Genome sequences of 36,000- to 37,000-year-old modern humans at Buran-Kaya III in Crimea

E. Andrew Bennett, Oğuzhan Parasayan, Sandrine Prat, Stéphane Péan, Laurent Crépin, Alexandr Yanevich, Thierry Grange, Eva-Maria Geigl

▶ **To cite this version:**

E. Andrew Bennett, Oğuzhan Parasayan, Sandrine Prat, Stéphane Péan, Laurent Crépin, et al.. Genome sequences of 36,000- to 37,000-year-old modern humans at Buran-Kaya III in Crimea. Nature Ecology & Evolution, 2023, 10.1038/s41559-023-02211-9 . hal-04280395

HAL Id: hal-04280395
https://hal.science/hal-04280395v1

Submitted on 11 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Genome sequences of 36-37,000 year-old modern humans at Buran-Kaya III in Crimea

E. Andrew Bennett[1,5,7], Oğuzhan Parasayan[1,6,7], Sandrine Prat[2], Stéphane Péan[3], Laurent Crépin[3], Alexandr Yanevich[4], Thierry Grange[1*], Eva-Maria Geigl[1*]

[1]Institut Jacques Monod, CNRS, Université Paris Cité, 75013 Paris, France.

[2]UMR 7194 (HNHP), MNHN/CNRS/UPVD, Alliance Sorbonne Université, Musée de l'Homme, Palais de Chaillot, 17 Place du Trocadéro, 75116 Paris, France.

[3]UMR 7194 (HNHP), MNHN/CNRS/UPVD, Muséum national d'Histoire naturelle, Alliance Sorbonne Université, Institut de Paléontologie Humaine, 1 rue René Panhard, 75013 Paris, France.

[4]Institute of Archaeology, National Academy of Sciences of Ukraine, 2 Volodymyra Ivasyuk Avenue, 04210 Kyiv, Ukraine.

[5]Present address: Key Laboratory of Vertebrate Evolution and Human Origins, Institute of Vertebrate Paleontology and Paleoanthropology, Center for Excellence in Life and Paleoenvironment, Chinese Academy of Sciences, Beijing 100044, China

[6]Present address: Institut Pasteur, Université Paris Cité, CNRS UMR2000, Human Evolutionary Genetics Unit, 75015 Paris, France

[7]These authors contributed equally

*Correspondence to: eva-maria.geigl@ijm.fr or thierry.grange@ijm.fr

Abstract

Populations genetically related to present-day Europeans first appeared in Europe at some point after 38-40,000 years ago, following a cold period of severe climatic disruption. These new migrants would eventually replace the pre-existing modern human ancestries in Europe, but initial interactions between these groups are unclear due to the lack of genomic evidence from the earliest periods of the migration. Here we describe the genomes of two 36-37,000-year-old individuals from Buran-Kaya III in Crimea as belonging to this newer migration. Both genomes share the highest similarity to Gravettian-associated individuals found several thousand years later in southwestern Europe. These genomes also revealed that the population turnover in Europe after 40,000 years ago was accompanied by admixture with pre-existing modern human populations. European ancestry prior to 40,000 years ago persisted not only at Buran-Kaya III, but is also found in later Gravettian-associated populations of western Europe and Mesolithic Caucasus populations.

MAIN TEXT

Introduction

Archeological and genetic data show anatomically modern humans (AMHs) to be present in Europe by at least 45,000 BP (years Before Present)[1], and possibly earlier[2]. Genomic analyses of AMH remains prior to 40,000 years ago have revealed several diverse and poorly characterized populations[1,3–5]. All of these showed evidence of admixture with Neanderthals either from the initial admixture present in all non-Africans or more recent, local events as in the genomes from Peștera cu Oase in Romania (Oase1)[6] and Bacho Kiro in Bulgaria[1]. Only the individuals from Bacho Kiro, who were found associated with Initial Upper Paleolithic archeological assemblages (IUP), have demonstrated some genomic relationship with modern human populations, specifically those found in East Asia[1]. At some point after 40,000 years ago, these early ancestries appear to vanish from Europe as a new genomic profile appears, a timing concurrent with changing climatic and environmental conditions after the Campanian Ignimbrite super-eruption in Italy, which blanketed southern and eastern Europe in ash 39,847-39,136 BP[7], at the beginning of the nearly 2,000-year-long Heinrich Stadial 4 cold period with catastrophic impacts on the local ecosystems[8–12]. This post-Campanian Ignimbrite (post-CI) wave of ancestry first appears after the Heinrich Stadial 4 and is documented in Europe in various forms during the Early and Mid Upper Paleolithic (EUP and MUP, respectively), such as in genomes from the sites of Kostenki14 (~38,000 cal BP)[13] and Sunghir (~34,000 cal BP)[14] in the eastern European Plain and in two ~35,000-year-old individuals, GoyetQ116-1 from Belgium[3] and BK1653[1] from the MUP deposits of Bacho Kiro, Bulgaria, both of whom also contain trace ancestry linked to the previous IUP inhabitants of Bacho Kiro[1]. Post-CI ancestry is also represented in central and western Europe beginning at least 32,000 years ago in individuals belonging to the "Věstonice" and "Fournol" genetic clusters, respectively, both of which are in direct association with Gravettian cultural assemblages[3,15]. Unlike the previous

occupants of Europe, this post-CI wave of ancestry shares detectable genetic drift with modern Europeans[3,13], but due to the sparseness of genomic data from EUP humans in Europe, the origins of this ancestry, the timing of its appearance in different regions, and the routes that populations bearing this ancestry took entering the European continent are unknown. Apart from a trace IUP Bacho Kiro ancestry component found in some members of the Fournol cluster and BK1653, interactions between post-CI incoming populations with the descendants of preceding groups are poorly defined.

Discovered in 1990, the rock shelter site of Buran-Kaya III in the Crimean Peninsula contains rich deposits of human activity dating from the Middle Paleolithic until the Middle Ages[16–19]. The assemblages of layers 6-2 to 5-2, dated to ca. 38-34,000 cal BP[17,19], have yielded backed microliths, microgravette points, bone tools, ochre, and body ornaments of ivory, shell, and teeth, as well as multiple human skull fragments, some with post-mortem cut marks[17]. The similarity of the cultural deposits to later Gravettian assemblages of Europe have led the excavating archeologists to assign this material to the earliest appearance of a Gravettian cultural layer, a designation that has recently been challenged on the basis of the considerable gap of time and geography between these artefacts and those appearing 5,000 years later in Gravettian sites of central Europe[20–23]. Other arguments consider that the assemblages of the EUP layers of Buran-Kaya III can be more parsimoniously associated with a cluster of non-Aurignacian contemporaneous EUP layers found in the Caucasus that also contain backed lithics[20,23,24], such as those at Ortvale Klde, layers 4d-4c[25], Mezmaiskaya, layers 1c-1a[26], and Dzudzuana, layer D[27]. Still others have proposed a dual approach and termed EUP layers of Buran-Kaya III along with those in the northern Caucasus sites as “proto-Gravettian”[28,29], and the similarities of these backed-lithic Caucasus sites with the Early Ahmarian lithics produced in the Levant has been used by some to suggest connections between these industries[27,30,31].

While a strict correlation between the continuity of cultural and technological practices and the genetic continuity of the creators of archeological material cannot always be expected, genomic characterization, where possible, of human remains associated with archeological assemblages provides additional keys to understanding the cultural and historical contexts of the material and the behaviors, adaptations and cultural contacts of ancient human groups. To define the population associated with the earliest appearance of this cultural assemblage in Europe, and to determine how the EUP occupants of Crimea fit into our current knowledge of the AMH settlement dynamics of Europe at the close of the Heinrich 4 period 38,000 years ago, we investigated the genetic identity of the human remains of the EUP layers 6-2 and 6- of Buran-Kaya III.

Results

Two cranial fragments (BuKa3A and BuKa3C), one with cut marks (BuKa3A), were aseptically excavated in 2009 from layers 6-1 (BuKa3A) and 6-2 (BuKa3C) of Buran-Kaya III (see Supplementary Note 1 and 2). Radiocarbon (AMS) dates from associated material ranged from 36,840 to 35,685 cal BP (31,900+240-220 ^{14}C BP) for BuKa3A and 37,415 - 36,245 cal BP (32,450+250-230 ^{14}C BP) for BuKa3C (Extended Data Fig. 1 and Table S1). Initial genetic characterization of the samples revealed extremely poor preservation of endogenous DNA, both in fragment length (average 36 bp) and in low quantity relative to environmental DNA (<0.1%). Applying various pretreatments and extraction buffers[32,33] (see Supplementary Material for details) to the samples, we identified conditions that considerably increased the endogenous DNA content, enough to allow genomic analysis (~0.5% for BuKa3A and ~4% for BuKa3C). Enrichment for mitochondrial DNA and shotgun sequencing resulted in 82- (BuKa3A) and 11-fold (BuKa3C) coverage of the mitogenomes. Reads mapping to the mitochondrion were used to determine the posterior probability for contamination with modern human sequences using the program Schmutzi[34], which estimated a contamination rate below

1% for BuKa3A, and haplocheck[35] detected no mitochondrial contaminants above 1% in either sample (see Methods). Both individuals were determined to be male using the ratios of reads mapping to chromosome X and either Y[36] or the average of autosomal-mapped reads[37]. The older sample, ~36,800 year-old BuKa3C, belonged to a mitochondrial haplogroup basal to haplogroup U, a lineage branching prior to the split that gave rise to U-haplogroups found throughout E/MUP Europe[3,13–15,38] (Extended Data Fig. 2). The mitochondrial haplogroup of BuKa3A was determined to belong to an early branch of the N lineage, N1. This assignment falls outside of the lineages previously reported for N in UP Europe, nearly all of which derive from later N branches (U and R haplogroups, Extended Data Fig. 2). The N1 of BuKa3A is notably distinct from the pre-CI N haplogroups identified from the 31-52,000-year-old woman from Zlatý Kůň in Czechia, the 40,000-year-old Oase1 from Romania and those of the 45,000-year-old occupants of Bacho Kiro Cave in Bulgaria, all of which belong to more basal branches that have no modern descendants[4,6,39]. In contrast to these, the BuKa3A mitochondrial DNA sequence contains three of the eight mutations beyond N1 to N1b, a modern rare haplogroup most highly concentrated in the Near East, yet appearing broadly from western Eurasia to Africa. Among ancient samples, the mitochondrial sequence of an 11,000-year-old Epipalaeolithic Natufian from the Levant (Natufian9)[40] is also a later derivative of this pre-N1b branch. The Y haplogroup for BuKa3C was determined to belong to a basal F branch that has not been found among modern humans younger than 38,000 years-old, but clustering instead with Y haplogroups previously reported in southeastern Europe for the IUP BK (Bacho Kiro)_F6-620[1] and 40,000 year-old Oase1[6] (Fig. S1). All three of these Y chromosomes carry ancestral bases at diagnostic sites that exclude their assignment to any modern F lineages, which are principally found in East Asia. A single transversion may also point to a similar basal F haplogroup for BuKa3A but the lower coverage of this sample does not allow a reliable placement beyond haplogroup CT.

Neanderthal admixture

Neanderthal settlements attributed to the Micoquian, Kiik-Koba type were also found at Buran-Kaya III (layer B) and have been dated through faunal bone fragments from 43,200 to 40,200 cal BP (37,700 ± 900 to 35,649 ± 282 ^{14}C BP)[17], demonstrating an earlier Neanderthal presence at the site just prior to the CI eruption. AMH remains dating to this period found in Romania and Bulgaria have documented local admixture with late Neanderthals in southeastern Europe[1,6], leading us to investigate whether admixture with local Neanderthals in Crimea could be detected in early AMHs living at most several thousand years after the last Neanderthals in the region[17,41]. Neanderthal ancestry calculated using the direct f4-ratio method[42] determined that BuKa3A and BuKa3C possess 3.8 (SE 3.3-5.3) and 2.2 (SE 1.2-3.2) percent Neanderthal ancestry, respectively (Extended Data Fig. 3). Although the Neanderthal ancestry was slightly higher in BuKa3A than BuKa3C, it is in line with values calculated for other post-CI individuals from Europe using the same method, indicating recent Neanderthal admixture did not occur with the AMHs ancestral to the population to which the Buran-Kaya III individuals belonged.

Eurasian relationships

To maximize the comparable data between BuKa3 and other Paleolithic genomes, single nucleotide polymorphisms (SNPs) were called from 929 high-coverage modern genomes[43] and filtered to remove singletons and variants occurring in repetitive regions. This left ~13 million genome-wide SNPs, of which 265,510 and 537,076 overlapped with BuKa3A and BuKa3C respectively. The combined ~740,000 SNPs that occurred in at least one BuKa3 individual were then called from 46 previously published shotgun or high-coverage captured western Eurasian genomes from the Paleolithic to the Neolithic (Table S2). Outgroup *f3*-statistics were carried out on sample combinations including a minimum of 10,000 shared SNPs. Overall,

both BuKa3 genomes were more similar to genomes younger than 38,000 years than to those older (Fig.1A, Table S3), placing them within the diversity entering Europe at the beginning of this time period. They were also more similar to present-day populations inside Europe than to those outside (Extended Data Fig. 4). In defining the relationship between the two Buran Kaya III individuals, low overlap between samples gave unsatisfactory standard errors with outgroup-*f3* analysis, but a clade-like relationship between both BuKa3 individuals was demonstrated with *f4*(BuKa3A, BuKa3C; *test*, Mbuti) resulting in a Z-score never significantly different from zero for all UP samples tested (Table S4). Among the Paleolithic genomes, both BuKa3 individuals were found to share the highest amount of drift with the several thousands of years younger Gravettian-associated Fournol and Věstonice clusters in central and western Europe. The two BuKa3 individuals, separated by 1,000 years, did not possess a completely uniform ancestry, and shared unequal amounts of drift with several UP individuals. In outgroup-*f3* analysis, BuKa3A had higher affinity than BuKa3C to the Fournol cluster, and shared more alleles with several pre-CI individuals, in particular Oase1, three IUP Bacho Kiro individuals, a 31-52,000-year-old woman from Zlatý Kůň in Czechia, and Tianyuan from East Asia[44] (Fig. 1A and B, Table S3). Downsampling BuKa3C to an equal SNP count with BuKa3A did not change these results (Table S3). The outgroup *f3* results, and in particular the higher than expected affinity of the East Asian Tianyuan to BuKa3A, suggested that, similar to Goyet116-1 and several members of the Fournol cluster, BuKa3A may also be admixed with ancestry present in Europe prior to the CI eruption and Heinrich Stadial 4. The outgroup-*f3* results were explored in greater detail through Multi-Dimensional Scaling (MDS) and plotted on a 3-D scatter plot (Fig. 1C, Extended Data Fig. 5, Fig. S2 and Supplementary Video1), which can be used to visualize the global genomic relationships of the ancient populations of Eurasia. In this analysis, the two BuKa3 samples sit within the plane of the E/MUP cluster of Europe composed of genomes dating between 38,000 and 27,000 years ago. This cluster is

further differentiated along the D3 axis with the Fournol ancestry at one extreme, and the Věstonice and Kostenki14/Sunghir ancestry at the other. BuKa3 is found between these two groups, and is also positioned the furthest of these group members along the D1 axis towards pre-CI individuals and the Mesolithic Caucasus Hunter Gatherer (CHG) genetic cluster, as well as Neolithic individuals from present-day northwestern Iran (Fig. 1C, Extended Data Fig. 5, Fig. S2 and Supplementary Video1).

Genetic relationships with post-CI populations

Among the genomes between 38,000 and 27,000 BP, which represent the wave of new ancestry arriving into post-CI Europe, BuKa3 was found to share the most alleles with 7,000 year younger Gravettian-associated individuals of the Fournol cluster, Serinyà and Fournol85 from northeastern Spain and southern France, respectively[15], as investigated using *f4*(*E/MUP, E/MUP*; BuKa3, Mbuti) (Table S5). A third Fournol cluster member, the 31,000-year-old Ormesson from northern France[15], did not show the same affinity. Additionally, the 29,000-year-old Paglicci12 from southern Italy[15] had a high affinity for one or both BuKa3 individuals when compared with several individuals from eastern Europe, especially Sunghir3, which appeared to have the least amount of shared drift with contemporary sequenced European genomes. Overall, the Fournol relationship was stronger with BuKa3A than BuKa3C, who could be discriminated from BuKa3A by sharing significantly more alleles with Kostenki14 than either Sunghir3, BK1653, or Muierii1 (a 34,000-year-old woman from Romania)[38]. Unlike BuKa3C, BuKa3A shared more alleles, also when compared to Sunghir3, with GoyetQ116-1 and 31,000-year-old infant twins from Krems-Wachtberg in Austria (being combined from Krems1_1 and Krems1_2), who were buried with a central European Gravettian assemblage[45], denoted here as Krems1_1-2. All of these relationships are supported by high outgroup-*f3* results (Fig. 1, Tables S3). Tests of *f4*(*E/MUP*, BuKa3; *E/MUP*, Mbuti) showed reciprocity between Serinyà and Fournol85 and both BuKa3 individuals when

compared to most E/MUP genomes from eastern Europe, including several from the Věstonice cluster, although BuKa3C shared significantly more alleles with Kostenki14 than any of the Fournol cluster (Fig. 2A, Table S5).

The Fournol cluster differentiates Gravettian-associated assemblages from western and southwestern Europe from those of the Věstonice cluster found in Central Europe. Its members date from 31,000 to 26,000 years ago, contemporary with those of the Věstonice cluster. Fournol ancestry has also been modeled as a sister group to the Aurignacian-associated GoyetQ116-1 ancestry appearing in Belgium, and both share traces of admixture with a common pre-CI ancestry related to that found in IUP Bacho Kiro[15]. We further investigated how the relationship of BuKa3 to the Fournol and Věstonice clusters compares with those of a range of ancestries from the Mesolithic to the Upper Paleolithic using *f4*-statistics (Fig. 2B, Table S6). We find that while BuKa3 falls outside of the Fournol group compared with other members, both BuKa3 individuals share more alleles with Fournol than with Věstonice. We also note that in addition to Bacho Kiro ancestry already reported in the Fournol cluster, gene flow related to Zlatý Kůň is also present, as well as an unexpected affinity to Mesolithic hunter gatherers from the Caucasus (Fig. 2B, Table S6). These results describe a broad genetic relationship among pre-CI populations of Europe with post-CI EUP populations in Crimea and later MUP populations of western and southwestern Europe. This relationship is absent or reduced in MUP populations in central and eastern Europe, but is present in 11,000 to 14,000-year-old individuals from the Caucasus.

Pre-CI gene flow into Buran-Kaya III

To date, few instances of the ancestry present in Europe prior to 40,000 years ago surviving after the CI and Heinrich Stadial 4 have been reported[1,3,15], and little genomic data exist from the earliest period of these admixtures that could assist in better characterizing these events. To investigate the scope of pre-CI ancestry observed in BuKa3, we compared them with other pre-

LGM individuals through *f4*-statistics. Overall, BuKa3 was found to be more genetically similar to contemporary and younger post-CI (EUP/MUP) groups than to older pre-CI ones, sharing significantly more alleles with individuals younger than 38,000 years when compared with any individual older than 40,000 years (*f4*(*EUP/MUP*, *pre-CI*; BuKa3, Mbuti), Z > 3) (Table S5). Several exceptions to this result, however, were found with BuKa3A when pre-CI is Oase1, indicating gene flow between Oase1 and BuKa3A. This signal was statistically more significant when using EUP/MUP genomes from eastern Europe. To verify whether this affinity could not be due to possible elevated Neanderthal content sharing between BuKa3A and Oase1, we recalculated the statistic after masking the increased Neanderthal content found in the Oase1 genome[5]. Excluding excess Neanderthal ancestry increased the shared drift between BuKa3A and Oase1, indicating that this affinity could be traced to shared non-Neanderthal ancestry (Table S7). Additional deviations from zero demonstrating BuKa3A affinity for post-CI ancestry concerned Tianyuan and Zlatý Kůň with respect to several MUP individuals (Table S5). Because BuKa3 clustered with post-CI groups, the high degree of shared drift among post-CI group members would mask low-level admixture signals between BuKa3 and pre-CI individuals, as would similar amounts of shared pre-CI admixture among post-CI groups. We thus explored the BuKa3 relationships with any two pre-CI individuals with the statistic *f4*(*pre-CI*, *pre-CI*; BuKa3, Mbuti). In this formulation, we would expect a Z-score result of zero if two pre-CI individuals formed an outgroup with respect to BuKa3 as would be expected if there were no specific genetic relationships between the lineages leading to pre- and post-CI settlement of Europe. In these tests, BuKa3A shared significantly more alleles with Oase1 than with BK_F6-620 from IUP Bacho Kiro (Z = |3.5|) and an elevated, although less significant amount with respect to BK_CC7-335 and the 45,000-year-old Ust'Ishim from Siberia Z = |2.2|, |2.9|, respectively) (Extended Data Fig. 6, Table S5). In *f4*-statistical tests *f4*(*UP*, BuKa3; *pre-CI*, Mbuti), which compare relative genomic affinity

between BuKa3 and any high-coverage genome older than 27,000 years with a pre-CI individual, Zlatý Kůň was found to share more alleles with BuKa3A than with several pre- and post-CI individuals (Table S5). Z-scores for *f4*(*UP*, BuKa3A; Zlatý Kůň, Mbuti) were negative for nearly all samples tested with Z-scores below -2 for several post- and pre-CI individuals. BuKa3C showed a similar trend, although less pronounced and in fewer contexts. Interestingly, the results of this test were positive when UP is Fournol85, suggesting a possible elevated Zlatý Kůň-related ancestry component, greater than that of BuKa3, may be present in the Fournol cluster (Table S5). These results agreed with the position of BuKa3 between contemporary post-CI individuals and a pre-CI group comprising Oase1 and Zlatý Kůň in MDS analysis of inverted *f3*-statistics (Extended Data Fig. 5, Fig. S2 and Supplementary Video1). The evidence of shared ancestry between BuKa3A and Oase1 was unexpected since Oase1 ancestry had not previously been detected in any subsequent Eurasian populations[1,3,6]. A recent analysis of UP genomes, however, could successfully model Oase1 as belonging to a population similar to the IUP Bacho Kiro group with additional Neanderthal admixture[5], and groups related to Bacho Kiro have contributed ancestry to modern East Asians and GoyetQ116-1[1,3]. Zlatý Kůň had also been previously described as not to have contributed gene flow to later groups[4]. More recent analyses, however, presented admixture models that can support gene flow of between 2-29% from Zlatý Kůň into Bacho Kiro[5], which would support our observations of a Zlatý Kůň-related ancestry persisting in a broad distribution of populations. These observations were directly tested through *f4*-statistics in the form *f4*(Sunghir3/Kostenki14, *test*; Zlatý Kůň, Mbuti). This formulation takes advantage of the observed genetic position of Sunghir3, which appeared to represent an extreme end of post-CI, pre-LGM European ancestry in the preceding tests, perhaps due to having less admixture with other post-CI individuals in Europe. We also included Kostenki14, who shows signals of admixture with certain post-CI individuals, to compare with the more distant Sunghir3. This statistic attempts to measure the relative amounts

of gene flow from Zlatý Kůň found in post-CI Europe, including also the Mesolithic CHG cluster whose particular genetic affinity with Zlatý Kůň is demonstrated in Fig. 2B. These results show levels of Zlatý Kůň-related ancestry present in multiple post-CI individuals relative to Sunghir3 (Fig. 3A, Table S8). This signal is highest among select members of the Gravettian-associated Fournol and Věstonice clusters having a geographic distribution encompassing southern parts of western Europe to the Danube River in Austria (Fig. 3A). The Zlatý Kůň-related ancestry identified in Věstonice cluster members Krems1_1-2 and Paglicci12 appears to be reduced in other Věstonice cluster members north of the Danube River at Dolni Věstonice in Czechia Fig. 3A, Table S8, revealing a variable presence of this component within the Věstonice cluster, possibly corresponding to a North-South geographical cline. The masking of these signals when testing with Kostenki14, particularly with individuals from northern and eastern Europe, is assumed to be due to low level admixture of Zlatý Kůň-related ancestry in Kostenki14, or previously described eastern European admixture in MUP populations of central Europe from a lineage closer to Kostenki14 than to Sunghir3[3,15,38].

In light of the diverse ancestries and potential admixtures indicated by the *f3* and *f4* analysis, we attempted to separately model the ancestries of each Buran Kaya III individual within a qpGraph framework. To increase the strength of the modeling results and reduce the potential number of admixture edges, we focused on a minimal set of representative high coverage shotgun-generated UP genomes. The best-fitting models using the present-day African Mbuti population as an outgroup had Z-scores for the worst f4-statistic residuals of 1.73 for BuKa3A and 2.96 for BuKa3C (Fig. 3B). Best-fitting models for both BuKa individuals shared a similar structure, with Zlatý Kůň representing the deepest branching modern human ancestry after the Out of Africa event (OoA), in agreement with previous analyses[4,5,15]. In our model, the European post-CI group, including the BuKa3 individuals, Kostenki14 and Sunghir3, all

diverge from a common lineage branching prior to that leading to the European pre-CI groups represented here by Zlatý Kůň. Unlike the other post-CI genomes from eastern Europe, the BuKa individuals were best modeled as receiving 6-16% ancestry from a population ancestral to Zlatý Kůň. Models without this admixture component were found to have less support (Z= 3.65 and 4.66 for BuKa3A and BuKa3C, respectively) (Extended Data Fig. 7).

Caucasus connections

A distinct ancestry was identified from 13,300-year-old remains at Satsurblia Cave in western Georgia, which was described, along with similar nearby 9,800-year-old remains at Kotias Klde, as Caucasus Hunter Gatherer (CHG)[46], also known as the Satsurblia genetic cluster[3]. CHG ancestry has been estimated to have separated from Western Hunter Gatherers (WHG) approximately 45,000 years ago and thought to have existed as an isolated population during the LGM, later contributing ancestry to several different populations that played roles in developing the modern European genomic identity[40,46,47]. CHG ancestry was later found to have occupied the Caucasus region from as far back as the MUP in two 26,000-year-old individuals from Dzudzuana Cave[48] and 25,000-year-old remains from Kotias Klde[49], both in Georgia. Given the geographical proximity and cultural links between the contemporary layers of these Caucasus sites and Buran-Kaya III, we additionally investigated whether continuity with the CHG ancestry could extend 10,000 years earlier at Buran Kaya III. Results of the *f4*(*test*, BuKa3; CHG, Mbuti) show that BuKa3A and BuKa3C share significantly more alleles with the CHG cluster than nearly all other post-CI individuals older than 30,000 years, with the compelling exception of Krems1_1-2, the Fournol cluster, and Paglicci12 (Fig. 4A, Table S9). This shared affinity towards ancestry first characterized in much more recent Caucasus hunter-gatherers appears to mirror Zlatý Kůň as an ancestral component that links the BuKa3 population with the Fournol cluster and the western and southernmost Věstonice cluster

members. We used a biplot of *f4*(Sunghir, *test*; Zlatý Kůň, Mbuti) results to examine whether the Zlatý Kůň-related ancestry component we identified among pre-LGM populations was not similar to a component found in CHGs. We found the amounts of these two ancestry components to be in direct correlation in post-CI, pre-LGM populations (Fig. 4B), although the ancestry of Zlatý Kůň and CHGs is not the same, since CHGs share much more gene flow with all post-CI genomes than with Zlatý Kůň, as *f4*(Zlatý Kůň, *test*; CHG, Mbuti) always gives a significantly negative Z-score.

Discussion

With these two individuals from Buran-Kaya III we genetically characterize one of the earliest facets of the distinct EUP wave of ancestry first appearing in Europe between 40,000 and 37,000 BP. In Crimea, this ancestry was found associated with a lithic industry containing backed bladelets, well established in the northern and southern slopes of the Caucasus that has been compared to Gravettian assemblages appearing later in central Europe[16,18,30]. Disparities in the ancestries of the two individuals, occupying the same place some 1,000 years apart, suggest a non-uniform nature of the relationships among these early migrant groups, and propose a model of complex interactions and exchange not only at the interface of this incoming ancestry, but continuing among their descended lineages. The earlier BuKa3C was minimally admixed with pre-CI populations, but carried a Y haplogroup associated with these pre-CI groups, whereas the later BuKa3A contained a greater amount of pre-CI ancestry. Both of them showed genetic affinity to both Zlatý Kůň and Oase1, as well as select Bach Kiro individuals

compared to some pre- and post-CI genomes. Although the BuKa3 individuals were found to be genetically distant from nearly all currently available UP genomes, the population most

genetically similar was found thousands of years later concentrated in southern parts of western Europe. To the east-west genetic differentiation of Gravettian-associated groups identified in Europe during the MUP (Fournol and Věstonice)[15], these data add a north-south element, characterized by increased Zlatý Kůň-related ancestry found in coastal northeastern Spain and southern France, and to some degree in southern Italy and Austria, a region corresponding roughly to the distribution of proto-Aurignacian sites both prior to and during the Heinrich Event 4 as described in Marín-Arroyo *et al.*[50]. The elevated pre-CI admixture in the more recent BuKa3A compared to the 1,000-years older BuKa3C is reminiscent of population dynamics following the Neolithic expansion into Europe, where incidents of admixture increase only some time after the initial appearance of new migrants[37]. The degree of admixture between these incoming AMHs with the pre-CI ancestry of the existing inhabitants has not previously been appreciated. Of the seven high-coverage genomes now available from Europe dating to between 38,000 and 34,000 years ago, all four located to the west of eastern Europe can be modeled as having admixed with pre-CI European AMH groups[1,3]. We do not presume these signals to come specifically from separate admixture events with different groups related to the source sample populations, but may represent an underlying diversity of pre-CI ancestry that may have been present in Europe with various relationships to these groups. We note that the statistical significance of these results are elevated in comparison to eastern European UP genomes. This may be indicative of a widespread low-level presence of a post-CI ancestral substrate in central and western European EUP and MUP populations. The regularity of this phenomenon suggests the possibility that the earliest AMH settlers in Europe, perhaps belonging to populations substantially reduced in size by the preceding climatic changes of the Heinrich 4 cold period, may have been absorbed into the incoming EUP migrations. An alternative hypothesis may be that the worsening climate caused the migration of some of these pre-CI populations to refuge in the Caucasus, where they later admixed with newly arriving

migrants as temperatures rose beginning with the Greenland Interstadial 8 38,000 years ago. The apparent absence of this ancestry in the populations at Sunghir in eastern Europe may indicate an area either devoid or sparsely inhabited by pre-CI European lineages at the time. These two eastern populations are nearly contemporary to that of Buran Kaya III and geographically closer to Crimea than to central and western Europe, which may aid in the identification of the eventual zones of interaction or networks EUP populations may have used as they entered into Europe. The timing of this EUP movement into Europe appears to coincide with the close of the lengthy Heinrich 4 period of extreme cold encompassing the CI super-eruption and suggests that climate-related environmental changes may have facilitated this migration. Eastern Europe in particular would have been less habitable between 40,000 and 38,000 years ago, and existing pre-CI populations within Europe may have contracted, as attested by the disappearance of European Neanderthals[12,41], the scarcity of human settlements, and the reduction of forest cover and ungulate populations during this time[51,52]. Gradually increasing temperatures towards the end of the Heinrich 4 Event and the onset of the warmer Greenland Interstadial 8 (~38,000–35,000 BP)[9,51] would favor northward population expansions from refuges in the south[31,53]. A better genetic characterization of UP populations from Europe and the Caucasus, including the surrounding regions of North Asia, the Balkans, and the Levant, will be required to better resolve the movements and interactions of these groups before and after this transition period, as well as to determine the extent and distribution of pre-CI European ancestry among the growing list of pre-LGM genomes.

An awareness of past population dynamics can deepen the interpretation of the existing archeological record and assist in reconstructing the development and dispersal of technologies and artifacts associated with these populations. In this light, it is intriguing to contrast the genetic similarities between the BuKa3 group and Gravettian-associated populations of

western and central Europe with the similarities between the backed lithic assemblages and cultural artifacts found at Buran-Kaya III and contemporary sites in the Caucasus with Gravettian assemblages appearing later in central Europe. With such a large gap in the archeological record between these sites, in both time and geography, it is difficult to estimate the degree to which these similarities may be due to a continuity of EUP practices at Buran-Kaya III being maintained through tradition in descendant communities over thousands of years. To add to this, given the interactions between MUP groups in Europe documented through genetic exchange, it is easy to imagine multiple paths through which technical innovations might have been communicated across Europe. Similar local and climate-specific functional requirements may also play a role in the correlation of lithic industries[29]. Addressing the links between central European Gravettian assemblages and the EUP backed bladelet industries of Buran-Kaya III and the northern Caucasus, and perhaps also by extension the Early Ahmarian of the Levant as previously suggested[31], may complement our growing understanding of EUP population movements into Europe.

Methods

Dating

All radiocarbon dates were recalibrated using the software OxCal v4.4.4 based on the IntCal20 calibration data set[54].

Sample handling, DNA extraction, library construction, and sequencing

Three human parietal fragments were excavated aseptically or semi-aseptically from layer 6-1 and 6-2 during the 2009 excavation season at Buran-Kaya III. All pre-amplification sample preparation was performed in the dedicated ancient DNA facility using decontamination and clean-room protocols as described in Bennett, *et al.*[55]. All buffers and solutions were prepared using water decontaminated by gamma-irradiation (8 kGy). The surface of the bone was wiped three times with a cotton tip soaked in water, then bleach, then water again, and a small part was removed using a flame-sterilized diamond-disc of a Dremel saw (Dremel, Europe, The Netherlands). Bone fragments were then ground to fine powder in a 6775 Freezer/Mill®Spex SamplePrep in liquid nitrogen as described previously[37]. Powder from each sampling was divided into two equal portions, one of which was subject to a bleach treatment with a 1:20 dilution of hypochlorite followed by a washing step with water, as described[32]. Phosphate buffer washes, described in Korlević et al., (2015)[32], were collected and combined for DNA purification. Both phosphate and non-phosphate buffer-treated samples, including reagent-only mocks, were then incubated in 1.5 mL LoBind microcentrifuge tubes (Eppendorf, Hamburg, Germany) with 1 mL 0.5M EDTA, pH 8.0 (Sigma-Aldrich, St. Louis, MO), with 0.25mg/mL proteinase K (Sigma-Aldrich), 0.14M beta-mercaptoethanol and 0.05% UV-irradiated Tween-20 (Sigma-Aldrich), at 37°C for 24 H. This step was repeated for another 48 H and the two fractions were pooled. Following incubation, all tubes were centrifuged at maximum speed for 10 min, and 1 mL supernatant was mixed with 10 times its volume of "2M70" binding buffer (2M guanidine hydrochloride and 70% isopropanol) in a 15 mL tube and passed through

QIAquick silica columns (Qiagen, Hilden, Germany) using 25 mL tube extenders (Qiagen) and a vacuum manifold (Qiagen) as described[55]. 2M70 binding buffer has been shown to retain the smaller DNA fragments lost during purification with traditional binding buffers[33]. Columns were washed twice with 1 mL PE Buffer (Qiagen) then transferred to a micro-centrifuge and dried by spinning 1 minute at 16,100 × g, turning tubes 180° and repeating. DNA was eluted in a total of 60 µl of 10mM TET (Tris-HCl pH 8.0, EDTA 1mM containing 0.05% Tween-20) performed in two elution steps of 30 µl each, by spinning 16,100 × g for 1 minute after a 5-minutes incubation. The second 1 mL of extract was purified according to the same protocol but using 5M40 "binding buffer" (5M guanidine hydrochloride and 40% isopropanol). 5M40 binding buffer has been shown to remove more inhibitors than the 2M70 binding buffer but to result in a higher loss of smaller DNA molecules[33].

For the first series of screening steps, single-stranded libraries were constructed using either 2 µL (for screening) or 6 µL (for mitochondrial capture) of the eluted DNA, including mocks of all treatments, water only samples, and a positive control oligo following the protocol of Gansauge, *et al.*[56] using the splinter oligonucleotide TL110, and eluting in 50 µL TET. Either 40 µL (for mitochondrial capture) or 4 µL (for screening) of each library was used for bar-coding amplification using dual-barcoded single-stranded library adapters[57] as primers in the following 100 µL volume reaction: 10 µL 10x PCR Buffer + $MgCl_2$ (Roche, Basel, Switzerland) 0.4 µM of each primer, 80 µM dNTPs (Roche), and 15 units of FastStart Taq (Roche). Reactions were heated 95°C for 5 minutes, followed by 35 cycles of 95°C 20 s, 53°C for 45 s, 68°C for 45 s, and then 68°C for 5 minutes. Heteroduplexes that could confound size selection were resolved by diluting the PCR product 1:5 in a 100 µL reaction containing 20 µL of the initial reaction, 8 µl PCR Buffer + $MgCl_2$, 0.4 µM of standard Illumina primers P5 and P7, and 80 µM dNTPs, and amplified a single cycle of 95°C for 1 minute, 60°C for 2 minutes and 68°C for 5 minutes. Products were then purified and size-selected using NucleoMag beads

(Macherey-Nagel, Düren, Germany) for two rounds of purification/size selection according to the supplied protocol at a ratio of bead solution 1.3 times the reaction volume and eluted in 30 µl TET. Purified libraries were quantified using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA), Bioanalyzer2100 (Agilent, Santa Clara, California, USA), Qubit 2.0 Fluorometer (Thermo Fisher Scientific), and qPCR reaction. 46 to 148 ng of DNA were enriched for human mitochondrial sequence in two rounds of capture using 1200 ng of biotinylated RNA baits reverse transcribed from human mitochondrial PCR products[37] following the protocol described in Massilani *et al.*[58] but with four changes: (1) for hybridization and wash steps, 60°C was used instead of 62°C. (2) DNA/RNA-bait solution was incubated 96 hours instead of 48 hours. (3) Elution of the bead-bound enriched DNA was performed with a 5 minute incubation in 30 µL TET at 95°C followed by a magnetic bead separation and the transfer of the eluate to a new tube rather than a 0.1N NaOH elution followed by silica column purification. (4) All post-capture amplifications were performed for 35 cycles followed by a heteroduplex resolution step as described above. Enriched DNA was then quantified as above, and products from all libraries were pooled in equimolar amounts and sequenced on an Illumina MiSeq using a v3 reagent kit for 2x76 cycles, substituting primer CL72 for the Read1 sequencing primer as described[57].

Eight additional libraries along with positive and negative controls were made from 3-8 µL each of the remaining extract BuKa3A_4B, which had the highest relative endogenous DNA content. It should be noted that this extract was derived from the portion of the cranial fragment which included the cranial suture. Libraries were prepared as described above except one library was first treated with USER enzyme (New England Biolabs, Ipswich, Massachusetts, USA) for 30 minutes to remove deaminated cytosine damage. Prior to the barcoding amplification, a 6-cycle amplification of pre-barcoded libraries was performed using 45 µL of each library with 45 µL OneTaq 2X Master Mix with Standard Buffer (New England Biolabs),

and 0.1μM internal primers CL72[57] and CL130[56] with the above PCR conditions. Three pairs of different dual-barcoded adapters were then used to amplify 20 μL of each amplified library, followed by purification and size selection as described above, which allowed the later pooling of two different enrichment protocols and non-enriched DNA from the same library. An average of 1.2 μg of DNA to 1 μg RNA-baits was used for each mitochondrial enrichment, as above, however, an additional alternative "touch-down" hybridization protocol consisting of 60°C for 12 h, 59°C for 12 h, 58°C for 12 h, 57°C for 12 h, and 56°C for 48 h was tested for each library. For selected libraries, an alternative wash protocol described in Fu *et al.*[44] was also tested. Neither of these alternative protocols had a substantial impact on the results obtained. Enriched and shotgun libraries were then pooled separately and size selected on an E-Gel SizeSelect 2% agarose gel (Thermo Fisher). Enriched libraries were then sequenced on an Illumina MiSeq, as above, and shotgun libraries were sequenced on an Illumina NextSeq 500/550 High Output Kit v2 (2x75 cycles). The shotgun libraries were further gel-purified on a 6% polyacrylamide gel and sequenced again on the NextSeq as above.

Additional libraries from fresh extractions of BuKa3A and BuKa3C samples were constructed using 6 μL of extract and the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, Ipswich, MA, USA) after a pretreatment with USER enzyme mix (NEB, Ipswich, MA, USA) as described by the manufacturer. An unindexed Y-shaped adapter was ligated, the ligated libraries were first amplified for 8 cycles with the Illumina Rd1 and Rd2 sequencing primers, and then 20% of the library was further amplified for 12 cycles with Illumina unique dual indexed adapter primers. The libraries were purified independently using 1.3 volume of SPRI beads (Macherey-Nagel™ NucleoMag™ NGS Clean-up and Size Select) for two successive rounds of purification followed after pooling of three successive rounds of purification with SPRI beads as above. Sequencing was performed on an Illumina NovaSeq 6000 sequencer in paired-end mode.

Processing of Sequencing Data

All data processing steps were performed using shell instructions and in-house written scripts. Demultiplexing and fastq conversion of the raw output of Illumina sequencers in the BCL format was performed using bcl2fastq v2.20.0.422 (Illumina). Sequence quality was assessed using FastQC v0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Paired-end reads were merged and trimmed for adapters using leeHom v1.1.5[59]. Reads larger than 28 base pairs (bp) were selected using awk and mapped onto the Human Reference Genome hs37d5 using BWA (Burrows-Wheeler Aligner) v0.7.17[60] with a seed length of 18. A minimum length of 28 bp was chosen following analysis to identify the size limit minimizing spurious mapping reads following the approach described in[61] (Fig. S3). Aligned sequences in BAM format were processed with Samtools v1.9[62] to create index files and flagstat reports, and to select and sort mapped reads. Picard v2.20.0 (http://broadinstitute.github.io/picard/) was used to remove duplicates. Multiple libraries were first merged by Samtools v1.9, followed by filtering for mapping quality (-q 20). Reads smaller than 35 bp containing indels were then removed using an awk script to minimize spurious alignments[63]. Damage patterns were analyzed using mapDamage v2.2.1[64] (Fig. S4) followed by end-clipping using BamUtil v1.0.14 (https://genome.sph.umich.edu/wiki/BamUtil:_trimBam). According to the damage pattern, three to four bases at the ends of each read were clipped for the UDG-treated libraries and 10 bases for non-UDG treated libraries. Library details are given in Table S10.

Sex Identification, Uniparental Haplogroups and Contamination Estimation

The sex of the BuKa individuals was determined by chromosome counting using Qualimap 2[65] and Samtools v.1.9 idxstats using reads mapped with a mapping quality higher than 20. Haplogrep2[66] and Haplocheck v1.3.3[35] were used to classify the mitochondrial haplogroups and to estimate the contamination level. Schmutzi v1.5.5.5[34] was additionally used to estimate

contamination of the higher coverage mitochondrial reads of BuKa3A. Mitochondrial haplogroups were verified manually. A Bayesian tree of 35 ancient, 1 modern and 2 canonical mitochondrial sequences, excluding the hypervariable regions, was constructed using MrBayes v3.2.6[67] using a HKY+i+G nucleotide substitution model, which gave the highest log-likelihood for the tree out of all models tested (GTR and HKY with all combinations of four invariant sites and gamma distributions) using JModelTest2 v2.1.10[68]. The chain was run for 1,100,000 iterations, subsampled every 200 after discarding a 9% burn-in period and visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree). A list of sources for the mitochondrial sequences is given in Table S11. Y chromosome haplogroups analysis was performed with a subset of reads mapping to the Y chromosome with a mapping quality 37 or above, and with the terminal two bases trimmed from each end. Y haplogroups were determined by placing alleles onto reference phylogenies to find the lowest supported downstream node using Pathphynder v1.a[69].

Dataset Preparation and Population Genetics Analyses

In order to prepare a SNP panel, the GRCh38/hg38 genomic coordinates of genotype calls coming from 929 publicly available genomes (ftp://ngs.sanger.ac.uk/production/hgdp)[43] were lifted over the assembly of GRCh37/hg19 by using Picard v2.20.0 (http://broadinstitute.github.io/picard/) based on the UCSC liftOver tool (http://genome.ucsc.edu/cgi-bin/hgLiftOver). The variants coming from repetitive regions of the human genome were filtered out with BEDTools v2.27.1[70] from genotype calls that were lifted over. The heterozygous positions from extracted Western Eurasian and Mbuti populations were filtered by removing the singletons from each population, which resulted in ~13 million biallelic autosomal variants. ~740,000 positions that overlapped with at least one of the BuKa3 genomes were used to curate a 740k SNP panel with a pseudo-haploid SNP calling approach of each ancient genome using pileupCaller v1.4.0.5

(https://github.com/stschiff/sequenceTools). For *f3*-statistical comparisons with present-day genomes shown in Extended Data Fig. 4, the curation of another SNP panel was carried out based on the intersection between the variant panel of HGDP 929 genomes and BuKa3 called by the HaplotypeCaller tool in GenomeAnalysisTK v4.1.968[71] which led to ~1.7 million biallelic autosomal positions after exclusion of heterozygous positions coming only from BuKa3. Concerns of capture bias in our results that could be caused by integrating both captured and whole genome shotgun datasets were addressed by a parallel statistical analysis whereby SNPs empirically demonstrated to be prone to such bias were removed[72]. This reduced the total number of overlapping SNPs by approximately 35,000, and gave similar results, demonstrating a low impact of capture-related bias in our results when including datasets generated with SNP-captured data. We thus present results here for the total overlapping SNPs, but the full results for all filtered and non-filtered tests are given in Table S3 and S5.

For the Neanderthal ancestry analysis, the gvcf files of three high-coverage Neanderthal genomes, Altai, Vindija33.19, and Chagyrskaya were obtained from the website of the MPI of Evolutionary Anthropology in Leipzig (http://cdna.eva.mpg.de/neandertal/). The three genomes were merged using bcftools v1.9[73]. SNPs that were invariant in any of the Neanderthals and those residing in repeated DNA (as defined by the hg19 repeat mask from https://www.repeatmasker.org/) were removed. The overlapping sites between these positions and the 2200k SNP panel (described in Hajdinjak *et al.*[1]), amounting to 1.3 M sites, were used to call chimp sequences downloaded from ENA PRJEB36949 and mapped onto the hs37d5 human genome using BWA mem v0.7.17[60] deduplicated and filtered for a mapping quality higher than 30. SNP calling was performed using bcftools v1.9[73] mpileup (options -B -q30 -Q30) and call (option -m). Five Mbuti genome sequences (SRR1157019, SRR1157021, SRR1157023, SRR1157024, and SRR1157055) were processed in the same way. All vcf files were merged and filtered to remove invariant, missing and multiallelic sites, which left ~800k

autosomal positions that were used to call ancient genomes, including the two BuKa individuals. All files were then merged together for the estimation of the Neanderthal ancestry. The outgroup-*f3* and *f4* statistics and admixture graph were computed using qp3P, qpDstat and qpGraph programs from the ADMIXTOOLS package v5.0[74], respectively. After modeling with qpGraph, the software package ADMIXTOOLS 2 v2.0.0[75] was used to verify that the selected models were optimal. The outgroup-*f3* statistics were calculated in a pairwise manner in the form of *f3(*Mbuti*; ind1, ind2)* for all individuals. A distance matrix coming from inverted pairwise values (1-*f3*) was used to compute Multi-Dimensional Scaling (MDS) analysis (cmdscale function in R v4.3.1). To estimate the Neanderthal ancestry a direct *f4*-ratio was used in the form of $\alpha = f4$(Vindija33.19, Chimp*; Ancient_test*, Mbuti)/*f4*(Vindija33.19, Chimp; Chagyrskaya, Mbuti).

Data and materials availability

Sequence data generated in this study are available through the EBI Sequence Read Archive PRJEB64496. The mitochondrial sequences are available through GenBank (accession number MK934322 and OR327029).

Code availability

No new code and sequence analysis methods were developed. Details on the analysis settings are provided in the Methods section.

Acknowledgments

The paleogenomic facility of the Institut Jacques Monod obtained support from the University Paris Diderot within the program “Actions de recherches structurantes”. The sequencing facility of the Institut Jacques Monod, Paris, is supported by grants from the University Paris Diderot, the Fondation pour la Recherche Médicale (DGE20111123014), and the Région Ile-de-France (11015901). Moreover, we acknowledge support from the French national research center CNRS and EUR G.E.N.E. (ANR-17-EURE-0013 ; IdEx #ANR-18-IDEX-0001

l'Université de Paris ; Programme d'Investissements d'Avenir) for supporting the PhD fellowship extension of OP. We thank the National Academy of Sciences of Ukraine for permission to excavate at Buran-Kaya III and all the team members of the excavations.

We also thank the French National Research Agency (ANR-05-JCJC-0240-01), the Fyssen Foundation, the Muséum national d'Histoire naturelle, the Centre National de la Recherche Scientifique (CNRS) for their financial supports (excavations and anthropological analyses). We thank Olivier Gorgé for assistance with some of the sequencing.

Author contributions

EMG, SaP and StP initiated the project. EMG and TG supervised the study. EAB, EMG and TG produced data. OP and EAB performed formal analyses. OP, EAB, TG and EMG interpreted the data. SaP, StP, LC and AY provided samples and information to the archaeological context. EAB wrote the paper with input from all authors.

Competing Interests Statement

The authors declare no competing interests.

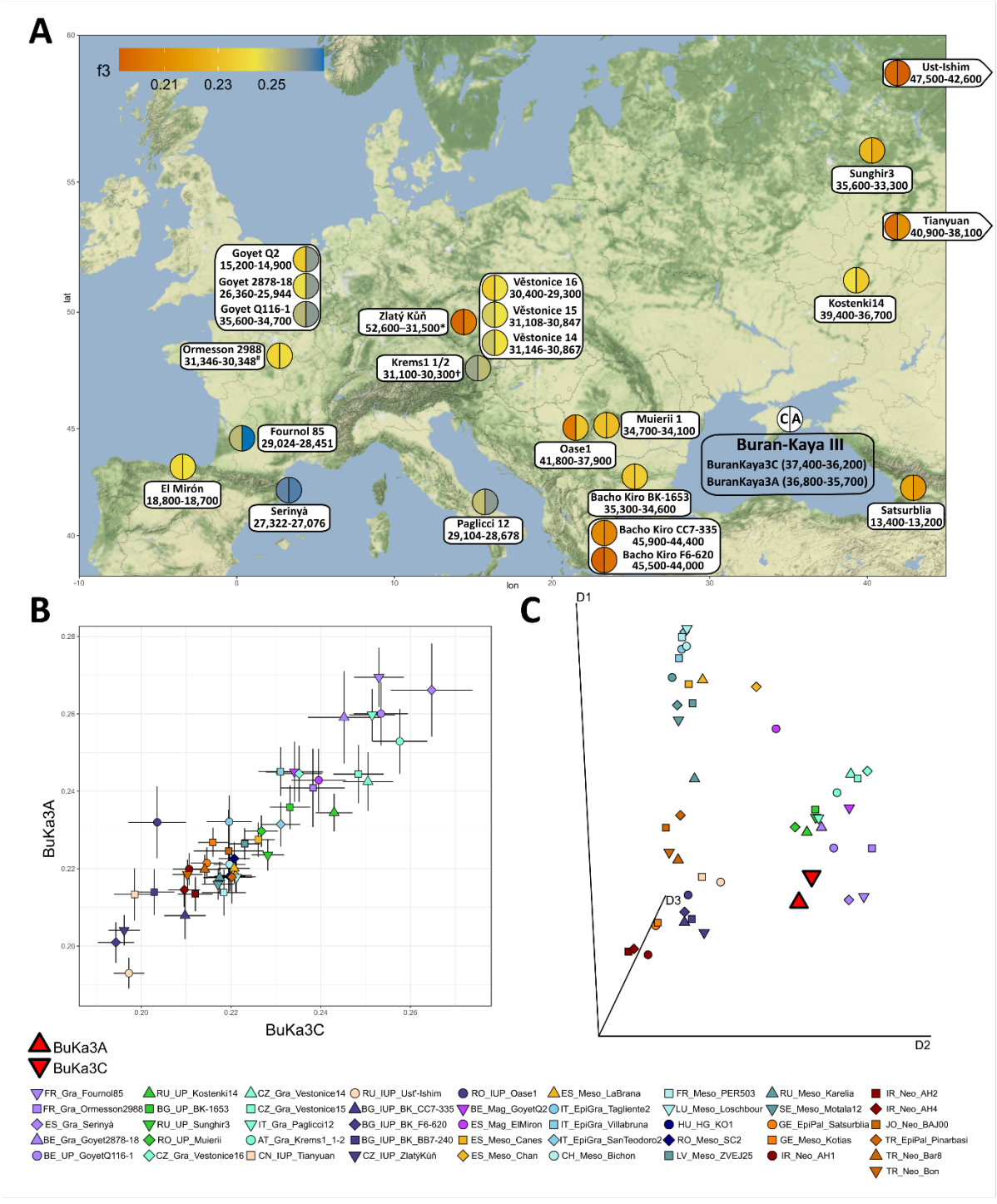
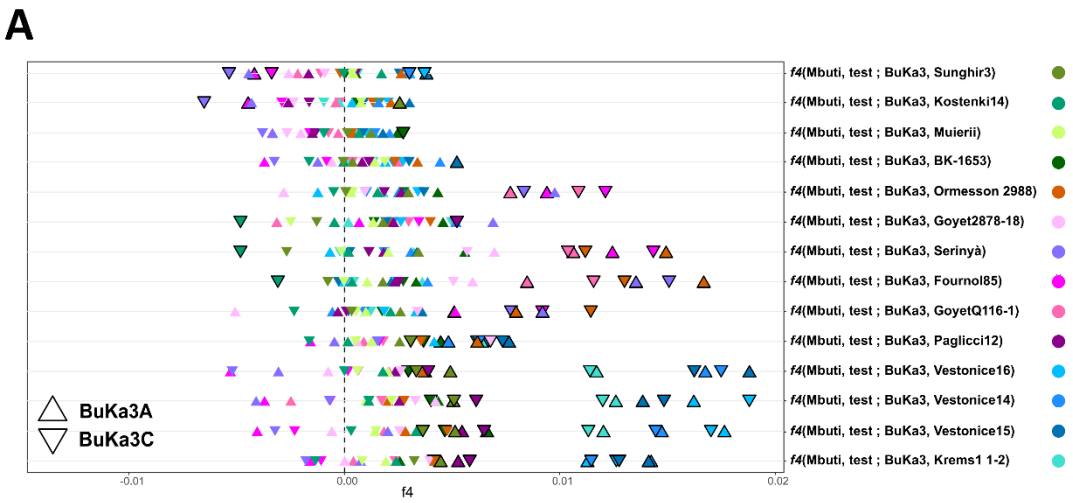


Figure 1. Genomic affinities between Buran-Kaya III and other ancient individuals. A, Heatmap showing the shared genetic drift calculated by f3-statistics in the form of *f3*(Mbuti; BuKa3, *test*). Higher *f3* values indicate greater allele sharing between Buran-Kaya III (BuKa3) and other individuals. The left half of the circles show *f3* values tested with BuKa3C and the right half with BuKa3A. Individuals that do not fit on the map due to geographical distance or

overlapping locations of archaeological sites are placed as close as possible to their correct locations. The map was created with ggmap[76] using Stamen Design map tiles[77]. **B,** Biplot of *f3*(Mbuti; BuKa3A/C, *test*) results. Error Bars = one standard error. **C,** Multidimensional scaling (MDS) of ancient individuals calculated using the pairwise genetic distance matrix coming from inverted *f3* values (1-*f3*) in the form of *f3*(Mbuti; *ancient1*, *ancient2*). The three dimensions displayed represent 8, 6 and 5% of the total variance. The error bars represent ±1 standard errors (s.e.), which were calculated using weighted block jackknife and a block size of 5 Mb.



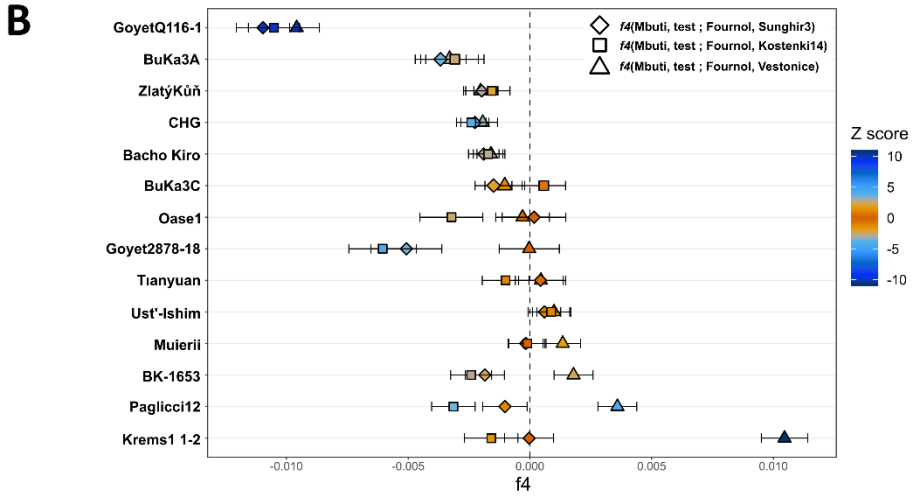


Figure 2. **Relationships among individuals and genetic clusters of MUP Europe. A,** Genomic affinities between BuKa3 and EUP/MUP individuals. Triangle color corresponds to *test* individual, as signified by the color legend to the right. Z-scores ≥ |3| are indicated by a bold outline. Fournol cluster members indicated by pink and violet shades, Věstonice cluster members by shades of blue and dark mauve. **B,** Genomic affinities of pre-LGM individuals and Mesolithic Caucausus Hunter Gatherers (CHG) to the Fournol cluster relative to Kostenki14, Sunghir3, and three members of the Věstonice cluster from Dolní Věstonice (Věstonice 14, Věstonice 15, Věstonice 16). Error bars = one standard error. The error bars represent ±1 standard errors (s.e.), which were calculated using weighted block jackknife and a block size of 5 Mb.

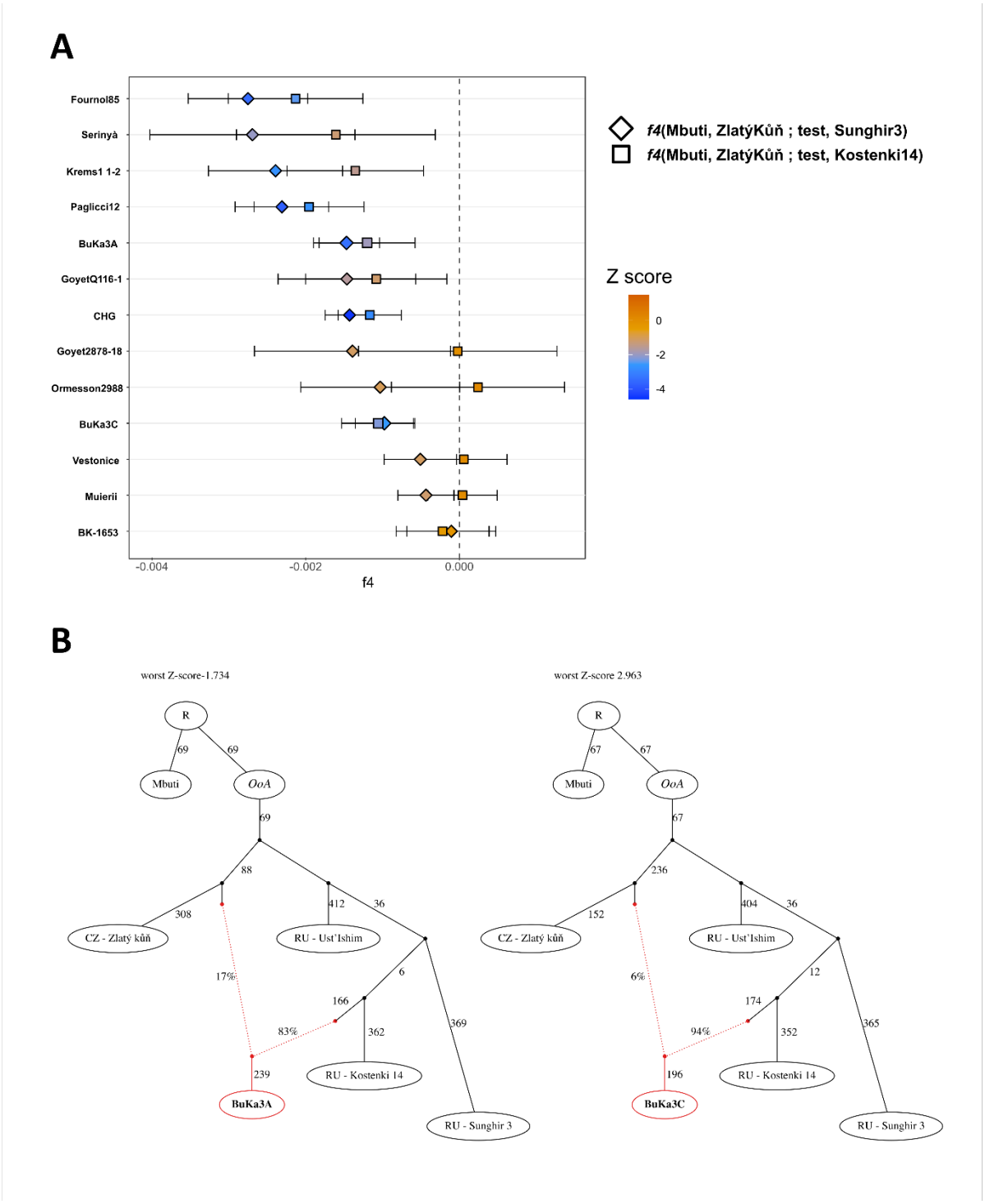


Figure 3. Detection of Zlatý Kůň gene-flow and Buran Kaya III population modeling. A, Shared alleles between Zlatý Kůň and ancestries from Paleolithic/Mesolithic Europe relative to Sunghir3 and Kostenki14. Error bars = one standard error. **B,** Population modeling of Buran Kaya III individuals with estimation of admixture from a a population along the Zlatý Kůň lineage using qpGraph. Only high-coverage, shotgun-generated data are featured using the

parameter allsnps:NO with 32,775 and 70,280 overlapping SNPs in all individuals for BuKa3A and BuKa3C models, respectively. The models have no significant outliers and the Z-score for the worst f4-statistic residuals is indicated. The error bars represent ±1 standard errors (s.e.), which were calculated using weighted block jackknife and a block size of 5 Mb.

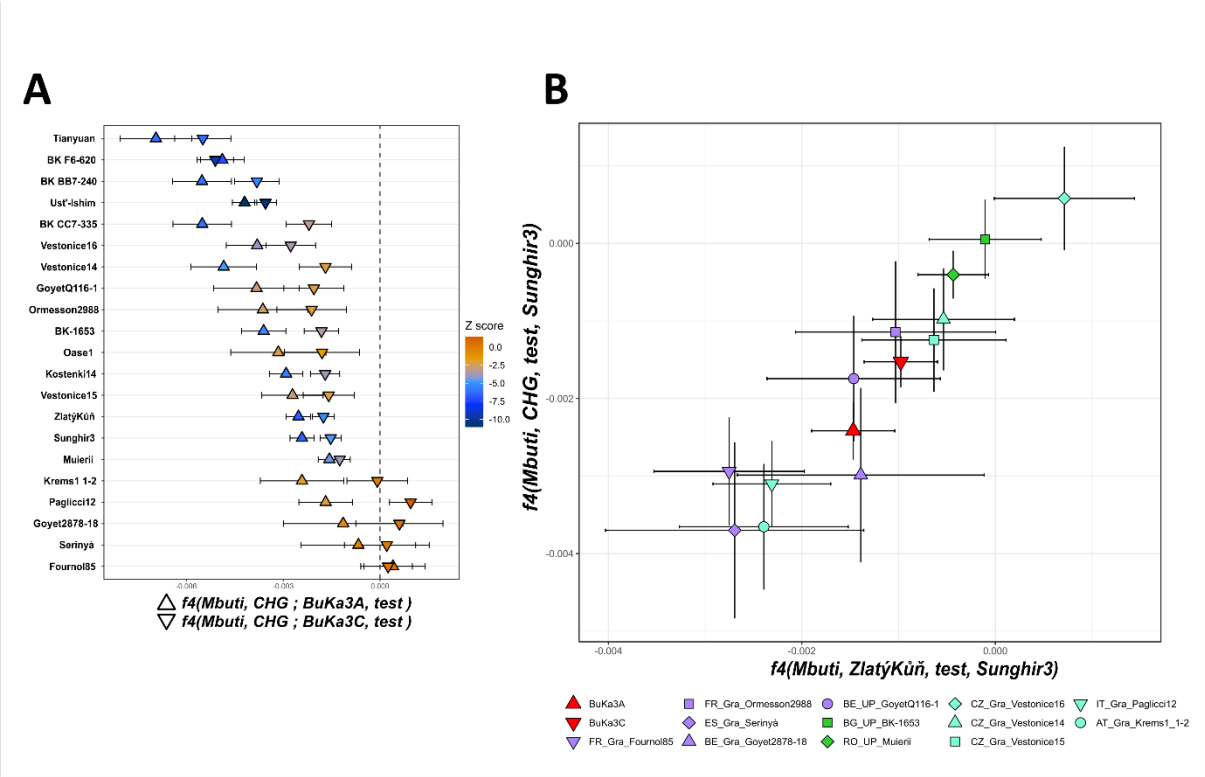
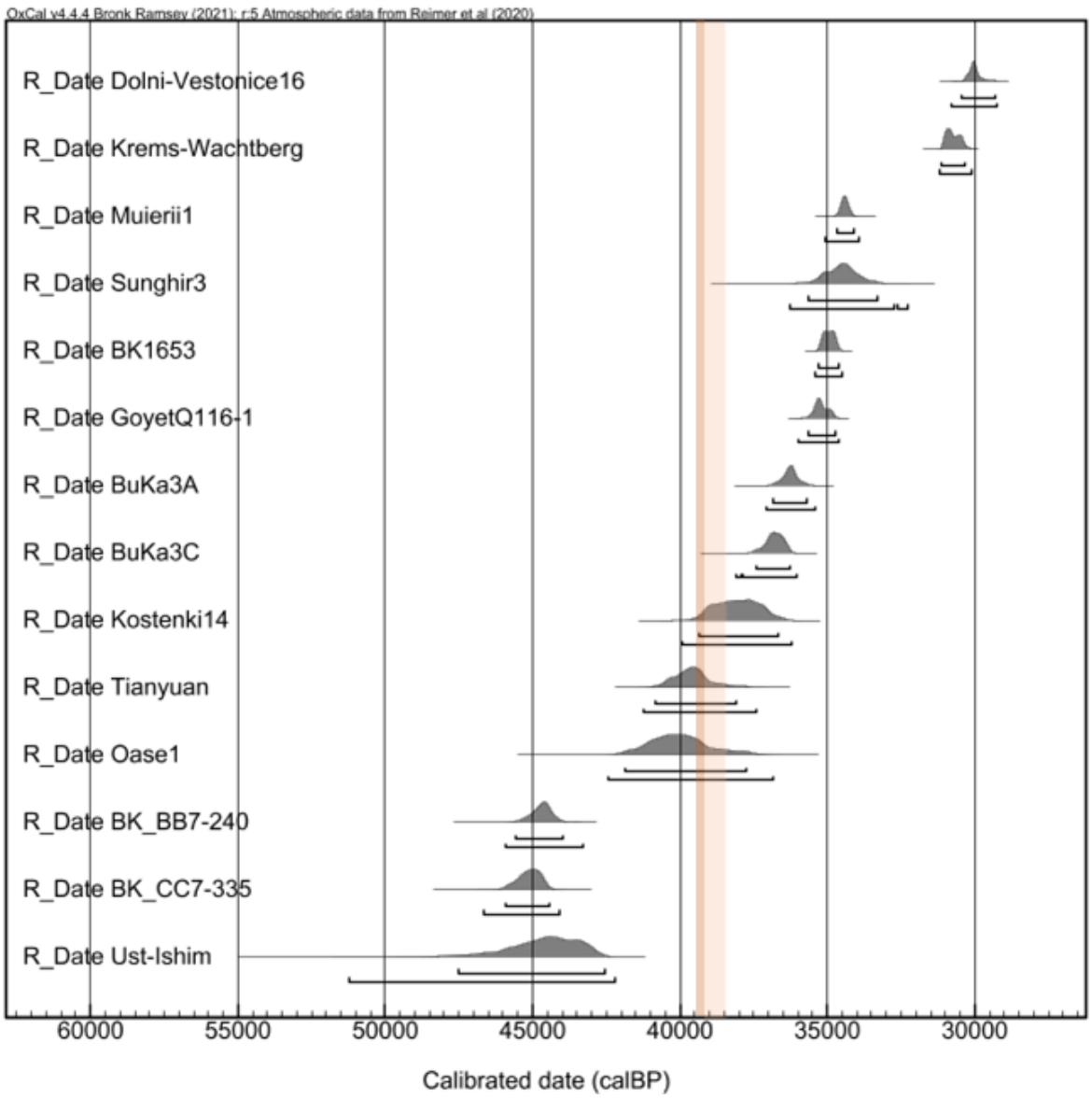
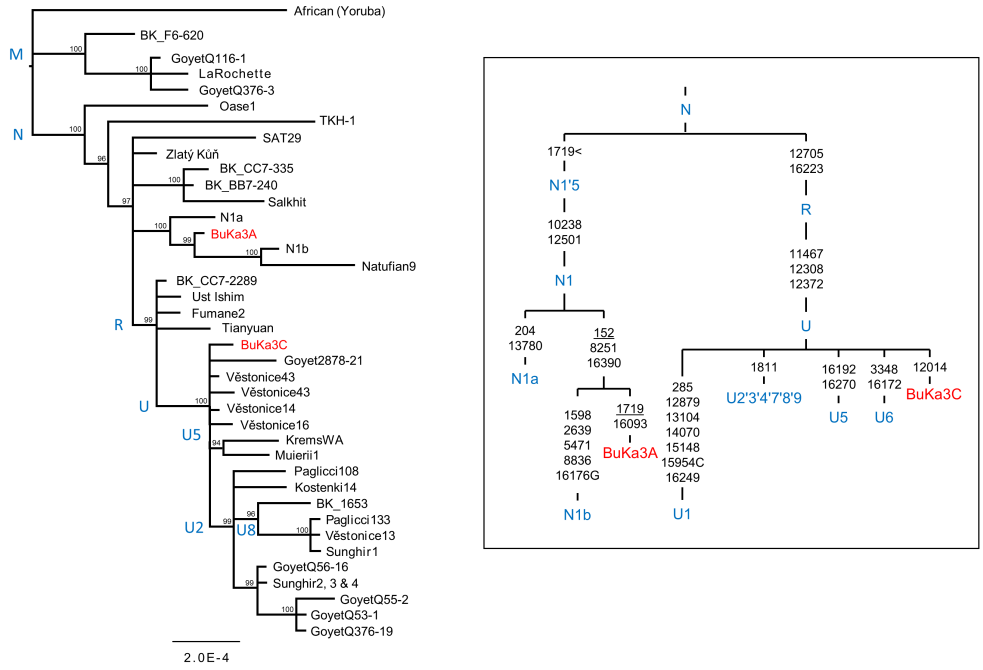


Figure 4. Genomic affinity to CHG ancestry relative to the two Buran Kaya III individuals and shared CHG/Zlatý Kůň-like ancestry. A, Allele sharing between pre-LGM individuals and CHG ancestry relative to BuKa3A and BuKa3C. Error bars = one standard error. **B,** Biplot showing the correspondance between CHG and Zlatý Kůň allele sharing amounts with EUP/MUP individuals relative to Sunghir3. Violet, Fournol cluster members; pale blue, Věstonice cluster members. Error bars = one standard error. Full results for A and B can be found in Table S6.



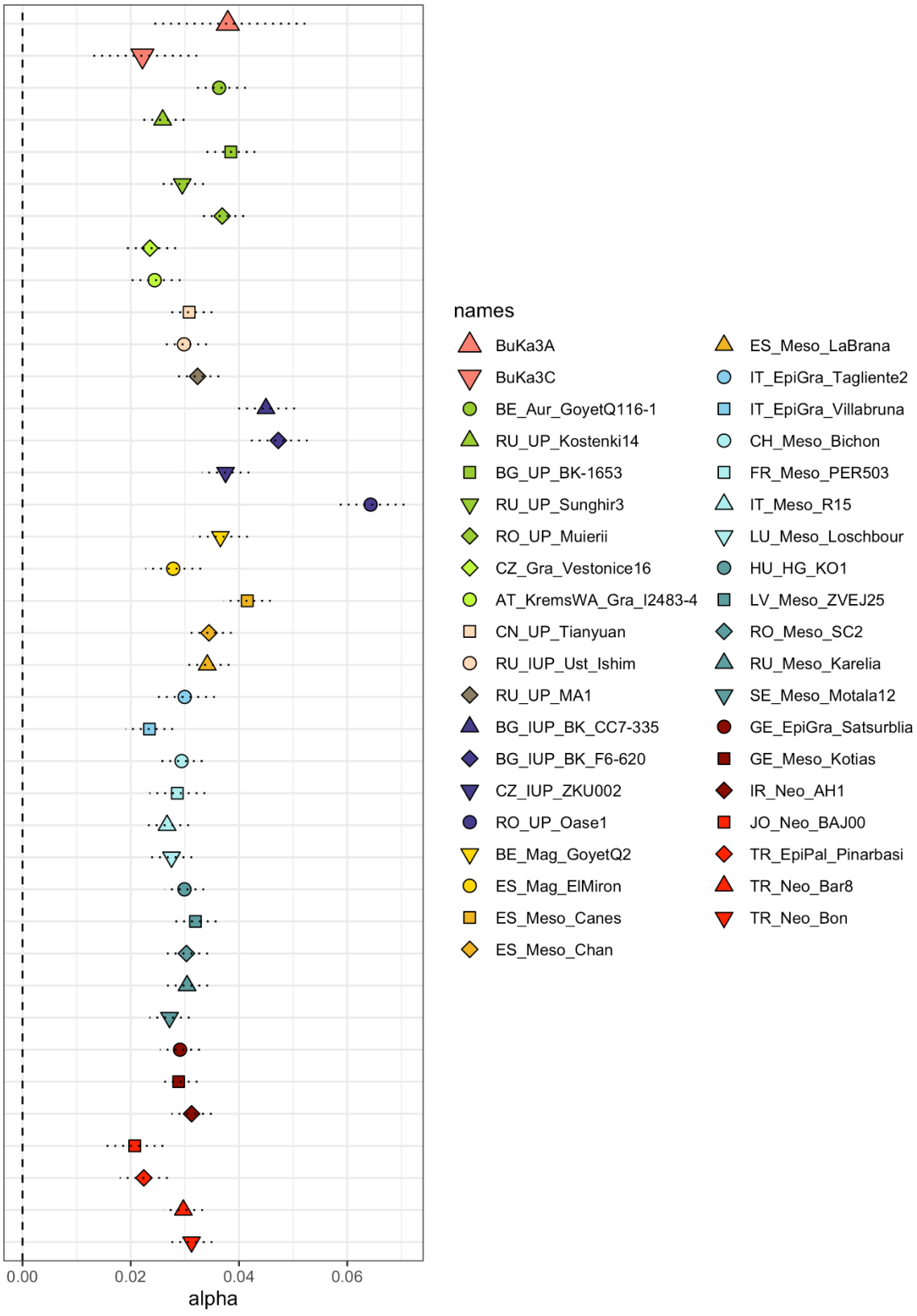
Extended Data Figure 1. Recalibration of comparative AMS 14C dates of Buran-Kaya III.

Buran-Kaya III (BuKa3) layers 6-1 and 6-2 with Upper Paleolithic samples. All dates were calibrated using the software OxCal v4.4.4 based on the IntCal20 calibration data set. Climatic events are represented by the orange band, the Campanian Ignimbrite (CI) volcanic super eruption event by the darker orange band, the Heinrich Event 4 by the lighter orange band.



Extended Data Figure 2. Bayesian phylogenetic tree of mitochondrial sequences.

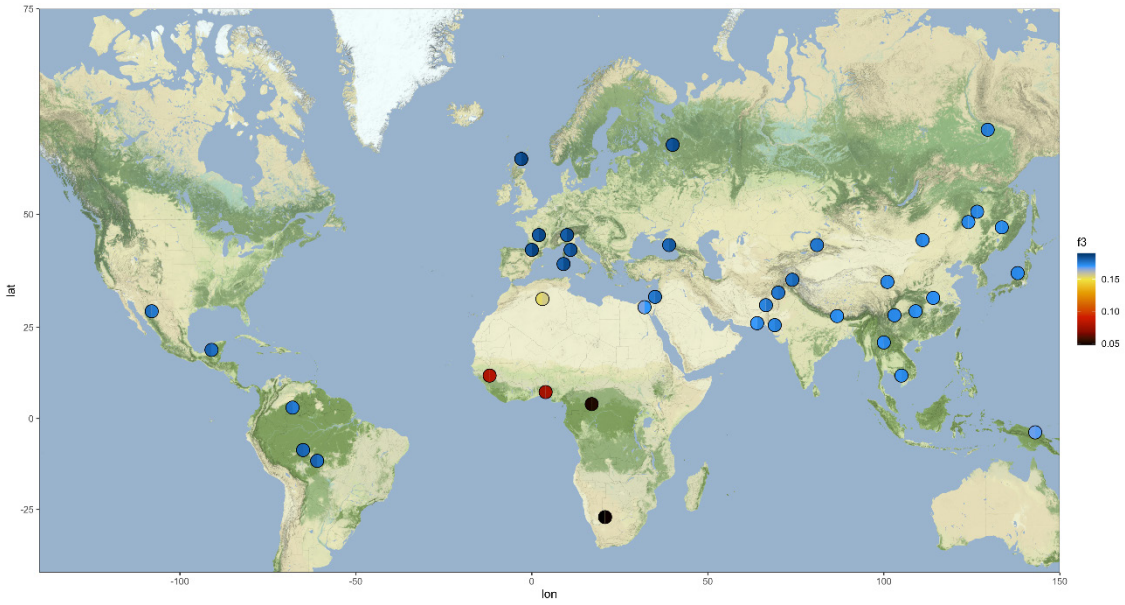
Excluding Hyper-Variable Regions, from Early and Mid-Upper Palaeolithic individuals including those from Buran-Kaya III (red). Posterior probability indicated at the nodes. Scale bar denotes substitutions per site. Sources for mitochondrial sequences are listed in Table S3. Additional ancient and canonical N sequences are included to give branching details. Inset gives detailed mutation information from mtphyl v.5.003 (https://sites.google.com/site/mtphyl/home) for both BuKa3 individuals. References for mitochondrial sequences are given in Table S11.



Extended Data Figure 3. Neanderthal ancestry calculated for ancient individuals.

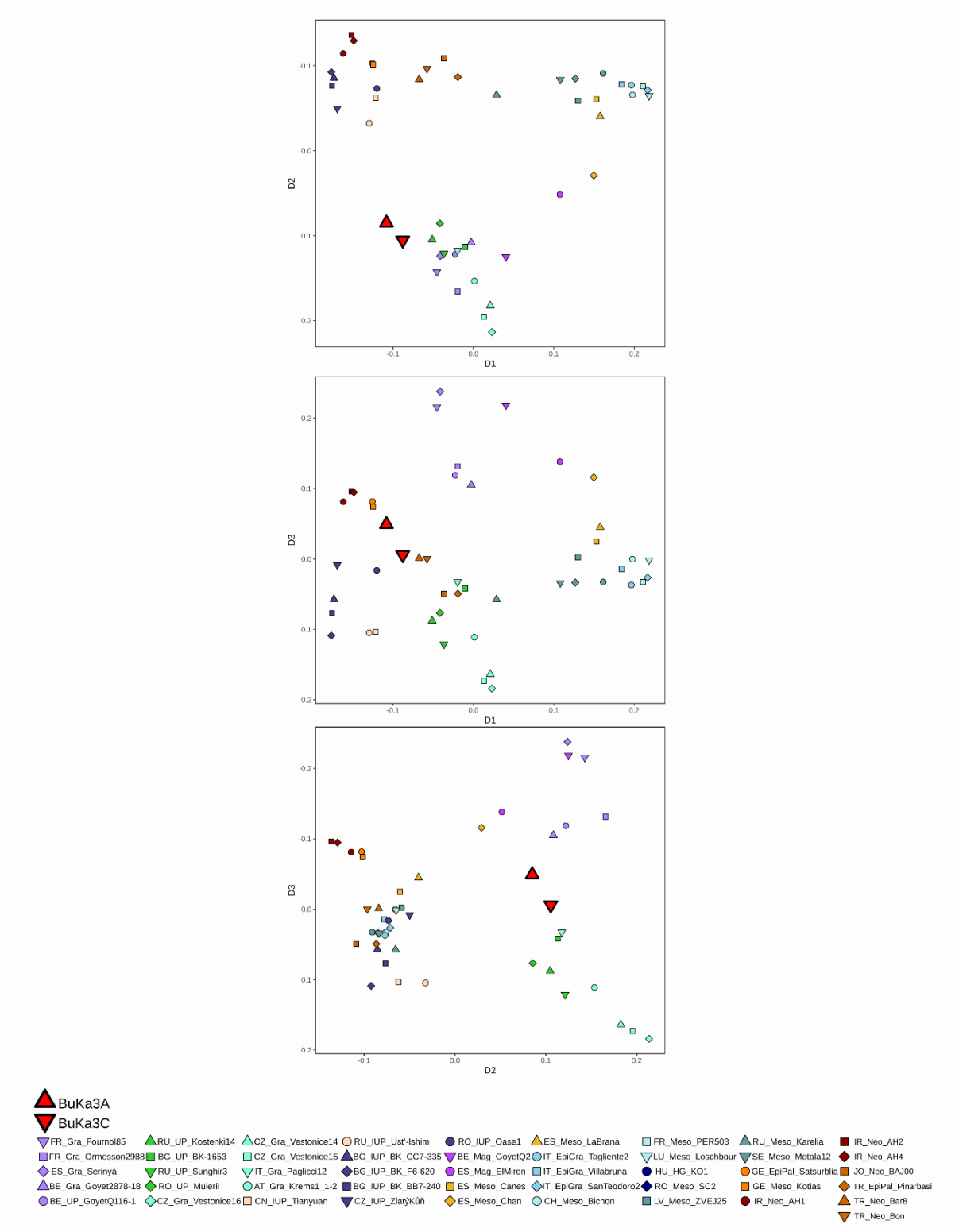
Neanderthal ancestry calculated using direct f4-ratio test[78]. α values calculated in the form of (Vindija33.19, Chimp; *test*, Mbuti)/(Vindija33.19, Chimp; Chagyrskaya, Mbuti). The error

bars represent ±1 standard errors (s.e.), which were calculated using weighted block jackknife and a block size of 5 Mb.



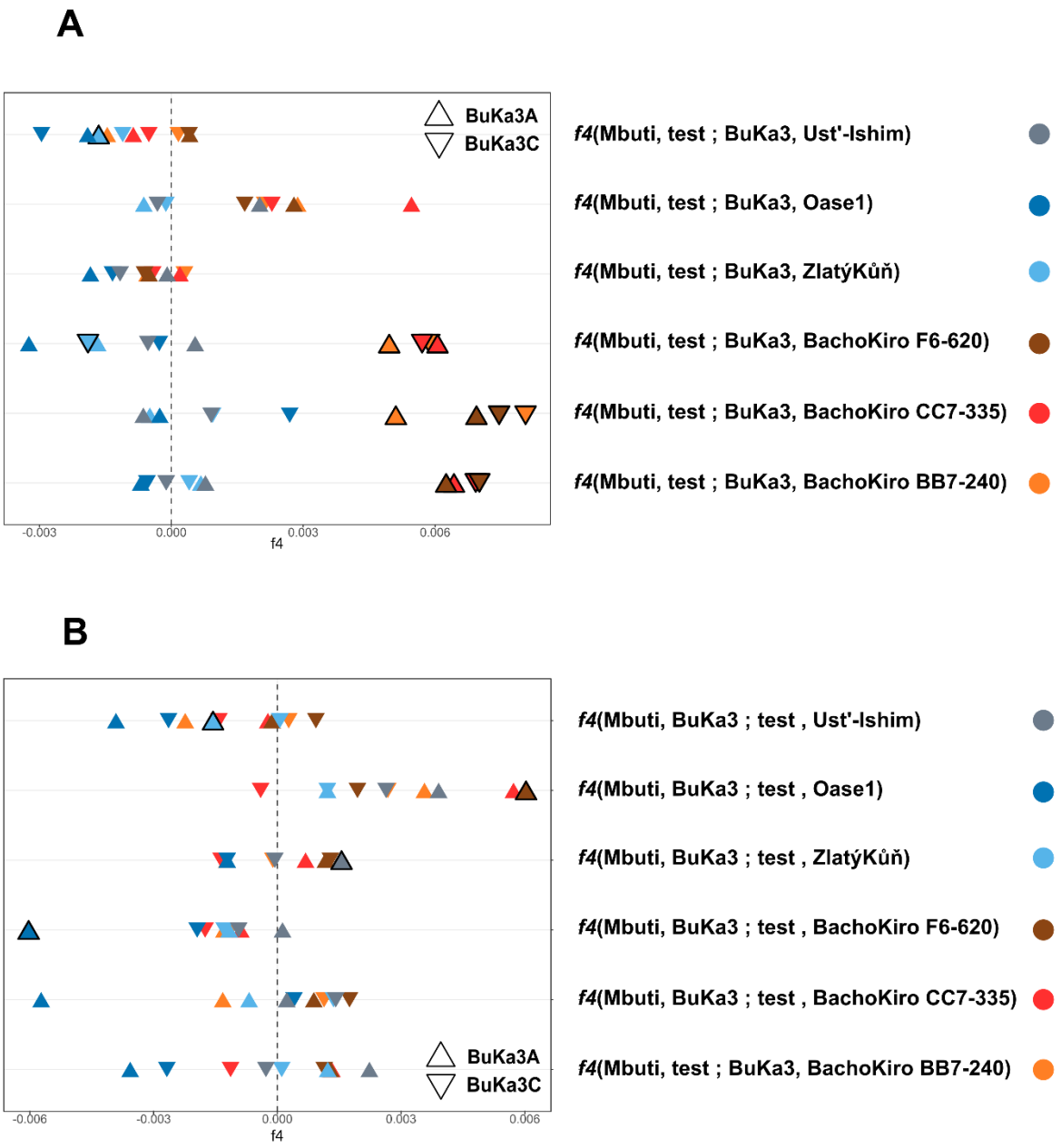
Extended Data Figure 4: Genetic affinities between Buran-Kaya III and present-day populations.

Heatmap showing the shared genetic drift calculated by *f3*-statistics in the form of *f3*(Mbuti; BuKa3, *modern*). Higher *f3* values indicate greater allele sharing between Buran-Kaya III (BuKa3) and populations from the HGDP dataset. The left half of the circles shows *f3* values tested with BuKa3C while the right side show those tested with BuKa3A. The map was created with ggmap[76] using Stamen Design map tiles[77].



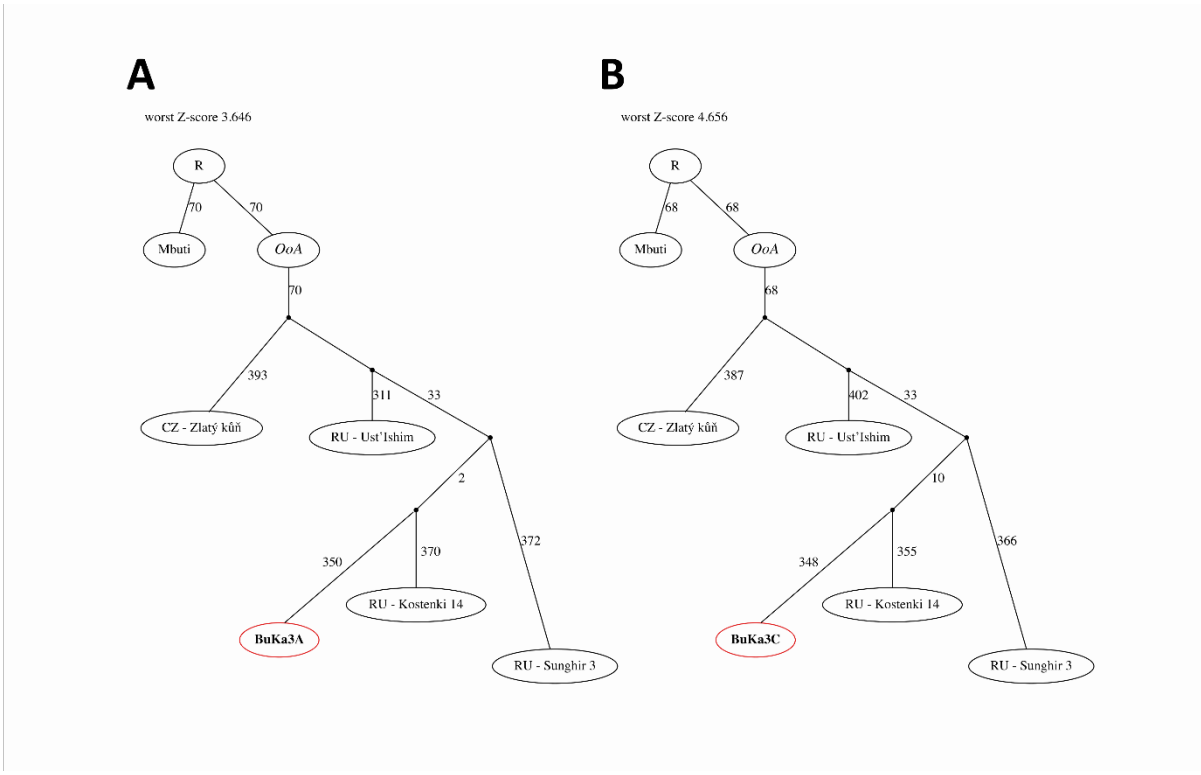
Extended Data Figure 5. 2D Multidimensional scaling (MDS).

MDS calculated using the pairwise genetic distance matrix coming from inverted *f3* values (1-*f3*) in the form of *f3*(Mbuti; *ancient1*, *ancient2*), for the first three dimensions representing 8, 6 and 5% of the total variance. See also Supplementary Video1 and Figure S2.



Extended Data Figure 6. Summary of f4 statistics involving BuKa3 and pre-CI genomes.

A. Results of *f4*(Mbuti, *pre-CI* ; BuKa3, *pre-CI*) **B.** Results of *f4*(Mbuti, BuKa3 ; *pre-CI*, *pre-CI*). Full results in Table S5.



Extended Data Figure 7. Modeling BuKa3 ancestry by qpGraph without admixture. Best-fitting qpGraph models of BuKa3A and BuKa3C without admixture (compare with Fig 3B). Buran-Kaya III individuals modeled with pre-LGM shotgun genomes using allsnps: NO parameters. The Z-score for the worst f4-statistic residuals is indicated.

References

1. Hajdinjak, M. *et al.* Initial Upper Palaeolithic humans in Europe had recent Neanderthal ancestry. *Nature* **592**, 253–257 (2021).
2. Slimak, L. *et al.* Modern human incursion into Neanderthal territories 54,000 years ago at Mandrin, France. *Sci. Adv.* **8**, eabj9496 (2022).
3. Fu, Q. *et al.* The genetic history of Ice Age Europe. *Nature* **534**, 200–205 (2016).
4. Prüfer, K. *et al.* A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat. Ecol. Evol.* **5**, 820–825 (2021).
5. Vallini, L. *et al.* Genetics and Material Culture Support Repeated Expansions into Paleolithic Eurasia from a Population Hub Out of Africa. *Genome Biol. Evol.* **14**, evac045 (2022).
6. Fu, Q. *et al.* An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524**, 216–219 (2015).

7. 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).
8. Fedele, F. *et al.* The Campanian Ignimbrite Factor: Towards a Reappraisal of the Middle to Upper Palaeolithic "Transition". in *Living under the Shadow: Cultural Impacts of Volcanic Eruptions* vol. 53 19–41 (2007).
9. Fitzsimmons, K. E., Hambach, U., Veres, D. & Iovita, R. The Campanian Ignimbrite Eruption: New Data on Volcanic Ash Dispersal and Its Potential Impact on Human Evolution. *PLoS ONE* **8**, e65839 (2013).
10. Black, B. A., Neely, R. R. & Manga, M. Campanian Ignimbrite volcanism, climate, and the final decline of the Neanderthals. *Geology* **43**, 411–414 (2015).
11. Heinrich, H. Origin and Consequences of Cyclic Ice Rafting in the Northeast Atlantic Ocean During the Past 130,000 Years. *Quat. Res.* **29**, 142–152 (1988).
12. Zilhão, J. Neandertals and moderns mixed, and it matters. *Evol. Anthropol. Issues News Rev.* **15**, 183–195 (2006).
13. Seguin-Orlando, A. *et al.* Genomic structure in Europeans dating back at least 36,200 years. *Science* **346**, 1113–1118 (2014).
14. Sikora, M. *et al.* Ancient genomes show social and reproductive behavior of early Upper Paleolithic foragers. *Science* **358**, 659–662 (2017).
15. Posth, C. *et al.* Palaeogenomics of Upper Palaeolithic to Neolithic European hunter-gatherers. *Nature* **615**, 117–126 (2023).
16. Prat, S. *et al.* The Oldest Anatomically Modern Humans from Far Southeast Europe: Direct Dating, Culture and Behavior. *PLoS ONE* **6**, e20834 (2011).
17. Péan, S. *et al.* The Middle to Upper Paleolithic Sequence of Buran-Kaya III (Crimea, Ukraine): New Stratigraphic, Paleoenvironmental, and Chronological Results. *Radiocarbon* **55**, 1454–1469 (2013).
18. Yanevich, A. Les occupations gravettiennes de Buran-Kaya III (Crimée) : contexte archéologique. *L'Anthropologie* **118**, 554–566 (2014).
19. Prat, S. *et al.* The First Anatomically Modern Humans from South-Eastern Europe. Contributions from the Buran-Kaya III Site (Crimea). *Bull. Mém. Société Anthropol. Paris* (2018) doi:10.3166/bmsap-2018-0032.
20. Hublin, J.-J. The modern human colonization of western Eurasia: when and where? *Quat. Sci. Rev.* **118**, 194–210 (2015).

21. Reynolds, N. & Green, C. Spatiotemporal modelling of radiocarbon dates using linear regression does not indicate a vector of demic dispersal associated with the earliest Gravettian assemblages in Europe. *J. Archaeol. Sci. Rep.* **27**, 101958 (2019).
22. Reynolds, N. Threading the Weft, Testing the Warp: Population Concepts and the European Upper Paleolithic Chronocultural Framework. in *Culture History and Convergent Evolution* (ed. Groucutt, H. S.) 187–212 (Springer International Publishing, 2020). doi:10.1007/978-3-030-46126-3_10.
23. Reynolds, N. The Gravettian is Dead: Against Equivocation and Reification in Chronocultural Studies of the Upper Palaeolithic. in *Les sociétés gravettiennes du Nord-Ouest européen : nouveaux sites, nouvelles données, nouvelles lectures / Gravettian societies in North-western Europe: new sites, new data, new readings* (eds. Touzé, O., Goutas, N., Salomon, H. & Noiret, P.) 309–321 (Presses Universitaires de Liège, 2021).
24. Demidenko, Y. E. Crimean Upper Paleolithic. in *Encyclopedia of Global Archaeology* (ed. Smith, C.) 1782–1791 (Springer New York, 2014). doi:10.1007/978-1-4419-0465-2_1864.
25. Cullen, V. L. *et al.* A revised AMS and tephra chronology for the Late Middle to Early Upper Paleolithic occupations of Ortvale Klde, Republic of Georgia. *J. Hum. Evol.* **151**, 102908 (2021).
26. Golovanova, L. V. *et al.* The Early Upper Paleolithic in the northern Caucasus (new data from Mezmaiskaya Cave, 1997 excavation). *Eurasian Prehistory* **4**, 43–78 (2012).
27. Bar-Yosef, O. *et al.* Dzudzuana: an Upper Palaeolithic cave site in the Caucasus foothills (Georgia). *Antiquity* **85**, 331–349 (2011).
28. Hoffecker, J. F. A New Framework for the Upper Paleolithic of Eastern Europe. (2012).
29. Hoffecker, J. F. & Holliday, V. T. Landscape Archaeology and the Dispersal of Modern Humans in Eastern Europe. in *The Upper Paleolithic of Northern Eurasia and America: Sites, Cultures, Traditions* 140–158 (Petersburg 2014).
30. Svoboda, J. A. The Gravettian on the Middle Danube. *PALEO Rev. Archéologie Préhistorique* **19**, 203–220 (2007).
31. Hoffecker, J. F. The early upper Paleolithic of eastern Europe reconsidered. *Evol. Anthropol. Issues News Rev.* **20**, 24–39 (2011).
32. Korlević, P. *et al.* Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. *BioTechniques* **59**, (2015).

33. Glocke, I. & Meyer, M. Extending the spectrum of DNA sequences retrieved from ancient bones and teeth. *Genome Res.* **27**, 1230–1237 (2017).
34. Renaud, G., Slon, V., Duggan, A. T. & Kelso, J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol.* **16**, (2015).
35. Weissensteiner, H. *et al.* Contamination detection in sequencing studies using the mitochondrial phylogeny. *Genome Res.* **31**, 309–316 (2021).
36. Skoglund, P., Storå, J., Götherström, A. & Jakobsson, M. Accurate sex identification of ancient human remains using DNA shotgun sequencing. *J. Archaeol. Sci.* **40**, 4477–4482 (2013).
37. Brunel, S. *et al.* Ancient genomes from present-day France unveil 7,000 years of its demographic history. *Proc. Natl. Acad. Sci.* **117**, 12791–12798 (2020).
38. Svensson, E. *et al.* Genome of Peştera Muierii skull shows high diversity and low mutational load in pre-glacial Europe. *Curr. Biol.* **31**, 2973-2983.e9 (2021).
39. Hublin, J.-J. *et al.* Initial Upper Palaeolithic Homo sapiens from Bacho Kiro Cave, Bulgaria. *Nature* **581**, 299–302 (2020).
40. Lazaridis, I. *et al.* Genomic insights into the origin of farming in the ancient Near East. *Nature* **536**, 419–424 (2016).
41. Higham, T. *et al.* The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature* **512**, 306–309 (2014).
42. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neandertal introgression. *Proc. Natl. Acad. Sci.* **116**, 1639–1644 (2019).
43. Bergström, A. *et al.* Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020).
44. Fu, Q. *et al.* DNA analysis of an early modern human from Tianyuan Cave, China. *Proc. Natl. Acad. Sci.* **110**, 2223–2227 (2013).
45. Teschler-Nicola, M. *et al.* Ancient DNA reveals monozygotic newborn twins from the Upper Palaeolithic. *Commun. Biol.* **3**, 650 (2020).
46. Jones, E. R. *et al.* Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nat. Commun.* **6**, (2015).
47. Jeong, C. *et al.* The genetic history of admixture across inner Eurasia. *Nat. Ecol. Evol.* **3**, 966–976 (2019).

48. Lazaridis, I. *et al. Paleolithic DNA from the Caucasus reveals core of West Eurasian ancestry*. http://biorxiv.org/lookup/doi/10.1101/423079 (2018) doi:10.1101/423079.

49. Allentoft, M. E. *et al.* Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).

50. Marín-Arroyo, A. B., Terlato, G., Vidal-Cordasco, M. & Peresani, M. Subsistence of early anatomically modern humans in Europe as evidenced in the Protoaurignacian occupations of Fumane Cave, Italy. *Sci. Rep.* **13**, 3788 (2023).

51. Fedele, F. G., Giaccio, B., Isaia, R. & Orsi, G. The Campanian Ignimbrite Eruption, Heinrich Event 4, and palaeolithic change in Europe: A high-resolution investigation. in *Geophysical Monograph Series* (eds. Robock, A. & Oppenheimer, C.) vol. 139 301–325 (American Geophysical Union, 2003).

52. Hoffecker, J. F. *Desolate landscapes: Ice-Age settlement in Eastern Europe*. (Rutgers University Press, 2002).

53. Marti, A., Folch, A., Costa, A. & Engwell, S. Reconstructing the plinian and co-ignimbrite sources of large volcanic eruptions: A novel approach for the Campanian Ignimbrite. *Sci. Rep.* **6**, 21220 (2016).

54. Reimer, P. J. *et al.* The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).

55. Bennett, E. A. *et al.* Library construction for ancient genomics: single strand or double strand? *BioTechniques* **56**, 289–298 (2014).

56. Gansauge, M.-T. *et al.* Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Res.* **45**, e79 (2017).

57. Gansauge, M.-T. & Meyer, M. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748 (2013).

58. Massilani, D. *et al.* Past climate changes, population dynamics and the origin of Bison in Europe. *BMC Biol.* **14**, 93 (2016).

59. Renaud, G., Stenzel, U. & Kelso, J. leeHom: adaptor trimming and merging for Illumina sequencing reads. *Nucleic Acids Res.* **42**, e141–e141 (2014).

60. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–60 (2009).

61. van der Valk, T. *et al.* Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* **591**, 265–269 (2021).

62. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–9 (2009).

63. de Filippo, C., Meyer, M. & Prüfer, K. Quantifying and reducing spurious alignments for the analysis of ultra-short ancient DNA sequences. *BMC Biol.* **16**, (2018).

64. 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).

65. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* btv566 (2015) doi:10.1093/bioinformatics/btv566.

66. Weissensteiner, H. *et al.* HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. *Nucleic Acids Res.* **44**, W58–W63 (2016).

67. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).

68. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9**, 772 (2012).

69. Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing Ancient DNA Sequences into Reference Phylogenies. *Mol. Biol. Evol.* **39**, msac017 (2022).

70. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

71. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

72. Rohland, N. *et al.* Three assays for in-solution enrichment of ancient human DNA at more than a million SNPs. *Genome Res.* **32**, 2068–2078 (2022).

73. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).

74. Patterson, N. *et al.* Ancient Admixture in Human History. *Genetics* **192**, 1065–1093 (2012).

75. Maier, R., Flegontov, P., Flegontova, O., Changmai, P. & Reich, D. *On the limits of fitting complex models of population history to genetic data*. http://biorxiv.org/lookup/doi/10.1101/2022.05.08.491072 (2022) doi:10.1101/2022.05.08.491072.

76. Kahle, D. & Wickham, H. ggmap: Spatial Visualization with ggplot2. *R J.* **5**, 144–161 (2013).

77. Map tiles by Stamen Design, under CC BY 4.0. Data by OpenStreetMap, under ODbL. http://maps.stamen.com/#terrain/12/37.7706/-122.3782.

78. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. The limits of long-term selection against Neandertal introgression. (2018) doi:10.1101/362566.