The distributions reveal four patterns to which different statistical models were applied: (1) single component distributions (OSL4–OSL8 and OSL10) with OD values of 16–19%. The CAM was applied to these distributions to obtain final D_e values. (2)

Single component distributions with a few conspicuous outliers (OSL9 and OSL12). The outliers were detected using the normalized median absolute deviation (nMAD) (*100, 101*). Re-normalized L_n/T_n ratios were converted to natural logarithms and the nMAD calculated using 1.4826 as the appropriate correction factor for a normal distribution. Log L_n/T_n ratios were rejected if nMAD values were >2.5 (*90, 102*). The CAM was then applied to all accepted L_n/T_n values to calculate the final D_e value for age determination. (3) Mixed distributions with discrete components (OSL2) to which the FMM was applied to estimate the weighted mean L_n/T_n values for each component. (4) Mixed samples with a continuous distribution of L_n/T_n values (OSL1 and OSL3), likely due to partial bleaching of the K-rich feldspar grains. For these two samples, we used the minimum age model (MAM) (*82*) to calculate the minimum re-normalized L_n/T_n values. An additional uncertainty (i.e., σ_b) of 20% was added in quadrature to the relative standard error of each individual L_n/T_n value (*91*); σ_b was estimated from the OD values obtained for single component distributions.

Final D_e values for age determination were obtained by projecting the modelled re-normalized L_n/T_n values for each sample onto its corresponding K-feldspar SGC. One of the features of the L_nT_n method is that the uncertainty associated with D_e estimates become asymmetric when the natural signal intensity lies on the non-linear region of the DRC. So all uncertainties associated with D_e are expressed as a 2σ range (95% confidence interval).

Optical ages and comparisons with radiocarbon ages

All quartz and K-feldspar age estimates are summarized in Tables S8–S10. The SAR quartz D_e values (Table S8) are statistically consistent with the L_n/T_n-derived quartz D_e values, but much less precise. This suggests that the upper end of the quartz dose distributions are not significantly truncated, but that the SAR quartz D_e values are at the limits of the technique. The final ages are, therefore, based on D_e values derived using the L_nT_n method for both quartz and K-feldspar. There are several advantages to measuring both minerals: (1) K-feldspar grains have an internal beta dose rate from the decay of radioactive ^{40}K and ^{87}Rb inside the mineral grains. As a result, K-feldspar grains are less affected by any changes or uncertainties in the external environmental dose rate, including time-dependent changes in moisture content or disequilibrium in the ^{238}U, ^{235}U and ^{232}Th decay series. (2) The OSL and pIRIR signals are reset by sunlight (‘bleached’) at different rates. The quartz OSL signal can be fully reset to zero in a matter of seconds, whereas a residual signal is still present for K-feldspar even after hours of sunlight exposure (*103*). A comparison between quartz and K-feldspar ages can, therefore, be used to test whether samples have been adequately bleached prior to burial (*104*). (3) K-feldspar pIRIR signals usually saturate at much higher doses than quartz OSL, which enables larger finite D_e values (and older ages) to be measured from K-feldspar grains, allowing dating of all layers at Baishiya Karst Cave.

Below, we discuss the age estimates for each sample and highlight the key features that provide confidence, or not, in the final ages for these samples. All ages are reported as a 95.4% confidence interval because the age uncertainties are asymmetric (see Tables S9 and S10).

OSL1 (middle of Layer 2): The L_n/T_n distributions for both quartz OSL and K-feldspar pIRIR show large overdispersion (Tables S9 and S10). We interpret this to be the result of post-depositional mixing that introduced younger grains into a population of well-bleached quartz grains (quartz OSL) and older grains into a population of partially-bleached feldspar grains (K-feldspar pIRIR). The most reliable age estimates are, therefore, obtained using the MAX for quartz—15.3–27.2 ka (L_nT_n OSL)—and the MAM for K-feldspar—12.4–20.9 ka. Both optical ages are statistically consistent, but they are inconsistent with even the youngest of the ^{14}C age estimates that, for this layer, range from 28.7 to 45.6 cal ka BP (Table S3); the uppermost bone sample (B3, 8 cm depth), however, yielded the oldest ^{14}C age (45.6–42.0 cal ka BP). The stratigraphic features (e.g., geometry and boundaries) suggest that this layer is intact and post-depositional intrusion of large particles (such as bones) is unlikely, so the large range of radiocarbon ages of the bones from this layer is best explained as a result of deposition of reworked materials from the inner parts of the cave. The relatively steep topography of the entrance chamber (Fig. S1C) could have facilitated the reworking and transportation of the materials down slope and subsequent re-deposition into the relatively flat area where T2 is located. Together, the optical and ^{14}C ages indicate that accurate and precise age estimates for the depositional age of the sediments and any associated artifacts are difficult to achieve.

OSL2 (middle of Layer 3): Similar to OSL1, the quartz and K-feldspar L_n/T_n distributions show different patterns and suggest a complex combination of post-depositional mixing and partial resetting of the K-feldspar pIRIR signal that cannot be resolved easily. The quartz OSL age of 20.1–49.3 ka (Tables S9) is based on the MAX D_e value and is poorly constraint due to a paucity of reliable data, but also the complexity of the distributions that clearly indicate large-scale intrusion of modern and historical-age grains into a matrix of older grains (Fig. S20). The K-feldspar age of 38.4–54.9 ka (Tables S10) is based on the dominant FMM component (70% proportion of grains) (Fig. S26B). Both ages are statistically consistent and also agree with six ^{14}C ages on bone fragments that range from 33.5 to 49.5 cal ka BP (Table S3); two of the youngest bones (B67 and B94) were, however, collected from the deepest part of the layer (near the boundary with Layer 4). Like Layer 2, the optical and ^{14}C ages for Layer 3 indicate that accurate and precise age estimates for the depositional age of the sediments and any associated artifacts are difficult to achieve.

OSL3 (bottom of top half of Layer 4): Similar to OSL1 and 2, the L_n/T_n quartz and K-feldspar distributions show different patterns, but this time, the distributions for different quartz groups (Fig. S21) are dominated by a single population of aliquots with only a few conspicuous younger outliers. This shows a dramatic decrease in the degree of post-depositional mixing of modern and historical age activities compared to Layers 2 and 3. This reduction in mixing is also apparent in the SAR D_e distribution for this sample (Fig. S17). The L_n/T_n K-feldspar distribution shows a larger and more continuous spread of values, which we interpret to be primarily the result of partial bleaching. The weighted mean age obtained for the quartz aliquots of 39.1–53.3 ka (Tables S9) is statistically consistent with the minimum age of 35.0–52.2 ka for the partially-bleached K-feldspar grains (Tables S10), and are also similar to the optical ages obtained for the overlying Layer 3. No radiocarbon samples were collected from a similar depth in Layer 4. One bone sample (B105) was, however, collected from a depth of 25 cm, at the boundary between Layers 3 and 4. This sample gave an age of >44,890 cal BP, spanning the confidence interval of the optical ages.

OSL4 (bottom of Layer 4): The L_n/T_n distributions for quartz (Fig. S2) and K-feldspar (Fig. S26) show much less spread, with only a few young outliers in the quartz distributions, consistent also with the SAR D_e distribution (Fig. S17). The weighted mean quartz and K-feldspar ages of 52.2–67.6 ka and 57.9–71.6 ka, respectively, are statistically consistent. A single bone sample collected from the same depth as OSL4 (36 cm) also returned an infinite ^{14}C age of >46,830 BP (B147). The sedimentological characteristics of the bottom half of Layer 4 show little evidence for post-depositional mixing; the sediment is compact and contain 2–5 mm-thick bedded white calcite lamellae.

OSL5 (layer 6): The L_n/T_n distributions for quartz (Fig. S23) and K-feldspar (Fig. S26E) show limited spread. Weighted mean ages for quartz and feldspar are 70.5–89.5 ka and 87.8–118.4 ka, respectively. The ages just overlap at the 95.4% CI. This difference may be due to the quartz aliquots being very close to the effective limit of the technique, resulting in underestimation of the quartz D_e values for groups with the lowest saturation characteristics, hence the increase trend in D_e with Group number (see Tables S9). The quartz age should, therefore, be considered a minimum age and the K-feldspar age the best estimate age for this sample. Both quartz and K-feldspar ages are consistent with the infinite ^{14}C ages obtained for two bone samples from Layers 5 and 6 (>48,500 BP and >54,800 BP).

OSL6–OSL12 (Layers 6–10): These samples were only measured with K-feldspar grains and the resulting L_n/T_n distributions are all dominated by a single component (Fig. S26); some of the samples contain a few outliers. Age estimates are presented in Table S9. Ages for samples from Layers 6–9 are statistically consistent with mean age estimates ranging between ~96 and 112 ka. The two samples from the ~70 cm deep Layer 10 (OSL11 and OSL12) are older with age estimates of 124.4–173.2 ka and 152.4–242.4 ka for the top and bottom of this layer.

Optical (quartz and K-feldspar) and ^{14}C age estimates for all samples are summarized in Fig. 2. Overall, there is broad agreement between the quartz OSL, K-feldspar pIRIR and ^{14}C ages, supporting the reliability of our chronology. Ages range from ~20 ka near the top (~10 cm depth below surface) in Layer 2 to ~200 ka at the bottom (~165 cm depth below surface) of the current excavation in Layer 10. The optical dating results for sediments provide information suggesting that different processes were responsible for deposition of the upper layers (especially Layers 2 and 3) compared to the lower layers (Layer 4–10). It also shows that the upper layers were affected by intrusion of modern and historical-age grains probably due to trampling and digging of the pit feature, and incorporation of re-worked older remains into younger deposits (e.g., reversed ^{14}C ages). Further work is required and is ongoing to fully understand the depositional processes responsible for sediment deposition and incorporation of other materials in Layer 2 and 3. Until then, any association of sediment with artifacts, faunal remains and DNA remains tenuous, and accurate and precise ages for these two layers cannot be assured. We observed little evidence for post-depositional mixing of sediments from Layer 4 to the base of the excavation in Layer 10. Furthermore, the flowstone layer at the bottom of Layer 3 (Fig. S4) could have limited possible post-depositional mixing processes. Hence, our single-grain data are supported by the stratigraphic observations, and we are confident in the age estimates obtained for the deeper layers (Layer 4 and those below).

Bayesian modelling of optical and ^{14}C ages

To establish a chronological framework for the sedimentary profile revealed in T2, we constructed a Bayesian age model that includes both the optical and ^{14}C ages. Bayesian analysis was conducted using OxCal v4.3 (*105*). All ^{14}C dates were calibrated using the IntCal13 calibration curve (*29*). The OSL and pIRIR ages were input as C_dates in calendar years before AD 1950 with their 1σ errors as likelihood estimates.

We used the sequence of stratigraphic layers as prior information under the assumption that a layer that is stratigraphically lower than another is older. Samples from the same stratigraphic layer were modelled as a *Phase*, assuming that the measured ages are unordered and uniformly distributed, so any mixing within a layer will not influence the model; this is especially important for Layers 2 and 3. For the layers with abrupt and clear boundaries, upper and lower *Boundaries* were placed for each layer to constrain the start and end ages for the layer. *Transitional* boundaries was placed between Layers 2/3 and Layers 9/10, where gradual stratigraphic changes were identified (Table S1), assuming continuous sediment accumulation. The samples, phases, intervals and boundaries for the whole stratigraphic sequence were arranged in a *Sequence* according to their relative stratigraphic order. The quartz OSL age for OSL5 and three infinite radiocarbon ages (B147, B211 and B218) were included using Oxcal's '*Before*' command; the radiocarbon limit (50,000 yr) was given as *a terminus ante quem* for the radiocarbon samples.

We used a general t-type outlier model (*106*) to assess the likelihood of each age being consistent with the modelled ages. The radiocarbon age of the uppermost ^{14}C sample (B3) was assigned an outlier probability of 100%. All other samples were assigned prior outlier probabilities of 5%, with the posterior outlier probability calculated during the modelling process.

There are two occasions where it is reasonable to assume that stratigraphic or chronological gaps may exist. The first is Layer 5, a ~12 cm-thick very rubbly layer (Table S3), for which no age estimate was obtained. We inserted a *Boundary* at the end of Layer 6 and the start of Layer 4 and separated these two boundaries with an *Interval*. The *Interval* duration is based on the difference between the measured optical ages for OSL5 (Layer 6) and OSL4 (base of Layer 4). The second is Layer 4 for which we obtained two ^{14}C ages (B147 and B105) and four optical age estimates from two samples (OSL3 and OSL4). The two samples for optical dating were collected from the same profile wall. OSL3 is from the middle of this layer (~31 cm deep) and OSL4 is from the bottom (~37 cm deep). Considering the diameter of sampling tubes (5 cm), there is only an ~2–3 cm gap in depth between the two samples. The age estimates for these two samples are, however, statistically different; the mean optical ages for OSL4 (~62 ka) is ~15 ka older than the mean optical ages for OSL3 (~45 ka).

We provide two different interpretations for the depositional history of Layer 4. In Scenario 1, we assume that the ages are correct, that the sedimentary characteristics, relative stratigraphic order of the samples (optical and ^{14}C) and information from the quartz and K-feldspar L_n/T_n distributions support that the layer as a whole is intact and that sedimentation started ~62 ka ago and ceased ~45 ka ago. This scenario suggests a very slow and gradual sedimentation rate; ~10 mm of sediment accumulation per thousand years. In Scenario 2, we make the same assumptions as in Scenario 1, except that we assume that the difference between the ages for the two samples indicate a depositional hiatus in between or erosional break of ~10–15 ka and that Layer 4 was deposited in two broad pulses. Further sedimentological and chronological work needs to be done to support or refute either of the scenarios. In the final Bayesian model, we used Scenario 2. The only difference is the inclusion of an *Interval* between OSL4 and OSL3 where the duration is based on the difference between the measured optical ages of these two samples and the *Interval* is bracketed with an end *Boundary* for Layer 4 lower and a start *Boundary* for Layer 4 upper. We tested the sensitivity of the inclusion of this interval by also running the model without the Layer 4 *Interval* and additional *Boundaries*. Changes in modelled age estimates is negligible, but the overall model agreement index is better using Scenario 2.

The Bayesian modelled chronology for the stratigraphic sequence from BKC is presented in (Fig. 2). The CQL code used to generate the model is listed in Table S11, and the corresponding data provided in Table S12. All modelled age ranges were calculated at 68.2% and 95.4% posterior probability. The age likelihood (prior to modelling) and posterior (mathematically modelled) distributions are shown in Fig. 2, using pink and blue shading, respectively. The uncertainties associated with the Boundary estimates were calculated using the OxCal platform and are based on the total unshared component of error only for optical ages from Layers 6–10. The shared error, ~3% on average, was combined, in quadrature, afterwards with the unshared errors of the modelled ages from Layers 6–10 to obtain their final modelled age uncertainties. For the upper layers (Layers 2–4), where both optical and radiocarbon ages were included, we used their full (shared and unshared) errors.

Twenty-seven samples were collected and dated from T2 in BKC and 32 age estimates obtained. All (except the one from H1) were included in the Bayesian age model in Fig. 2. All samples show good consistency relative to each other and to the stratigraphic prior applied. In addition to the ^{14}C age for B3 that was assigned an outlier probability of 100%, the model identified only two further ages with posterior outlier probabilities of >10% (K-feldspar pIRIR ages for OSL1 and OSL8; Table S12). The overall model agreement index is 67%, and the individual agreement indices for most ages are greater than 90% (mean value is 94%) with only two smaller than 60% (pIRIR ages for OSL3 and OSL8; Table S12). These suggest that the modelled results generally agree well with our experimental data and stratigraphic interpretations.

Sedimentary DNA extraction and analysis

Eight sediment samples from the middle of each layer, except Layers 1 and 5, were collected for sedimentary DNA analysis. Masks and gloves were worn during sample collection to prevent contamination.

DNA extraction, library preparation and mtDNA capture

Fifty ul of DNA extract was prepared from between 108 and 238 mg of each sediment sample from the different layers of the Baishiya site using the method described in Dabney et al. (*107*) (Table S18). In total, 40 libraries were produced, using 10 ul each of 35 extracts and five negative controls, using a single stranded library preparation protocol (*108*). For extracts from a subset of the sediment samples, several libraries were generated based on positive ancient hominin results obtained in the first screening (Table S18). Each library was spiked with small copies of a control oligonucleotide to evaluate the efficiency of library preparation, as failure to find copies of the control oligonucleotide would suggest the extract was inhibited (*109*). The total number of molecules in each library and of the spiked-in control oligonucleotides were quantified using qPCR (*110*). Libraries from Layers 7 and 9 were inhibited, as we find either low or no copies of the spiked-in control oligonucleotide (Table S18).

Each library was amplified for 35 cycles using AccuPrime Pfx DNA polymerase (Life Technologies)[4]. The sample-specific barcodes were introduced into both the P7 and P5 library adaptors during amplification (*111*). PCR products were purified with the MinElute PCR purification kit (Qiagen, Hilden, Germany). Library concentrations were determined using a NanoDrop 2000 spectrophotometer.

Hybridization capture was performed in two rounds using a modified on-bead hybridization protocol (*15, 112*) with a human mtDNA probe set (*16*) and 242 mammalian mtDNA probe sets (*15*), respectively. After pooling the captured libraries, the heteroduplices of the library pool were removed in a one-cycle PCR reaction using the Herculase II Fusion DNA polymerase (Agilent Technologies) (*113*) with primers IS5 and IS6 (*114*). The concentration of DNA in the pools was assessed using a DNA-1000 chip (Agilent Technologies).

Sequencing and data processing

The libraries were sequenced for 76 cycles from each end of the library insert, and two additional 7 base pair index reads were included using the double-index sequencing (*111*) on an Illumina MiSeq platform (MS-102-3001 MiSeq Reagent Kit v3,150-cycle). After spiking in an indexed control PhiX 174 library, the resulting yield was 0.27-0.86% control reads (index 5'-TTGCCGC-3'). After base calling using Bustard (Illumina), we used *leeHom* (*115*) to trim and merge the paired sequences that overlapped by at least 11 base pairs. We restricted the downstream analyses to the fragments with the expected indices for each library.

All fragments captured using the human mtDNA probe set were aligned to the revised Cambridge Reference Sequence (rCRS, NC_012920.1 (*116*)) using *bwa* (version: 0.5.10-evan.10-1-gfb1ff83) (*117*) with parameters: －n 0.01 －o 2 －l 16500. After mapping fragments with a minimum map quality of MAPQ$\geqslant$25 and with length $\geqslant$35 base pairs were retained. PCR duplicates were removed by collapsing sequences starting and ending at identical coordinates using *bam-rmdup* (https://github.com/mpieva/biohazard-tools version: 0.6.3, Table S18).

Identification of hominin fragments and ancient DNA authenticity

Unique fragments were realigned to the non-redundant mammalian mtDNA database from the National Center for Biotechnology Information (NCBI) using the Nucleotide Basic Local Alignment Search Tool (BlastN) (*118*). Since multiple different species are returned using blastn for any given matched fragment, we used the lowest common ancestor (LCA) algorithm in MEGAN (*119*) (version 6.15.2) to assign fragments to taxa at the family level. The blast output files were imported and phased through MEGAN with parameters MinScore=35.0, TopPercent=10.0, and MinSupportPercent=0.1 to retain the fragments with a minimal alignment score of 35, which includes the top 10% of the highest scores (*15*). Taxa represented by at least 1% of the total number of identifiable fragments are shown in Fig. S27.

To evaluate whether DNA from ancient hominids can be found in the different layers, only fragments assigned to *Hominidae* were retained (Table S19). The frequencies of terminal C to T substitutions compared to the reference genome for all and for putatively deaminated fragments are shown in Table S19. We find that DNA fragments in the libraries from Layers 2, 3, 4, and 7 show a C to T substitution frequency significantly higher than 10% on at least one of their ends when testing all fragments (Table S19). The percentage of Cs that appear as

Ts varied between 20.5% and 51.8% at the 5'-ends of the molecules and between 11.9% and 52.6% at their 3'-ends (Table S13, Table S19). As an elevated frequency of terminal C to T substitutions is a prominent feature of ancient DNA (*120, 121*), we conclude that these libraries contain authentic ancient hominid DNA fragments.

Lineage inferences using hominin branch-specific variants

To identify branch-specific variants for modern humans, Neandertals, Denisovans and the Sima de los Huesos individual we used mtDNA consensus sequences created for each hominin group using a panel of 54 modern humans, 23 Neandertals, 4 Denisovans and the Sima de los Huesos individual, and aligned these to the chimpanzee mtDNA (*15, 19, 122*). We then identified the base pair carried by DNA fragments from each library at these informative sites. We ignored T nucleotides on forward strands and A nucleotides on reverse strands within the first and last three alignment positions to reduce the effect of deamination. The number and proportion of DNA fragments covering the variants specific to each branch are shown in Table S19. We also repeated this analysis for deaminated fragments only by restricting the analysis to fragments with C→T substitutions in the first three positions at the 5'-end and the last three positions at the 3'-end. Libraries from Layers 2, 3, 4, and 7 contain DNA fragments matching variants unique to the Denisovan branch, as listed in Table S13 and Table S19.

Estimating present-day human DNA contamination

For the libraries containing authentic ancient hominin mtDNA, we estimated the extent of contamination with modern human mtDNA. To do this, we used all fragments (i.e. prior to filtering only for the fragments assigned to hominids) to increase statistical power, since even though animal DNA fragments are present, there is no reason they would match more often to a present-day human variant compared to an archaic one. We used a likelihood-based method that has been used previously for estimating present-day contamination in ancient mitochondrial DNA (*123*). The likelihood-based method simultaneously estimates contamination and error by comparing mtDNA fragments with a panel of 311 global present-day humans, whose mtDNA haplotypes are treated as the contaminating population. For all libraries in Layer 2, Layer 4, and Layer 7 except five, we estimated less than 5% of present-day human DNA (Table S13). The library L5832 has the lowest contamination rate estimate of present-day human DNA, and it has 60% (6/10) of the mtDNA fragments that carry lineage-specific substitutions matching the modern human state. Since it is too low coverage (0.66-

fold) to use to detect present-day human DNA, to exclude the possibility of modern human contamination in library L5832, we only used deaminated fragments for further analyses.

There are two libraries （L5793, L5794） from Layer 4 (with a depth about 32 cm) with a high proportion of fragments matching the modern human state (Table S19) even after restricting to deaminated DNA fragments. This might indicate the presence of ancient modern human DNA in Layer 4 sediments. However, the extent of C to T substitutions on DNA fragments that carry Denisovan variants is about five times higher than on DNA fragments that carry modern human variants (Table S14), which suggests that the only ancient DNA present matches the Denisovan state. To avoid the possibility of modern human contamination in subsequent analyses, we excluded these two libraries.

Reconstructing mtDNA consensus genomes

We built two kinds of consensuses from each layer. One is aimed to build an mtDNA consensus sequence from the authentic ancient hominin mtDNA fragments obtained from each of the four layers (see "decision" column in Table S13). To do so, we merged all hominid fragments from libraries with low levels of contamination from present-day human mtDNA (i.e., lower than 5%) and only deaminated fragments from libraries with present-day contamination above 5% from the same layer (Table S15). We also built another consensus mtDNA sequence for each of the four layers with only deaminated fragments for all libraries excluding the two libraries where modern human mtDNA fragments were slightly deaminated (albeit much less so than Denisovan mtDNA fragments). The deaminated fragments were those with C→T substitutions in the first three positions at the 5'-end and in the last three positions at the 3'-end. For Layer 4, additional consensus mtDNA sequences were created for each extract to see if the phylogenetic tree is consistent across different extracts. Majority base calling was used to determine the consensus allele, where more than two-thirds of the overlapping fragments supported the base call and at least two fragments covered the position. To reduce the impact of ancient DNA damage, any T appearing in the first three or last three positions of the fragments was masked as N. The average coverage and the total number of bases called for each consensus can be found in Table S15.

Determining whether sequences originated from more than one mtDNA genome

1) Using variants among Denisovan mtDNA sequences

To test whether more than one mtDNA sequence was obtained from any particular layer, we looked at positions in the mtDNA genome where there was at least 10-fold coverage and where the support for the mtDNA consensus was lower than 80 (i.e., where two different bases were observed, each constituting no more than 80% of observations) (*15*). We then tested whether these positions are known mtDNA variants among Denisovans. In the data from Layer 4, nine such positions are present (four after restricting to DNA fragments with evidence of deamination), suggesting that at least two individuals may have been sampled (Table S15).

2) Using match to the consensus allele at specific positions in Layer 4 mtDNA

We detected three positions where the Layer 4 consensus mtDNA sequence (built with low levels of contamination from present-day human mtDNA (i.e., lower than 5%) and only deaminated fragments from libraries with present-day contamination above 5%) differed from the sequences found in 54 modern humans, four Denisovans, and Sima. We identified 80, 112, and 9 distinct DNA fragments overlapping these three positions. Generally, enriching for ancient DNA using deaminated fragments would increase the number of fragments carrying the consensus allele relative to using all fragments. Instead, we find that support for the deaminated fragments from Layer 4 is 42-86%, lower than the support for the set of all hominid fragments for Layer 4 (67-76%, Table S16). The range for the fragments overlapping these positions is 67-80% for the set of all fragments, and 43-71% for the set of deaminated fragments only. This pattern cannot be explained by modern human contamination, and it suggests that at least one other individual with a different sequence can be found for these positions, suggesting that DNA from more than one individual is present in Layer 4.

3) Using a maximum-likelihood method

We next made use of a previously described maximum-likelihood approach, which allows co-estimation of the number of mtDNA components present in a sample (i.e. by testing models with one, two, or at least three components), as well as the proportion of each component and the divergence between each component when multiple components are identified (*15*). To reduce the chance of residual contamination by present-day human mtDNA being identified as a separate component, we aligned the DNA fragments assigned to *Hominidae* from each library to the reference Denisovan mtDNA genome (NC_013993.1, (*1*)) and retained only fragments showing a C to T substitution at their first and/or last alignment position. We analyzed data from each library generated from Layers 2, 3, 4, and 7 separately, as well as after merging all data generated per layer. To identify variable positions within each dataset, Ts on fragments sequenced in the forward orientation and As on fragments sequenced in the reverse orientation were ignored, as were bases with a quality lower than 30. Fragments with indels to the reference genome were ignored as well. For each dataset, the analysis was carried out using all variable positions observed, and after removing positions where the four Denisovan mtDNA sequenced to date differ from a worldwide panel of 311 present-day human mtDNA genomes, as an observed variant at the latter positions is more likely to originate from a contaminant. This was tested using a likelihood ratio test as described in (*15*).

In Layers 2, 3 and 7, we detected a single mtDNA component in the data for each layer (Table S17). We note, however, that the observed coverage at variable positions in each of these three datasets is lower than 2.5-fold, the threshold under which the model is underpowered to detect multiple components (*15*). In the merged data from Layer 4, and in four out of the fifteen libraries made from sediment collected in that layer, we detected two mtDNA components. In the merged dataset from Layer 4, we estimated the frequency of the minor component to be 10.2%, and the divergence between the two detected components to 0.5%.

Reconstructing phylogenetic trees

Since the coverage of fragments from Layers 2, 3, and 7 are low, to maximize the amount of overlapping data with published Denisovan data, we reconstructed three separate phylogenetic trees for the consensus mtDNA from Layers 2, 3, and 7. For each of those, we co-analyzed that layer's consensus mtDNA genome with Layer 4's consensus mtDNA genome, as well as with mtDNA sequences of the four Denisovans (*Denisova 2*, *Denisova 3*, *Denisova 4* and *Denisova 8*) from Denisova Cave and a Middle Pleistocene mtDNA sequence from *Sima*. We applied the software MrBayes (*124*) and ran 30,000,000 iterations of the Markov Chain Monte Carlo with the first 3,000,000 iterations discarded as burn-in. We used a General Time Reversible sequence evolution model with a Gamma distribution (GTR+G) determined by the best-fit model approach of jModeltest2 (*125*). In all trees, we find that the consensus mtDNA sequences from all four layers cluster with the younger Denisovans (*20*) (*Denisova 3* and

Denisova 4, 100% posterior support) for both low contaminated fragments (Fig. S28A-C) and deaminated fragments only (Fig. S28D-F). For each extract of Layer 4, the consensus mtDNA from each extract in Layer 4 shows the same pattern as for the merged Layer 4 consensus mtDNA genome (Fig. S29).

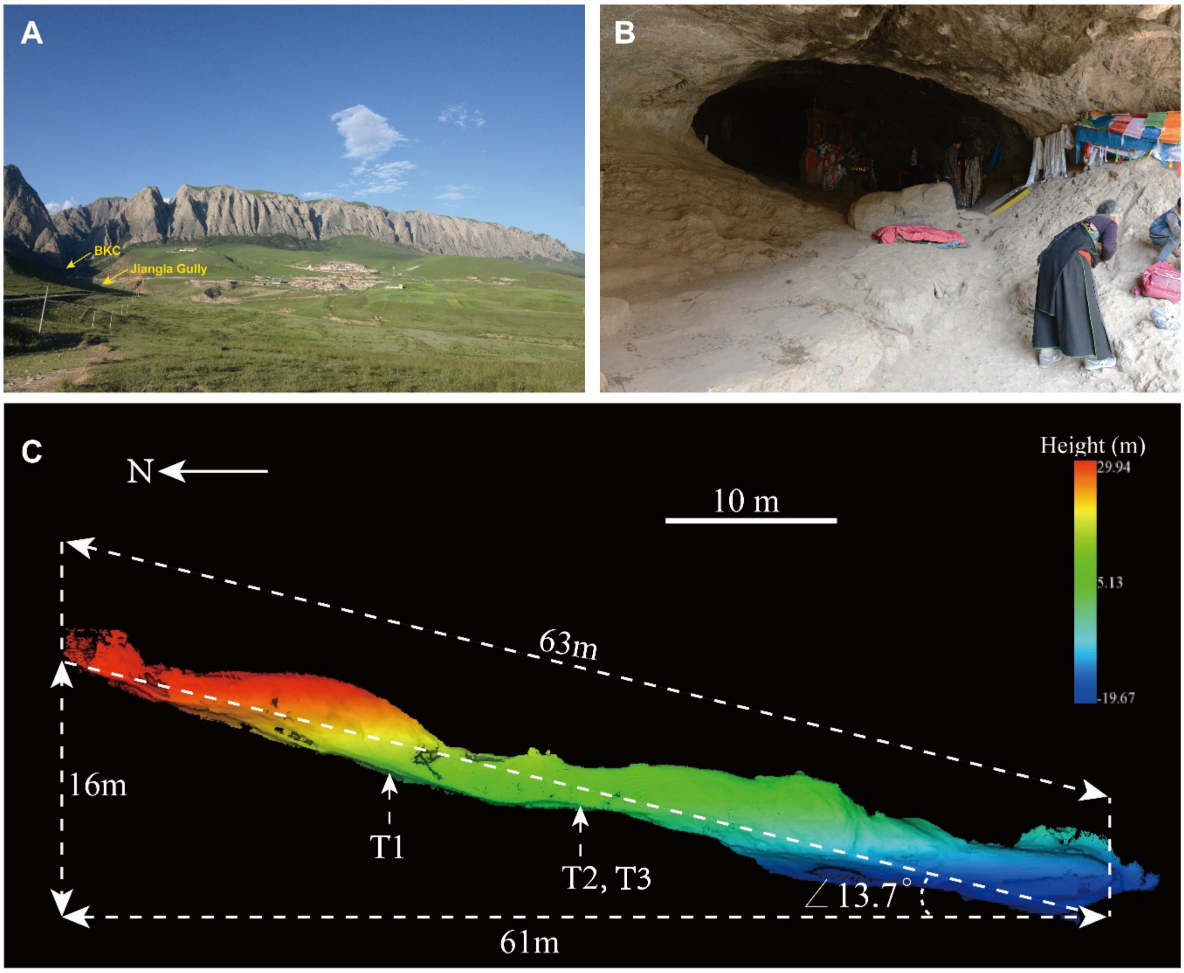

Fig. S1. Baishiya Mountain Massif and Baishiya Karst Cave. (A) View of the Baishiya Mountain massif from the south. The surrounding alpine meadow grasslands can be seen in A, as well as the tributary stream (Jiangla Gully) of the Yangqu River in the left of A. (B) The entrance of the Baishiya Karst Cave. (C) Cross-section of the entrance chamber.

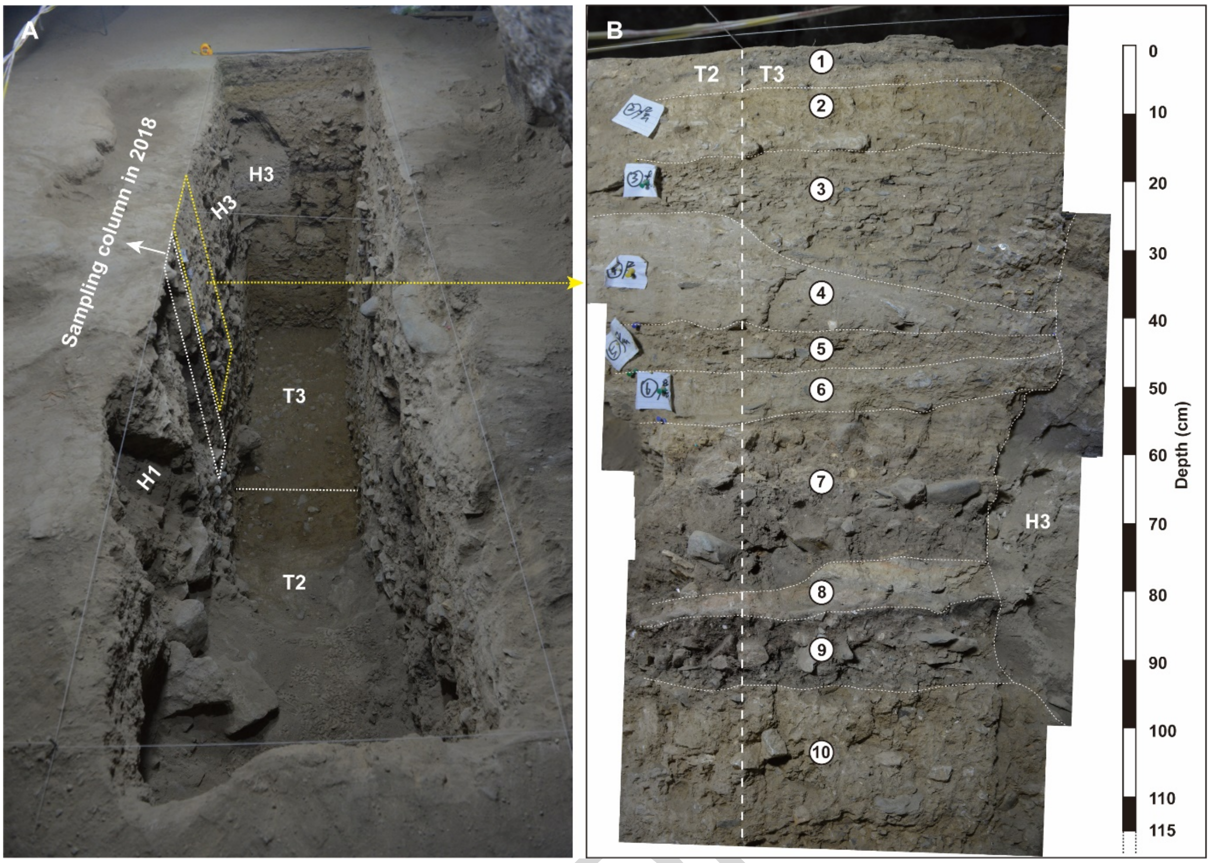

Fig. S2. Excavation units in 2018 and 2019. (A) Excavation units T2 and T3 and the main sampling column. (B) Southwest wall of T2 and T3. Two historic pit features (H1 and H3) are indicated. Layer numbers are denoted by circled numbers in B.

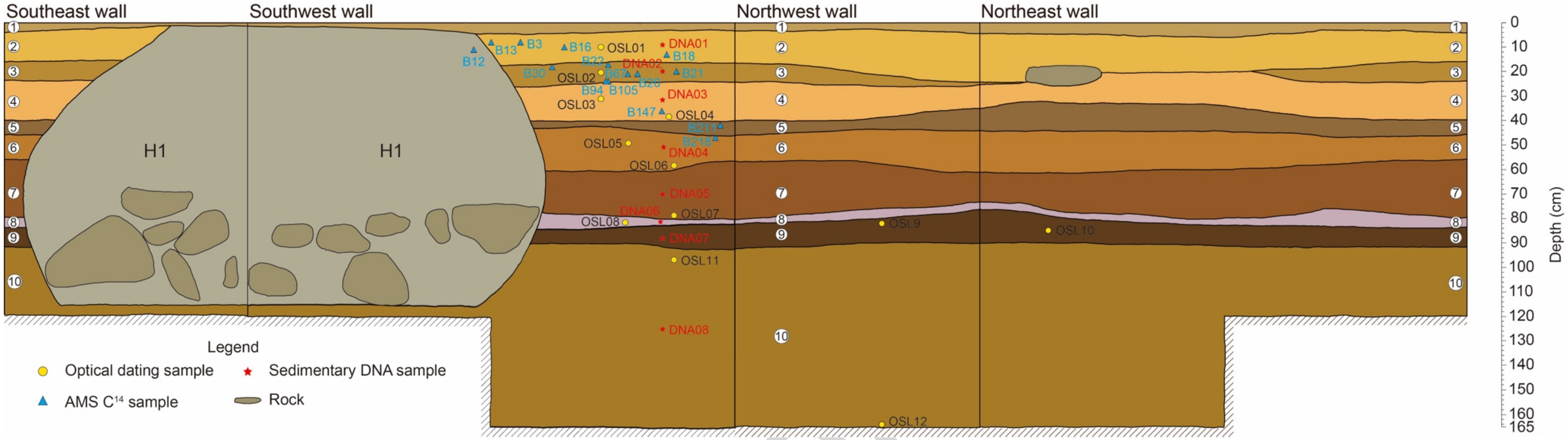

Fig. S3. Stratigraphic drawing for the four wall profiles of T2, together with locations and codes for optical dating, radiocarbon dating and sedimentary DNA samples.

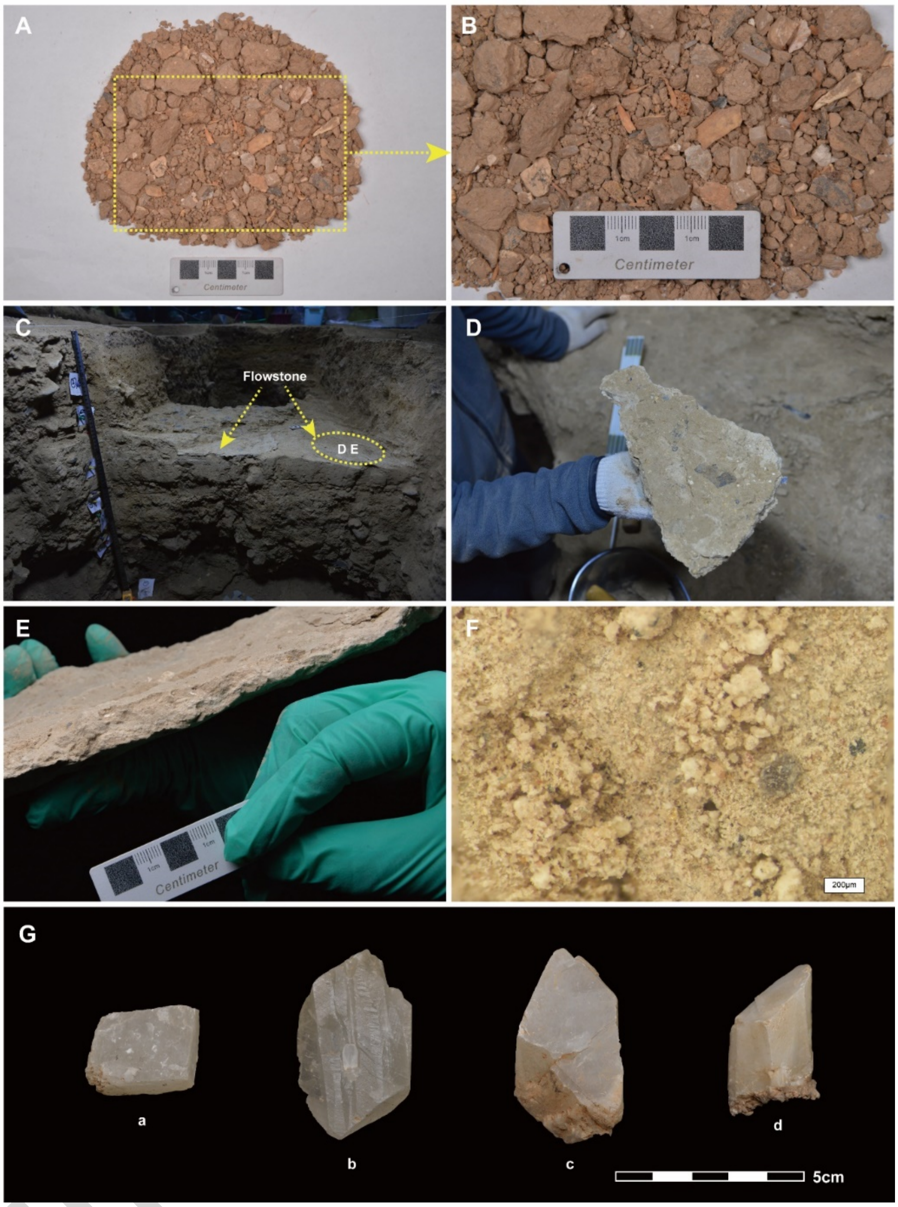

Fig. S4 Sediments from Layer 5, flowstone at the bottom of Layer 3 and calcite crystals from the excavation. (A) Sediments from Layer 5 showing sub-rounded soil aggregates, calcite crystals and other clasts. (B) Detailed view of the sediments shown in A. (C) Flowstone at the bottom of Layer 3 in T3. (D) One piece of the flowstone collected in T3 in 2019. (E) Cross section of the collected flowstone sample. (F) Microscopic photograph of the collected flowstone sample using the OLYMPUS SZX2-ILLTQ Optical Microscope (Magnification is 25.2 times) showing that this flowstone consists mainly of carbonate crystal particles. (G) Different types of calcite crystals found during the excavation. Hombohedral calcite crystal (a, Layer 10), scalenohedral calcite crystal (b and c, Layer 10) and prismatic calcite crystal (d, Layer 5).

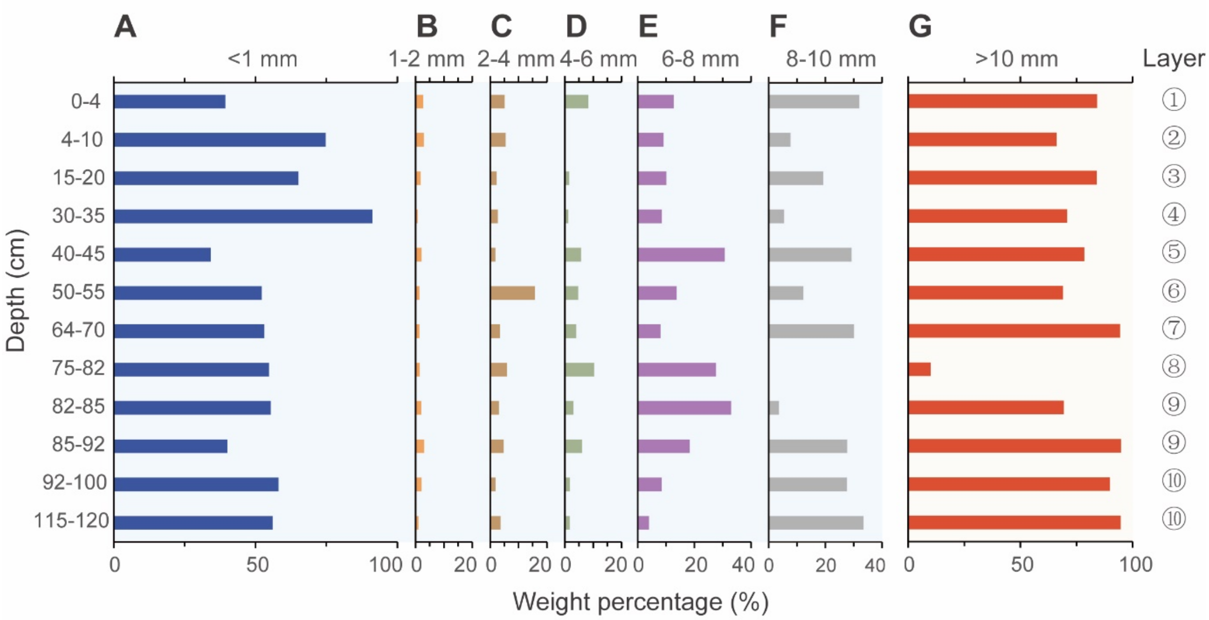

Fig. S5 Weight percentages of different fractions for samples from each layer. (A) <1 mm fraction. (B) 1–2 mm fraction. (C) 2–4 mm fraction. (D) 4–6 mm fraction. (E) 6–8 mm fractions. (F) 8–10 mm fraction. (G) >10 mm fraction. A-F were obtained from sieving results of a small amount of sediments smaller than 10 mm from selected samples. G was obtained from sieving results of the whole selected samples. More information about the different fractions can also been seen in Fig. S6 -10.

Fig. S6 Optical microscopic images of clasts (1-10 mm) from sieved samples for each layer.
The black scale bars in the lower right corner of each photograph represent 1mm.

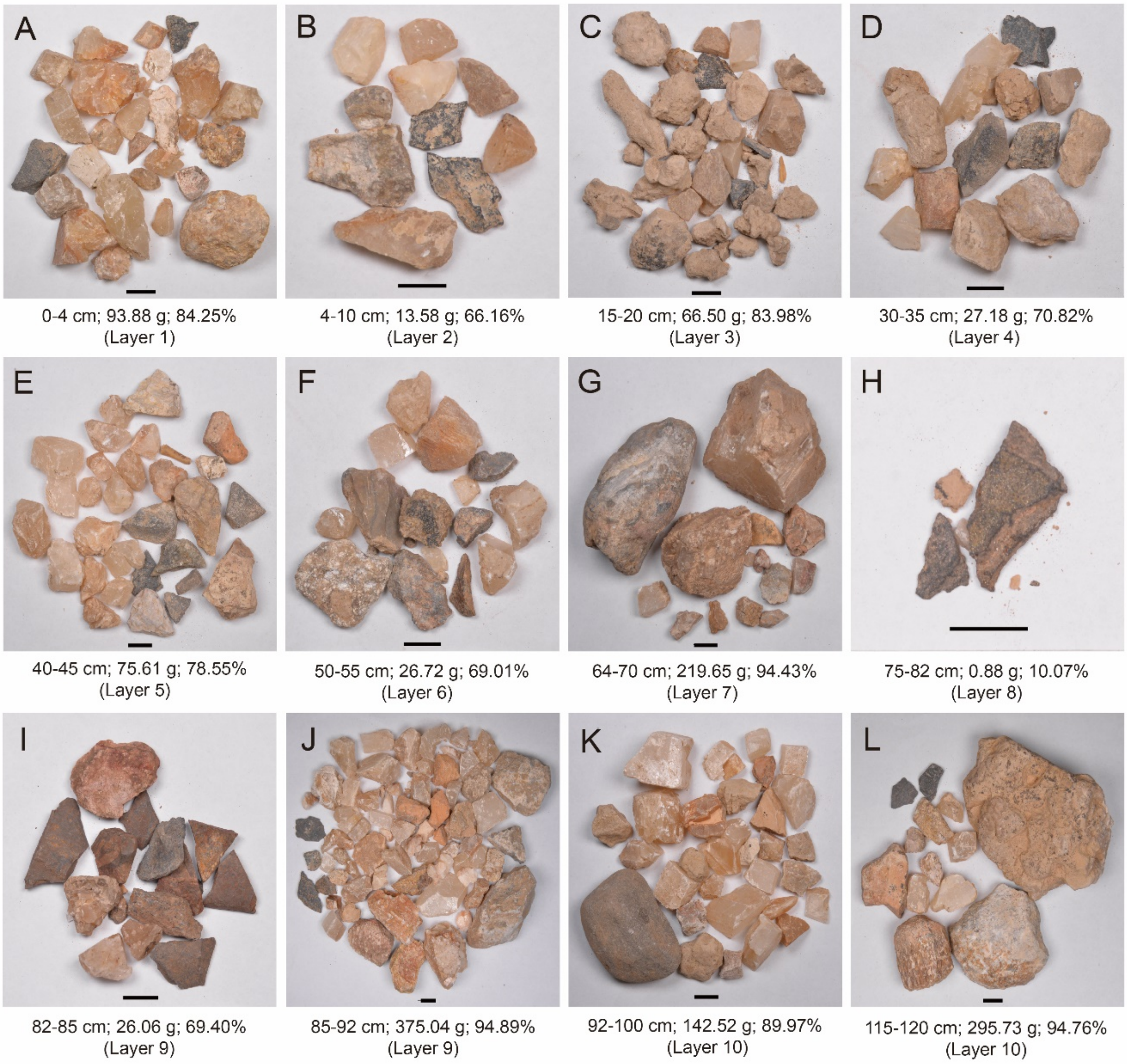

Fig. S7 Photos of clasts (> 10 mm) from sieved samples for each layer. The black scale bars at the bottom of each photograph represent 10 mm.

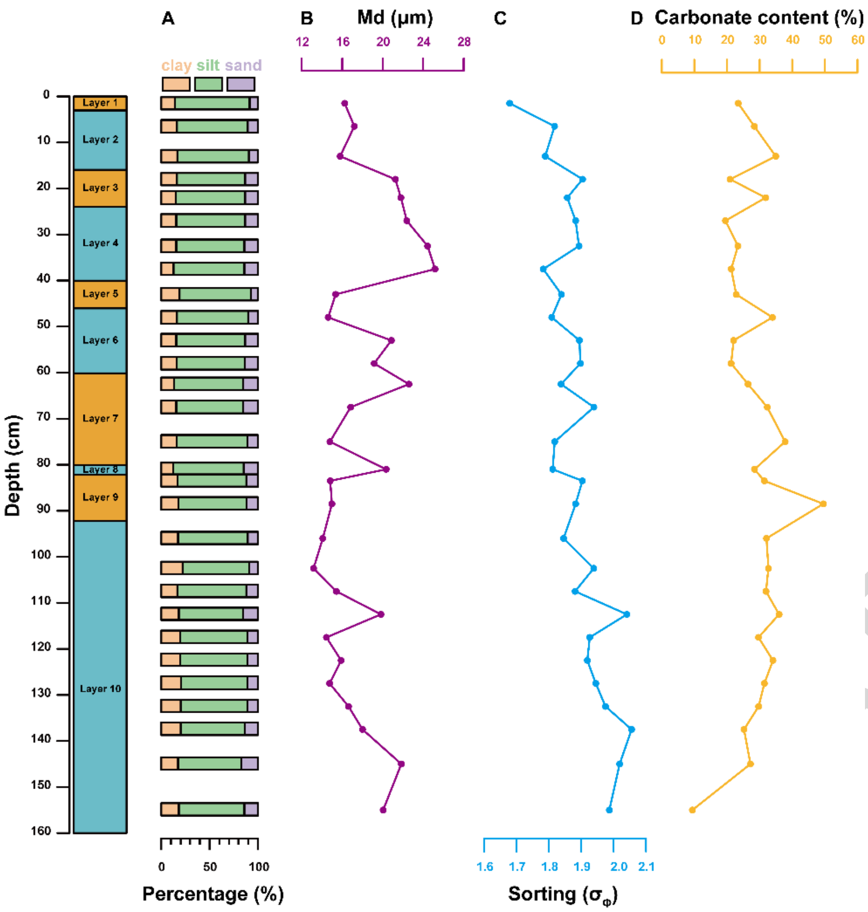

Fig. S8 Grain size and carbonate content results. (A) Percentages of sand, silt and clay in each measured sample. (B) Median grain size. (C) Sorting index (graphic standard deviation). (D) Carbonate content. The flowstone layer at the bottom of Layer 3 is indicated by shaded red box in the lithology graph. The stratigraphic layers are coloured alternately for illustration purpose only, and do not reflect any sedimentary features.

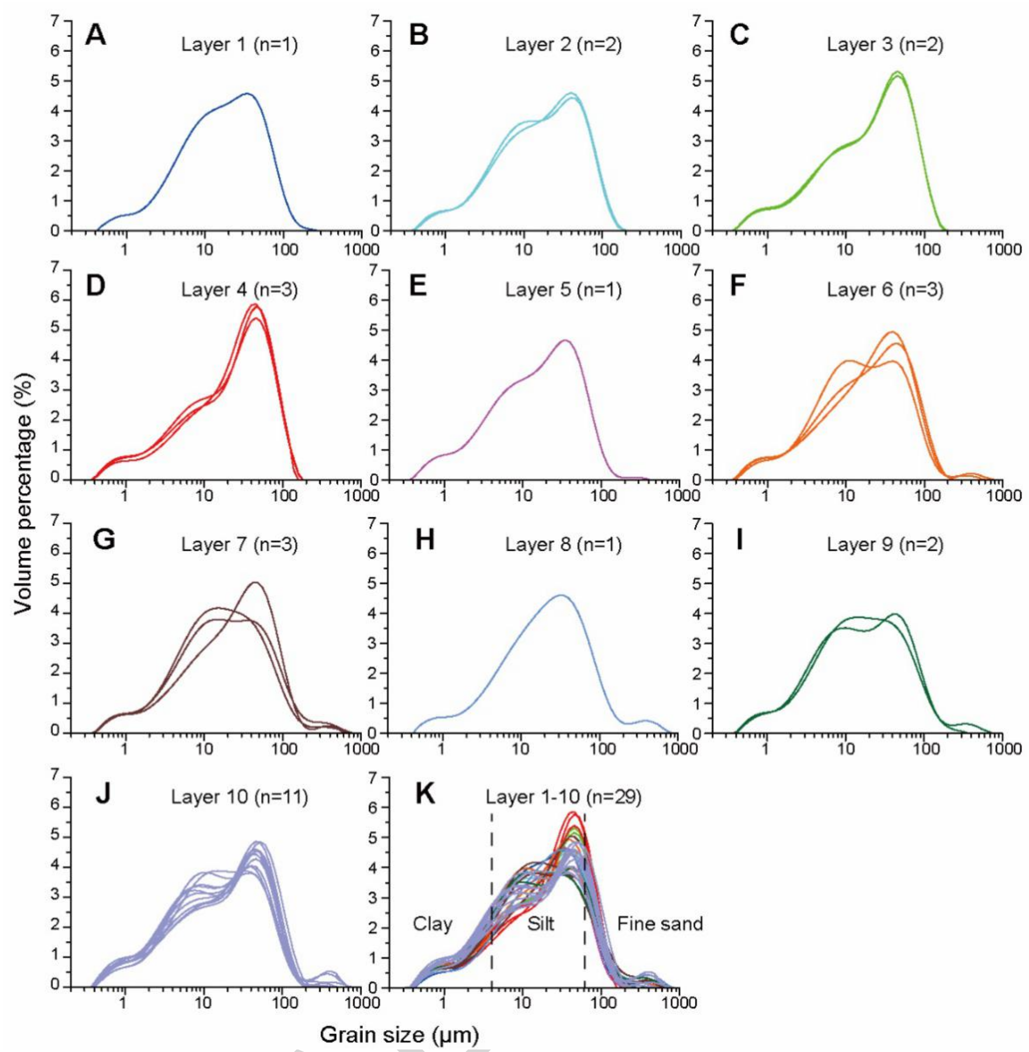

Fig. S9 Grain size frequency distributions of fine sediments (<1 mm) for samples from each layer. Distributions of clay, silt and fine sand are indicated in K.

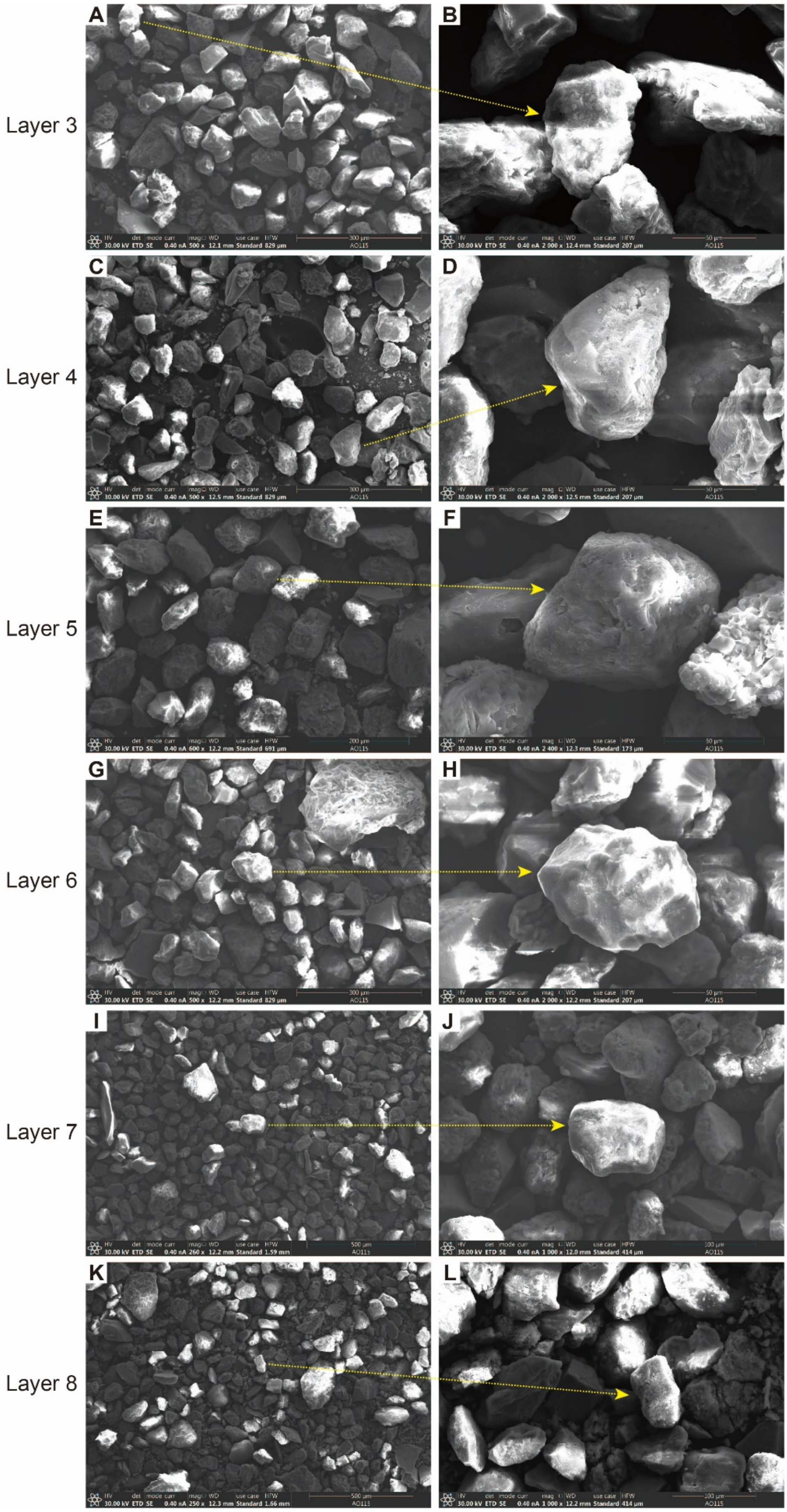

Fig. S10 SEM images of fine sediments for selected samples from Layer 3-8.

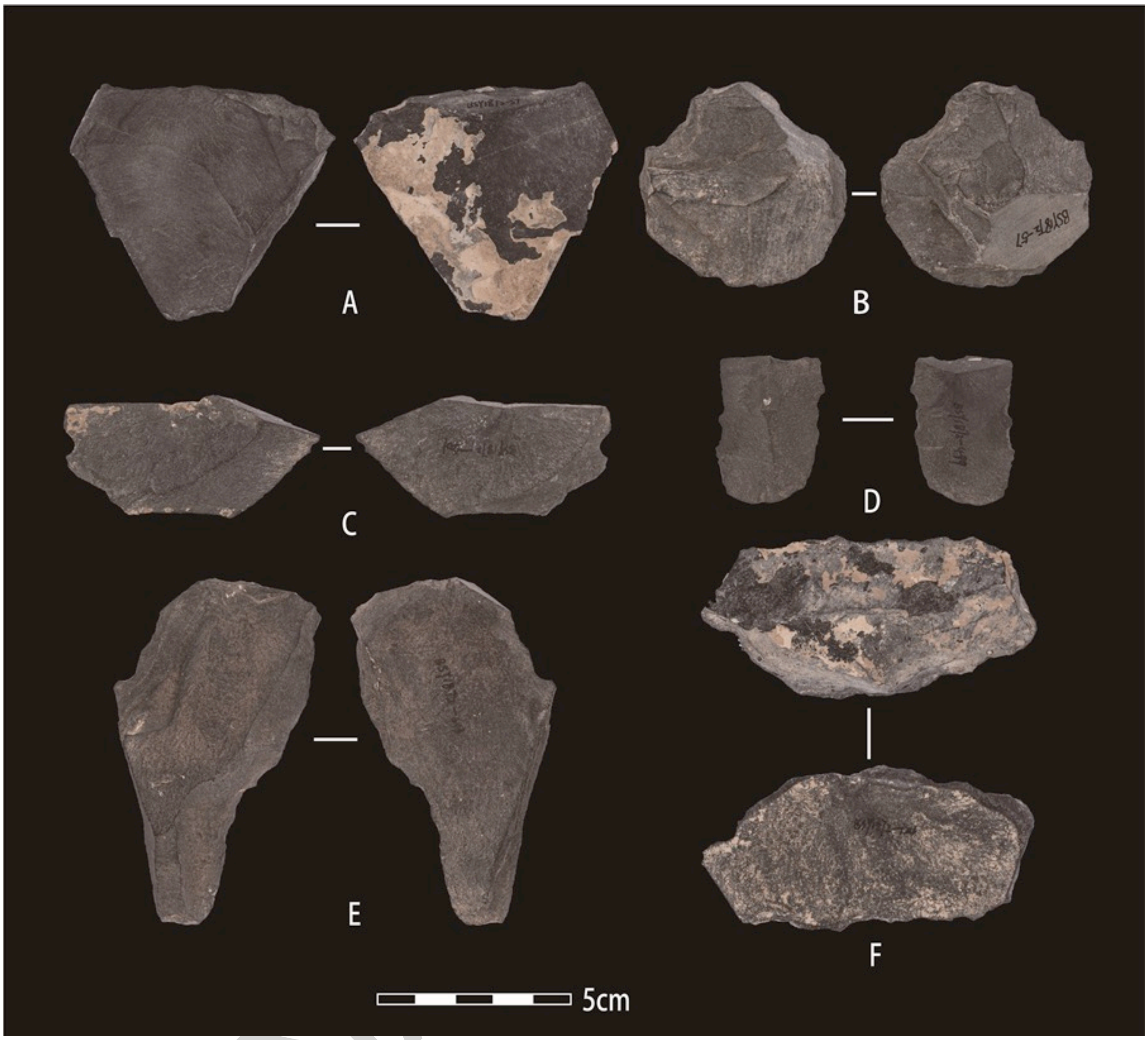

Fig. S11. Selected stone artifacts from T2. (A) Flake (Layer 2). (B) Discoidal core (Layer 2). (C) Flake (Layer 4). (D) Flake (Layer 6). (E) Flake tool (Layer 7). (F) Flake (Layer 3).

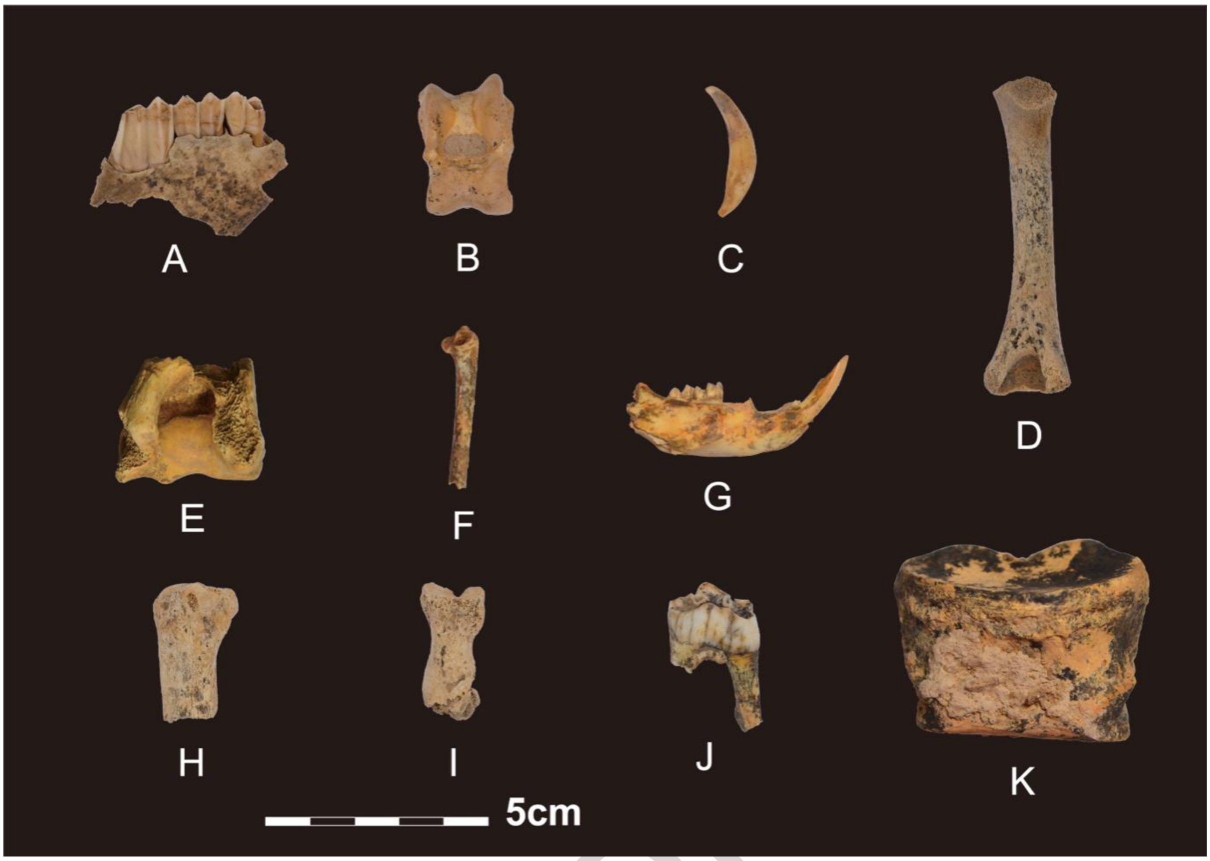

Fig. S12. Selected animal fossils from T2. (A) Gazelle (*Procapra sp.*) maxilla (Layer 2). (B) Left gazelle (*Procapra sp.*) astragalus (Layer 2). (C) Fox (*Vulpes sp.*) canine (Layer 3). (D) Unfused left gazelle (*Procapra sp.*) humerus (Layer 3). (E) Left small bovids humerus (Layer 4). (F) Right bird ulna (Layer 5). (G) Right Himalayan marmot (*Marmota himalayana*) mandible (Layer 7). (H) Hyena (*Crocuta sp.*) metacarpal II (Layer 10). (I) Hyena (*Crocuta sp.*) phalanx I (Layer 10). (J) Left large bovids mandibula P4 (Layer 10). (K) Left rhinoceros (*Coelodonta sp.*) phalanx I (Layer 10).

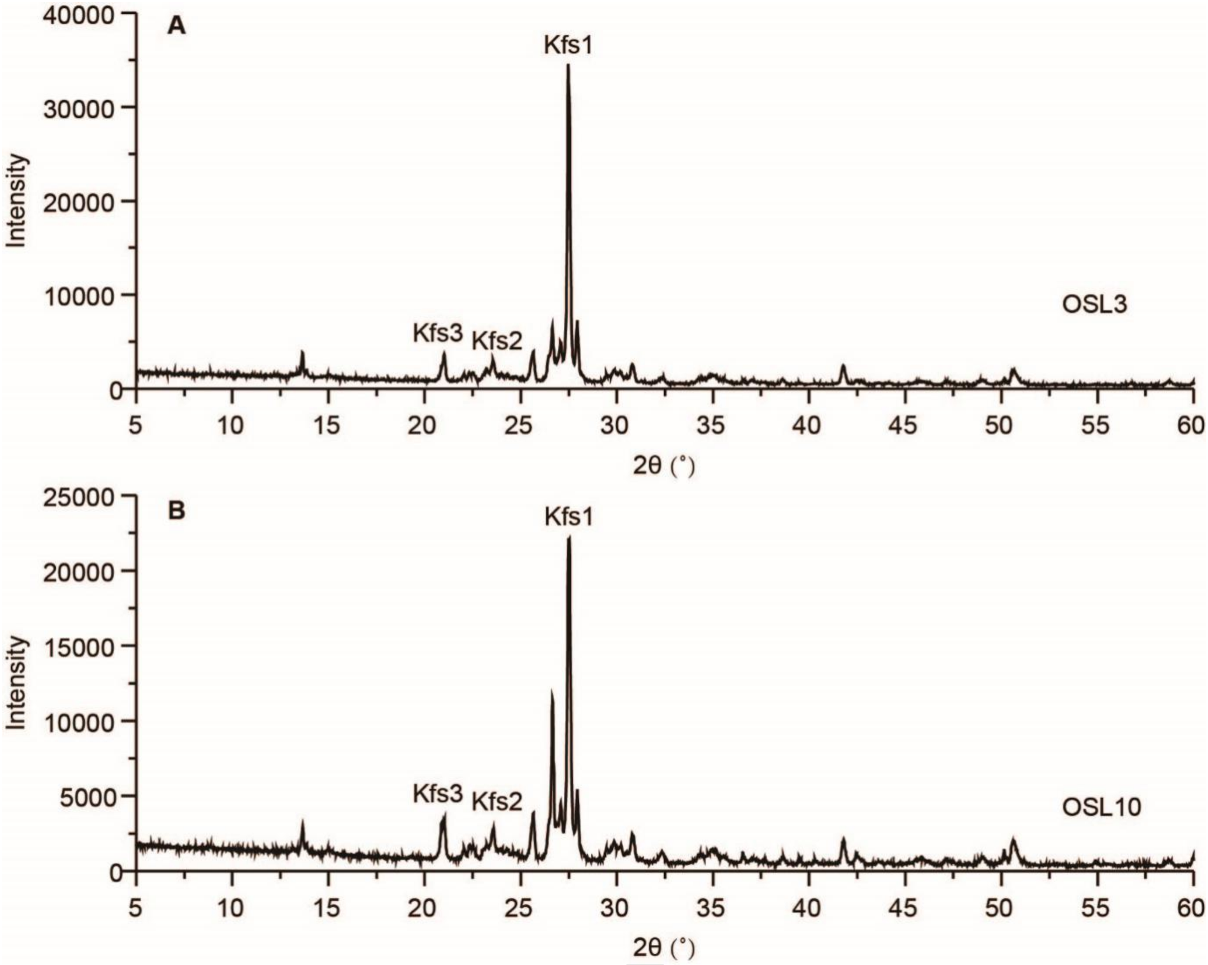

Fig. S13. XRD results of density-separated K-rich feldspar fractions of OSL3 (A) and OSL10 (B). The standard error is ~5%. Kfs represents the strongest K-feldspar peaks observed.

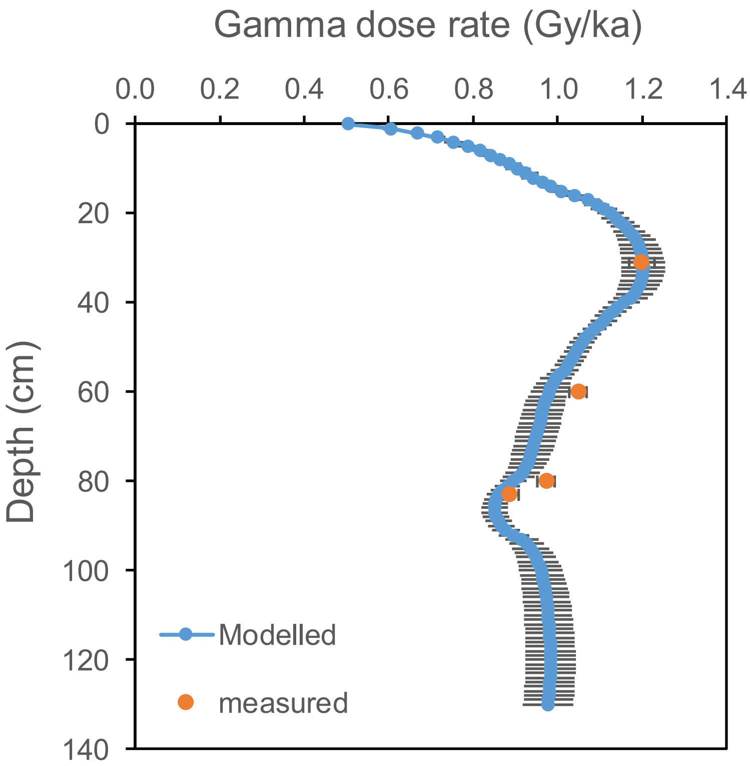

Fig. S14. Modelled and in-situ gamma dose rate gradient for the southwest profile of T2.

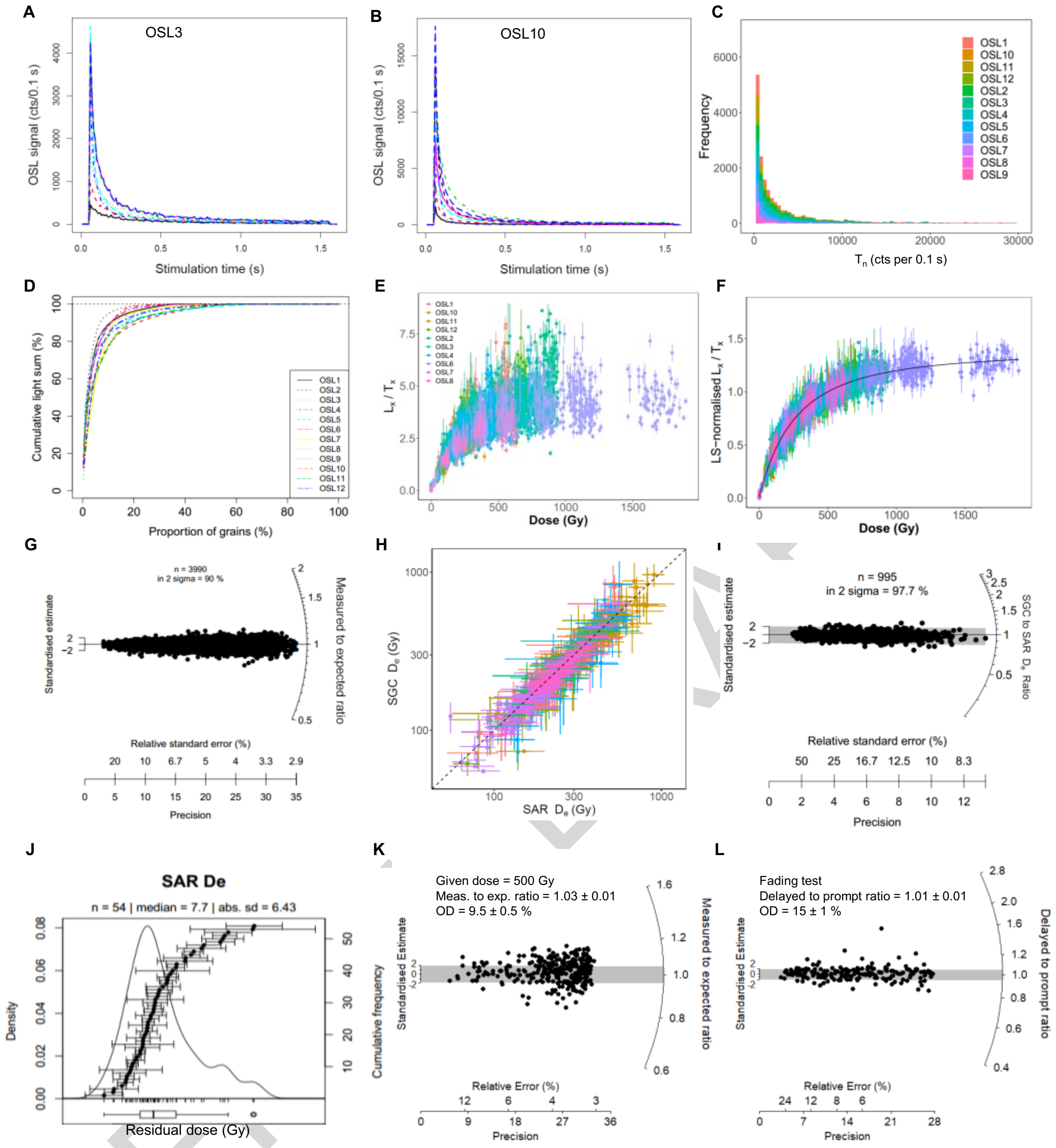

Fig. S15. Luminescence characteristics of K-feldspar grains. Representative pIRIR decay curves of 6 K-feldspar grains from OSL3 (A) and OSL10 (B). (C) Histogram showing the pIRIR T_n signal intensity distribution for individual K-feldspar grains from different samples. (D) Cumulative light sum plot of single-grain K-feldspar pIRIR signals for different samples. (E) Single-grain pIRIR dose response curves for different samples shown in different colours. (F) LS-normalized SGC based on the data shown in E. (G) Radial plot showing the ratio of measured and expected signals obtained from the SGC curve in F. (H) Comparison between SAR D_e and SGC D_e values. The broken line represents the 1:1 ratio. (I) Radial plot showing the ratios between the SGC and SAR derived D_e values. (J) Kernel density curve and cumulative frequency plot showing the residual doses of OSL6. (K) Radial plot showing the distribution of signal-recovery ratios obtained for dose recovery tests for given dose of 500 Gy. (L) Delayed to prompt signal ratios (‘fading ratio test’) for individual grains from sample OSL6.

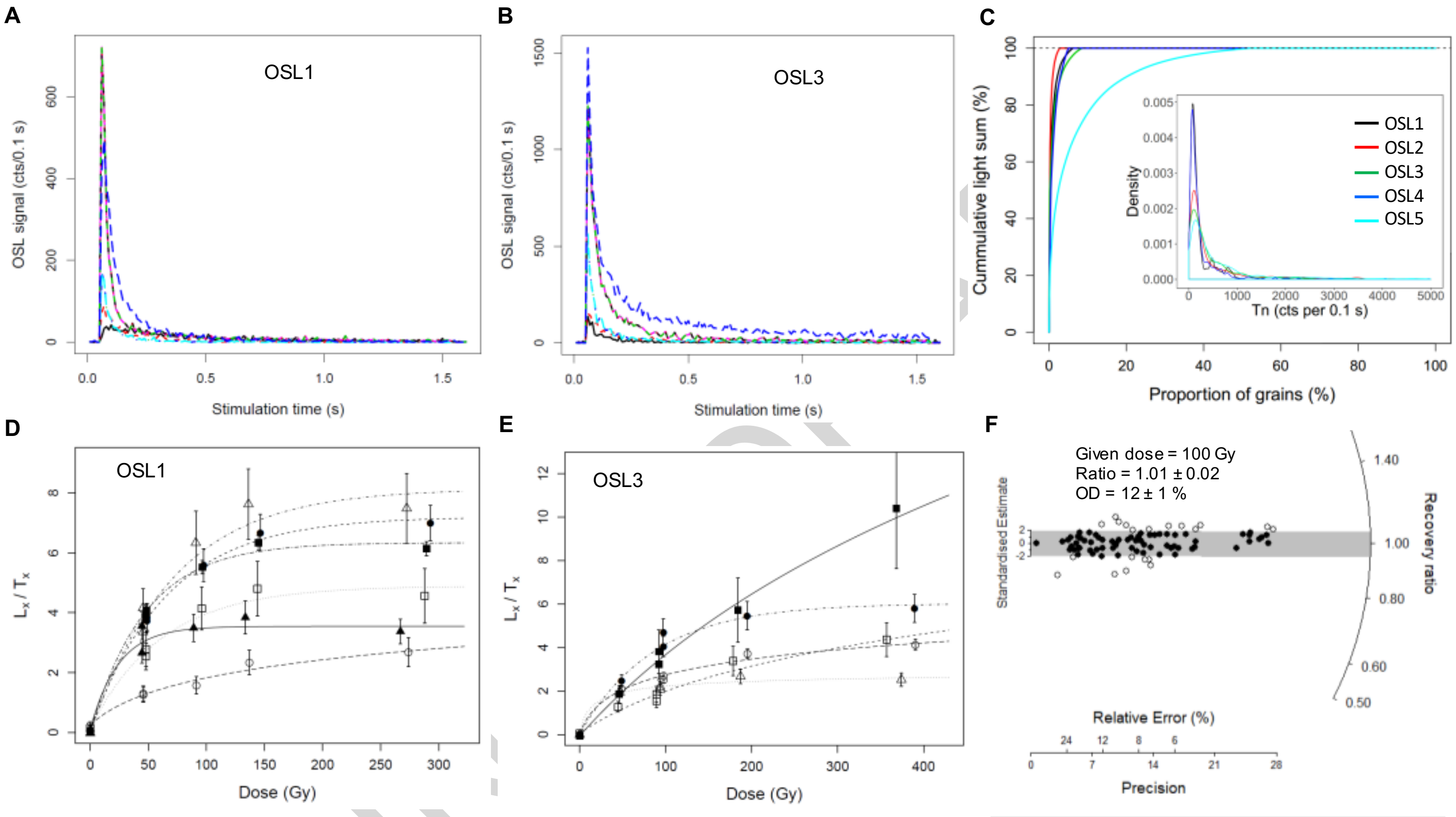

Fig. S16. Luminescence characteristics of quartz. Representative OSL decay curves of 6 quartz aliquots from OSL1 (A) and OSL3 (B). (C) Cumulative light sum plot of small-aliquot quartz for different samples. The inset shows the density probability distribution of T_n intensity of the OSL signals for individual aliquots. (D–E) Quartz OSL dose response curves of different aliquots from OSL1 and OSL3. (F) Radial plot showing the distribution of signal-recovery ratios obtained from the simplified dose recovery test.

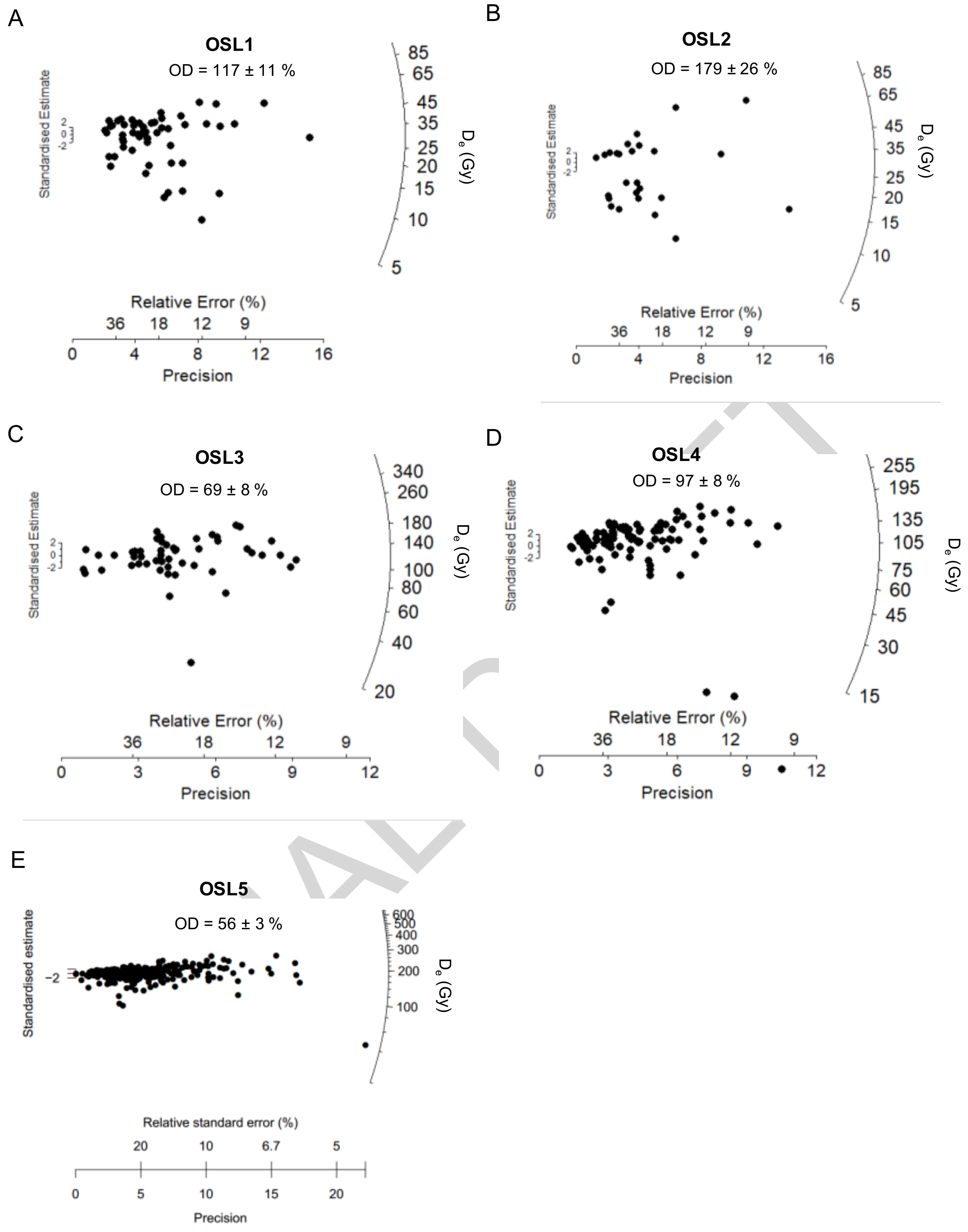

Fig. S17. SAR D_e distributions for all quartz samples.

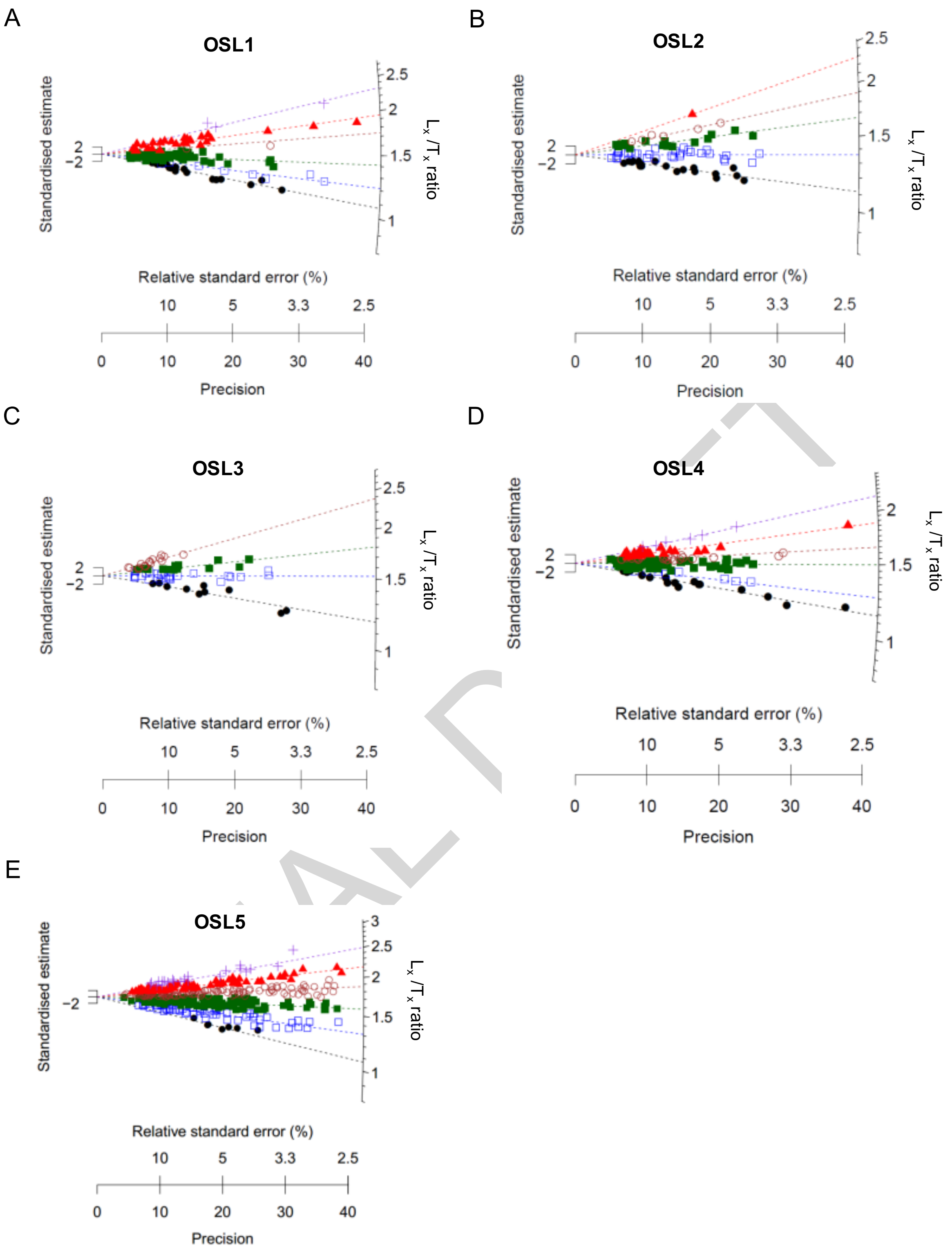

Fig. S18. Radial plots showing the distributions of the ratios of L_x/T_x values between two regenerative doses of 300 and 80 Gy for all accepted aliquots from different quartz samples. The different colours and symbols represent different groups of grains identified using the FMM.

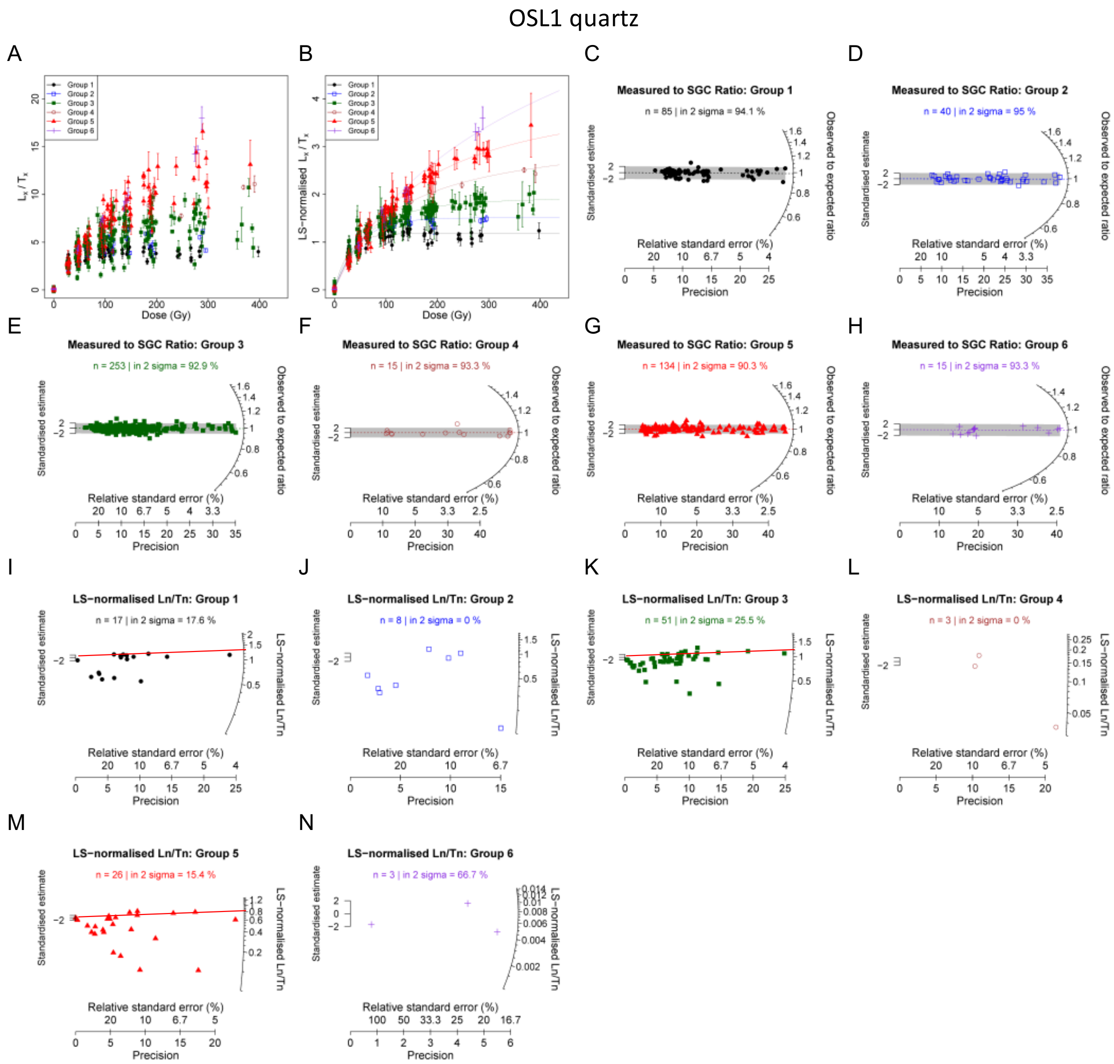

Fig. S19. Quartz OSL SGC results and L_n/T_n distributions for all the grains that passed rejection criteria 1–3 (Table S7) for sample OSL1. (A) Comparisons of all DRCs that pass the rejection criteria 1–3. Different colours and symbols represent different groups of grains identified using FMM. (B) Comparison of the LS-normalized L_x/T_x values for different groups. The data set for each group was fitted using a GOK function (full lines) and then normalized to unity at 50 Gy. (C–H) Radial plots showing the ratios between the LS-normalizsed L_x/T_x and expected values based on the best-fit SGC shown in (B); the shaded band captures 2σ range from unity. The total number of grains (n) and percentage falling inside the 2σ band are shown for each group. (I–N) Radial plots showing the LS-normalized natural signals (L_n/T_n) of different groups. The full lines represent the best estimate weighted mean L_n/T_n values for individual groups (see text for description of age models used). All error bars in panels A and B represent 1σ standard errors. All figures and data analyses are based on building functions in R packages "Luminescence" (*126*) and "numOSL" (*76*).

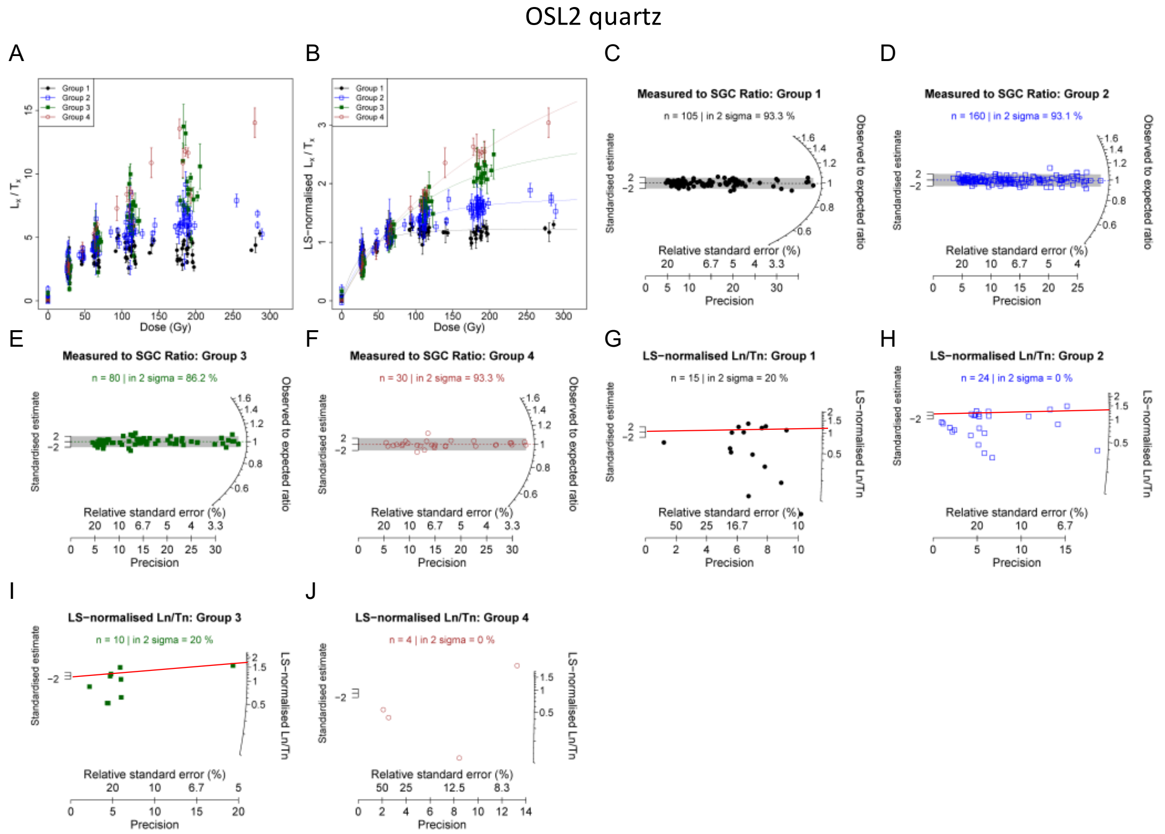

Fig. S20. Quartz OSL SGC results and L_n/T_n distributions for sample OSL2. (A–J) See Fig. S11.

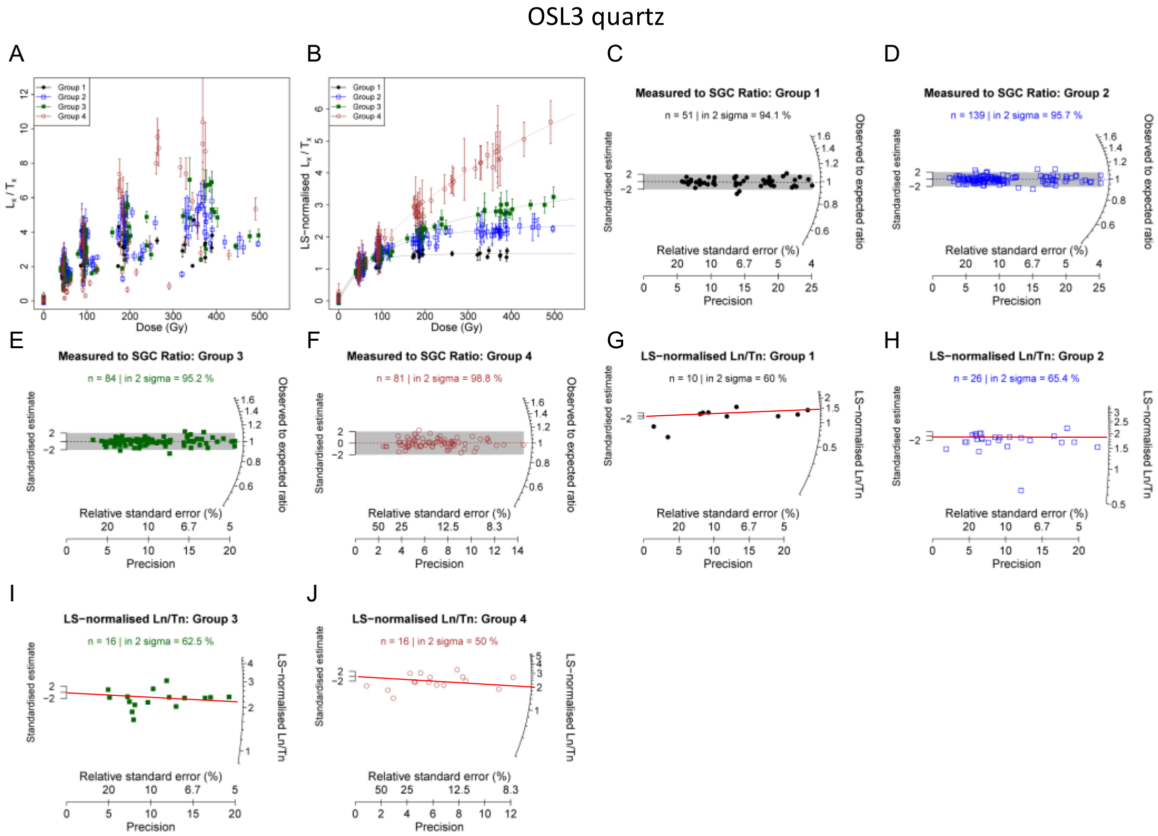

Fig. S21. Quartz OSL SGC results and L_n/T_n distributions for sample OSL3. (A–J) See caption for Fig. S11.

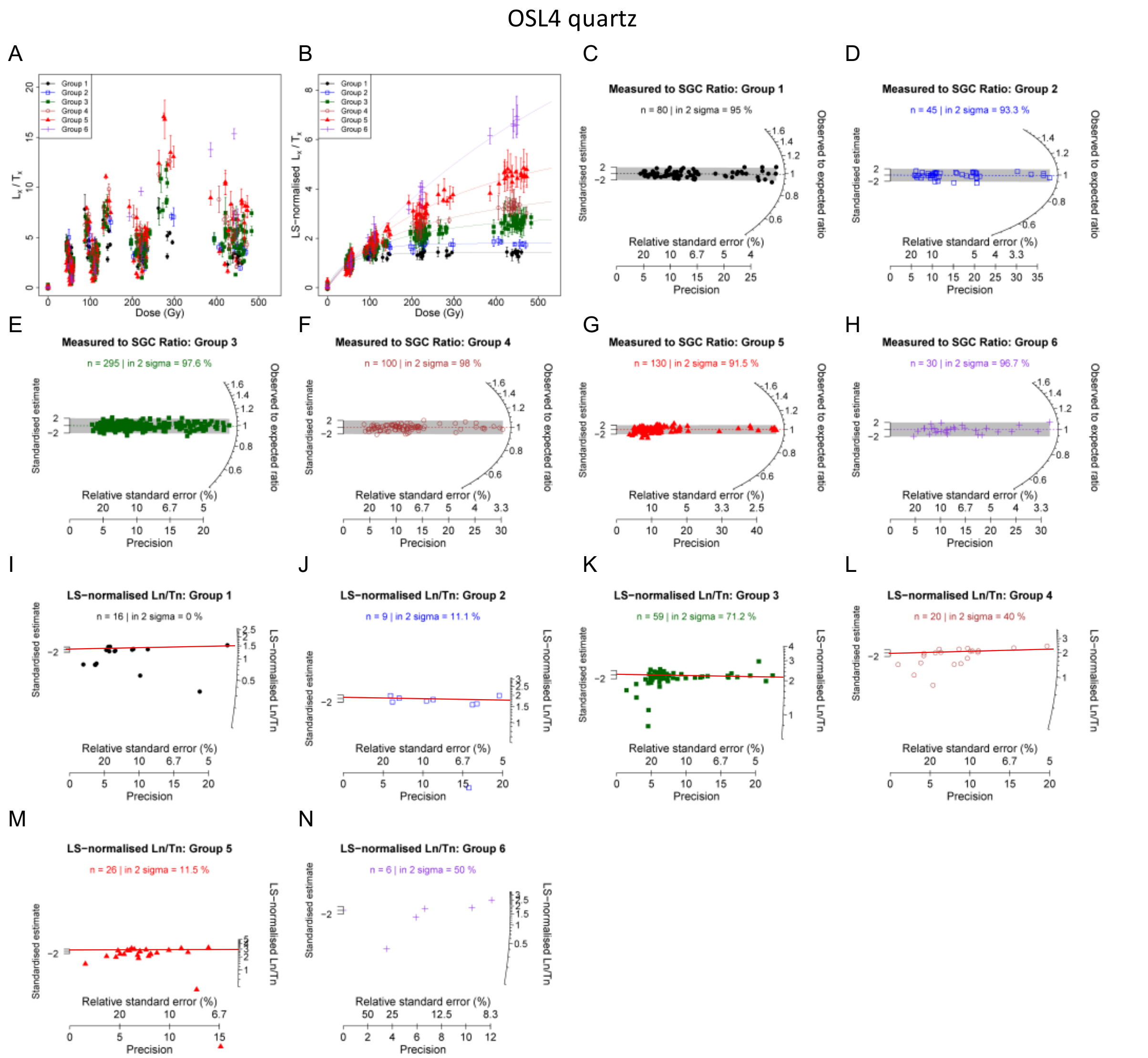

Fig. S22. Quartz OSL SGC results and L_n/T_n distributions for sample OSL4. (A–N) See caption for Fig. S11.

OSL5 quartz

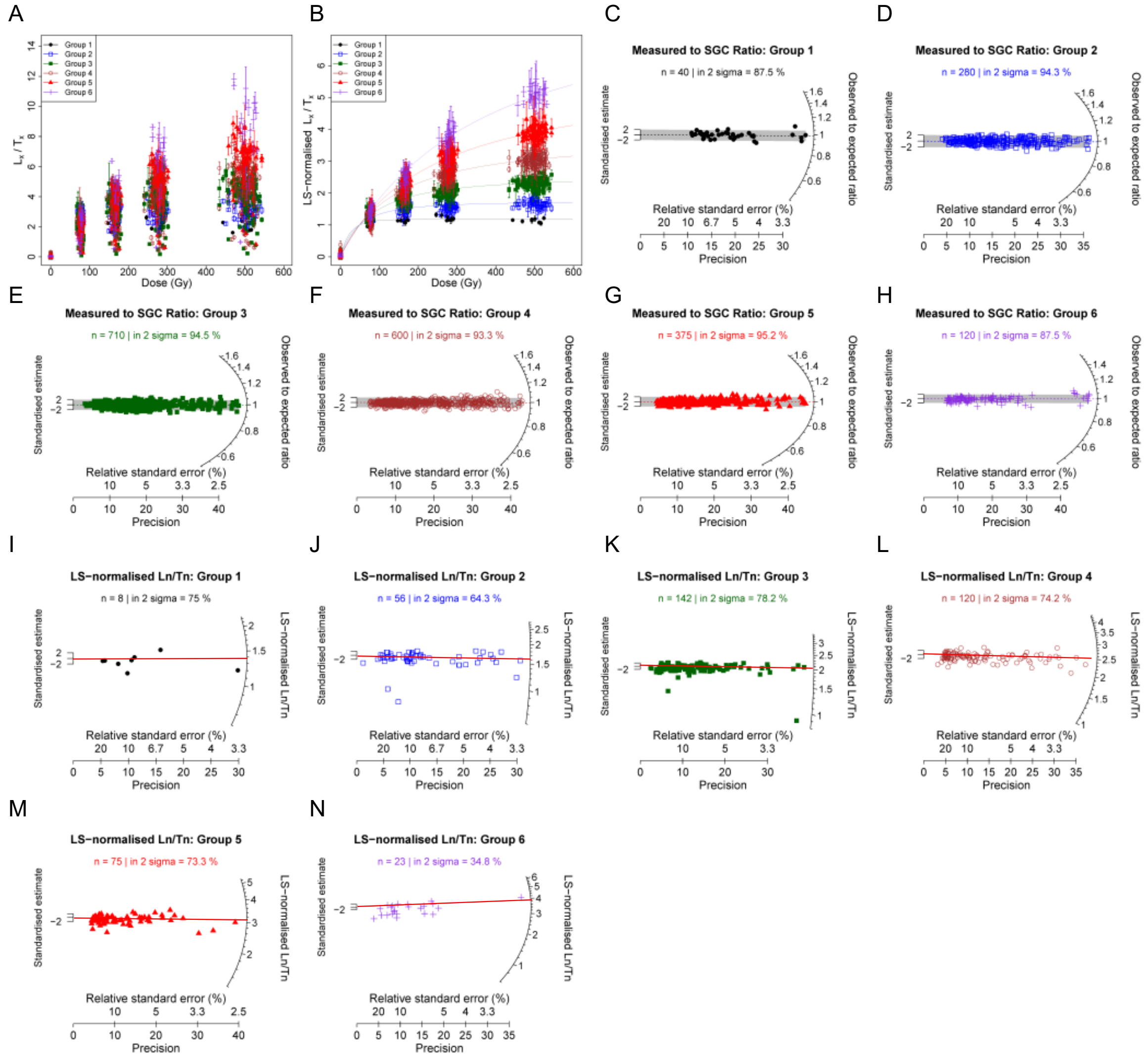

Fig. S23. Quartz OSL SGC results and L_n/T_n distributions for sample OSL5. (A–N) See caption for Fig. S11.

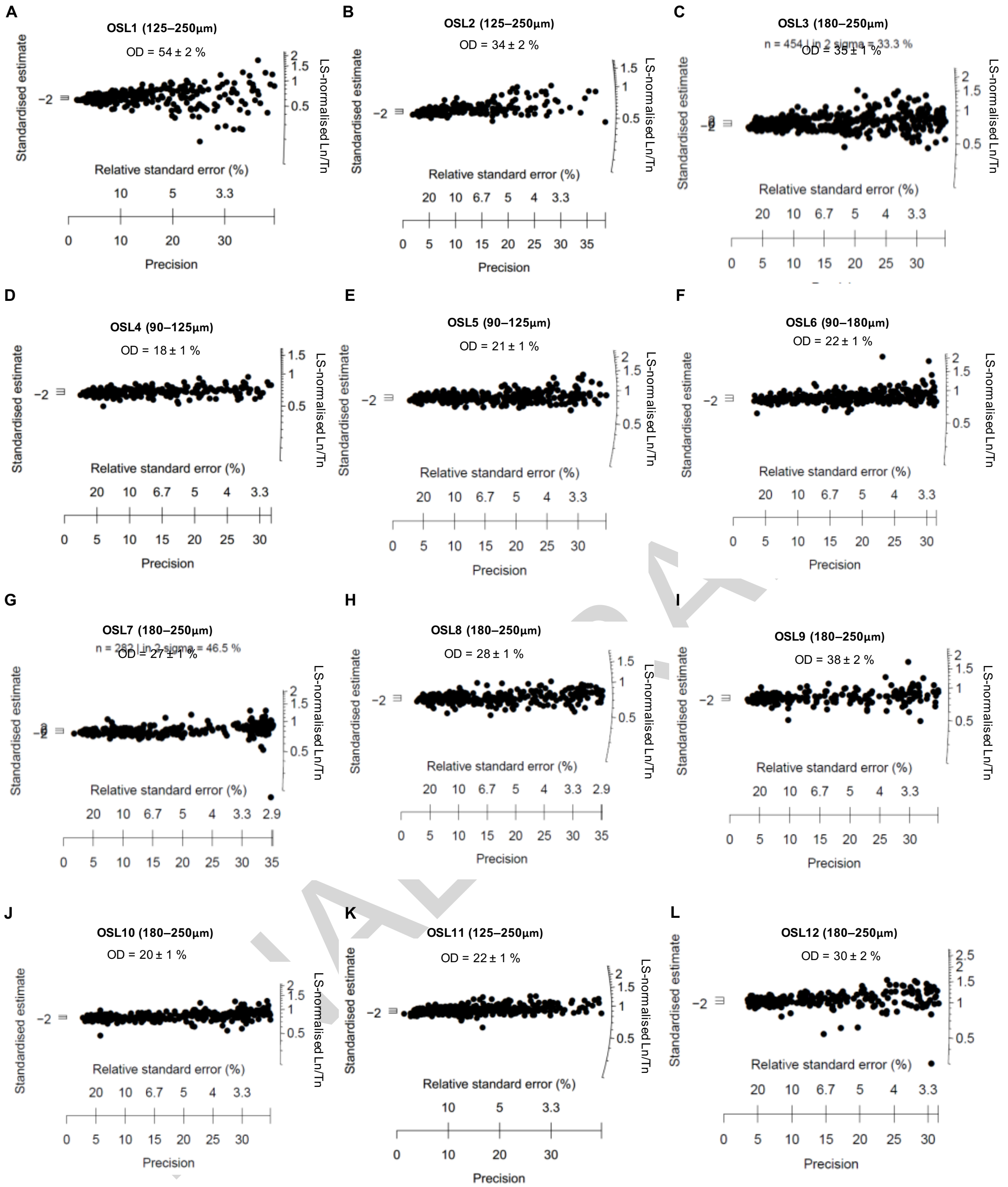

Fig. S24. pIRIR LS-normalized L_n/T_n distributions and over-dispersion (OD) values for K-feldspar grains from all samples. A LS-normalised L_n/T_n ratio of 1 is equivalent to 500 Gy.

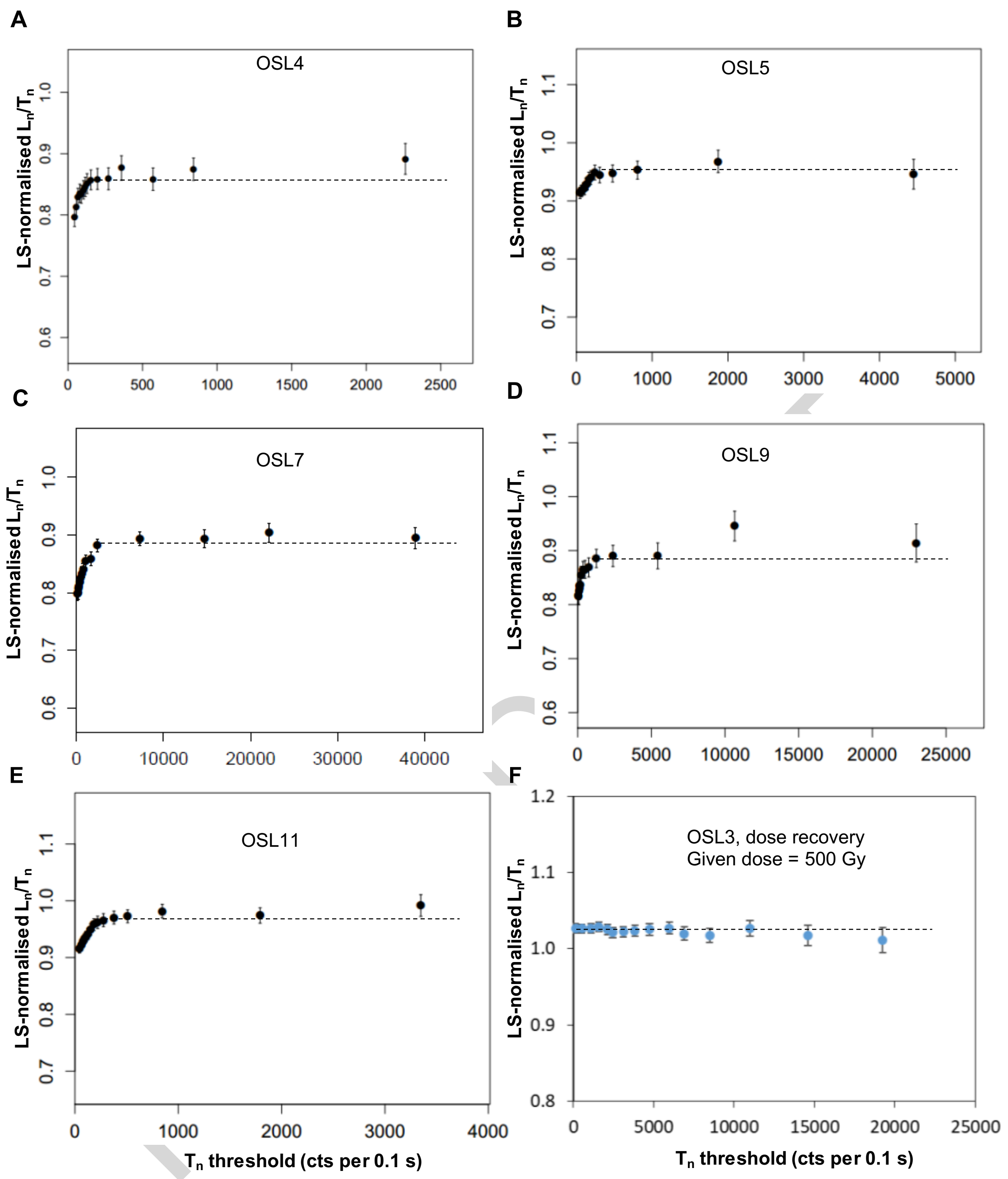

Fig. S25. CAM re-normalized L_n/T_n ratios plotted as a function of T_n threshold for five natural samples (A–E) and dose recovery (laboratory-irradiated) sample (F). The dashed lines represent the normalized L_n/T_n ratio 'plateau'. All error bars are at 1σ.

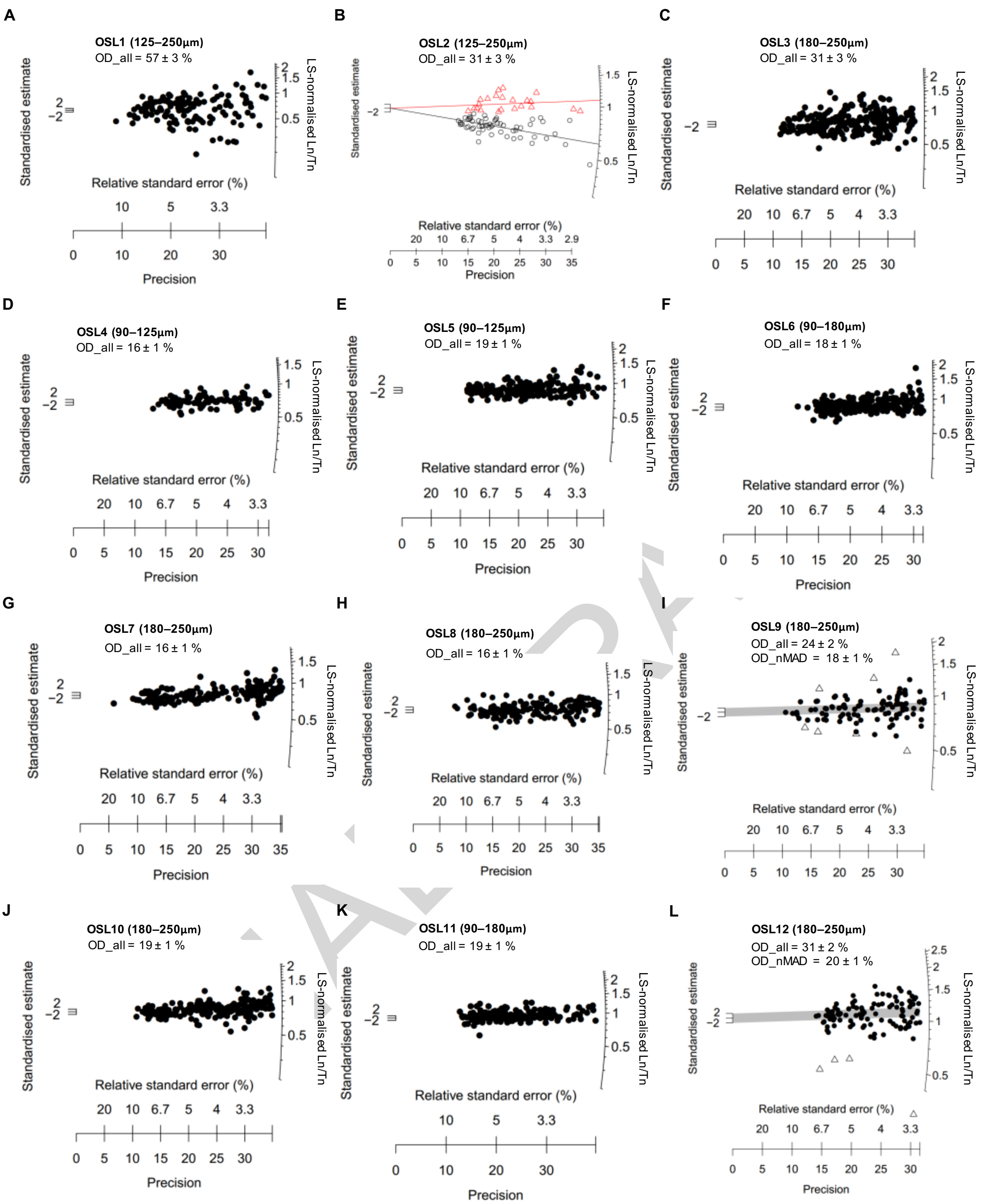

Fig. S26. Distributions of re-normalized L_n/T_n ratios for K-feldspar grains after applying a T_n threshold of 500 cts/0.1 s of optical stimulation time. The CAM was applied to the distributions in D–H, J and K to obtain a weighted mean L_n/T_n value. The nMAD was applied to the distributions in OSL9 (I) and OSL12 (L) to identify and reject a small number of intrusive grains (open circles) prior to application of CAM. Panels A and C show widely dispersed and continuous L_n/T_n distributions, so no reliable L_n/T_n value could be obtained. The distribution in B consists of two discrete components (shown in different colours), for which weighted mean L_n/T_n values were estimated using the FMM. A LS-normalised L_n/T_n ratio of 1 is equivalent to 500 Gy.

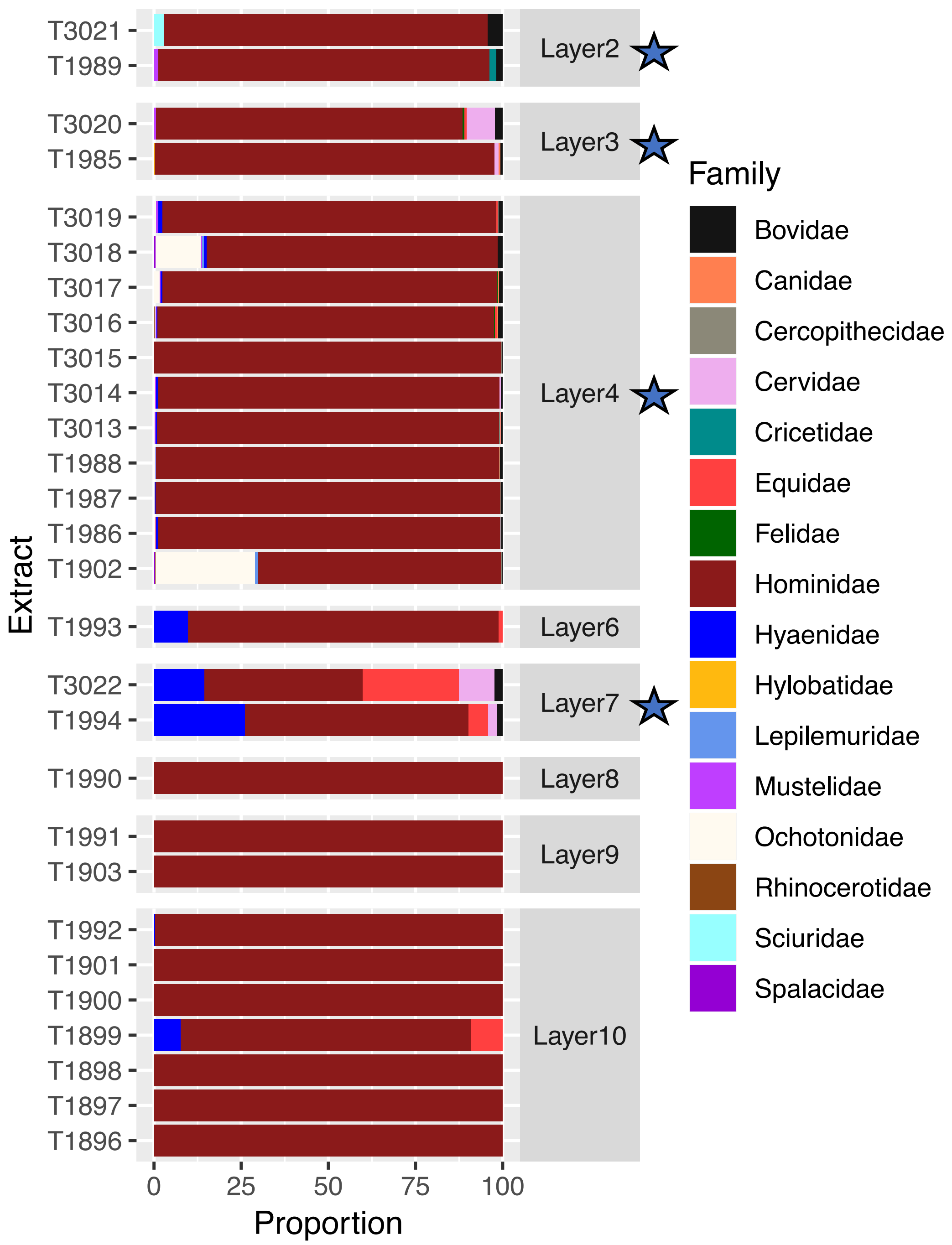

Fig. S27. The proportion of DNA fragments in each extract that were assigned to one of the listed mammalian families. The frequencies of terminal C to T substitutions when compared to the human reference genome in DNA fragments assigned to *Hominidae* for all and for putatively deaminated fragments are shown in Table S19. DNA fragments in the libraries from Layers 2, 3, 4, and 7 show C to T substitution frequencies significantly higher than 10% on at least one of their ends when testing all fragments (Table S19). The stars indicate layers from which ancient hominid mtDNA fragments according to this criterion were retrieved. All libraries from the same extract are combined (i.e. T1902 and T1994).

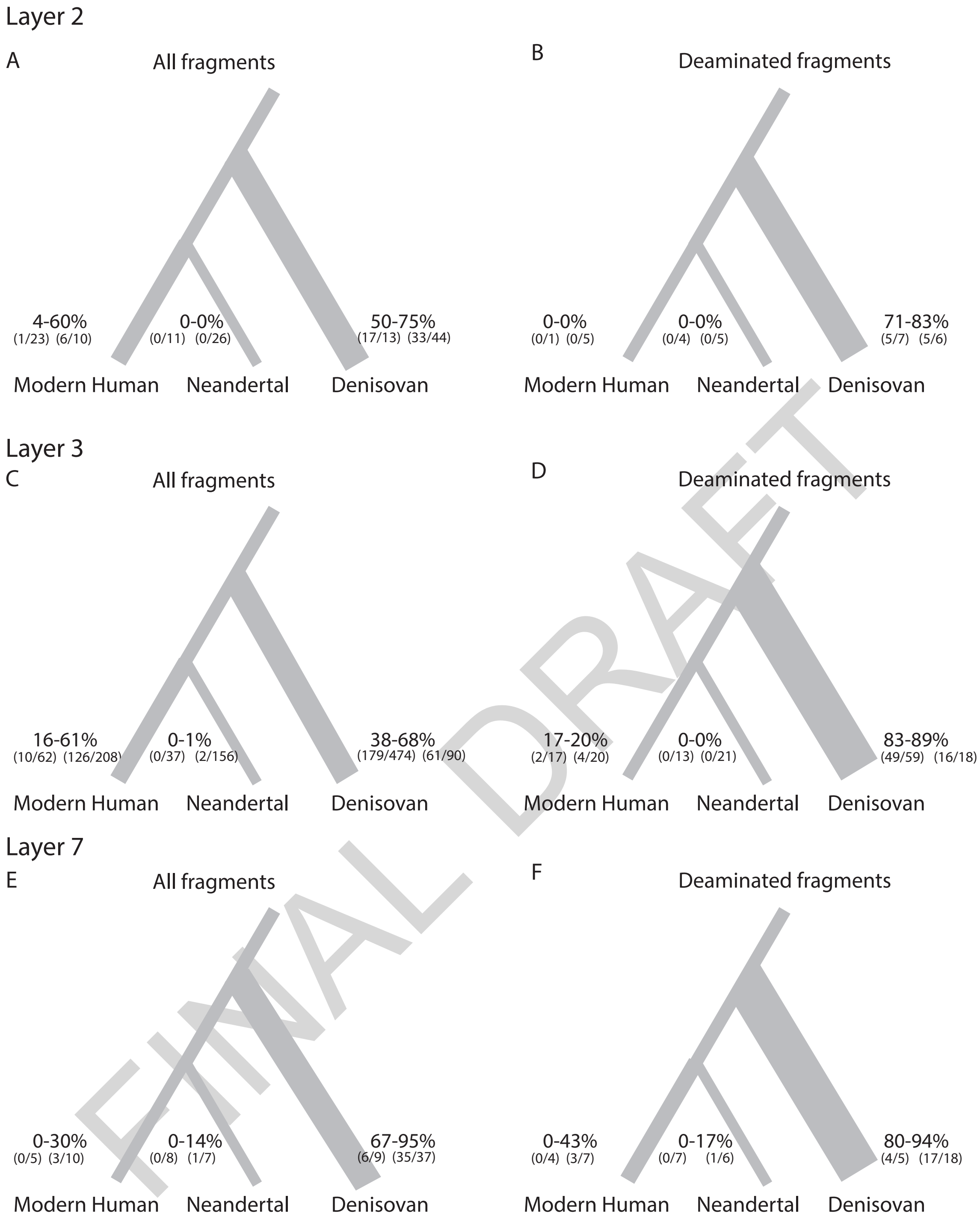

Fig. S28. Lineage inferences using modern human, Neandertal and Denisovan branch-specific substitutions for all fragments (A, C &E) and deaminated fragments (B, D &F) from Layer 2, 3 and 7. Ranges for the percentage of lineage-matching sites for all libraries from each layer are given. The fractions give the absolute number of sequenced fragments carrying derived lineage-specific alleles over the total number of fragments covering positions where such alleles occur.

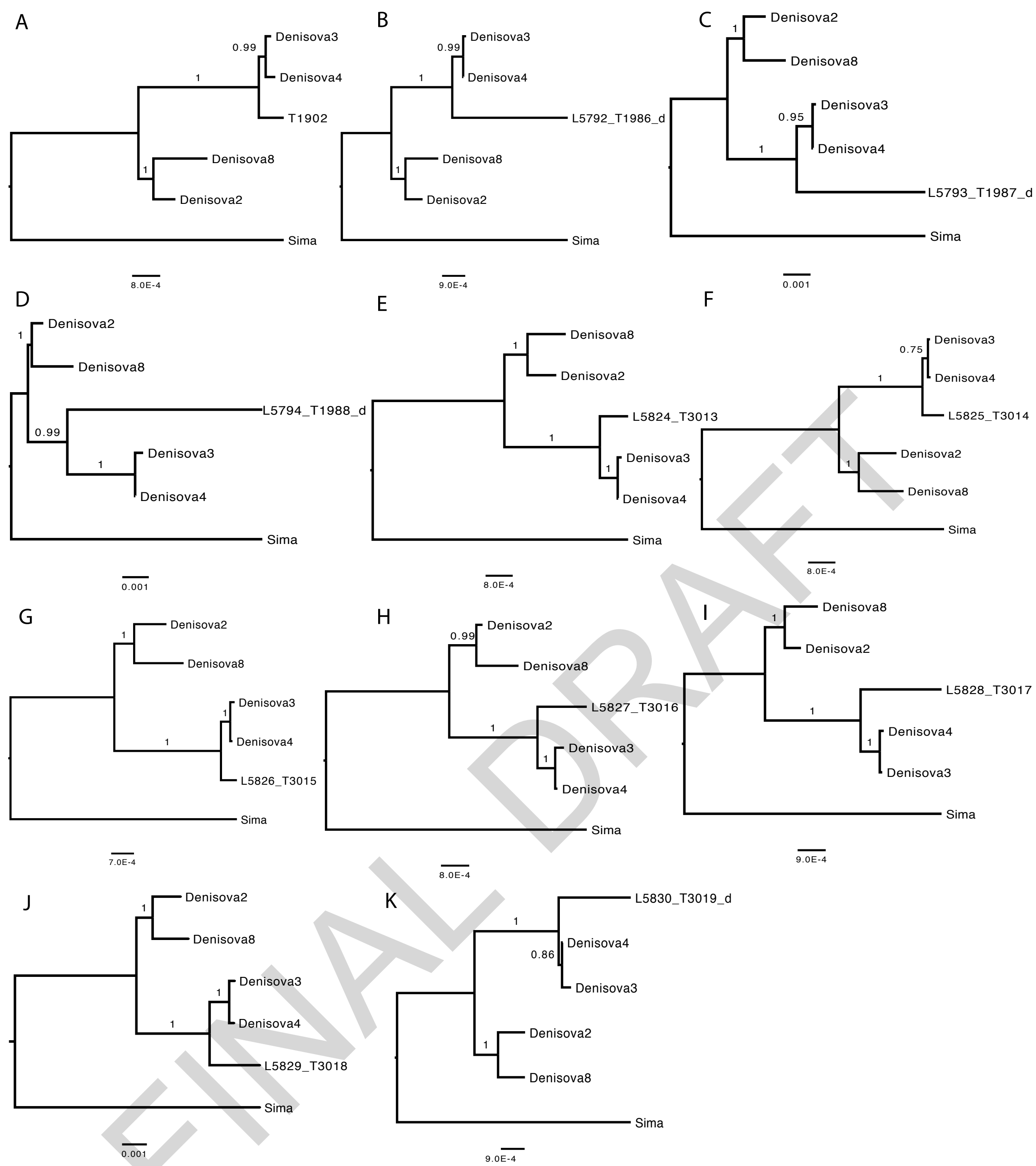

Fig. S29. MtDNA phylogenetic trees for sediment samples from 11 extracts in Layer 4. These trees also include the mitochondrial genomes of four Denisovans and an individual from *Sima de los Huesos* (*Sima*). The phylogeny was estimated with a Bayesian approach under a GTR+Gamma model of sequence evolution. "_d" indicate the consensus is created with the deaminated fragments only. The "T" number indicates this consensus built from which extract from Layer 4.

Table S1. Description of stratigraphic layers of T2. Type of boundary is defined by its width according to Garrison (*127*): abrupt (< 3 cm), clear (3–5 cm), gradual (5–10 cm) and diffuse (>10 cm).

Layer	Thickness (cm)	Description	Munsell color	Lower boundary
1	1–6	Silts with abundant charcoal, calcite crystals, fresh tree branches and grasses, sparse stone artifacts and bone fragments (Fig. S6, S7A, S8A and S9A), and a piece of plastic, suggesting that it is a highly mixed deposits disturbed by recent human activity. The top 1–2 cm is very compact, possibly due to recent human trampling. The lower 2–6 cm is loose and porous.	dry 7.5YR 6/4, light brown	Abrupt
2	12–20	Silts with abundant stone artifacts and bone fragments, and many calcite crystals that presumably fell from the cave roof (Fig. S6, S7B, S8A and S9B). This layer is relatively compact.	dry 10YR 6/4, light yellowish brown	Gradual
3	1–9	Silts with abundant stone artifacts, bone fragments, charcoal, pebbles, limestone fragments, calcareous nodules and calcite crystals (Fig. S6, 7C, S8A and S9C). Cemented pieces of clay aggregates with limestone fragments, stone artifacts and bone fragments, and voids among them are common (Fig. S2B). Medium platy structures of fine sediments could be observed in the lower part of this layer (Fig. S2B). Discontinuous flowstone layers (1–2 cm thick) occur at the bottom of the layer (Fig. S4C–F). This layer is compact and weakly stratified and sub-horizontally bedded.	dry 7.5YR 6/4, light brown	Abrupt
4	6–16	Silts with sparse charcoal, abundant stone artifacts and bone fragments, many calcareous nodules and some calcite crystals ((Fig. S6, S7D and S9B). It is highly compact. White lamella bedding of carbonate deposition with thicknesses of 2–5 mm was observed in this layer (Fig. S2B).	dry 7.5YR 6/6, reddish yellow	Abrupt
5	2–12	Silts with gravels, limestone fragments, stone artifacts, bone fragments, and calcite crystals (Fig. S6, S7E, S8A and S9E). It also contains many small (less than 5 mm) rounded soil aggregates and sub-rounded calcite crystal and limestone fragments (Fig. S4A and B). It is relatively loose because of the larger voids among these different inclusions.	dry 7.5YR 5/3, brown	Abrupt
6	9–17	Silts with abundant stone artifacts and animal bones, sparse charcoal, many calcareous nodules, many pebbles (~ 5 cm) and cobbles (~10 cm), and calcite crystals (Fig. S6, S7F, S8A and S9F). It is compact and stony.	dry 7.5YR 6/4, light brown	Clear
7	12–24	Silts with abundant stone artifacts, animal bones, limestone fragments and cobbles (Fig. S6, S7G, S8A and S9G). Sediments are mainly fine and compact. It is comparatively less stony than Layer 6 with fewer cobbles (~10 cm).	dry 5YR 3/2, dark reddish brown	Abrupt
8	2–7	White to red silts (Fig. S2B, SA and S9H). This layer contains some small pieces of charcoal, sparse stone artifacts and animal bones, some limestone fragments, cobbles and gravels (Fig. S6, S7H). The brownish sediments are similar to those from Layer 7, but more compact. These sediment features are suggestive of burning.	dry 7.5YR 5/3, brown	Abrupt
9	6–14	Silts with abundant limestone fragments, pebbles, cobbles, calcite crystals, stone artifacts and bone fragments (Fig. S6, S7I-J, S8A and S9I). It is blocky, very stony with many cobbles (10–15 cm), and relatively compact.	dry 7.5YR 2.5/3, very dark brown	Gradual
10	28–75	Silts with abundant stone artifacts, animal bones, cobbles, limestone fragments, calcareous nodules and calcite crystals (Fig. S6, S7K-L, S8A and SJ-K). Sediments are highly compact and stony with many cobbles together with stone artifacts.	dry 7.5YR 4/3, brown	Not reached

Table S2. XRD analysis of measured samples from each layer.

Depth (cm)	Layer	Semiquant (%)								
		Quartz	Calcite	Feldspar	Clay minerals	Dolomite	Magnetite	Halite	Aragonite	Titanomagnetite
0-4	1	39	27	24	8	1	—	—	—	—
4-10	2	39	26	24	9	2	—	—	—	—
15-20	3	42	12	34	9	3	—	—	—	—
30-35	4	44	19	27	7	3	—	—	—	—
40-45	5	36	26	29	6	1	2	—	—	—
50-55	6	46	9	39	6	—	—	—	—	—
64-70	7	35	40	17	5	—	3	—	—	—
75-82	8	37	38	12	6	—	—	1	6	—
82-85	9	46	23	24	6	—	—	—	—	1
85-92	9	8	30	7	17	—	38	—	—	—
92-100	10	42	36	15	5	—	2	—	—	—
115-120	10	35	27	24	9	—	4	—	—	—

Table S3. AMS radiocarbon dating results from BKC. The prefix “OxA” for laboratory numbers refer to samples dated by the University of Oxford. Laboratory numbers with the prefix “LZU” indicate those samples prepared in Lanzhou University and measured at Peking University. Calibrated age ranges were rounded to the closest decade.

Laboratory Number	Sample Number	Depth (cm)	Layer	Material	^{14}C age (BP)	±	1σ（68.2%） (cal BP)		2σ（95.4%） (cal BP)		mu	sigma	median
							from	to	from	to			
OxA-39180	B12	11	H1	Bone	833	19	760	710	780	700	740	25	740
OxA-39179	B3	8	2	Bone	39,700	1,100	44,430	42,650	45,630	42,040	43,730	940	43,630
OxA-39181	B13	8	2	Bone	28,410	280	32,790	31,870	33,190	31,530	32,350	440	32,350
OxA-39182	B16	10	2	Bone	25,200	220	29,500	28,970	29,810	28,730	29,270	270	29,250
OxA-38709	B18	13	2	Bone	27,820	270	31,890	31,270	32,470	31,150	31,700	350	31,630
OxA-38710	B22	17	3	Bone	37,150	830	42,310	40,980	42,910	40,160	41,590	690	41,620
OxA-39185	B30	18	3	Bone	42,400	1,600	47,560	44,400	49,470	43,510	46,250	1,540	46,100
OxA-39184	B21	20	3	Bone	41,600	1,600	46,670	43,570	48,950	42,830	45,580	1,580	45,380
OxA-39183	B20	21	3	Bone	32,120	440	36,500	35,500	37,330	35,020	36,100	560	36,050
LZU-19161	B67	21	3	Bone	29,610	130	33,910	33,660	34,040	33,540	33,790	120	33,790
LZU-19162	B94	24	3	Bone	30,220	140	34,400	34,070	34,570	33,940	34,250	160	34,240
OxA-38711	B105	25	4	Bone	44,000	2,000	49,070	45,880		44,890	47,370	1,460	47,390
LZU-19163	B147	36	4	Bone	> 46,110								
OxA-38712	B211	42	5	Bone	> 48,500								
OxA-38713	B218	47	6	Bone	> 54,800								

Table S4. Summary of the water content, external and internal dose rates for quartz (Q) and K-feldspar (KF) for all samples.

Sample	Depth (cm)	Grain size (μm)	Mineral	Water content (%)	External dose rate (Gy/ka)		Internal dose rate (Gy/ka)	Total dose rate (Gy/ka)
					Beta	Gamma		
OSL1	10	90–125	Q	7.0 ± 1.8	1.68 ± 0.07	0.91 ± 0.03	0	2.59 ± 0.07
		125–250	KF		1.61 ± 0.06		0.81 ± 0.14	3.33 ± 0.16
OSL2	20	90–125	Q	6.0 ± 1.5	2.19 ± 0.08	1.12 ± 0.03	0	3.31 ± 0.09
		125–250	KF		2.10 ± 0.08		0.81 ± 0.14	4.03 ± 0.16
OSL3	31	180–250	Q	5.0 ± 1.3	1.86 ± 0.07	1.20 ± 0.03	0	3.06 ± 0.07
		180–250	KF				0.92 ± 0.10	3.98 ± 0.12
OSL4	35	90–125	Q	5.0 ± 1.3	1.93 ± 0.07	1.20 ± 0.05	0	3.13 ± 0.09
		90–125	KF				0.81 ± 0.14	3.62 ± 0.10
OSL5	50	90–125	Q	5.0 ± 1.3	1.87 ± 0.07	1.05 ± 0.03	0	2.92 ± 0.08
		90–125	KF	5.0 ± 1.3		1.05 ± 0.03	0.49 ± 0.05	3.41 ± 0.09
OSL6	60	90–180	KF	7.0 ± 1.8	1.93 ± 0.08	1.05 ± 0.02	0.60 ± 0.10	3.58 ± 0.13
OSL7	80	180–250	KF	16.0 ± 4.0	1.72 ± 0.08	0.97 ± 0.02	0.92 ± 0.10	3.62 ± 0.13
OSL8	83	180–250	KF	21.0 ± 5.3	1.37 ± 0.08	0.89 ± 0.02	0.92 ± 0.10	3.18 ± 0.13
OSL9	84	180–250	KF	15.0 ± 3.8	1.48 ± 0.07	0.87 ± 0.03	0.92 ± 0.10	3.27 ± 0.12
OSL10	86	180–250	KF	10.0 ± 2.5	1.27 ± 0.05	0.85 ± 0.05	0.92 ± 0.10	3.04 ± 0.12
OSL11	98	125–250	KF	15.0 ± 3.8	1.41 ± 0.07	0.95 ± 0.05	0.81 ± 0.14	3.17 ± 0.16
OSL12	165	180–250	KF	12.5 ± 3.1	1.87 ± 0.08	1.11 ± 0.05	0.92 ± 0.10	3.91 ± 0.14

Table S5. Experimental procedures for small aliquot quartz (a) and single-grain K-feldspar (b).

Step	(a) Small aliquot quartz OSL	(b) Single-grain K-feldspar pIRIR
1	Give regenerative dose, D_i [a]	Give regenerative dose, D_i [a]
2	Preheat at 260°C for 10 s	Preheat at 320°C for 60 s
3	Green laser stimulation (2 s) at 125 °C	IR diode stimulation at 200°C for 200 s
4	Give test dose, D_t	Single-grain IR laser stimulation (2 s) at 275°C
5	Preheat at 160°C for 5 s	Give test dose, D_t
6	Green laser stimulation (2 s) at 125 °C	Preheat at 320°C for 60 s
7	Blue LED stimulation at 280°C for 40 s	IR diode stimulation at 200°C for 200 s
8		Single-grain IR laser stimulation (2 s) at 275°C
9		IR diode bleaching at 325°C for 200 s
10		Return to step 1

[a] For the 'natural' sample, i = 0 and D_0 = 0. The whole sequence is repeated for several regenerative doses, including a zero dose and a repeat dose.

Table S6. Number of single K-feldspar grains measured, rejected and accepted for each sample, together with the reasons for their rejection. Note that, for the samples analysed using the SGC method, criteria related to construction of full dose response curves (criteria 2–4) are not available.

Sample	Method	No. of measured	Rejection criteria				Accepted
			(1) T_n too dim [a]	(2) Recuperation > 5%	(3) Poor DRC [b]	(4) Infinite D_e	
OSL2	SAR	200	15	2	17	0	166
OSL3	SAR	200	84	3	37	10	66
OSL4	SAR	200	5	5	25	4	161
OSL6	SAR	300	27	0	90	8	175
OSL7	SAR	100	48	2	21	4	25
OSL8	SAR	100	39	6	14	0	41
OSL10	SAR	200	90	5	62	2	41
OSL12	SAR	200	86	0	47	10	57
OSL1	SGC	1000	462				538
OSL2	SGC	200	6				194
OSL3	SGC	800	407				393
OSL4	SGC	200	0				200
OSL5	SGC	400	78				322
OSL6	SGC	600	336				264
OSL6	SGC	200	20				180
OSL7	SGC	500	241				259
OSL8	SGC	600	299				301
OSL8	SGC	600	148				452
OSL9	SGC	700	450				250
OSL10	SGC	1000	583				417
OSL11	SGC	400	82				318
OSL12	SGC	400	217				183

[a] Initial T_n signal is <3σ above the corresponding background, or the relative standard error of the net T_n intensity is >25%.

[b] The DRC has a figure-of-merit (FOM) value (*74, 75*) of >10% or a reduced chi squared (RCS) value of >5 (*76*).

Table S7. Number of single quartz aliquots measured, rejected and accepted for each sample, together with the reasons for their rejection.

Sample	No. of measured	Rejection criteria					Accepted	Proportion of saturated
		T_n too dim [a]	Recuperation > 5%	Poor DRC [b]	Infinite D_e	Zero dose		
OSL1	4000	3756	9	125	10	40	60 (1.5%)	9%
OSL2	5500	5240	4	180	12	38	26 (0.4%)	16%
OSL3	4700	4376	0	255	13	5	51 (1.1%)	19%
OSL4	3000	2738	3	121	25	14	99 (3.3%)	18%
OSL5	1600	895	8	267	87	3	340 (21.3%)	20%

[a] Initial T_n signal is <3σ above the corresponding background, or the relative standard error of the net T_n intensity is >25%.

[b] The DRC has a figure-of-merit (FOM) value (*74, 75*) of >10% or a reduced chi squared (RCS) value of >5 (*76*).

Table S8. Summary of quartz SAR D_e and age estimates using the maximum age model (MAX) and minimum age model (MAM). Note that these are not the final D_e and age estimates for the quartz samples (see text for discussions).

Sample	Over-dispersion (%)	Age model [a]	No. of grains [b]	D_e (Gy)	Age (ka)
OSL1	117 ± 11	MAX	60 (16%)	65.5 ± 15.7	25.3 ± 6.1
		MAM	60 (30%)	1.8 ± 0.4	0.7 ± 0.2
OSL2	179 ± 26	MAX	26 (14%)	124.7 ± 53.6	37.7 ± 16.2
		MAM	26 (11%)	0.7 ± 0.3	0.2 ± 0.1
OSL3	69 ± 8	MAX	51 (59%)	153.3 ± 36.4	50.1 ± 12.0
OSL4	97 ± 8	MAX	99 (75%)	166.4 ± 7.4	53.2 ± 3.0
OSL5	56 ± 3	MAX	340 (83%)	223.0 ± 5.2	76.5 ± 3.1

[a] The MAM was applied after removing aliquots with negative and positive D_e values consistent with zero at 2σ. MAM was not applied to OSL3–OSL5, as the number of aliquots with small D_e values are too few.

[b] The percentage of aliquots that provide meaning weight to the calculation of the MIN and MAX is shown in parentheses.

Table S9. Quartz L_n/T_n D_e values and age estimates for different DRC groups of all samples. D_e values are based on a statistical analysis of the L_n/T_n values of accepted aliquots in each group, using a range of different age models (see text for details). The D_e and age uncertainties are asymmetric due to the non-linearity of the SGCs, so only the mean and ±2σ ranges (95% CI) are shown. Age uncertainties include a 2% systematic error to allow for any bias associated with calibration of the laboratory beta source.

Sample	DRC Group	No. of accepted grains [a]	Over-dispersion (%)	Age model [b]	D_e (Gy) [c]	Combined D_e (Gy) [e]			Age (ka)		
						mean	-2σ	+2σ	mean	-2σ	+2σ
OSL1	1	17 (15.7%)	144 ± 27	MAX (59%)	saturated	55	40	70	21.3	15.3	27.2
	2	8 (7.4%)	- [d]	- [d]	-						
	3	51 (47.2%)	98 ± 11	MAX (37%)	69 ± 10						
	4	3 (2.8%)	- [d]	- [d]	-						
	5	26 (24.1%)	133 ± 23	MAX (26%)	44 ± 10						
	6	3 (2.8%)	- [d]	- [d]	-						
OSL2	1	15 (27.8%)	137 ± 27	MAX (33.3%)	saturated	101	67	163	30.5	20.1	49.3
	2	24 (44.4%)	175 ± 26	MAX (45.8%)	101 ± 22						
	3	10 (18.5%)	188 ± 43	MAX (30%)	>79						
	4	4 (7.4%)	- [d]	- [d]							
	5	1 (1.9%)	- [d]	- [d]							
OSL3	1	10 (14.7%)	138 ± 33	FMM (90%)	saturated	140	121	162	45.7	39.0	53.3
	2	26 (38.2%)	58 ± 8	FMM (85%)	130 ± 12						
	3	16 (23.5%)	33 ± 7	CAM (100%)	177 ± 27						
	4	16 (23.5%)	90 ± 17	FMM (88%)	133 ± 12						
OSL4	1	16 (11.8%)	127 ± 24	FMM (69%)	saturated	186	166	209	59.5	52.1	67.7
	2	9 (6.6%)	- [d]	- [d]	-						
	3	59 (43.4%)	68 ± 7	FMM (86%)	192 ± 16						
	4	20 (14.7%)	124 ± 20	FMM (70%)	175 ± 16						
	5	26 (19.1%)	66 ± 10	FMM (58%)	193 ± 23						
	6	6 (4.4%)	- [d]	- [d]	-						
OSL5	1	8 (1.9%)	- [d]	- [d]	-	231	209	258	79.2	70.3	89.6
	2	56 (13.2%)	53 ± 5	FMM (95%)	198 ± 35						
	3	142 (33.5%)	22 ± 2	FMM (96%)	216 ± 13						
	4	120 (28.3%)	13 ± 1	nMAD CAM (90%)	222 ± 7						
	5	75 (17.7%)	18 ± 2	FMM (95%)	246 ± 12						
	6	23 (5.4%)	52 ± 8	FMM (54%)	276 ± 53						

[a] Percentage grains in each DRC group is shown in parentheses.

[b] Percentage grains included in age models for D_e estimation is shown in parentheses.

[c] 'Saturated' means that the weighted mean LS-normalized L_n/T_n value is statistically consistent with the saturation level of the corresponding SGC at 2σ.

[d] The number of grains are insufficient (N = <10) to produce statistically significant results.

[e] The final D_e was calculated from the weighted mean D_e values of all groups that produced finite D_e values.

Table S10. K-feldspar L_n/T_n D_e values and age estimates for all samples. D_e values are based on a statistical analysis of the L_n/T_n values for grains with T_n signals >500 cts/0.1 s, using a range of different age models (see text for details). The D_e and age uncertainties are asymmetric due to the non-linearity of the SGC, so only the mean and ± 2σ ranges (95% CI) are shown. Best estimate ages are shown in bold (see the text for justification). Age uncertainties include a 2% systematic error to allow for any bias associated with calibration of the laboratory beta source.

Sample	Layer	Depth (cm)	Grain size (µm)	Dose rate (Gy/ka)	Model	D_e (Gy)			Age (kyr)		
						mean	-2σ	+2σ	mean	-2σ	+2σ
OSL1	2	10	125–250	3.33 ± 0.16	MAM	54	45	63	**16.2**	**12.4**	**20.9**
OSL2	3	20	125–250	4.03 ± 0.16	FMM-1	185	167	203	**45.8**	**38.4**	**54.9**
					FMM-2	658	461	1041	163.2	105.8	280.9
OSL3	4	31	180–250	3.98 ± 0.12	MAM	171	148	195	**42.9**	**35.0**	**52.2**
OSL4	4	35	90–125	3.62 ± 0.10	CAM	233	221	245	**64.3**	**57.9**	**71.6**
OSL5	6	50	90–125	3.41 ± 0.09	CAM	347	315	382	**101.8**	**87.8**	**118.4**
OSL6	6	60	90–180	3.58 ± 0.13	CAM	369	352	386	**102.9**	**91.7**	**116.0**
OSL7	7	80	180–250	3.62 ± 0.13	CAM	356	336	377	**98.3**	**86.6**	**112.2**
OSL8	8	83	180–250	3.18 ± 0.13	CAM	307	291	324	**96.5**	**84.9**	**110.5**
OSL9	9	84	180–250	3.27 ± 0.12	nMAD CAM	347	319	379	**106.3**	**90.8**	**125.4**
OSL10	9	86	180–250	3.04 ± 0.12	CAM	341	324	360	**112.3**	**99.1**	**128.0**
OSL11	10	98	125–250	3.17 ± 0.16	CAM	463	436	493	**145.9**	**124.4**	**173.2**
OSL12	10	165	180–250	3.91 ± 0.14	nMAD CAM	742	637	880	**190.0**	**152.4**	**242.4**

Table S11. CQL code for Bayesian modelling.

```
Plot(){
 Outlier_Model("General",T(5),U(0,4),"t");
 Sequence(){
  Boundary("Start layer 10");
  C_date("OSL12 (KF IRSL)", calBP(189990),
14340, 18420)
  {
   z=155;
   Outlier(0.05);
  };
  C_date("OSL11 (KF IRSL)", calBP(145900),
6610, 6870)
  {
   z=98;
   Outlier(0.05);
  };
  Boundary("Transition 10/9");
  Phase("Layer 9")
  {
   C_date("OSL10 (KF IRSL)", calBP(112260),
3990, 4120)
   {
    z=86;
    Outlier(0.05);
   };
   C_date("OSL09 (KF IRSL)", calBP(106310),
5160, 5550)
   {
    z=84;
    Outlier(0.05);
   };
  };
  Boundary("End Layer 9");
  Boundary("Start Layer 8");
  C_date("OSL08 (KF IRSL)", calBP(96530),
3530, 3630)
  {
   z=83;
   Outlier(0.05);
  };
  Boundary("End Layer 8");
  Boundary("Start Layer 7");
  C_date("OSL07 (KF IRSL)", calBP(101660),
4020, 4230)
  {
   z=80;
   Outlier(0.05);
  };
  Boundary("End Layer 7");
  Boundary("Start Layer 6");
  C_date("OSL06 (KF IRSL)", calBP(102860),
3170, 3270)
  {
   z=60;
   Outlier(0.05);
  };
```

```
  C_date("OSL05 (KF IRSL)", calBP(101750), 5100,
5590)
  {
   z=50;
   Outlier(0.05);
  };
  Before("Minimum age estimate")
  {
   Date("OSL05 (Qtz OSL)", N(calBP(79210),
4455))
  };
  Before("Radiocarbon limit")
  {
   Date("B218 (C14)", N(calBP(50000), 100))
  };
  Boundary("End layer 6");
  Before("Radiocarbon limit")
  {
   Date("B211 (C14)", N(calBP(50000), 100))
  };
  Interval("Interval (layer 5)", N(37420, 6068));
  Boundary("Start layer 4");
  Phase("Layer 4 bottom")
  {
   C_date("OSL04 (KF IRSL)", calBP(64330),
2850, 2890)
   {
    z=37;
    Outlier(0.05);
   };
   C_date("OSL04 (Qtz OSL)", calBP(59480),
3910, 4310)
   {
    z=37;
    Outlier(0.05);
   };
   Before("Radiocarbon limit")
   {
    Date("B147 (C14)", N(calBP(50000), 100))
   };
  };
  Boundary("End lower layer 4");
  Interval("Interval (middle layer 4)", N(15000,
4000));
  Boundary("Start upper layer 4");
  Phase("Layer 4 top")
  {
   C_date("OSL03 (KF IRSL)", calBP(42930),
3290, 3400)
   {
    z=31;
    Outlier(0.05);
   };
```

```
C_date("OSL03 (Qtz OSL)", calBP(45720), 3410,
3860)
     {
      z=31;
      Outlier(0.05);
     };
     R_date("B105 (C14)", 44000, 2000)
     {
      z=25;
      Outlier(0.05);
     };
    };
    Boundary("End Layer 4");
    Boundary("Start Layer 3");
    Phase("Layer 3")
    {
     R_date("B94 (C14)", 30220, 140)
     {
      z=24;
      Outlier(0.05);
     };
     R_date("B67 (C14)", 29610, 130)
     {
      z=21;
      Outlier(0.05);
     };
     R_date("B20 (C14)", 32120, 440)
     {
      z=21;
      Outlier(0.05);
     };
     C_date("OSL02 (KF IRSL)", calBP(45800),
3030, 3160)
     {
      z=20;
      Outlier(0.05);
     };
     C_date("OSL02 (Qtz OSL)", calBP(30520),
5250, 9430)
     {
      z=20;
      Outlier(0.05);
     };
R_date("B21 (C14)", 41600, 1600)
     {
      z=20;
      Outlier(0.05);
     };
     R_date("B30 (C14)", 42400, 1600)
     {
      z=18;
      Outlier(0.05);
     };
R_date("B22 (C14)", 37150, 830)
     {
      z=17;
      Outlier(0.05);
     };
    };
    Boundary("Transition 3/2");
    Phase("Layer 2")
    {
     R_date("B18 (C14)", 27820, 270)
     {
      z=13;
      Outlier(0.05);
     };
     R_date("B16 (C14)", 25200, 220)
     {
      z=10;
      Outlier(0.05);
     };
     C_date("OSL01 (KF IRSL)", calBP(16230),
1590, 1590)
     {
      z=10;
      Outlier(0.05);
     };
     C_date("OSL01 (Qtz OSL)", calBP(21270),
3010, 3010)
     {
      z=10;
      Outlier(0.05);
     };
     R_date("B13 (C14)", 28410, 280)
     {
      z=8;
      Outlier(0.05);
     };
     R_date("B3 (C14)", 39700, 1100)
     {
      z=8;
      Outlier(1);
     };
    };
    Boundary("End layer 2");
  };
};
```

Table S12. Bayesian age model estimates ('Modelled age ranges') for T2, at 68.2% and 95.4% probabilities. The modelled transitional, start and end boundary ages for each stratigraphic layer (or combination of layers) and interval durations are highlighted in bold and italics. All ages are given in thousands of years and rounded off to the closest century.

Sample	MODELLED AGE RANGES (thousands of years) 95.4% probability		68.2% probability		Outlier probabilities		Convergence (%)	Agreement index (%)
	from	*to*	*from*	*to*	*prior*	*posterior*		
Start layer 10	***156.9***	***224.8***	***168.1***	***199***			***49.4***	
OSL12 (KF IRSL)	153.2	201.8	162.8	185.9	0.05	0.08	93.8	71.1
OSL11 (KF IRSL)	133.3	160.7	139.9	153.5	0.05	0.05	96	100.4
Transition 10/9	***109***	***149.2***	***112.9***	***133.2***			***90.6***	
OSL10 (KF IRSL)	107	120.7	109.9	116.7	0.05	0.04	98.8	109.5
OSL09 (KF IRSL)	105.9	120.8	109	116.2	0.05	0.06	98.6	78.6
End Layer 9	***103.4***	***116.9***	***106.4***	***113.2***			***98.5***	
Start Layer 8	***100.7***	***113.2***	***103.2***	***109.2***			***98.1***	
OSL08 (KF IRSL)	99.8	111	102.1	107.2	0.05	0.19	98.7	23.2
End Layer 8	***99***	***109.9***	***101.3***	***106.4***			***98.7***	
Start Layer 7	***97.9***	***107.7***	***100.1***	***104.8***			***98.8***	
OSL07 (KF IRSL)	97.3	106.5	99.5	103.9	0.05	0.03	99.1	126.6
End Layer 7	***96.6***	***105.8***	***98.8***	***103.2***			***98.9***	
Start Layer 6	***95.2***	***104.1***	***97.5***	***101.8***			***97.6***	
OSL06 (KF IRSL)	94.7	103.3	96.9	101.1	0.05	0.06	98.1	71.3
OSL05 (KF IRSL)	93.5	102.5	95.9	100.3	0.05	0.04	97.4	106.8
End Layer 6	***90.6***	***102***	***94.4***	***99.7***			***78.7***	
Interval (Layer 5)	***24.3***	***39.1***	***28.6***	***35.9***			***99.7***	***90.4***
Start Layer 4	***59.9***	***72***	***62.3***	***67.9***			***75.6***	
OSL4 (KF IRSL)	58.9	67.7	61	65.4	0.05	0.03	94.6	108.8
OSL4 (Qtz IRSL)	57.8	67.5	60.3	65	0.05	0.04	94.5	102.6
End lower layer 4	***55.2***	***65.8***	***58.4***	***63.5***			***83.5***	
Interval (middle layer 4)	***6.6***	***17.6***	***9.7***	***15.2***			***99.8***	***100.6***
Start upper layer 4	***45.9***	***52.6***	***47.1***	***50.1***			***95***	
OSL03 (KF IRSL)	45.3	50.1	46.4	48.9	0.05	0.05	98.9	58.9
OSL03 (Qtz OSL)	45.4	50.4	46.4	49	0.05	0.03	99	121.9
B105 (C14)	45.5	49.8	46.5	48.9	0.05	0.02	99	112.2
End layer 4	***44.7***	***49.2***	***45.8***	***48.1***			***98.2***	
Start layer 3	***43.8***	***48.1***	***44.8***	***46.9***			***97.4***	
B94 (C14)	33.9	34.6	34.1	34.4	0.05	0.01	98.3	103.4
B67 (C14)	33.6	45.4	33.7	40.5	0.05	0.2	0	85.3
B20 (C14)	35.1	37.8	35.6	36.6	0.05	0.03	99	101.7
OSL02 (KF IRSL)	38.5	46.7	41.4	45.3	0.05	0.05	98.4	99.2
OSL02 (Qtz OSL)	33.2	45.6	34.6	42.8	0.05	0.05	98.2	105.3
B21 (C14)	40.9	46.6	43.4	45.5	0.05	0.05	98.5	110.3
B30 (C14)	39.3	46.8	43.7	45.7	0.05	0.07	98.1	95.3
B22 (C14)	39.9	43	40.9	42.3	0.05	0.03	98.7	101.8
Transition 3/2	***32.1***	***34.2***	***32.6***	***33.7***			***71.6***	
B18 (C14)	31.2	32.5	31.4	32	0.05	0.02	98.8	104.2
B16 (C14)	23.7	30	29	29.5	0.05	0.05	95.9	99.6
OSL01 (KF IRSL)	14.1	31.5	15.8	24.2	0.05	0.22	94.6	76
OSL01 (Qtz OSL)	16.3	30.2	19.1	26	0.05	0.06	95.7	94.4
B13 (C14)	31.5	33.1	31.8	32.7	0.05	0.02	98.3	105.2
B3 (C14)	17.8	33.5	25.1	32.5	1	1	81.5	102
End layer 2	***5.4***	***28.4***	***12***	***20.6***			***85.2***	

Table S13. Library information for sediment samples from Layers 2, 3, 4, and 7, in which authentic ancient hominin mtDNA was detected.

Indexed library ID	Sample ID	Extract ID	Layer	Unique_hominid_sequences	Proportion of damaged	5'CT%	3'CT%	D_support %	D_support _deam%	H_support %	H_support _deam%	Contamination%	95% CI	Coverage	Coverage_deam	Decision
L5795	C4092	T1989	Layer 2	381	23	44.6	25.8	75.0(33/44)	83.3(5/6)	4.3(1/23)	0.0(0/5)	4.4	2.9-6.2	1.13	0.24	all
L5832	C4092	T3021	Layer 2	218	19	34	17.4	50.0(17/34)	71.4(5/7)	60.0(6/10)	0.0(0/1)	2.5	3.6-1.7	0.66	0.12	damage
L5791	C4090	T1985	Layer 3	2357	14	20.5	15.6	37.8(179/474)	83.1(49/59)	60.6(126/208)	20.0(4/20)	11.4	9.3-13.6	7.37	0.95	damage
L5831	C4090	T3020	Layer 3	723	24	40.9	34.7	67.8(61/90)	88.9(16/18)	16.1(10/62)	16.7(2/12)	1.7	3.1-1.0	2.19	0.52	all
L5641	C4091	T1902	Layer 4	1193	26	49.1	33.1	92.0(206/224)	96.1(49/51)	4.0(4/99)	0.0(0/21)	1.6	1.2-2.1	3.75	0.95	all
L5788	C4091	T1902	Layer 4	1819	30	43.7	38.7	92.9(287/309)	95.7(89/93)	7.7(10/130)	0.0(0/32)	2	1.6-2.4	5.54	1.61	all
L5789	C4091	T1902	Layer 4	1253	28	45.9	32.6	85.1(166/195)	88.2(45/51)	10.9(10/92)	7.4(2/27)	2.1	1.7-2.6	4.01	1.13	all
L5790	C4091	T1902	Layer 4	1071	27	40.9	30.8	89.5(188/210)	94.3(50/53)	9.0(8/89)	4.8(1/21)	1.8	1.4-2.3	3.33	0.87	all
L5820	C4091	T1902	Layer 4	1034	24	38.4	32.9	90.4(161/178)	100.0(36/36)	7.9(5/63)	0.0(0/9)	1.8	2.3-1.4	3.21	0.75	all
L5792	C4091	T1986	Layer 4	5069	14	27.1	13.3	43.8(452/1031)	75.8(94/124)	54.6(249/456)	11.1(5/45)	19.7	16.6-22.9	17.23	2.28	damage
L5793	C4091	T1987	Layer 4	6898	16	30.6	13.5	48.7(617/1266)	81.4(158/194)	58.3(381/653)	26.0(19/73)	16.2	13.7-18.8	23.28	3.54	Excluded
L5794	C4091	T1988	Layer 4	5368	13	23.3	11.9	31.0(368/1186)	74.3(104/140)	66.5(338/508)	23.5(12/51)	14.9	12.6-17.7	19.53	2.41	Excluded
L5824	C4091	T3013	Layer 4	4851	26	47.3	30.2	92.9(706/760)	93.0(174/187)	7.3(30/411)	3.8(3/79)	3.8	4.6-3.1	15.47	4.03	all
L5825	C4091	T3014	Layer 4	4314	29	45.2	34.3	94.6(601/635)	95.7(178/186)	6.0(20/334)	3.4(2/59)	3.6	4.4-2.8	13.8	3.97	all
L5826	C4091	T3015	Layer 4	20320	27	47.3	33.1	92.8(3130/3372)	94.3(798/846)	3.9(61/1564)	1.7(6/345)	4.9	6.1-3.8	63.09	17.2	all
L5827	C4091	T3016	Layer 4	2685	29	47.6	32.5	92.8(375/404)	96.1(98/102)	9.8(20/204)	6.8(3/44)	4.3	5.7-3.2	8.41	2.42	all
L5828	C4091	T3017	Layer 4	1997	27	46.4	34.2	80.9(301/372)	90.0(81/90)	13.9(23/166)	6.7(2/30)	4.1	5.0-3.4	6.38	1.71	all
L5829	C4091	T3018	Layer 4	1814	25	42.7	29.7	87.7(256/292)	96.1(74/77)	12.5(16/128)	0.0(0/20)	4.5	5.7-3.6	5.74	1.39	all
L5830	C4091	T3019	Layer 4	2927	30	42.2	41	93.1(448/481)	97.9(137/140)	6.4(14/219)	2.8(2/71)	7.8	9.4-6.3	9.24	2.7	damage
L5800	C4097	T1994	Layer 7	81	46	50	41.9	66.7(6/9)	80.0(4/5)	20.0(1/5)	33.3(1/3)	1.1	0.2-3.6	0.2	0.09	all
L5821	C4097	T1994	Layer 7	191	46	46.9	39.4	90.6(29/32)	94.4(17/18)	30.0(3/10)	0.0(0/4)	0.5	1.7-0.1	0.54	0.25	all
L5822	C4097	T1994	Layer 7	218	36	45.6	37.7	94.6(35/37)	90.9(10/11)	0.0(0/5)	0.0(0/2)	2.9	4.3-1.9	0.59	0.21	all
L5823	C4097	T1994	Layer 7	431	37	51.8	52.6	92.8(64/69)	92.9(26/28)	6.2(2/32)	5.9(1/17)	1.4	2.2-0.8	1.29	0.49	all
L5833	C4097	T3022	Layer 7	251	33	49.1	33.8	82.1(32/39)	92.3(12/13)	13.6(3/22)	42.9(3/7)	0.9	19.3-0.1	0.69	0.23	all

D_support%: Percentage of fragments matching the Denisovan variant; D_support_deam% - percentage of putatively deaminated fragments matching the Denisovan variant; H-support%- percentage of fragments matching the modern human variant; D_support_deam% - percentage of putatively deaminated fragments matching the modern human variant. The red color indicates the possible modern human contamination.

Table S14. The frequencies of nucleotide substitutions of libraries on different lineage.

	5'CT	3'CT	5'CT_95%CI	3'CT_95%CI	cond5'CT	cond3'CT	cond5'CT_95%CI	cond3'CT_95%CI
All sequences								
L5793	30.6	13.5	28.1-33.0	12.0-15.1	42.4	20	29.6-55.9	13.4-28.1
L5794	23.3	11.9	20.9-25.8	10.3-13.6	36.4	17.6	20.4-54.9	8.4-27.1
Denisovan-vs-modern human								
L5793_denisovan	46.1	19.6	41.3-50.9	16.1-23.4	38.1	17	18.1-61.6	7.6-30.8
L5793_human	9.2	1.7	6.7-11.9	0.9-3.0	50	6.2	1.3-98.7	0.2-30.2
L5794_denisovan	49.1	21	42.2-55.3	16.4-26.3	50	11.1	11.8-88.2	0.9-24.3
L5794_human	6.9	3.2	4.9-9.1	2.0-4.4	0	0	0.0-52.2	0.0-28.5
Denisova+Sima-vs-modern human								
L5793_denisovan+sima	45.5	19.9	38.7-52.3	15.3-25.2	36.4	18.2	10.9-69.2	5.2-40.3
L5793_human	9.2	2	5.9-12.8	0.7-3.5	NA	0	N/A	0.0-28.5
L5794_denisovan+sima	47.3	21.7	37.7-57.0	15.4-29.1	66.7	22.2	9.4-99.2	0.3-48.2
L5794_human	7.8	2.7	5.1-11.2	1.4-4.7	0	0	0.0-84.2	0.0-36.9

CT: the frequencies of C to T substitution

Table S15. Summary of mtDNA information for the samples for low contaminated fragments and only deaminated fragments (with "_d") of Layers 2, 3, 4, and 7.

Layer	Coverage	Total bases in consensus	Proportion of mtDNA genome	Number of Pos >=10x	Number of position consensus support <80%&>=10X	Total number of the 93 sites variable in the 4 Denisovans covered
Layer 2	1.3	4,643	28%	13	0	0
Layer 3	3.2	11,302	68%	286	2	0
Layer 4	144	16,545	100%	16,452	54	9
Layer 7	3.3	9,519	57%	1,015	21	0
Layer 2_d	0.37	1,203	7.30%	1	0	0
Layer 3_d	1.5	6,022	36.30%	23	0	0
Layer 4_d	40	16,474	99.10%	15,573	31	4
Layer 7_d	1.3	4,282	25.80%	215	5	0

Table S16. Summary of specific mutations in Layer 4 that differed from global present-day human, Denisovan and Neandertal mtDNA.

			Low contamination fragments						Deaminated fragments					
rCRs_position	**Reference**	**4_D_S_layer4L_d**	**Allele**	**Coverage**	**Allele**	**Coverage**	**Total Coverage**	**Consensus Support**	**Allele**	**Coverage**	**Allele**	**Coverage**	**Total Coverage**	**Consensus Support**
6975	t	Tttttcn	c	61	t	19	80	76%	c	15	t	8	23	65%
7091	a	aaaaagn	g	81	a	31	112	72%	g	10	a	14	24	42%
16260	c	Ccccctt	t	6	c	3	9	67%	t	6	c	1	7	86%

4_D_S_layer4L_d: The order of the consensus covered at the position: 4 published Denisovans, Sima, Layer 4 mtDNA consensus from low contamination fragments and Layer 4 mtDNA consensus from deaminated fragments.

The following tables are provided as online Excel files.

Table S17. The number of mtDNA components present in data from Layers 2, 3, 4, and 7, estimated using a maximum-likelihood approach. For each dataset, the average coverage at sites covered by DNA fragments and the number and proportion of variable sites observed are shown. The log-likelihood of models with one, two or three mtDNA components (k) are presented under “logL1k”, “logL2k” and “logL3k”, respectively. The “best model” noted is the one with the lowest Akaike Information Criterion (AIC) value. When that model contains multiple components, the estimated fraction of the minor component and its divergence from the major one are shown. The p-value of a likelihood ratio test to evaluate whether the best model is significantly more supported than the second best one is noted under “relL”. For each dataset, the analysis was performed using all observed variable positions, and after excluding positions where the four Denisovan mtDNA genomes sequenced to date differ from a panel of 311 present-day human mtDNA genomes (“-fixed diff”). In the “interpretation” column, an asterisk indicates that the observed coverage is lower than 2.5-fold, thereby reducing the power of the model to detect multiple components.

Table S18. Sediment samples analyzed in this study. Details on the sample information, amounts of material used for DNA extraction, and the number of molecules in each extract are presented.

Table S19. The authenticity of mtDNA fragments and the hominin lineage assignment. For each library, the frequency of each nucleotide substitution for the ancient DNA in these sequences are reported. + or ++ indicate the libray are significantly have ancient DNA character. The percentage and number of sequences matching variants specific to each branch of a phylogenetic tree relating four hominin groups were reported, using all sequences in a samples and after retaining only those exhibiting the first and last three C to T substitutions. CT%- the frequencies of C to T substitution; H_support- percentage of fragments matching the modern human variant; H_support_deam - percentage of putatively deaminated fragments matching the modern human variant; N_support-percentage of fragments matching the Neandertal variant; N_support_deam - percentage of putatively deaminated fragments matching the Neandertal variant D_support% - percentage of fragments matching the Denisovan variant; D_support_deam% - percentage of putatively deaminated fragments matching the Denisovan variant; S_support - percentage of fragments matching the Sima variant; S_support_deam- percentage of putatively deaminated fragments matching the Sima variant. All of the 95% confidence interval are gotten by an exact binomial teste.