A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia

Kay Prüfer, Cosimo Posth, He Yu, Alexander Stoessel, Maria Spyrou, Thibaut Deviese, Marco Mattonai, Erika Ribechini, Thomas Higham, Petr Velemínský, et al.

HAL Id: hal-03192787

https://amu.hal.science/hal-03192787

Submitted on 8 Apr 2021

Check for updates

OPEN

A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia

Kay Prüfer [1,2,10 ✉], Cosimo Posth [1,3,10 ✉], He Yu [1], Alexander Stoessel[1,2,4], Maria A. Spyrou [1,3], Thibaut Deviese[5,6], Marco Mattonai [7], Erika Ribechini [7], Thomas Higham[5], Petr Velemínský[8], Jaroslav Brůžek [9] and Johannes Krause[1,2 ✉]

Modern humans expanded into Eurasia more than 40,000 years ago following their dispersal out of Africa. These Eurasians carried ~2–3% Neanderthal ancestry in their genomes, originating from admixture with Neanderthals that took place sometime between 50,000 and 60,000 years ago, probably in the Middle East. In Europe, the modern human expansion preceded the disappearance of Neanderthals from the fossil record by 3,000–5,000 years. The genetic makeup of the first Europeans who colonized the continent more than 40,000 years ago remains poorly understood since few specimens have been studied. Here, we analyse a genome generated from the skull of a female individual from Zlatý kůň, Czechia. We found that she belonged to a population that appears to have contributed genetically neither to later Europeans nor to Asians. Her genome carries ~3% Neanderthal ancestry, similar to those of other Upper Palaeolithic hunter-gatherers. However, the lengths of the Neanderthal segments are longer than those observed in the currently oldest modern human genome of the ~45,000-year-old Ust'-Ishim individual from Siberia, suggesting that this individual from Zlatý kůň is one of the earliest Eurasian inhabitants following the expansion out of Africa.

Only three genomes have been recovered from individuals that fall close in time to the settlement of Europe and Asia more than 40 thousand years ago (ka)[1,2]. A complete genome has been produced from the ~45,000-year-old remains of Ust'-Ishim, a Siberian individual who showed no genetic continuity to later Eurasians[3]. This contrasts with the ~40,000-year-old East Asian individual from Tianyuan whose genome is more closely related to many present-day Asians and Native Americans than to Europeans[4]. From Europe, only the partial genome of an individual called Oase 1 and dated to ~40 ka has been recovered, and this showed no evidence of shared ancestry with later Europeans[5]. However, Oase 1 carried more Neanderthal ancestry (6–9%) than other modern human genomes sequenced to date, owing to admixture with Neanderthals that occurred within the six generations before the individual lived.

Here, we study genome sequences generated from a largely complete ancient skull that was discovered alongside other skeletal elements in 1950 inside the Koněprusy cave system in present-day Czechia[6,7] (Fig. 1, Extended Data Fig. 1 and Supplementary Section 1). All skeletal elements were found to originate from one adult female individual called Zlatý kůň (Golden Horse) after the hill on top of the cave system. Archaeological investigations ascribed the stone and bone tools retrieved from the cave to the early Upper Palaeolithic. However, the artefacts in association with this individual could not be confidently assigned to any specific cultural technocomplex[6,8]. The remains were first thought to be at least 30,000 years old in accordance with morphological and stratigraphic information and

the type of associated faunal remains[8,9]. Moreover, damage on the left side of the frontal human bone was interpreted as biting and gnawing by hyenas, which went extinct from central Europe around 24 ka[10,11]. Whereas direct radiocarbon dating resulted in a much younger date of ~15 ka (12,870 ± 70 years BP; GrA-13696)[12], a recent craniometric analysis that included a virtual reconstruction of the Zlatý kůň skull supports that the individual lived before the last glacial maximum[13].

In an attempt to clarify the age of Zlatý kůň, we radiocarbon dated a cranial bone fragment, resulting in a significantly older date of ~27 ka (23,080 ± 80 years BP; MAMS-36077) compared with the first direct date. A third date, comprising a solvent pre-wash treatment followed by ultrafiltration on the same bone fragment, produced a younger date of ~19 ka (15,537 ± 65 years BP; OxA-38602)[14]. The large discrepancies between the three direct dates suggest that the Zlatý kůň specimen is highly contaminated and that radiocarbon dating on bulk collagen may be unreliable (Supplementary Section 2 and Extended Data Fig. 2). We therefore extracted the amino acid hydroxyproline from leftover collagen to attempt to date a compound-specific fraction from the bone[15]. This yielded the oldest determination of ~34 ka (29,650 ± 650 years BP; OxA-38022). However, we suspect this is also artificially young due to the presence of trace exogenous contaminants derived from animal glue, as supported by genetic analysis (discussed below and in Supplementary Section 2). We therefore conclude that the hydroxyproline determination reflects a minimum age, with the true age likely to be much older.

[1]Max Planck Institute for the Science of Human History, Jena, Germany. [2]Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany. [3]Institute for Archaeological Sciences, Archaeo- and Palaeogenetics, University of Tübingen, Tübingen, Germany. [4]Institute of Zoology and Evolutionary Research, Friedrich Schiller University Jena, Jena, Germany. [5]Oxford Radiocarbon Accelerator Unit, Research Laboratory for Archaeology and the History of Art, School of Archaeology, University of Oxford, Oxford, UK. [6]Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE), Aix-Marseille University, CNRS, IRD, INRAE, Collège de France, Aix-en-Provence, France. [7]Department of Chemistry and Industrial Chemistry, University of Pisa, Pisa, Italy. [8]Department of Anthropology, National Museum, Prague, Czech Republic. [9]Department of Anthropology and Human Genetics, Faculty of Science, Charles University, Prague, Czech Republic. [10]These authors contributed equally: Kay Prüfer, Cosimo Posth. ✉e-mail: pruefer@eva.mpg.de; cosimo.posth@uni-tuebingen.de; krause@eva.mpg.de
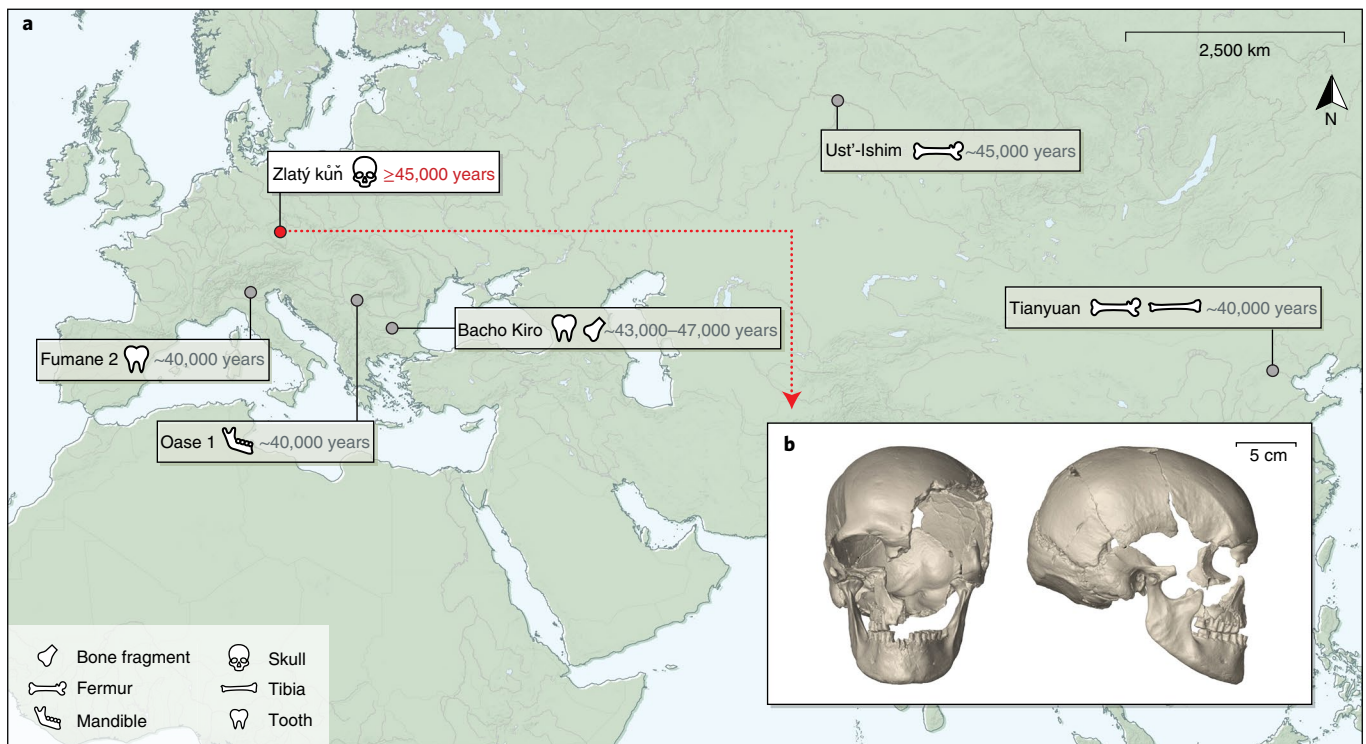
Fig. 1 | The Zlatý kůň fossil. a, Locations of the Koněprusy cave, where the Zlatý kůň human remains were found, and of other fossils with an age of at least ~40,000 years that yielded genome-wide data (Ust'-Ishim, Oase 1 and Tianyuan) or mtDNA (Fumane 2 and Bacho Kiro). **b**, Micro-computed tomography-based virtual reconstruction of the Zlatý kůň skull in frontal and lateral view. The map in **a** was created with QGIS[47] using Natural Earth[48] vector data.

We extracted DNA from ~15 mg bone powder from the Zlatý kůň petrous portion of the temporal bone and first enriched for and sequenced the mitochondrial genome (mtDNA) to ~150-fold coverage (Methods). Around 4% of the mtDNA sequences were estimated to stem from human contamination (Supplementary Section 3). The reconstructed mtDNA belongs to haplogroup N and its branch length, measured as the number of accumulated substitutions, is similar to those of the currently oldest sequenced modern human mtDNA genomes (Fig. 2a and Extended Data Fig. 3), including the recently published mtDNAs from Bacho Kiro, a cave in Bulgaria with remains dating to 43–47 ka[1]. Bayesian tip dating suggests that Zlatý kůň lived ~43 ka (95% highest posterior density = 31.5–52.6 ka).

To study the nuclear genome, we sequenced ~20 million DNA fragments after targeted enrichment with oligonucleotide probes for 1.24 million single nucleotide polymorphisms (SNPs)[16]. A total of ~678,000 targeted SNPs (54%) were covered at least once after genome-wide enrichment (capture dataset). In addition, we sequenced ~4 billion random DNA fragments from the same DNA library of Zlatý kůň, resulting in ~3.8-fold genomic coverage (shotgun dataset). In line with the sex assignment based on morphology[13], the X chromosome and autosomes showed similar coverage, indicating that Zlatý kůň is female (Extended Data Fig. 4). The presence of Y chromosomal sequences suggested that up to 4% of the nuclear DNA sequences in the shotgun dataset originate from male contamination. Estimates based on linkage disequilibrium[17] suggest that nuclear contamination is <1% in the capture dataset and ~2% in the shotgun dataset (Supplementary Section 4). The majority of the Zlatý kůň shotgun sequences (~3.2-fold out of ~3.8-fold total) have been generated from a single-stranded DNA library that allows for the quantification of contamination with an explicit model of DNA damage in the DNA molecules[18]. This model yielded an estimate of contamination of 0.1% (s.e. = ±2.0%) (Supplementary Section 4).

We used the non-human fraction of the shotgun data to further investigate whether the use of animal glue could have influenced our attempts at radiocarbon dating of the Zlatý kůň skull. Searching a metagenomic database, we found that the highest proportion of non-human mammalian shotgun sequences aligned to bovids (Supplementary Section 2). We were able to reconstruct ~95% of the bovid mtDNA from the shotgun sequences of the single-stranded library and found that it falls within the most common modern European cattle haplogroup[19] in a phylogenetic analysis (Extended Data Fig. 5). Low levels of substitutions that are indicative of ancient DNA damage suggest that the cattle sequences do not derive from present-day laboratory contaminants (Extended Data Fig. 5). Taken together, these results suggest that the Zlatý kůň skull has been preserved with glue from cattle that penetrated into the sequenced petrous bone.

To gain insight into the genetic relationship of Zlatý kůň to present-day and ancient individuals, we calculated summary statistics based on the sharing of alleles (f_3, f_4 and D statistics[20]) with our capture and shotgun datasets. We first compared Zlatý kůň with present-day European and Asian individuals using an African population (Mbuti) as an outgroup and found that Zlatý kůň shares more alleles with Asians than with Europeans (Extended Data Fig. 6). A closer relationship to Asians has also been observed for other Upper Palaeolithic and Mesolithic European hunter-gatherers compared with present-day Europeans and can be explained by ancestry in present-day Europeans from a deeply divergent out-of-Africa lineage referred to as basal Eurasian[21]. European hunter-gatherers generally do not carry basal Eurasian ancestry, whereas such ancestry is widespread among ancient hunter-gatherers from the Caucasus, Levant and Anatolia[22–24]. When we tested European hunter-gatherers without basal Eurasian ancestry against ancient and present-day Asians, we found that none of these comparisons indicate a closer
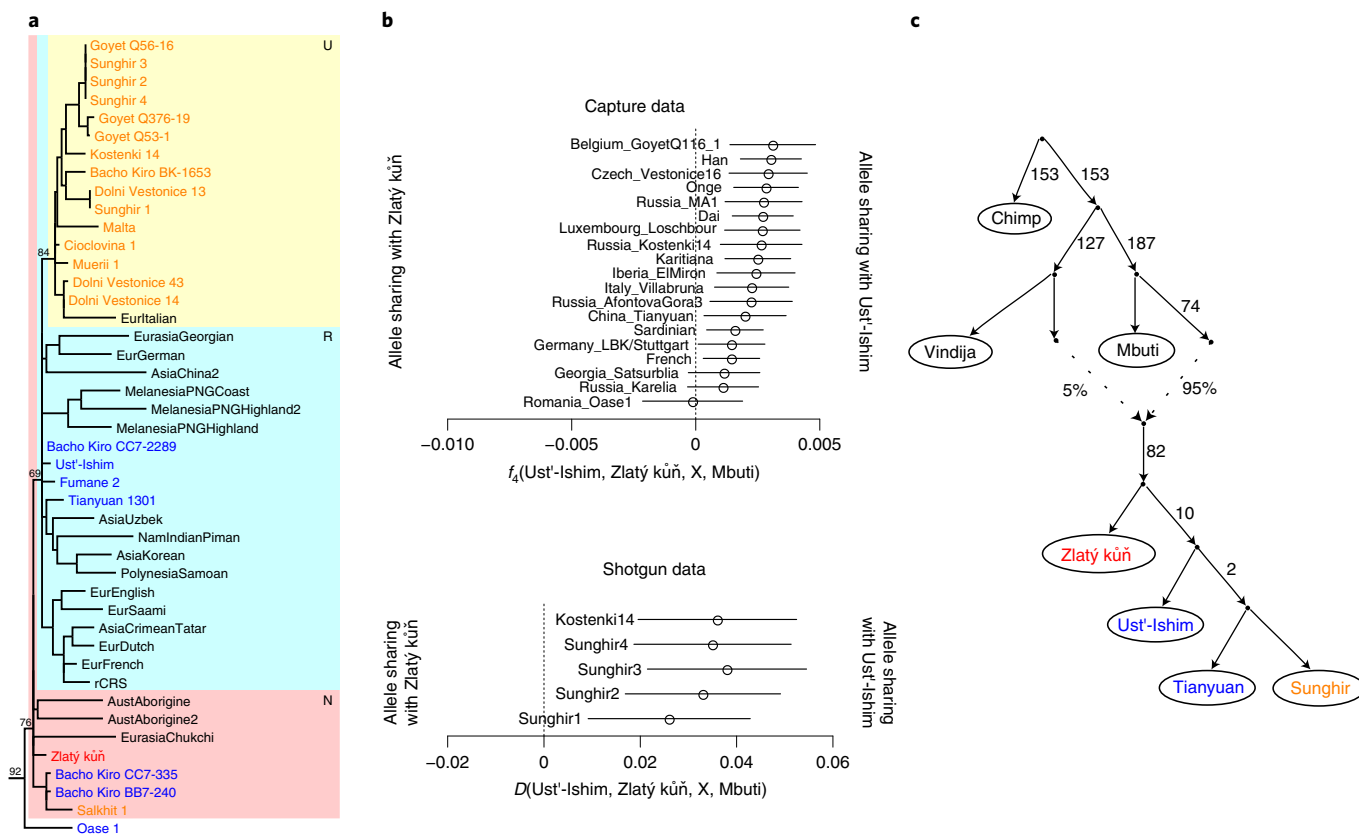
Fig. 2 | Genetic relationship with present-day and ancient humans. a, mtDNA haplogroup N of a maximum-parsimony phylogenetic tree of mtDNA from Zlatý kůň (red font), Upper Palaeolithic individuals ~40 ka or older (blue) or between ~37 and ~24 ka (orange), and present-day individuals (black) (the entire tree is presented in Extended Data Fig. 3). **b**, Analysis of nuclear sequences showing that Ust'-Ishim shares more alleles with European and Asian hunter-gatherers and later Eurasians than does Zlatý kůň. The error bars represent two standard errors. **c**, Admixture graph of the relationship inferred from the nuclear capture dataset. Zlatý kůň diverges earlier than Ust'-Ishim and the ancestors of later Eurasian populations represented here by the Upper Palaeolithic Tianyuan and Sunghir genomes (highest outlier $|Z| = 3$). A single gene flow event from Neanderthals into the ancestor of all tested modern humans fits the data. Colours for individuals follow the same scheme as in panel **a**.

relationship of Zlatý kůň with either group (Supplementary Sections 5 and 9 and Extended Data Fig. 7). This suggests that Zlatý kůň falls basal to the split of the European and Asian populations.

To date, only two ancient Eurasian genomes have been produced from individuals who, like Zlatý kůň, appear to fall basal to the split of Europeans and Asians: Ust'-Ishim and Oase 1. To test whether Zlatý kůň derives from the same population as Ust'-Ishim, we tested for a closer relationship to it compared with other ancient Eurasian hunter-gatherers[24–26]. Interestingly, we found that Ust'-Ishim shares more ancestry with later Eurasian individuals (Fig. 2b). This suggests that Zlatý kůň was part of a population that split earlier from the population that later gave rise to Ust'-Ishim and other Eurasian populations (Fig. 2c). Due to the limited data for Oase 1, we are unable to clarify whether Zlatý kůň and Oase 1 derive from the same or separate populations.

Around 6–9% of the genome of Oase 1 is derived from Neanderthals, compared with 2–3% in present-day and ancient Eurasians[5,27,28]. To test whether a higher contribution is also present in Zlatý kůň, we calculated Neanderthal ancestry on the shotgun dataset as the excess of shared alleles with a Neanderthal as opposed to an African and normalized this quantity by the expected sharing between two Neanderthals as opposed to an African (f_4 ratios[20]; Supplementary Section 6). Zlatý kůň is estimated to carry 3.2% (s.e. = ±0.32%) Neanderthal ancestry, which is the highest value among six Upper Palaeolithic and one Mesolithic Eurasian hunter-gatherers with genome-wide data (range = 3.0–2.1%).

However, this difference was not significant at a level of two standard errors for five out of seven comparisons (Fig. 3a).

To study the distribution of Neanderthal ancestry along the genome, we first determined 430,075 sites on autosomes where the genome of a high-coverage European Neanderthal carries on both chromosomes a variant that is not observed in more than 99.9% of present-day Africans and great ape outgroups (Supplementary Section 7). Of the 166,721 sites that are covered by Zlatý kůň shotgun data, 4,480 (2.7%) carried the Neanderthal allele. Neanderthal sites in the Zlatý kůň genome cluster into segments where they occur at high frequency (~50%; Fig. 3b) and we used this clustering to label segments of likely Neanderthal ancestry with a hidden Markov model (Supplementary Section 7). One of the Neanderthal segments falls within a large region on chromosome 1 that shows little to no evidence of Neanderthal ancestry in present-day humans[29] (Extended Data Fig. 8). This suggests that this desert of Neanderthal ancestry had not been fully formed at the time Zlatý kůň lived.

Recombination will break long Neanderthal segments into shorter segments over time. To gain more insight into the timing of Neanderthal admixture in Zlatý kůň, we scaled the length of the Neanderthal segments using either an African American map (AA map)[30] or the Decode recombination map (deCODE map)[31] and compared the genetic length of the 100 longest segments in Zlatý kůň with those identified in other early Eurasian hunter-gatherers using the same method (Extended Data Fig. 9). We found that Zlatý kůň carries segments that are on average longer than those of all
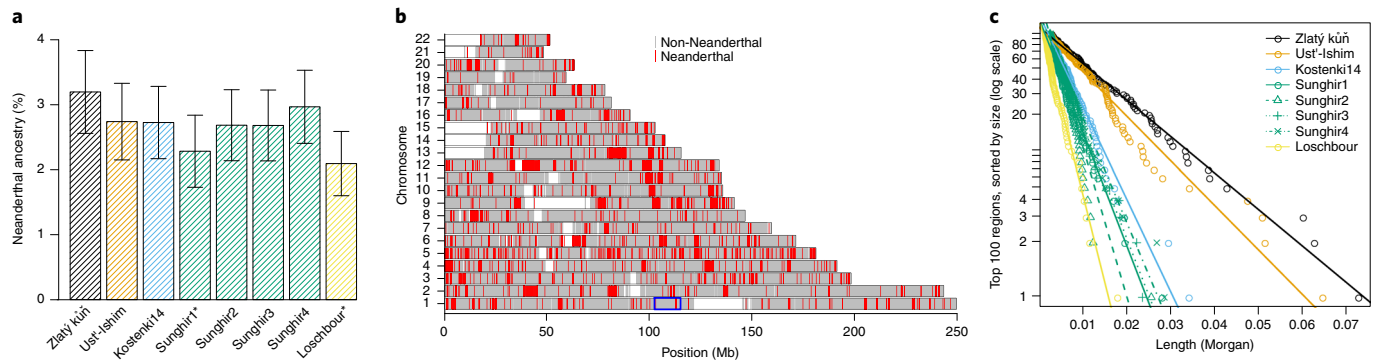
Fig. 3 | Neanderthal ancestry in Zlatý kůň and ancient Eurasian genomes. a, Estimate of Neanderthal ancestry in ancient Eurasian hunter-gatherer genomes. The error bars indicate two standard errors. Individuals whose names are marked with an asterisk fall outside of the error bars for Zlatý kůň. **b**, Segments of Neanderthal ancestry in Zlatý kůň. The blue box shows the location of a desert of Neanderthal ancestry in present-day non-Africans. **c**, Length of the 100 largest Neanderthal segments in the genomes of Zlatý kůň and other Upper Palaeolithic and Mesolithic Eurasian hunter-gatherers. The y axis is logarithmic and the lines indicate a linear fit. The colours are as in **a**.

other Eurasian hunter-gatherers (Fig. 3c). Assuming that recombination breaks Neanderthal ancestry into shorter segments every generation (see ref. [5] and Supplementary Section 8), we estimated that the last admixture with Neanderthals occurred ~70–80 generations before Zlatý kůň lived (AA map: 74 generations (95% confidence interval (CI) = 61–89); deCODE map: 78 generations (95% CI = 64–94)). In contrast, the genome of the currently oldest sequenced modern human, Ust'-Ishim, yielded significantly higher estimates of 94 generations (95% CI = 77–113) and 99 generations (95% CI = 81–119) using the AA map and deCODE map, respectively. Estimating ages using segments called by admixfrog (a software for inferring local Neanderthal ancestry[32]) yielded comparable results (Supplementary Section 8).

To estimate the time of admixture independent of calling Neanderthal ancestry segments, we applied a previously published method that is based on the correlation of the state of Neanderthal informative sites over increasing distances[33]. Applying this method to the Zlatý kůň shotgun dataset yielded an estimate of 63 generations since admixture (s.e. = ±0.6), whereas Ust'-Ishim was estimated to have lived 84 generations after the admixture (s.e. = ±1.3).

Most of the Neanderthal ancestry in present-day and ancient humans probably originates from a common admixture event with a group of Neanderthals who were more closely related to European Neanderthals than to a Neanderthal from the Altai Mountains[28]. To test whether the Neanderthal ancestry in Zlatý kůň shows the same relationship, we used D statistics to compare the sharing of alleles with the high-coverage genomes of a European and an Asian Neanderthal in addition to five published low-coverage genomes from late Neanderthals from Europe[34]. Zlatý kůň showed no significant difference compared with Ust'-Ishim in its sharing of alleles with late Neanderthals, in line with the similar proportion of Neanderthal ancestry in these two hunter-gatherers (Supplementary Section 6 and Extended Data Fig. 10).

Assuming a common Neanderthal admixture event, these results suggest that Zlatý kůň is of approximately the same age as the ~45,000-year-old Ust'-Ishim individual or up to a few hundred years older. However, if a second Neanderthal admixture event affected Ust'-Ishim after the initial common Neanderthal admixture, as was previously suggested[33], Zlatý kůň could be even several thousands of years older than Ust'-Ishim. We have not found support for a second Neanderthal admixture event in the Zlatý kůň data (Supplementary Section 8).

The genetic identity of the modern humans who colonized Eurasia before ~40 ka remains largely unknown. Here, we sequenced and analysed the genome of an early European and determined that she was part of a population that formed before the populations that gave rise to present-day Europeans and Asians split from one another. Our estimated age of ~45,000 years or even older could make Zlatý kůň the oldest European individual with a largely preserved skull[13]. As for Ust'-Ishim and Oase 1, Zlatý kůň shows no genetic continuity with modern humans who lived after ~40 ka. It is possible that this discontinuity is linked to the Campanian Ignimbrite eruption ~39 ka that severely affected the climate in the Northern Hemisphere and that may have reduced the viability of Neanderthals and early modern humans in large parts of western Eurasia[35,36]. Whether the modern humans who lived before the turnover event, such as the Oase 1 and Bacho Kiro individuals[3,5], belonged to the same early European population can only be resolved with further genome-wide data from those individuals[37]. Future genetic studies of these and other early European individuals will thus help to reconstruct the history of these first modern humans who expanded into Eurasia after the out-of-Africa event and before the major dispersal that gave rise to modern-day non-African populations.

Note added in proof: We refer readers to related work by M. Hajdinjak et al.[37] who analysed nuclear sequences from Bacho Kiro individuals that dated to around 45,000 years ago.

Methods

Laboratory procedures and shotgun sequencing. All laboratory procedures were conducted in the dedicated ancient DNA facilities of the Max Planck Institute for the Science of Human History in Jena, Germany. The Zlatý kůň cranium was sampled from the base of the cranium with a dentist drill after removal of a thin layer of bone powder. Two aliquots of ~15 mg bone powder were sampled and one was used to extract DNA using an established protocol[38]. A double-stranded DNA library with partial uracil-DNA glycosylase treatment was generated from 25% of DNA extract[39]. After quantification, the library was double indexed[40] and quantified again to establish the number of PCR cycles necessary to reach the amplification plateau[41]. The resulting library was diluted and shotgun sequenced on two lanes of an Illumina HiSeq 4000 platform for 2 × 50 cycles. An additional 30% of the same extract was used to generate a single-stranded library[42] on an automated liquid handling system (Bravo; Agilent Technologies). After indexing, the library was amplified for 30 cycles followed by a reconditioning PCR to remove heteroduplexes. The resulting library was diluted and shotgun sequenced on an entire flow cell (eight lanes) of an Illumina HiSeq 4000 platform for 1 × 75 cycles.

Basic processing and sequence alignment of shotgun data. A total of ~580 million paired-end reads from the double-stranded library matched the correct indices, allowing for up to one mismatch per index, and were further processed with EAGER[43]. Adapter sequences were trimmed, filtered for a minimum length of 30 base pairs (bp) and mapped to the hg19 human reference genome using BWA[44] with the following parameters: -n 0.01 -l 16500. Approximately 10% of sequences mapped with an average fragment length of 49 bp. These sequences were filtered for a minimum mapping quality of 30 and duplicates were removed using Dedup[43]. Sequences showed C to T exchanges to

the human reference, indicative of ancient DNA damage, with a frequency of ~13% at both terminal positions[45,46]. The first and last two bases of shotgun sequences were masked (set to N) to reduce the effect of damage-associated substitutions on subsequent analyses.

An additional ~3.355 billion single-end reads were produced from the single-stranded library on a dedicated run. All sequences were processed using EAGER with the same parameters as above, except the mapping quality filter was set to 25. Around 15% of reads mapped to the human reference, with an average fragment length of 46 bp and 30% of C to T substitutions at the 5′ molecule termini.

Target enrichment, sequencing and processing. The indexed library was further amplified to perform targeted enrichment of both the complete mtDNA (mtDNA capture)[5] and ~1.24 million nuclear SNPs (1240K capture)[16] followed by HiSeq paired-end sequencing and index filtering, resulting in 600,000 and 20 million reads, respectively. MtDNA capture and 1240K capture data were mapped against the mtDNA reference sequence (the Revised Cambridge Reference Sequence (rCRS)) using CircularMapper[43] and the human reference genome (GRCh37/ hg19), respectively, with the same parameters as above and the mapping quality filter set to 30.

A random allele was drawn from the 1240K capture data using PileupCaller in pseudohaploid mode (https://github.com/stschiff/sequenceTools). The calling of transversions among the 1.24 million target SNPs considered full sequences, whereas 2 bp at both termini of sequences were trimmed before calling transition SNPs. Finally, the resulting calls were merged with a large genotype dataset of ancient and present-day individuals for the same set of ~1.24 million SNPs (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data; version 37.2).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
All of the sequence data generated and analysed during the current study are available in the European Nucleotide Archive under study accession number PRJEB39040.

References
1. Hublin, J.-J. et al. Initial Upper Palaeolithic *Homo sapiens* from Bacho Kiro Cave, Bulgaria. *Nature* **581**, 299–302 (2020).
2. Benazzi, S. et al. Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature* **479**, 525–528 (2011).
3. Fu, Q. et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).
4. Yang, M. A. et al. 40,000-year-old individual from Asia provides insight into early population structure in Eurasia. *Curr. Biol.* **27**, 3202–3208.e9 (2017).
5. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).
6. Prošek, F. et al. The excavation of the 'Zlatý kůň' cave in Bohemia. The report for the 1st research period of 1951 (part 2) [in Czech]. *Československý kras.* **5**, 161–179 (1952).
7. Vlček, E. The Pleistocene man from the Zlatý kůň cave near Koněprusy [in Czech]. *Anthropozoikum* **6**, 283–311 (1957).
8. Prošek, F. The research in the Golden Horse Cave near Koněprusy [in Czech]. *Archeologické Rozhl.* **4**, 206–209 (1952).
9. Vlček, E. in *Catalogue of Fossil Hominids. Part II: Europe* (eds. Oakley, K. et al.) 47–64 (British Museum (Natural History), 1971).
10. Vlček, E. Other findings of the Pleistocene man's remains on Zlatý kůň near Koněprusy [in Czech]. *Archeologické Rozhl.* **9**, 305–310 (1957).
11. Diedrich, C. G. & Zak, K. Prey deposits and den sites of the Upper Pleistocene hyena *Crocuta crocuta spelaea* (Goldfuss, 1823) in horizontal and vertical caves of the Bohemian Karst (Czech Republic). *Bull. Geosci.* **81**, 237–276 (2006).
12. Svoboda, J. A., van der Plicht, J. & Kuželka, V. Upper Palaeolithic and Mesolithic human fossils from Moravia and Bohemia (Czech Republic): some new ¹⁴C dates. *Antiquity* **76**, 957–962 (2002).
13. Rmoutilová, R. et al. Virtual reconstruction of the Upper Palaeolithic skull from Zlatý kůň, Czech Republic: sex assessment and morphological affinity. *PLoS ONE* **13**, e0201431 (2018).
14. Brock, F., Higham, T., Ditchfield, P. & Ramsey, C. B. Current pretreatment methods for AMS radiocarbon dating at the Oxford Radiocarbon Accelerator Unit (ORAU). *Radiocarbon* **52**, 103–112 (2010).
15. Deviese, T., Comeskey, D., McCullagh, J., Ramsey, C. B. & Higham, T. New protocol for compound-specific radiocarbon analysis of archaeological bones. *Rapid Commun. Mass Spectrom.* **32**, 373–379 (2018).
16. Mathieson, I. et al. Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* **528**, 499–503 (2015).
17. Nakatsuka, N. et al. ContamLD: estimation of ancient nuclear DNA contamination using breakdown of linkage disequilibrium. *Genome Biol.* **21**, 199 (2020).
18. Peyrégne, S. & Peter, B. M. AuthentiCT: a model of ancient DNA damage to estimate the proportion of present-day DNA contamination. *Genome Biol.* **21**, 246 (2020).
19. Verdugo, M. P. et al. Ancient cattle genomics, origins, and rapid turnover in the fertile crescent. *Science* **365**, 173–176 (2019).
20. Patterson, N. et al. Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
21. Lazaridis, I. et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
22. Feldman, M. et al. Late Pleistocene human genome suggests a local origin for the first farmers of central Anatolia. *Nat. Commun.* **10**, 1218 (2019).
23. Lazaridis, I. et al. Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
24. Fu, Q. et al. The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
25. Sikora, M. et al. Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).
26. Seguin-Orlando, A. et al. Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
27. Green, R. E. et al. A draft sequence of the Neandertal genome. *Science* **328**, 710–722 (2010).
28. Prüfer, K. et al. A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).
29. Vernot, B. et al. Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235–239 (2016).
30. Hinch, A. G. et al. The landscape of recombination in African Americans. *Nature* **476**, 170–175 (2011).
31. Kong, A. et al. Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* **467**, 1099–1103 (2010).
32. Peter, B. M. 100,000 years of gene flow between Neandertals and Denisovans in the Altai mountains. Preprint at *bioRxiv* https://doi.org/10.1101/2020.03.13.990523 (2020).
33. Moorjani, P. et al. A genetic method for dating ancient genomes provides a direct estimate of human generation interval in the last 45,000 years. *Proc. Natl Acad. Sci. USA* **113**, 5652–5657 (2016).
34. Hajdinjak, M. et al. Reconstructing the genetic history of late Neanderthals. *Nature* **555**, 652–656 (2018).
35. Black, B. A., Neely, R. R. & Manga, M. Campanian Ignimbrite volcanism, climate, and the final decline of the Neanderthals. *Geology* **43**, 411–414 (2015).
36. Giaccio, B., Hajdas, I., Isaia, R., Deino, A. & Nomade, S. High-precision ¹⁴C and ⁴⁰Ar/³⁹Ar dating of the Campanian Ignimbrite (Y-5) reconciles the time-scales of climatic-cultural processes at 40 ka. *Sci. Rep.* **7**, 45940 (2017).
37. Hajdinjak, M. et al. Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature* https://doi.org/10.1038/s41586-021-03335-3 (2021).
38. Dabney, J. et al. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc. Natl Acad. Sci. USA* **110**, 15758–15763 (2013).
39. Rohland, N., Harney, E., Mallick, S., Nordenfelt, S. & Reich, D. Partial uracil–DNA–glycosylase treatment for screening of ancient DNA. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20130624 (2015).
40. Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* **40**, e3 (2012).
41. Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).
42. Gansauge, M.-T., Aximu-Petri, A., Nagel, S. & Meyer, M. Manual and automated preparation of single-stranded DNA libraries for the sequencing of DNA from ancient biological remains and other sources of highly degraded DNA. *Nat. Protoc.* **15**, 2279–2300 (2020).
43. Peltzer, A. et al. EAGER: efficient ancient genome reconstruction. *Genome Biol.* **17**, 60 (2016).
44. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
45. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc. Natl Acad. Sci. USA* **104**, 14616–14621 (2007).
46. Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F. & Orlando, L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* **29**, 1682–1684 (2013).
47. QGIS v.3.12 (QGIS, accessed 21 February 2020); https://qgis.org/en/site/
48. Natural Earth vector map data (Natural Earth, accessed 16 March 2020); https://www.naturalearthdata.com/downloads/

49. Bronk Ramsey, C. Bayesian analysis of radiocarbon dates. *Radiocarbon* **51**, 337–360 (2009).

50. Reimer, P. J. et al. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).

Acknowledgements

Author contributions

K.P., C.P., H.Y., M.A.S. and J.K. analysed the genetic data. A.S. performed the micro-computed tomography scanning and created the virtual reconstruction of the skull. T.D., M.M., E.R. and T.H. performed the radiocarbon dating and elemental analysis. P.V. and J.B. provided archaeological data. K.P., C.P. and J.K. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-021-01443-x.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-021-01443-x.

Correspondence and requests for materials should be addressed to K.P., C.P. or J.K.

Peer review information *Nature Ecology & Evolution* thanks Carles Lalueza-Fox for their contribution to the peer review of this work.

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

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

OxCal v4.4.2 Bronk Ramsey (2020); r:5 Atmospheric data from Reimer et al (2020)

R_Date OxA-38022

R_Date OxA-38602

R_Date MAMS-36077

R_Date GrA-13696

45000　40000　35000　30000　25000　20000　15000

Calibrated date (calBP)
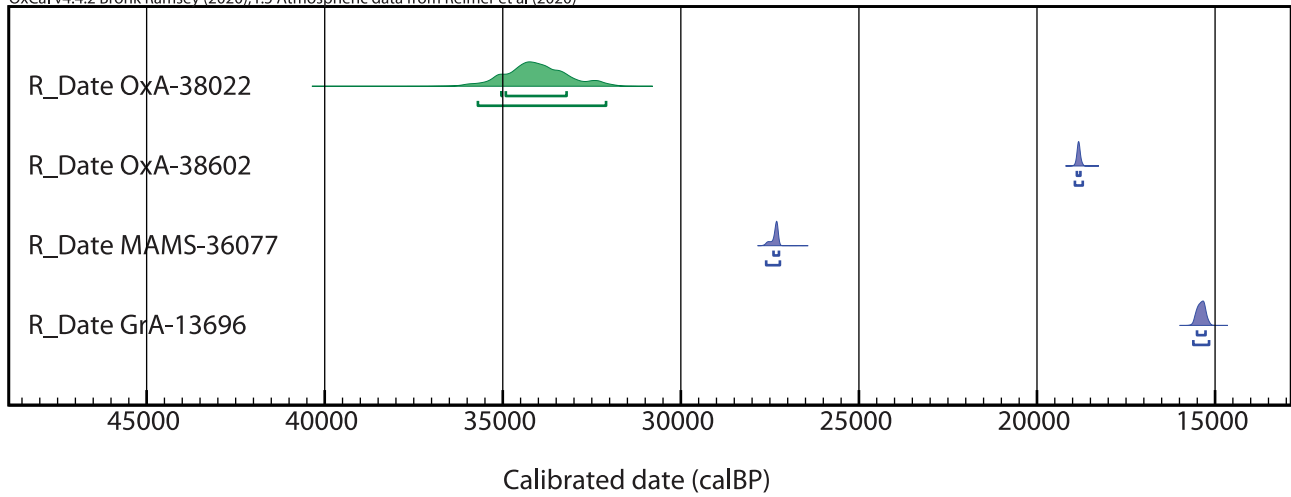
Extended Data Fig. 2 | Four direct radiocarbon dates calibrated in years before present (calBP) with 68.3% and 95.4% probability. OxA-38022 refers to the HYP dating performed in Oxford. The dates were calibrated using OxCal 4.4[46] and the IntCal20 calibration curve[49,50].
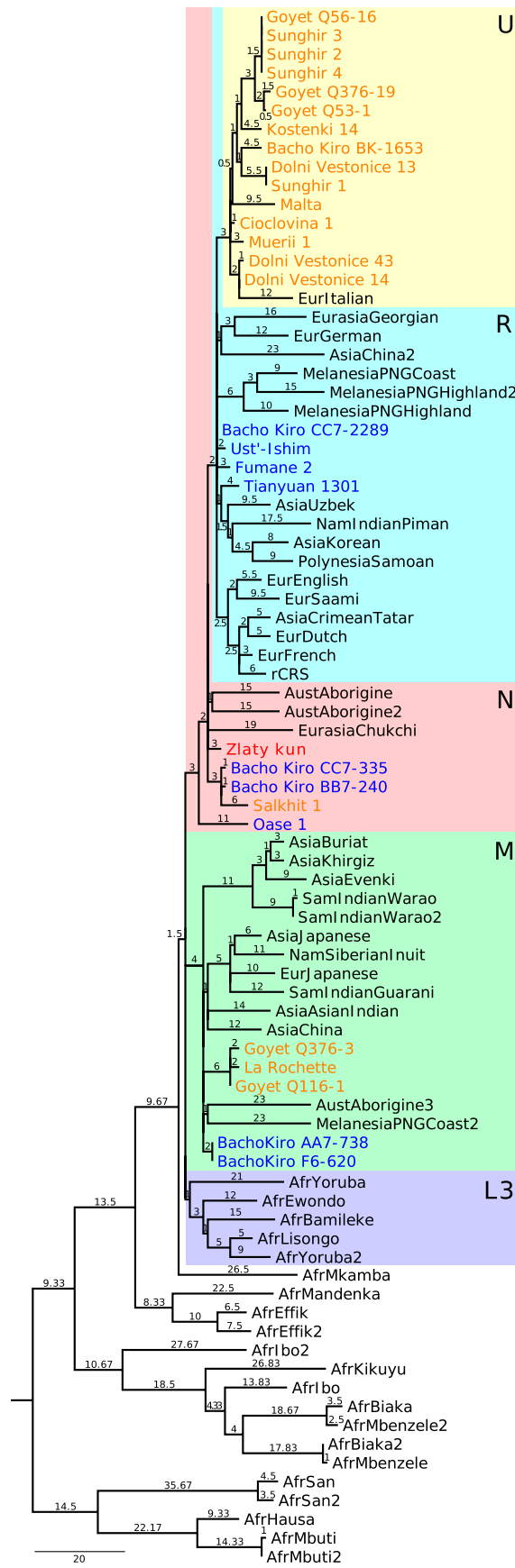
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Maximum Parsimony tree of 54 modern-day worldwide mtDNA sequences, 28 Palaeolithic Eurasian mtDNAs, Zlatý kůň mtDNA and a Neanderthal mtDNA as outgroup. Text color relates to radiocarbon date, that is red (undetermined), blue (45–40 kya) and yellow (40–24 kya). Numbers show the number of substitutions along lineages; capital letters on the right side of the tree indicate mtDNA haplogroups.
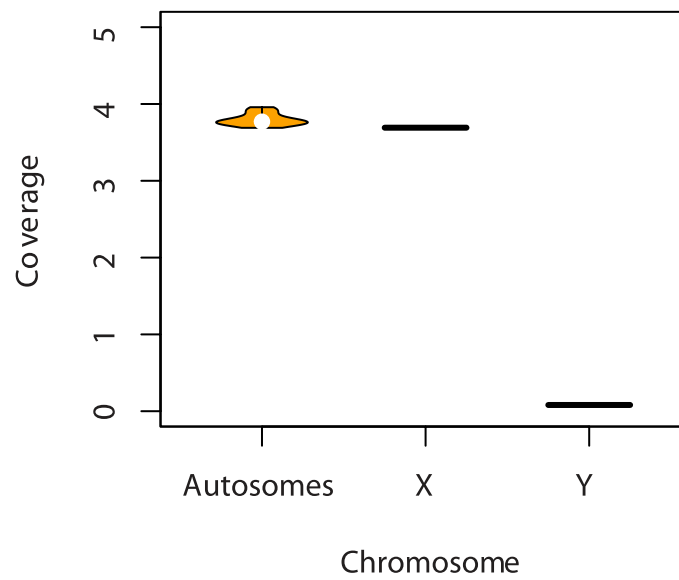
Extended Data Fig. 4 | Coverage on sex chromosomes and autosomes. Average coverage by shotgun sequences for chromosome X, Y (horizontal dashes) and autosomes (violin plot).
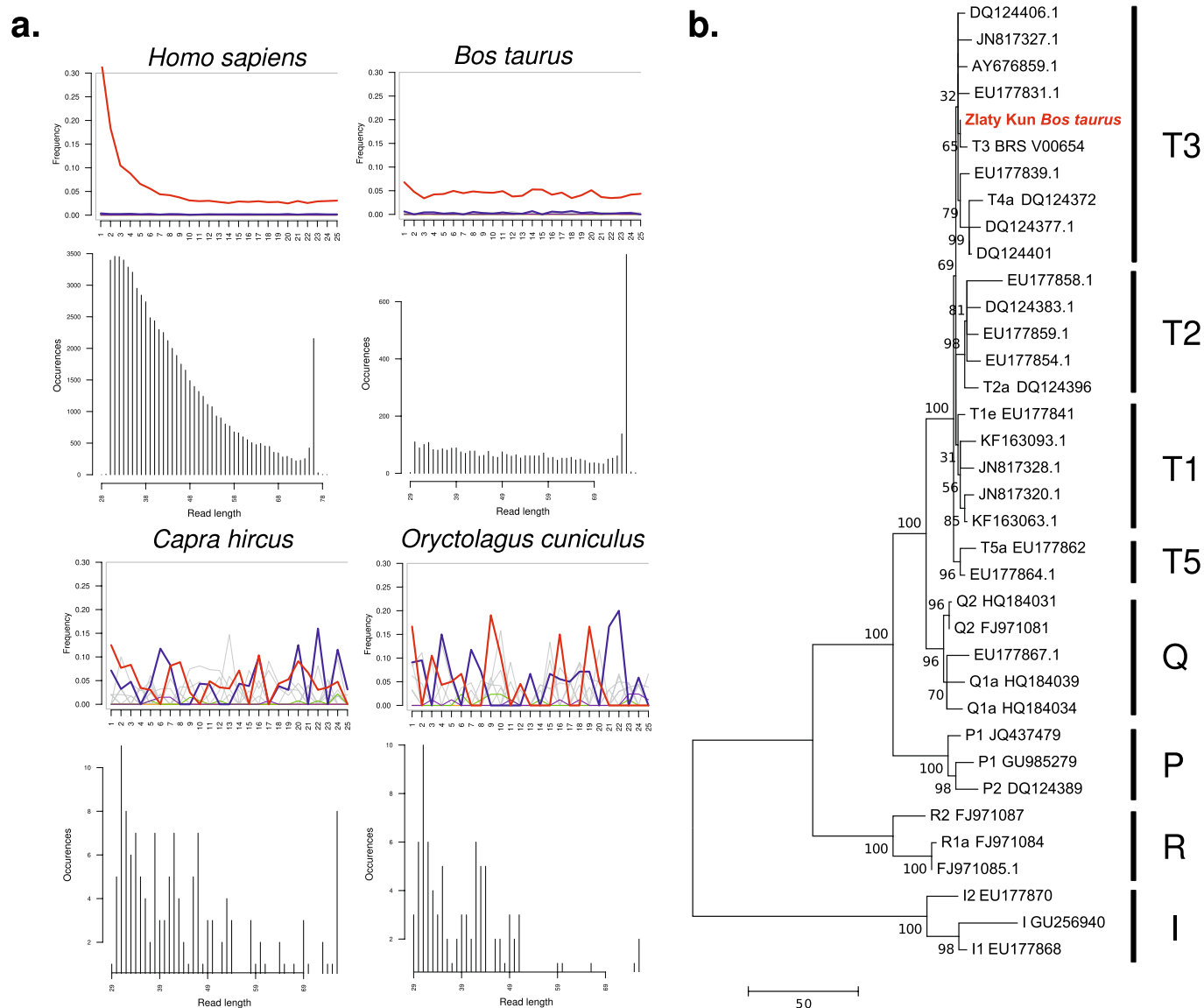
Extended Data Fig. 5 | Characteristics of *Bos taurus* mtDNA sequences and genome. a, C-to-T substitutions at the 5′ end and read length distributions of mtDNA sequences mapping specifically to the mtDNA references of *Homo sapiens*, *Bos taurus*, *Capra hircus* and *Oryctolagus cuniculus*. Note that a single-end 75-cycle chemistry kit was used for the retrieval of the reported sequencing reads and, therefore, DNA fragments of length 76 bp or above have not been sequenced at their full length and are binned here together. **b**, Maximum parsimony tree of 35 present-day and ancient cattle mtDNAs and the *Bos taurus* mtDNA reconstructed from the deep shotgun sequencing of the DNA extracted from the Zlatý kůň petrous bone. Cattle haplogroups are listed on the right side of the tree.
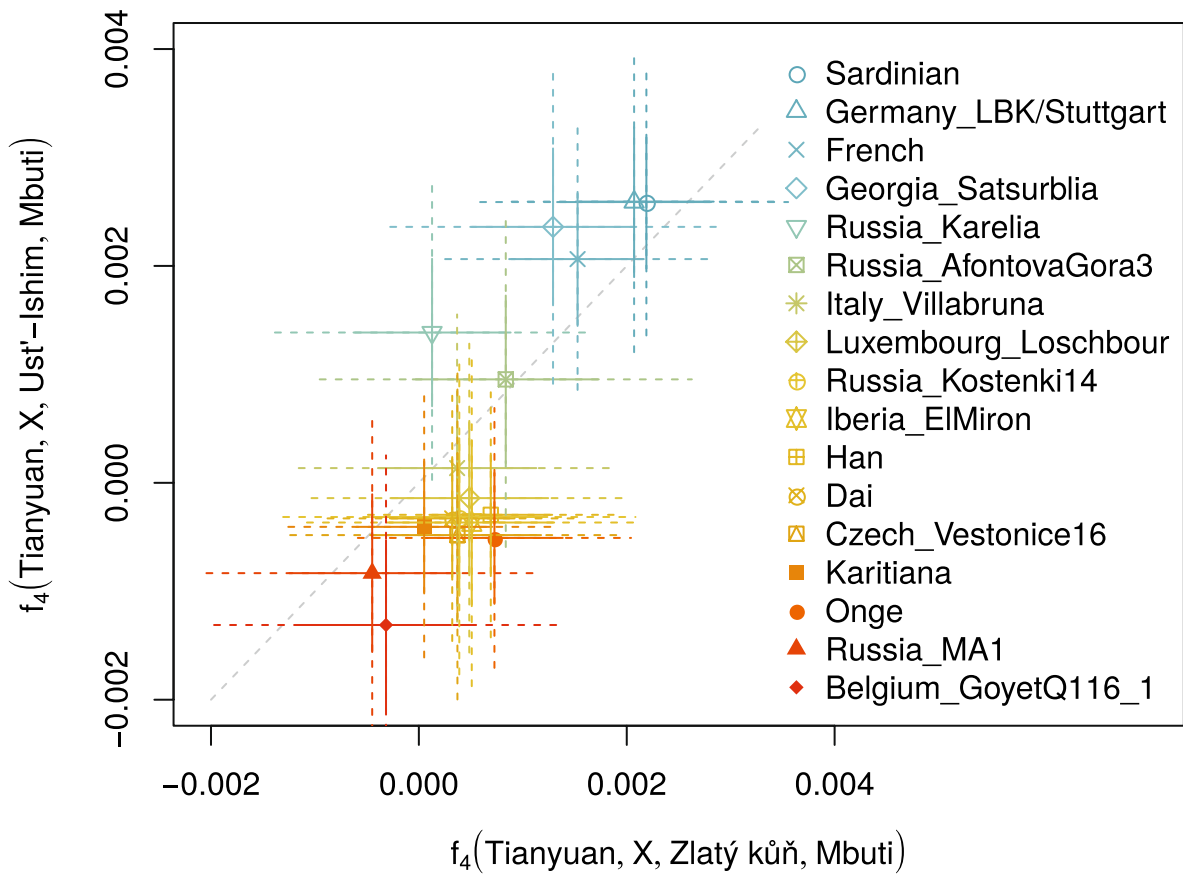
Extended Data Fig. 6 | Comparable signal of Basal Eurasian ancestry in Zlatý kůň and Ust'-Ishim. Zlatý kůň's asymmetric sharing of alleles with some Europeans compared to Asians (x-scale) is also observed to a comparable degree in Ust'-Ishim (y-scale). The ancient genome from Tianyuan is used as a proxy for Asians. The solid and dash error bars correspond to one and two standard errors, respectively.
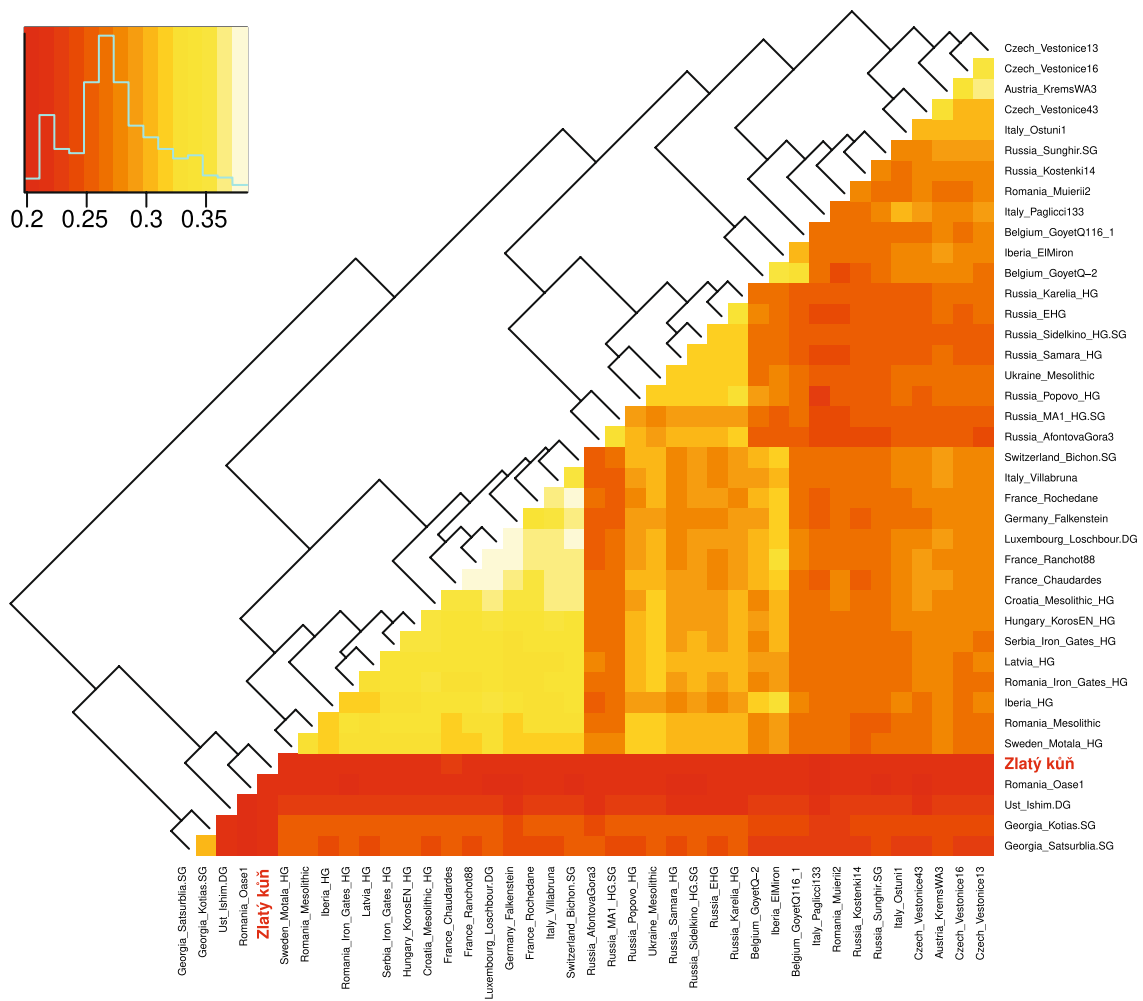
Extended Data Fig. 7 | Heat-map matrix of pairwise f3-outgroup statistics with Mbuti as outgroup. Zlatý kůň shows no closer affinity to Ust'-Ishim and Oase 1 or to any later hunter-gatherer individuals.
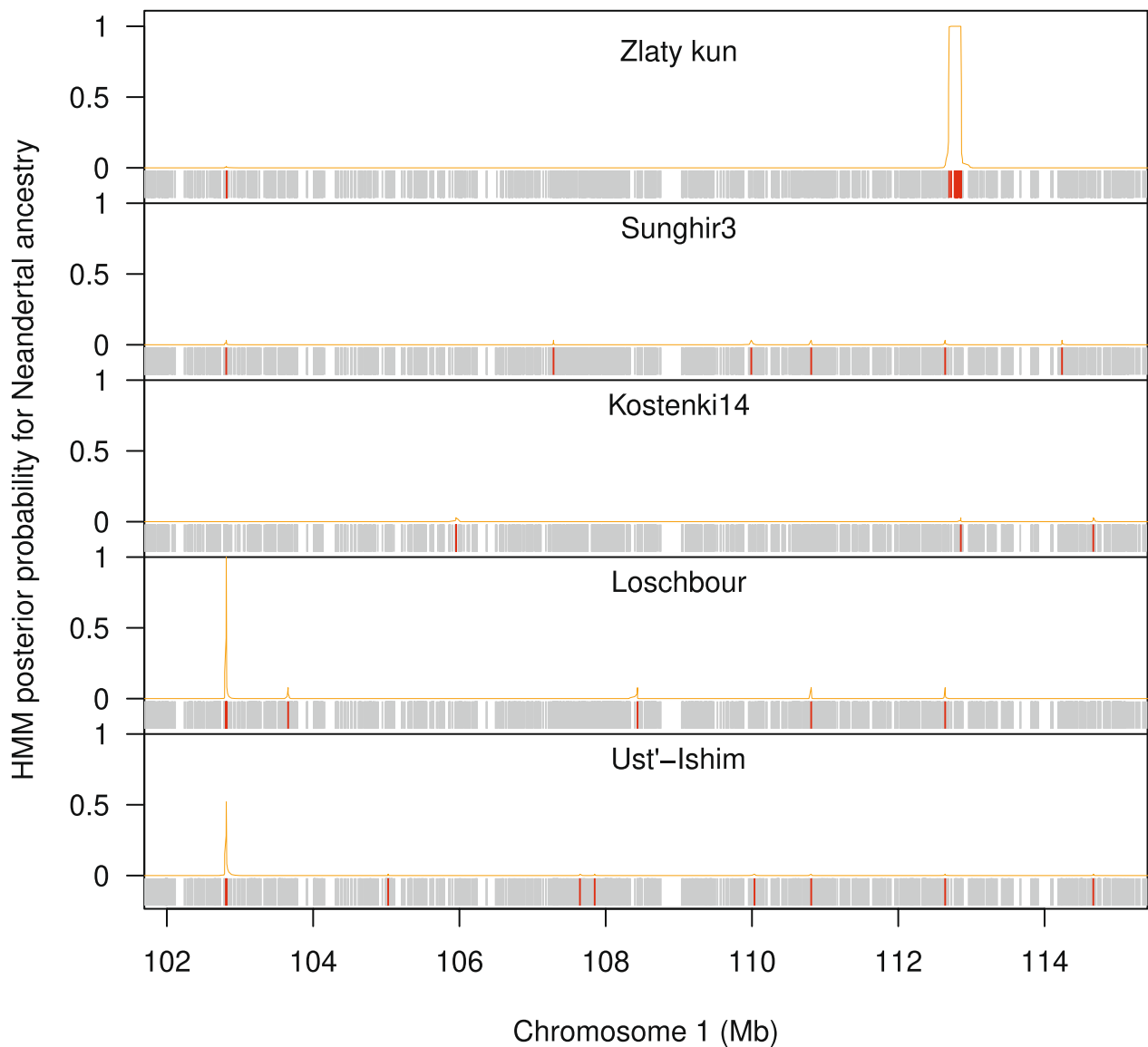
Extended Data Fig. 8 | Neanderthal desert region on chromosome 1 in Zlatý kůň and 5 high-coverage ancient genomes. Each panel shows Neanderthal-shared variants in red and non-shared variants in gray. The orange line indicates the posterior decoding probabilities for Neanderthal ancestry from the hidden Markov model approach. The Neanderthal ancestry segment in Zlatý kůň (chr1:112696231-112855758) is supported by 17 Neanderthal shared variants. A short region, spanning only 6 bp (chr1:102812075-102812081) is called in Loschbour and contains four sites that are Neanderthal shared. The short length and high density of sites do not support a true Neanderthal origin.
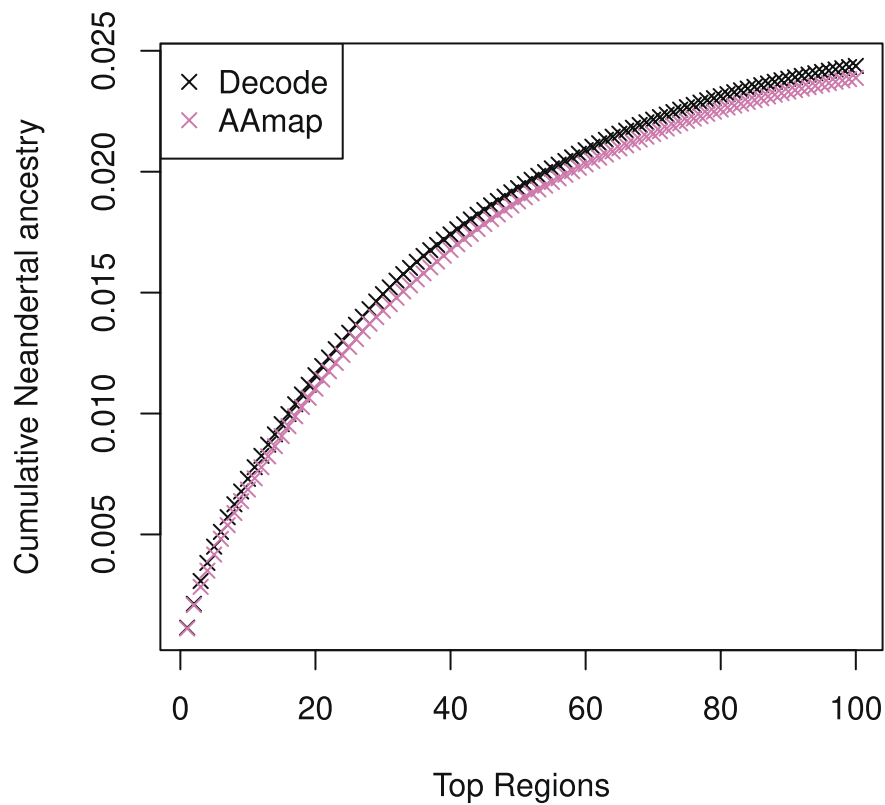
Extended Data Fig. 9 | Fraction of the Zlatý kůň genome covered by the top 100 called introgressed regions sorted by length. The fraction of the genome covered by Neanderthal ancestry (y-scale) was calculated by dividing the cumulative length in Morgan of introgressed regions up to and including the nth longest region shown on the x-scale by two times the total length of the autosomes in Morgan. Results are shown for the African American or Decode recombination maps.
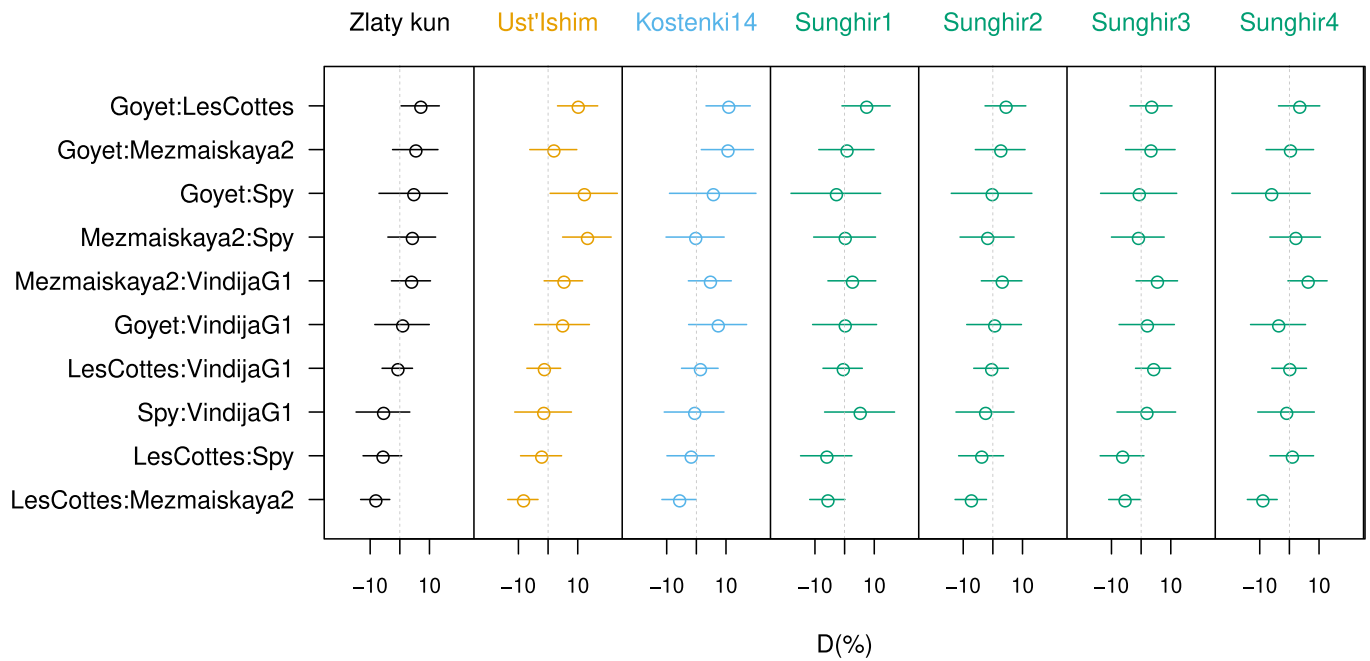
Extended Data Fig. 10 | Pairwise comparison of low-coverage Neanderthals in their sharing of alleles with European and Asian hunter-gatherers. Shown are the D-statistics of the form D(Mbuti, Hunter-Gatherer, Neanderthal1, Neanderthal2). The Hunter-Gatherer corresponds to the columns; rows show the two compared Neanderthals in order Neanderthal1: Neanderthal2. Error bars correspond to two standard errors.

Corresponding author(s): Kay Prüfer, Cosimo Posth, Johannes Krause

Last updated by author(s): 16.02.2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	bcl2fastq v2.20 dnaclust v3.0.0 fastqc v0.11.4 EAGER v1.92.59 AdapterRemoval v2.3.1 dedup v0.12.2 mapDamage v2.0.9 BWA v0.7.12 samtools v1.3 pileupCaller v8.6.5

Data analysis	schmutzi AuthentiCT Haplofind muscle MEGA6 BEAST v1.8.2 ModelGenarator v.85 contamLD admixtool 5.1 (qp3Pop, qpDstat, qpGraph, qpF4Ratio) Neanderthal-dating admixfrog 0.6.2-dev3 R version 3.3.0 and version 3.4.4

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All sequence data generated and analysed during the current study are available in the European Nucleotide Archive under the study accession number PRJEB39040.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to determine ancient DNA sample size a priori.
Data exclusions	No data was excluded.
Replication	Replication is achieved by comparing the results obtained with the shotgun and the nuclear capture datasets.
Randomization	No randomization was performed since the newly generated data derive from a single ancient individual.
Blinding	No blinding was performed. The data was not grouped for analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☐	☒ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Palaeontology and Archaeology

Specimen provenance
: The specimen was provided by the National Museum in Prague, Czechia.

Specimen deposition
: The specimen is deposited at the Department of Anthropology of the National Museum in Prague, Czechia.

Dating methods
: Three new radiocarbon dates were obtained from the right zygomatic bone: 1) Collagen ultrafiltration dating at the Klaus Tschira Archaeometry Center in Mannheim, Germany; 2) Amino acid hydroxyproline dating at the Oxford Radiocarbon Accelerator Unit in Oxford, UK; 3) Sample pretreatment (solvent wash, decalcification in acid, base wash, re-acidification and gelatinization) and ultrafiltration dating at the Oxford Radiocarbon Accelerator Unit in Oxford, UK.

☒ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight
: The Department of Anthropology of the National Museum in Prague, Czechia approved and provided guidance on the study protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.